



**HAL**  
open science

# Individual-based models in adaptive dynamics, asymptotic behaviour and canonical equation: the case of small and frequent mutations

Vincent Hass

► **To cite this version:**

Vincent Hass. Individual-based models in adaptive dynamics, asymptotic behaviour and canonical equation: the case of small and frequent mutations. Mathematics [math]. Université de Lorraine, 2023. English. NNT: 2023LORR0165 . tel-04295339

**HAL Id: tel-04295339**

**<https://hal.univ-lorraine.fr/tel-04295339>**

Submitted on 20 Nov 2023

**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.



**UNIVERSITÉ  
DE LORRAINE**

**BIBLIOTHÈQUES  
UNIVERSITAIRES**

## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)  
*(Cette adresse ne permet pas de contacter les auteurs)*

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

# Modèles individu-centrés en dynamiques adaptatives, comportement asymptotique et équation canonique : le cas des mutations petites et fréquentes

## THÈSE

présentée et soutenue publiquement le 26 Septembre 2023

pour l'obtention du

**Doctorat de l'Université de Lorraine**

(Mention Mathématiques Appliquées)

par

Vincent HASS

### Composition du jury

<i>Président :</i>	Viet Chi TRAN	Prof.	Université Gustave Eiffel
<i>Rapporteurs :</i>	Jochen BLATH	Prof.	Goethe-Universität Frankfurt
	Michèle THIEULLEN	MCF-HDR	Sorbonne Université
<i>Examineurs :</i>	Sylvain BILLIARD	MCF-HDR (Biologie)	Université de Lille
	Pascal MOYAL	Prof.	Université de Lorraine
	Léa POPOVIC	Prof.	Concordia University of Montréal
<i>Directeur :</i>	Nicolas CHAMPAGNAT	DR	Inria Nancy Grand-Est

Mis en page avec la classe thesul.

## Remerciements

*“Nous transmettons ce que nous savons, mais surtout, nous transmettons ce que nous sommes.”*

Christophe ANDRÉ

Cette thèse représente près de cinq années de ma vie. Aussi, j’aimerais remercier toutes les personnes qui, d’une manière ou d’une autre, ont permis l’aboutissement de ce travail et ont rendu cette période aussi agréable.

Je voudrais tout d’abord remercier Nicolas CHAMPAGNAT pour m’avoir accompagné durant cette aventure de thèse et même bien avant ! Merci pour m’avoir aussi bien conseillé sur les choix de Masters recherche parisiens durant ma prépa agrég à Nancy. Merci de m’avoir recommandé pour me permettre d’intégrer la prestigieuse formation de *Mathématiques pour les Sciences du Vivant* portée, en partie, par le laboratoire de mathématiques d’Orsay et l’École Polytechnique. Merci d’avoir guidé mes premiers pas en recherche, pour la confiance que tu m’as accordé, pour m’avoir partagé tes conseils et ton expérience avec enthousiasme et pour m’avoir fait découvrir de belles mathématiques. J’ai appris beaucoup à tes côtés, tant sur le plan scientifique que sur le plan humain et je t’en suis très reconnaissant : un grand merci pour toutes ces années.

Je tiens aussi à remercier Michèle THIEULLEN et Jochen BLATH pour m’avoir fait l’honneur de rapporter ma thèse. Merci pour l’intérêt que vous avez porté à mon travail, pour vos questions et vos commentaires très enrichissants qui m’ont permis d’améliorer la qualité de ce manuscrit. Merci également à Sylvain BILLIARD, Pascal MOYAL, Léa POPOVIC et Viet Chi TRAN d’avoir accepté d’être dans mon jury. J’ai une pensée particulière pour Chi et Pascal qui ont fait partie de mes comités de suivi de thèse, et suivent donc l’avancée de mes travaux depuis la première année.

Je tiens à remercier également l’Inria Nancy Grand-Est et l’Université de Lorraine pour avoir financé ces cinq années de recherche. Merci également à l’Institut Élie Cartan de Lorraine où il y fait bon travailler et à la Chaire de Modélisation Mathématique et Biodiversité (MMB) pour les multiples rencontres entre mathématiciens et biologistes notamment à Aussois.

Je remercie toutes les personnes qui m’ont permis de m’épanouir en tant que jeune enseignant-chercheur en mathématiques. Mes premiers remerciements chaleureux vont à Hervé LECOMPTE ! “*Verstanden ?*” C’est grâce à toi que cette incroyable aventure mathématiques a pu voir le jour et qui m’a transmis ce goût pour l’enseignement : je t’en suis très reconnaissant. Merci pour tes cours à la fois limpides et exigeants, animés avec beaucoup de plaisir et d’amusement. Merci de m’avoir fait découvrir le “*Super Mario de la calculatrice*” : j’en rigole encore ! Merci pour ta bienveillance, ta bonne humeur, ta grande sympathie et ces nombreux fous rires : merci beaucoup pour ces belles années lycée !

J’ai une pensée particulière pour Pierre-Adrien TAHAY pour m’avoir parlé avec enthousiasme de la Classe Préparatoire Universitaire (CPU) Maths-Physique et pour Didier SCHMITT qui m’a permis d’intégrer cette formation à Nancy. Un grand merci à Anne DE ROTON qui a été ma boussole, ma référence pour mon entrée dans cette nouvelle vie d’étudiant : merci pour ton infinie gentillesse ! Quelle joie, quand j’ai appris durant mon doctorat, que je partageais avec

toi le cours de “Compléments d’Analyse” à la faculté des sciences ! Je remercie ensuite tous les enseignants qui ont entretenu mon goût pour les mathématiques et à poursuivre dans cette voie. Ne pouvant garantir que la liste est exhaustive, je présente toutes mes excuses aux personnes que je risque éventuellement d’oublier. Merci à Vincent BANSAYE, Wolfgang BERTRAM, Pierre-Emmanuel CHAPUT, François CHARGOIS, Philippe CHASSAING, Julien CHIQUET, Lucilla CORRIAS, Benoît DANIEL, Anne DE ROTON, Aurélia DESHAYES, David DOS SANTOS FERREIRA, Bruno DUCHESNE, Sylvain FAURE, Lucas FRESSE, Olivier GARET, Françoise GÉANDIER, Anne GÉGOUT-PETIT, Alain GENESTIER, Jean-Sébastien GIET, Christophe GIRAUD, Carl GRAHAM, Jean-François GROSJEAN, Caroline GRUSON, Oussama HIJAZI, Khalid KOUFANY, Estelle KUHN, Aline KURTZMANN, Anna KOWALSKA-CHASSAING, Béatrice LAROCHE, Vladimir LATOCHA, Émilie LEBARBIER, Manfred MADRITSCH, Régine MARCHAND, Irène MARCOVICI, Tristan MARY-HUARD, Julien MAUBON, Bertand MAURY, Séraphin MEFIRE, Damien MÉGY, Sylvie MÉLÉARD, Michel MÉO, Anne-Hélène MONSORO-BURQ, Yann PETOT, Marie-Anne POURSAT, Frédéric ROBERT, Stéphane ROBIN, Hans Henrik RUGH, Thomas STOLL, Gérald TENENBAUM, Matei TOMA, Jérémie UNTERBERGER et Amandine VÉBER. Mention spéciale à Régine pour m’avoir fait découvrir la beauté des probabilités et m’avoir tant inspiré par tes méthodes pédagogiques : merci beaucoup ! Autre mention spéciale à Sylvie pour m’avoir initié à la finesse des processus stochastiques, pour ton investissement sans faille et les rencontres organisées par la Chaire MMB que tu portes. Il s’avère que j’ai eu la chance de continuer à côtoyer beaucoup d’entre eux encore pendant ma thèse, et j’ai été très ému par leurs réguliers encouragements et marques de sympathie.

Je souhaite aussi remercier les collègues, responsables de cours, avec qui j’ai eu le privilège de travailler et qui m’ont permis d’effectuer mes enseignements dans les meilleures conditions. Un grand merci à Bernadetta ADDIS, Anne DE ROTON, Antoine HENROT, Antoine LEJAY, Ilaria LUCARDESI, Rémi PEYRE, Bruno PINÇON, Didier SCHMITT et Denis VILLEMONAIS. Un grand merci aux nombreux chargés de TD avec qui j’ai pris plaisir à discuter et à faire évoluer mes pratiques d’enseignement. Sur le même plan, j’aimerais également remercier toutes les personnes qui rendent ce laboratoire particulièrement accueillant et nos conditions de travail exceptionnelles, notamment Nathalie BENITO, Isabelle BLANCHARD, Élodie CUNAT, Emmanuelle DESCHAMPS, Virginie LAMOUREUX, Laurence QUIROT, Paola SCHNEIDER toujours là pour nous aider dans les démarches administratives et logistiques, Didier GEMMERLÉ qui m’a toujours dépanné avec ma machine et le babycluster, Souad BOUTAGUERMOUCHET, Sabrina FERRY et Aurore TRANCHINA, les chaleureux membres de l’équipe de probabilités et statistiques et de TOSCA/BIGS/SIMBA. Je remercie également Benjamin, Clémence, Jimmy, Nicolas, Rémi, Pierre-Adrien et Tom pour avoir co-géré avec moi le séminaire des doctorants de l’IECL. Une mention spéciale à Pierre-Adrien pour m’avoir guidé au mieux dans les méandres administratives de l’Éducation Nationale : merci pour tous tes conseils et d’avoir été cette référence vers qui me tourner !

J’ai une pensée affectueuse pour Mike pour avoir animé avec brio cette formation d’anglais à l’Inria avec Armelle, Jacqueline et Sabrina. Merci également au groupe des *Jorange* pour avoir rendu nos diverses formations doctorales plus rigolotes et pour nos apéros dinatoires.

Je tiens également à remercier celles et ceux qui m’ont entouré pendant mes années de doctorat et d’ATER aux Mines de Nancy. Un grand merci à mes co-bureau : Fatma, Paul, Rodolphe, Thomas et Valentin. Un merci particulier à Valentin pour ta grande gentillesse, ta bienveillance et nos nombreuses discussions : cela a constitué pour moi un soutien et un réconfort. Merci au joyeux groupe avec qui on allait manger les midis à la cantine de l’Inria : merci pour cette convivialité, nos fous rires et discussions enrichissantes. Mention spéciale pour Rodolphe qui nous

envoyait un petit message sur la mailing list pour nous réunir le midi même quand lui-même n'était pas présent ! Merci également à Coralie pour ses nombreux conseils, sa bonne humeur et d'avoir été présente pour moi quand certaines périodes étaient moins faciles pour moi. Merci également à Sara pour toutes ces petites discussions dans les couloirs du 3<sup>e</sup> étage, ton soutien, tes petits mots gentils et ta sympathie. Je tiens également à te remercier Virgile, tout d'abord pour ton Mouvement Brownie du jeudi matin, nos discussions en pâtisserie et surtout ta manière de voir les choses avec tant d'émerveillement et de bienveillance : merci pour ces moments d'apaisement et de réconfort.

Je souhaite aussi remercier les doctorants/post-doctorants avec qui je suis parti en conférence et celles et ceux avec qui j'ai eu la chance de créer du lien. Je pense notamment à Apolline, Édouard, Émilie, Florian, Gaëtan, Guillaume, Gwenaël, Josué, Mathilde, Maxime, Nicolàs, Pierre-Adrien, Ségolène, Ulysse et Yohan.

Un grand merci également à Jean-Baptiste, Philippe et Tristan pour nos nombreuses séances de badminton à Aiguillettes et Brabois avec Arnaud, Richard et Thierry. Un double merci à Jean-Baptiste et Philippe pour nos courses à pied, notamment celle de la Saint Nicolas et pour m'avoir initié au trail sur les plateaux de Maxéville.

Je tiens également à remercier Christophe, Clémence, Coralie, Dimitry, Éloïse, Johann, Pierre, Pierre-Adrien, Rémi, Rodolphe, Thomas, . . . pour nos soirées jeux. Merci Coralie pour nous avoir tant fait rire avec ta fusée et ton dentiste au Pictionary en ligne !

Enfin, mention spéciale à tous les doctorants qui ont donné vie à nos pauses : Anouk, Christophe, Clémence, David, Dimitry, Fabien, Florian, Hassan, Ibtissem, Iury, Jérémy, Jimmy, Jocelyn, Johann, Marco, Nicolas, Nicolàs, Nassim, Philippe, Pierre, Pierre-Adrien, Robin, Rodolphe, Thomas, Tom, Valentin, Virgile, Yann, Youssef, . . .

J'ai une pensée affectueuse pour mes étudiants de la faculté des sciences, de Télécom Nancy et de l'École des Mines avec qui j'ai eu beaucoup de plaisir à faire cours. Mention spéciale pour mes étudiants de la filière FIGIM de l'École des Mines que j'ai pu accompagner un peu plus longtemps et avec qui j'ai passé de bons moments.

Je remercie mes amis qui m'accompagnent pour certains depuis quelques temps déjà. Un grand merci à Alexis, Camille, Claire, Clémence, Évelyne, Gauthier, Grâce, Hélène, Ju', Juju, P'tit Juju, Matthieu, M'sieur Clément, Poti, Valentin et Valou pour votre fidélité. C'est toujours un plaisir de se retrouver.

Je tiens également à remercier ma belle famille pour leur accueil et leur soutien.

Enfin, je tiens à remercier de tout cœur les membres de ma famille pour leur soutien et de m'avoir toujours encouragé. Un grand merci à toi Shirley, ma gaïsböck préférée, pour nos franches rigolades, nos souvenirs passés, nos bêtises, nos heures au téléphone à refaire le monde, à en débattre, à mieux le comprendre et à en plaisanter ! Merci pour ton écoute bienveillante, de m'aider à relativiser et à voir les choses sous un meilleur angle. Merci d'être là ! Merci à mes sœurs, Béatrice et Nathalie pour tous nos moments de qualité passés ensemble et d'avoir toujours été au petit soin pour moi. Un grand merci à Camille et à Lana pour tous ces moments de rigolades, de jeux, de balades et bien sûr des bêtises qu'on peut raconter ! Un grand merci à toi aussi mémère pour toutes ces petites bulles d'évasion passées à la maison.

Un grand merci à toi Papa, pour ton calme, ta persévérance et ton humour, et qui m'accompagne toujours par la pensée. Un ÉNORME merci à toi Maman, pour ton soutien sans faille, ta disponibilité et ton amour inconditionnel depuis toujours : tu es pour beaucoup dans les différents succès que j'ai pu connaître dans mes études. Ce travail de thèse est en partie le tien aussi. Merci d'être une maman aussi incroyable ! Un énorme merci à vous deux pour l'éducation

que j'ai reçue et les valeurs que vous m'avez transmises.

Mes derniers remerciements sont pour toi Justine. La vie m'a donné la chance de te rencontrer, le temps m'a permis de mieux te connaître et chaque jour passé avec toi est un cadeau très précieux. Merci pour ta présence, ta bienveillance, ton soutien, ton écoute, ton énergie positive, tes bêtises, d'être solaire, d'être si inspirante, ... : merci d'être toi mon amour ! Merci d'avoir créé un NOUS !

Un grand merci à vous tous !



*Je dédie cette thèse  
à ma maman,  
à mon papa,  
à ma mémère.*



# Table of contents

<b>Une brève introduction générale</b>	<b>xi</b>
1 État de l'art biologique en dynamiques adaptatives . . . . .	xi
2 État de l'art mathématiques en dynamiques adaptatives . . . . .	xiv
3 Mes contributions . . . . .	xvii
<b>A brief general introduction</b>	<b>xxi</b>
1 Biological state of the art in adaptive dynamics . . . . .	xxi
2 Mathematical state of the art in adaptive dynamics . . . . .	xxiv
3 My contributions . . . . .	xxvii
<b>1 A detailed introduction</b>	<b>1</b>
1.1 Adaptive dynamics . . . . .	2
1.1.1 Darwinian evolutionary biology . . . . .	3
1.1.2 Biological context of adaptive dynamics . . . . .	5
1.1.3 Mathematical context . . . . .	14
1.1.4 Canonical Equation . . . . .	22
1.2 Mathematical theories . . . . .	42
1.2.1 FLEMING-VIOT processes . . . . .	42
1.2.2 Slow-fast asymptotic analysis . . . . .	56
1.3 My contributions . . . . .	63
1.3.1 From Chapter 2 . . . . .	63
1.3.2 From Chapter 3 . . . . .	70
1.3.3 Prospects . . . . .	78
<b>2 Existence, uniqueness and ergodicity for the centered Fleming-Viot process</b>	<b>85</b>
2.1 Introduction . . . . .	86
2.2 Existence for the centered FLEMING-VIOT process . . . . .	89
2.2.1 Centered FLEMING-VIOT martingale problem and main result . . . . .	89

2.2.2	Sketch of proof of Theorem 2.2.3 . . . . .	91
2.2.3	Some extensions to the centered FLEMING-VIOT martingale problem . . .	93
2.2.4	Some properties of the centered FLEMING-VIOT process . . . . .	98
2.3	Uniqueness for the centered FLEMING-VIOT process . . . . .	102
2.3.1	Main result . . . . .	102
2.3.2	Notations and outline of the uniqueness proof . . . . .	102
2.3.3	Proof of Theorem 2.3.1 from Theorem 2.3.4 and Lemma 2.3.5 . . . . .	105
2.4	Ergodicity for the centered FLEMING-VIOT process . . . . .	106
2.4.1	MORAN's models and FLEMING-VIOT's processes . . . . .	106
2.4.2	Backward construction of MORAN's process and KINGMAN's coalescent .	110
2.4.3	Centered variables and centering effects . . . . .	112
2.4.4	Coupling arguments with two distinct initial conditions . . . . .	113
2.4.5	Proof of Theorem 2.4.1 . . . . .	115
2.4.6	Characterisation of the invariant probability measure . . . . .	115
2.5	Proof of Theorem 2.2.3 . . . . .	122
2.5.1	DOOB's semi-martingale decomposition of $\langle \text{id}, Y_t - Y_s \rangle, s \leq t$ . . . . .	122
2.5.2	DOOB's semi-martingale decomposition of $\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_s \rangle}, Y_t - Y_s \rangle, s \leq t,$ $j \in \{0, 1, 2\}$ . . . . .	123
2.5.3	Expressions of the terms of (2.11) involving $F'$ . . . . .	124
2.5.4	Expressions of terms of (2.11) involving $F''$ . . . . .	127
2.5.5	Error terms . . . . .	128
2.5.6	Extension to test functions $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ and $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . . . . .	129
2.5.7	$L^2$ -martingale and quadratic variation . . . . .	129
2.5.8	Technical result for Lemma 2.5.2 . . . . .	130
2.5.9	Proof of Lemma 2.5.2 . . . . .	132
2.6	Proof of the results of Section 2.3 . . . . .	133
2.6.1	Study of a semi-group . . . . .	133
2.6.2	Proof of Lemma 2.3.5 . . . . .	141
2.6.3	Proof of Theorem 2.3.4 . . . . .	144
<b>3</b>	<b>Convergence of IBMs with small and frequent mutations to the CEAD</b>	<b>149</b>
3.1	Introduction . . . . .	150
3.2	A trait structured IBM . . . . .	154
3.2.1	Parameters and assumptions of the model . . . . .	155
3.2.2	Main result . . . . .	156
3.2.3	Sketch and outline of the proof . . . . .	157
3.3	Infinitesimal generator approximation . . . . .	162

---

3.3.1	Generators and martingales . . . . .	163
3.3.2	Asymptotic expansions . . . . .	164
3.3.3	Generators and martingales in the case of polynomials in $\mu$ . . . . .	167
3.3.4	Proof of Lemma 3.3.4 . . . . .	171
3.4	Moments estimates . . . . .	173
3.4.1	Estimates of the moment of order 6 . . . . .	174
3.4.2	Some inequalities on the moment of order 2 . . . . .	177
3.4.3	Convergence to $+\infty$ of the stopping time $\tau^K$ and support concentration property . . . . .	178
3.4.4	Proof of Lemma 3.4.9 . . . . .	179
3.5	Tightness on the torus . . . . .	188
3.6	Characterisation of the occupation measure limit on the torus . . . . .	194
3.6.1	Proof of Proposition 3.6.3 . . . . .	195
3.6.2	Proof of Lemma 3.6.4 . . . . .	197
3.7	Characterisation of the limiting values of the slow component . . . . .	205
3.7.1	Convergence of the slow component on the torus . . . . .	205
3.7.2	End of the proof of Theorem 3.2.1 . . . . .	205
	<b>Annexes</b>	<b>207</b>
	<b>A Technical lemmas involved in the existence proof</b>	<b>207</b>
A.1	Approximation lemma . . . . .	207
A.2	Lemma of convergence . . . . .	209
A.3	Control of error terms . . . . .	209
	<b>B R code for IBM simulation</b>	<b>211</b>
B.1	R-code for the IBM for constant population with swap events . . . . .	211
B.2	R-code for the IBM with fluctuate population size . . . . .	213
	<b>Bibliography</b>	<b>217</b>



# Une brève introduction générale

Cette introduction générale de ma thèse se structure suivant trois grands axes. En section 1, je présente l'état de l'art biologique des dynamiques adaptatives en se focalisant sur les hypothèses biologiques qui régissent son cadre, ses succès et tout spécialement celui de l'équation canonique. En section 2, je présente les deux principales approches mathématiques étudiant l'équation canonique des dynamiques adaptatives ainsi que l'objectif de cette thèse. Enfin, en section 3, je présente les contributions de cette thèse ainsi que l'organisation générale du manuscrit.

## 1 État de l'art biologique en dynamiques adaptatives

Cette thèse, qui se place à l'interface entre mathématiques et biologie de l'évolution, part de motivations biologiques en dynamiques adaptatives. Ce domaine récent de la biologie théorique a été développé dans la seconde moitié des années 1990 par METZ ET AL [171], DIECKMANN et LAW [72] et GERITZ ET AL [112, 114]. Cette théorie étudie les liens entre *écologie* et *évolution*. On entend par écologie, l'ensemble des interactions entre les individus et leur environnement en incluant les autres individus de la population. L'évolution est à prendre au sens darwinien, c'est-à-dire reposant sur trois mécanismes de base. Premièrement, l'*hérédité* qui permet la transmission des caractéristiques phénotypiques individuelles d'une génération à l'autre. Deuxièmement, une *source de variation* parmi ces caractéristiques : dans les modèles étudiés dans cette thèse, cette source de variation provient exclusivement de mutations. Enfin, un *mécanisme de sélection (naturelle)* pouvant résulter de la compétition inter ou intra espèces : les individus les mieux adaptés à leur environnement sont plus susceptibles de survivre et de se reproduire et donc de transmettre leurs caractéristiques aux générations futures que ceux qui le sont moins.

**Hypothèses biologiques.** Pour fixer le cadre des dynamiques adaptatives, les biologistes font quatre hypothèses biologiques simplificatrices.

- (1) **Population asexuée.** Cette hypothèse a pour objectif de simplifier le schéma de reproduction des individus en le supposant clonal : une cellule (ou une plante, certains animaux, certains insectes, ...) mère donne deux cellules filles avec le même patrimoine génétique. Dans une population asexuée, les mutations qui sont des altérations permanentes du génome, sont considérées en dynamiques adaptatives comme agissant directement au niveau du phénotype. Pour être visible, i.e. ayant un effet phénotypique, les mutations doivent affecter une partie codante du génome et altérer les propriétés chimiques des protéines encodées [3]. Il est biologiquement raisonnable de supposer que les mutations sont rares.
- (2) **Mutations rares.** Cette hypothèse est fondamentale pour introduire une séparation d'échelle de temps entre une échelle écologique, rapide, correspondant aux interactions entre les individus et une échelle évolutive, plus lente, correspondant à l'arrivée des mutations. Puisque les mutations sont rares, le processus de sélection a suffisamment de temps

pour éliminer les traits désavantageux dans la population, avant l'arrivée d'une nouvelle mutation produisant un nouveau phénotype dans la population.

- (3) **Mutations (d'amplitude) petites.** Cette autre hypothèse évolutive suppose que l'amplitude des mutations est petite au niveau des phénotypes, autrement dit, le phénotype d'un individu mutant est légèrement différent de celui de son géniteur.
- (4) **Grande (taille de) population.** Cette hypothèse écologique permet de considérer des densités de population et de supposer leur dynamique déterministe. Noter toutefois que les mutations arrivent aléatoirement.

Les hypothèses biologiques (2), (3) et (4) permettent d'introduire trois paramètres d'échelle :

- $p \in [0, 1]$  est égal à la probabilité d'occurrence d'une mutation qui arrive au cours de la vie des individus. Noter que lorsque  $p \rightarrow 0$ , les mutations sont considérées *rare*s alors que pour  $p = 1$ , elles sont considérées *fréquentes*.
- $\sigma \in ]0, +\infty[$  est lié à la taille des mutations. Noter alors que  $\sigma \rightarrow 0$  signifie que les mutations ont des effets phénotypiques *petits*.
- $K \in \mathbb{N}^*$  correspond à la taille de la population. Remarquer que lorsque  $K \rightarrow +\infty$ , la (taille de) population est considérée *grande*.

Sous les quatre hypothèses biologiques précédentes, les dynamiques adaptatives ont connu un certain nombre de succès.

**Succès.** Les succès des dynamiques adaptatives viennent notamment de leur capacité à proposer de nombreux outils [74, 212, 68] pouvant s'adapter à une large variété de situations écologiques : interactions compétitives ou mutualistes, de type proie-prédateur, hôte-parasite ou encore plante-insecte.

L'un des outils centraux de la théorie des dynamiques adaptatives est celui du *Trait Substitution Sequence* (TSS). Il s'agit d'un processus stochastique introduit dans [171, 72] et mathématiquement étudié dans [14, 37, 47]. Ce modèle markovien de sauts décrit les invasions successives de mutants *avantageux* qui peuvent envahir et remplacer le trait résident d'une population monomorphique (i.e. que tous les individus de cette population portent le même phénotype). Ce modèle s'obtient sous les hypothèses (1), (2) et (4) et une hypothèse supplémentaire interdisant la coexistence à long terme de deux traits distincts, appelée principe "*Invasion Implique Fixation*" (IIF). En supposant de plus l'hypothèse (3), DIECKMANN et LAW ont proposé *heuristiquement* dans [72] une approximation du TSS sous forme d'une Équation Différentielle Ordinaire (EDO). Cette équation, appelée *Équation Canonique des Dynamiques Adaptatives* (abrégée par CEAD) est un outil important des dynamiques adaptatives et l'objet central de cette thèse.

**Équation Canonique des Dynamiques Adaptatives.** L'équation canonique permet de décrire, sur le long terme, le comportement du trait phénotypique dominant dans la population comme illustré sur les figures suivantes. Ces illustrations numériques ont été générées grâce au package *IBMPopSim* de GIORGI ET AL [116] et testées sur une version du modèle biologique [140] de KISDI.

Dans les Figures 1 et 2, on considère une population caractérisée par un seul trait phénotypique quantitatif et dont la population initiale est monomorphique de trait phénotypique de valeur 1.

Lorsque  $K$  et  $\sigma$  ne sont pas trop grands et  $p$  pas trop petit, comme pour la Figure 1, la



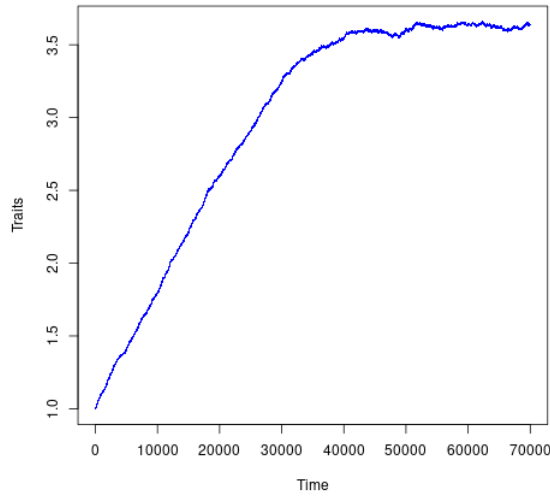


FIGURE 1 – Simulation numérique de la distribution des traits phénotypiques d'une certaine population pour les paramètres  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.0005$ .

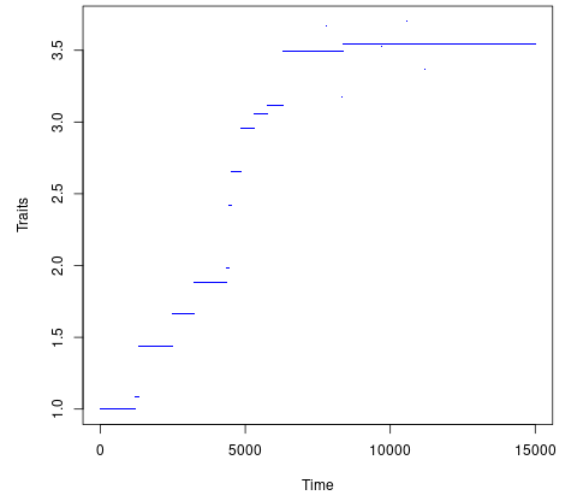


FIGURE 2 – Simulation numérique de la distribution des traits phénotypiques d'une certaine population pour les paramètres  $K = 300$ ,  $p = 0.00003$ ,  $\sigma = 0.2$ .

population évolue suivant un scénario relativement stable. Initialement concentrée autour du trait phénotypique de valeur 1, la population se déplace progressivement vers une valeur de trait phénotypique proche de 3,5 correspondant à celle du trait optimal, c'est-à-dire celui qui est le mieux adapté pour survivre et se reproduire. Une fois ce point atteint, la population se stabilise dans son état d'équilibre. L'équation canonique décrit la trajectoire de la courbe entre l'instant initial  $t = 0$  et l'instant  $t \simeq 40000$  où la population est à l'équilibre.

La Figure 2 traite le cas des mutations rares, hypothèse selon laquelle la théorie des dynamiques adaptative est fondée. Sur cette figure, on observe que l'évolution procède par sauts et décrit les invasions successives de mutants remplaçant le précédent trait résident. Autrement dit, la Figure 2 illustre la convergence vers une approximation du TSS où les mutations sont petites, c'est-à-dire vers l'équation canonique. Noter que dans cette situation, le nombre de traits coexistants au même moment est de l'ordre de deux ou trois, alors que dans le Figure 1 la population se concentre autour d'une certaine valeur de trait avec une petite variance phénotypique autour de celui-ci.

L'objectif de cette thèse est de décrire la dynamique d'évolution du trait phénotypique dominant dans la population comme solution de l'équation canonique lorsque les mutations sont *fréquentes* et dans le régime asymptotique de grande population ( $K \rightarrow +\infty$ ) et de petites mutations ( $\sigma \rightarrow 0$ ). La Figure 1 illustre précisément cette situation.

L'équation canonique est constituée de deux termes comme schématisée en Figure 3. L'un des termes est un gradient de *fitness* décrivant la force de sélection qui pousse la population à accroître localement sa fitness. Le second terme décrit l'effet des mutations.

$$\text{CEAD} = \boxed{\text{Gradient de fitness}} \times \boxed{\text{Effets des mutations}}$$

FIGURE 3 – Forme générale de l'équation canonique.

**Fitness.** La fitness  $\text{Fit}(y, x)$  mesure la valeur sélective d'un individu  $y$  donné dans un environnement donné incluant le reste de la population considérée de trait  $x$ . Cet individu peut être n'importe quel individu mutant (fictif) pouvant naître dans la population considérée. La manière de construire un *paysage de fitness* dépend du contexte écologique [172] et évolue, dans le cas des dynamiques adaptatives, en même temps que le changement évolutif de la population. La Figure 4 illustre la variation de la fitness en fonction du trait résident.

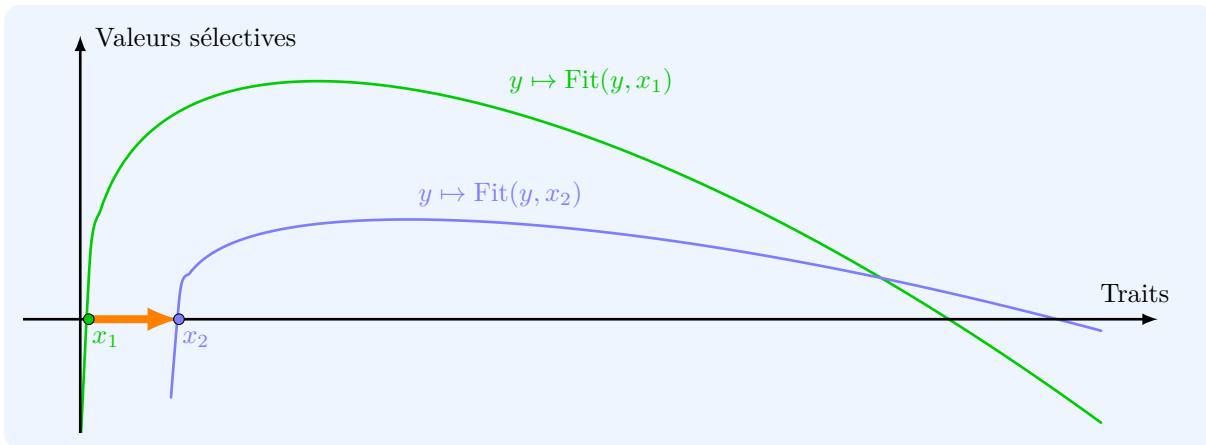


FIGURE 4 – Exemple de paysage de fitness (2D) en dynamiques adaptatives.

La flèche représente la mutation du trait résident  $x_1$  en le trait  $x_2$  dont la valeur est proche de  $x_1$  à cause de l'hypothèse (3). Cette mutation va ainsi changer le paysage adaptatif. La Figure 4 capture donc l'aspect dynamique du paysage de fitness en dynamiques adaptatives.

## 2 État de l'art mathématiques en dynamiques adaptatives

Ce domaine récent des dynamiques adaptatives soulève de nombreuses questions mathématiques originales. En particulier, celle d'étudier le comportement en temps long du trait phénotypique dominant d'une population donnée comme solution de l'équation canonique sous différents régimes asymptotiques. On présente ci-dessous deux approches mathématiques historiques portant sur l'équation canonique et appelées respectivement "*approche EDP des dynamiques adaptatives*" et "*approche stochastique des dynamiques adaptatives*". Ces deux approches sont basées sur un *modèle individu-centré* (abrégé par IBM).

**Modèles individu-centrés.** Un IBM est un modèle dans lequel l'ensemble des événements de naissance, mort, compétition, mutation de chaque individu est décrit en tout temps.

On considère une population structurée en trait phénotypique, de taille constante égale à  $K$ ,

où les capacités de survie et de reproduction de chaque individu sont caractérisées par un trait quantitatif  $x \in \mathbb{R}$ , i.e. une caractéristique globale soumise à sélection, e.g la capacité à absorber des nutriments.

On s'intéresse à approcher, sur le long terme, la dynamique d'une grande population. L'état de la population au temps  $t \geq 0$  peut être décrit par la mesure ponctuelle finie sur  $\mathbb{R}$ , mise à l'échelle par  $K$ , par

$$\nu_t^{K,p,\sigma} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i(t)}, \quad (1)$$

où  $x_i(t)$  désigne le trait phénotypique du  $i^{\text{ième}}$  individu au temps  $t$ . Les deux approches historiques utilisent différentes combinaisons des hypothèses biologiques **(2)**, **(3)** et **(4)** conduisant à différentes versions de l'équation canonique. Ces approches considèrent un paramètre d'échelle (en temps) sous lequel la distribution de la population sur l'espace des traits se concentre sur les masses de DIRAC, i.e. des sous-populations dans lesquelles tous les individus ont le même trait. L'équation canonique correspond donc au mouvement des masses de DIRAC. Noter qu'en 2001 [14] est la première référence qui fournit un formalisme mathématiques rigoureux pour établir l'équation canonique heuristiquement proposée par DIECKMANN et LAW [72].

**Approche EDP des dynamiques adaptatives.** Une version de [102] permet d'étudier le comportement de cet IBM sous une asymptotique de grande population ( $K \rightarrow +\infty$ ). Il est prouvé que cet IBM converge en loi vers un processus déterministe obtenu comme solution faible d'une certaine EDP. L'approche EDP des dynamiques adaptatives consiste à appliquer une limite de petites mutations ( $\sigma \rightarrow 0$ ) à ce processus déterministe, pour obtenir sur l'échelle de temps en  $1/\sigma$  une nouvelle version de l'équation canonique de DIECKMANN-LAW décrite par une équation d'HAMILTON-JACOBI avec contrainte. Il est également considéré dans cette approche que les mutations arrivent fréquemment au niveau individuel. Cette idée originale a été proposée en 2005 dans [75] par DIECKMANN, JABIN, MISCHLER et PERTHAME. Cette approche peut être résumée par la Figure 5 ci-dessous et correspond précisément à la flèche de droite :



FIGURE 5 – Résumé de l'approche EDP des dynamiques adaptatives.

Des extensions à cette approche sont notamment présentées dans [194, 160, 42, 174].

**Approche stochastique des dynamiques adaptatives.** Cette approche conduit, en deux étapes, vers la première version de l'équation canonique proposée par DIECKMANN-LAW. Cette approche stochastique peut être résumée par la Figure 6 ci-dessous :

- **Étape 1.** Cette étape consiste à prouver que l'IBM ci-dessus converge vers le TSS sous la double asymptotique de grande population ( $K \rightarrow +\infty$ ) et de mutations rares ( $p \rightarrow 0$ ) satisfaisant  $Kp \rightarrow 0$ . Noter que la limite de mutations rares permet de séparer les événements de mutation sur une échelle évolutive longue et la limite de grande population permet d'étudier la compétition entre les individus, entre deux événements de mutation sur une échelle écologique plus courte. Cette idée originale de séparation d'échelle de

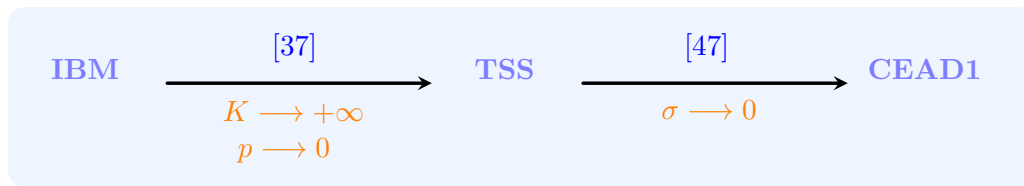


FIGURE 6 – Résumé de l’approche stochastique des dynamiques adaptatives.

temps, entre événements de mutation et interactions écologiques, a été proposée en 2006 par CHAMPAGNAT dans [37] et correspond à la flèche de gauche dans la Figure 6. Le TSS agit donc sur l’échelle de temps des mutations  $1/Kp$ . Noter que la Figure 2 illustre la convergence de l’IBM vers le TSS en supposant les mutations petites.

- **Étape 2.** Cette étape consiste à prouver la convergence du TSS vers la version originale de l’équation canonique sous une asymptotique de petites mutations ( $\sigma \rightarrow 0$ ) pour le scaling en temps approprié de  $1/\sigma^2$  (contre  $1/\sigma$  pour l’approche EDP) à cause des particularités du modèle biologique. Ce résultat de 2011 [47], dû à CHAMPAGNAT et MÉLÉARD, fournit une justification mathématique de l’équation canonique de [72] et correspond à la flèche de droite de la Figure 6. L’échelle de temps impliquée pour observer le phénomène de l’équation canonique à partir de l’IBM initial est  $1/Kp\sigma^2$ .

Des extensions pour l’Étape 1 sont notamment obtenues dans [209, 43, 19, 154, 55, 18, 24, 25]. D’autres extensions pour l’Étape 2 sont, par exemple, proposées dans [166, 51, 43, 55]. Dans [6], BAAR ET AL étudient la convergence – en une étape – de l’IBM vers l’équation canonique sous la triple asymptotique ( $K \rightarrow +\infty, p \rightarrow 0, \sigma \rightarrow 0$ ).

Les approches EDP et stochastique sont intéressantes parce qu’elles permettent de prédire le destin évolutif de communauté écologique mais elle restent néanmoins controversées par les biologistes.

**Controverses biologiques.** En 2005, un numéro spécial de *Journal of Evolutionary Biology*, consacré aux dynamiques adaptatives, comportait notamment un article critique décrivant les hypothèses, outils et les principales controverses de cette théorie [212].

- **Pour l’approche EDP.** Dans [195], une critique biologique est formulée à l’encontre de l’approche EDP à cause du *tail problem* : des populations exponentiellement petites qui devraient en réalité s’éteindre, peuvent avoir une influence forte sur la dynamique d’évolution de la population. En particulier, la vitesse d’évolution est trop rapide. Des modifications de l’équation d’HAMILTON-JACOBI ont été proposées dans [195, 173, 129] pour résoudre ce problème.
- **Pour l’approche stochastique.** La principale controverse dans [212] concerne l’approche stochastique puisqu’elle est basée sur une hypothèse biologique non-réaliste : les mutations sont supposées *trop rares*. En conséquence, le phénomène de l’équation canonique a lieu sur une échelle de temps  $1/Kp\sigma^2$  beaucoup trop longue. Les approches [169, 49] tentent de résoudre ce problème.

**Objectif de la thèse.** L’objectif de cette thèse est de corriger la controverse biologique sur les mutations rares en proposant des modèles probabilistes plus réalistes où les mutations ne

sont plus supposées rares. Plus précisément, notre objectif est d'évaluer mathématiquement les conséquences d'une nouvelle hypothèse biologique de mutations fréquentes ( $p = 1$ ) sur l'équation canonique en se plaçant dans un régime asymptotique de grande population ( $K \rightarrow +\infty$ ) et de petites mutations ( $\sigma \rightarrow 0$ ) comme schématisé dans la Figure 7.

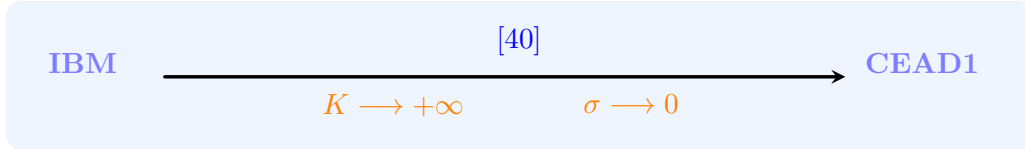


FIGURE 7 – Résumé de notre approche.

Remarquer que les deux limites  $\sigma \rightarrow 0$  et  $K \rightarrow +\infty$  doivent être prises simultanément. En effet, en prenant  $\sigma \rightarrow 0$  sans  $K \rightarrow +\infty$ , cela conduirait, comme dans [37, Remarque 3], à l'extinction de la population avant qu'une évolution significative soit observée. À l'inverse, en prenant  $K \rightarrow +\infty$  puis  $\sigma \rightarrow 0$  cela conduit à l'approche EDP présentée ci-dessus.

D'autre part, noter qu'en un certain sens, notre étude peut être vue comme une approche intermédiaire entre les approches stochastiques et EDP. L'échelle de temps pour observer l'équation canonique dans [47] est  $1/Kp\sigma^2$  et dans notre cas  $1/K\sigma^2$  avec  $p = 1$  d'où la consistance entre ces deux approches. On partage aussi avec l'approche EDP l'hypothèse de mutations fréquentes au niveau individuel. Ainsi, notre approche se place sur une échelle de temps évolutive intermédiaire satisfaisant

$$\frac{1}{\sigma} \ll \frac{1}{K\sigma^2} \ll \frac{1}{Kp\sigma^2}$$

où  $f(N) \ll g(N)$  signifie que  $f(N)/g(N)$  tend vers 0 quand  $N \rightarrow +\infty$ . De plus, notre approche supprime le principal inconvénient de l'approche EDP (tail-problem) où l'évolution est trop rapide et celui de l'approche stochastique (mutations rares) où l'évolution est trop lente. Ainsi, notre approche semble donc être la mieux adaptée pour résoudre ces diverses controverses biologiques. Ce qui est nouveau ici, c'est la convergence directe de l'IBM vers l'équation canonique dans le régime proposé. Une nouvelle approche implique donc une nouvelle démonstration. Ici, notre preuve est fondée sur une *analyse asymptotique lent-rapide* que l'on développe en section suivante.

### 3 Mes contributions

**Notre modèle.** Dans cette thèse, on étudie un modèle individu-centré décrivant l'évolution darwinienne d'une population asexuée et structurée en trait phénotypique. On considère une grande population de taille constante, caractérisée par un taux de ré-échantillonnage modélisant les pressions de sélection, et un taux de mutation où les mutations sont supposées fréquentes et arriver au cours de la vie. Dans ce modèle, l'état de la population à temps fixé est donné par une mesure sur l'espace des traits et l'évolution de la population est décrite en temps continu par le processus markovien  $(\nu_t^{K,p,\sigma})_{t \geq 0}$  de la forme (1).

**Résultat principal.** Le résultat principal de cette thèse consiste à prouver la convergence de cet IBM vers l'équation canonique de DIECKMANN-LAW, sous la double asymptotique de grande population ( $K \rightarrow +\infty$ ) et de petites mutations ( $\sigma \rightarrow 0$ ) et en supposant les mutations

fréquentes ( $p = 1$ ). Ce résultat n'est valide que pour une certaine gamme de paramètres  $\sigma$  (comme une fonction de  $K$ ) qui doivent être suffisamment petits mais pas trop non plus. Ce résultat est basé sur une *analyse asymptotique lent-rapide* avec utilisation d'une *méthode de moyennisation* nécessitant d'identifier et caractériser la composante limite rapide.

**Analyse asymptotique lent-rapide.** Notre composante lente correspond à la dynamique du trait phénotypique moyen  $(z_t^K)_{t \geq 0} := \left( \int_{\mathbb{R}} x \nu_{t/K\sigma^2}^{K,1,\sigma}(dx) \right)_{t \geq 0}$  qui opère sur l'échelle de temps  $1/K\sigma^2$ . Notons qu'il s'agit bien du trait moyen puisque  $\nu_t^{K,1,\sigma}$  est une mesure de probabilité. Notre composante rapide  $(\mu_t^K)_{t \geq 0}$  agit, quant à elle, sur l'échelle de temps  $K$  et correspond à la dynamique de la distribution des traits, centrée par rapport à la valeur du trait moyen et dilatée en  $1/\sigma\sqrt{K}$ , sur l'échelle de temps en  $1/K\sigma^2$  :

$$\mu_t^K := \frac{1}{K} \sum_{i=1}^K \delta_{\frac{1}{\sigma\sqrt{K}}(x_i(t) - z_t^K)}.$$

- **Étude de la composante rapide.** Le comportement en temps long de la composante rapide est non-classique et correspond à celui d'une diffusion à valeurs mesures originale que l'on interprète comme un processus de *FLEMING-VIOT recentré*. Ce processus est défini comme le processus de FLEMING-VIOT original translaté par moins sa moyenne empirique dont nous réalisons une étude détaillée dans cette thèse. Nous caractérisons le processus de FLEMING-VIOT recentré par un *problème de martingale* dont la preuve d'existence est basée sur le problème de martingale du processus de FLEMING-VIOT original et des développements asymptotiques. La preuve d'unicité, quant à elle, est fondée sur une *version faible* de la *méthode de dualité* et en supposant que la condition initiale du processus de FLEMING-VIOT recentré ait tous ses moments finis. Cette version non-classique de la relation de dualité permet de montrer que deux solutions au problème de martingale de FLEMING-VIOT recentré ont les mêmes marginales finies 1–dimensionnelles ce qui est suffisant pour conclure à l'unicité de la solution d'après le résultat [94, Théorème 4.4.2] de ETHIER-KURTZ. Enfin, on prouve par des arguments de couplage que ce processus de FLEMING-VIOT recentré bénéficie de *propriétés ergodiques (fortes)* et d'une caractérisation explicite de sa *mesure invariante* en exploitant une variante du *modified look-down* de DONNELLY-KURTZ [78, 80].
- **Méthode de moyennisation.** La mise en œuvre de notre méthode de moyennisation, inspirée par KURTZ [144] est basée sur des arguments de *compacité-unicité* suivant trois grandes étapes.

La première étape consiste à prouver la compacité des lois du couple  $(z^K, \Gamma^K)_{K \in \mathbb{N}^*}$  où  $\Gamma^K$  désigne la mesure d'occupation de la composante rapide  $\mu^K$ . La deuxième étape consiste à établir un problème de martingale pour tous points d'accumulation de la famille des lois de  $(z^K, \Gamma^K)_{K \in \mathbb{N}^*}$ . La dernière étape identifie ces points d'accumulation en deux temps. Dans un premier temps, on détermine la valeur limite  $\Gamma$  de  $\Gamma^K$  en terme de la limite de  $z^K$  et en exploitant des arguments d'ergodicité, d'où notre étude ergodique du processus de FLEMING-VIOT recentré. Noter qu'un argument de DAWSON [60], basé sur une extension de notre relation de dualité affaiblie, est aussi requis dans l'identification de  $\Gamma$ . Dans un second temps, on établit l'unicité au problème de martingale de la composante lente sachant que  $\Gamma$  est donnée.

Cette méthode de moyennisation requiert, dans notre situation, l'introduction de temps d'arrêt pour contrôler les moments de la composante rapide et prouver qu'ils

convergent en probabilité vers  $+\infty$  en utilisant des résultats de grandes déviations sur les marches aléatoires et des estimées de sortie de domaines. Notre méthode requiert également de réduire le problème initial sur la droite réelle au cas du tore pour prouver la tension de la composante lente. Puisque la composante lente limite est dirigée par l'équation canonique qui est une EDO non-explosive, on choisit le tore suffisamment grand pour conclure la preuve dans le cas de la droite réelle.

**Contributions scientifiques.** Cette thèse a donné lieu à deux contributions principales :

N. CHAMPAGNAT and V. HASS. *Existence, uniqueness and ergodicity for the centered Fleming-Viot process*. À paraître dans *Stochastic Processes and Their Applications* (<https://doi.org/10.1016/j.spa.2023.09.006>).

et

N. CHAMPAGNAT and V. HASS. *Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics*. Preprint (<https://arxiv.org/abs/2303.09936>). 2023. Soumis à *Annals of Applied Probability*.

qui correspondent respectivement aux chapitres 2 et 3 de ce manuscrit.

**Organisation de cette thèse.** Le chapitre 1 de thèse est une version *détaillée* de cette introduction et il se découpe en trois sections. La première section est consacrée à un développement détaillé de la biologie des dynamiques adaptatives en se focalisant sur les équations canoniques. L'approche heuristique de DIECKMANN-LAW y est précisée et mise en comparaison avec d'autres approches en biologie de l'évolution. Le phénomène de l'équation canonique des dynamiques adaptatives est également situé sur un axe temporel en lien avec d'autres événements éco-évolutifs. Dans cette première partie, on construit également un modèle individu-centré qui servira de base pour décrire les heuristiques mathématiques des deux approches historiques mais aussi pour notre étude. Des extensions mathématiques à l'équation canonique sont discutées ainsi que d'autres approches mathématiques avec un autre scaling temporel. Enfin, les deux principales controverses biologiques aux approches EDP et stochastiques, qui sont décrites, permettent de justifier notre étude comme une approche intermédiaire plus réaliste.

La deuxième section du chapitre 1 présente de façon pédagogique les théories mathématiques utilisées dans cette thèse. Elles fournissent notamment l'état de l'art et les méthodes pour le processus de FLEMING-VIOT original ainsi que celles pour l'analyse asymptotique lent-rapide avec méthode de moyennisation.

La troisième section du chapitre 1 est consacrée à présenter, plus en détail, les contributions de cette thèse dont une partie vise à étendre les méthodes de FLEMING-VIOT original au cas du FLEMING-VIOT recentré et à l'implémentation de notre méthode de moyennisation stochastique inspirée par KURTZ. On y présente également l'ensemble des difficultés et les solutions mises en œuvre pour obtenir nos différents résultats. Différentes perspectives naturelles y sont proposées.

Le chapitre 2 s'intéresse à des questions d'existence, d'unicité et d'ergodicité du processus de FLEMING-VIOT recentré. Le chapitre 3 est consacré à la convergence de notre modèle individu-centré vers l'équation canonique dans le régime asymptotique de grande population et de petites mutations tout en supposant que les mutations arrivent fréquemment au cours de la vie des individus. Ce manuscrit s'achève avec deux appendices qui regroupent tout d'abord les lemmes

techniques pour la preuve d'existence du processus FLEMING-VIOT recentré puis les codes sources permettant de simuler sous différents régimes la convergence de l'IBM vers l'équation canonique.

Bonne lecture !



# A brief general introduction

This general introduction to my thesis is structured along three main axes. In Section 1, I present the biological state of the art of adaptive dynamics by focusing on the biological assumptions that govern its framework, its successes and especially that of the canonical equation. In Section 2, I present the two main mathematical approaches to the canonical equation of adaptive dynamics and the goal of this thesis. Finally, in Section 3, I present the contributions of this thesis and the general organisation of the manuscript.

## 1 Biological state of the art in adaptive dynamics

This thesis, which is at the interface between mathematics and evolutionary biology, starts from biological motivations in adaptive dynamics. This recent field of theoretical biology was developed in the second half of the 1990s by METZ ET AL [171], DIECKMANN and LAW [72] and GERITZ ET AL [112, 114]. This theory studies the links between *ecology* and *evolution*. By ecology, we mean all the interactions between individuals and their environment, including other individuals in the population. Evolution is to be understood in the Darwinian sense, i.e. based on three basic mechanisms. Firstly, *heredity*, which allows the transmission of individual phenotypic characteristics from one generation to the next. Secondly, a *source of variation* among these characteristics: in the models studied in this thesis, this source of variation comes exclusively from mutations. Finally, a *mechanism of (natural) selection* that may result from inter- or intra-species competition: individuals that are better adapted to their environment are more likely to survive and reproduce and thus transmit their characteristics to future generations than those that are less so.

**Biological assumptions.** To set the framework for adaptive dynamics, biologists make four simplifying biological assumptions.

- (1) **Asexual population.** This assumption aims to simplify the reproduction scheme of individuals by assuming it to be clonal: a mother cell (or a plant, some animals, some insects, ...) gives two daughter cells with the same genetic inheritance. In an asexual population, mutations, which are permanent alterations of the genome, are considered in adaptive dynamics as acting directly on the phenotype. To be visible, i.e. to have a phenotypic effect, mutations must affect a coding part of the genome and alter the chemical properties of the encoded proteins [3]. It is biologically reasonable to assume that mutations are rare.
- (2) **Rare mutations.** This assumption is fundamental to introduce a time scale separation between an ecological scale, fast, corresponding to interactions between individuals, and an evolutionary scale, slower, corresponding to the arrival of mutations. Since mutations are rare, the selection process has enough time to eliminate disadvantageous traits in

the population, before the arrival of a new mutation producing a new phenotype in the population.

- (3) **Small (amplitude) mutations.** This other evolutionary assumption supposes that the amplitude of mutations is small at the level of phenotypes, i.e. the phenotype of a mutant individual is slightly different from that of its progenitor.
- (4) **Large population (size).** This ecological assumption allows to consider population densities and to assume that their dynamics are deterministic. However, mutations occur randomly.

Biological assumptions (2), (3) and (4) allow the introduction of three scaling parameters:

- $p \in [0, 1]$  is equal to the probability of occurrence of a mutation that occurs during the life of the individuals. Note that when  $p \rightarrow 0$ , mutations are *rare*, whereas when  $p = 1$ , they are *frequent*.
- $\sigma \in (0, +\infty)$  is related to the size of the mutations. Note then that  $\sigma \rightarrow 0$  means that mutations have *small* phenotypic effects.
- $K \in \mathbb{N}^*$  corresponds to the population size. Note that when  $K \rightarrow +\infty$ , the population (size) is *large*.

Under the previous four biological assumptions, adaptive dynamics have many successes.

**Successes.** The successes of adaptive dynamics come in particular from their capacity to propose numerous tools [74, 212, 68] that can be adapted to a wide variety of ecological situations: competitive or mutualistic interactions, of the predator-prey, host-parasite or plant-insect.

One of the central tools of the theory of adaptive dynamics is the *Trait Substitution Sequence* (TSS). This is a stochastic process introduced in [171, 72] and mathematically studied in [14, 37, 47]. This Markovian jump model describes the successive invasions of *advantageous* mutants that can invade and replace the resident trait of a monomorphic population (i.e. all individuals of this population carry the same phenotype). This model is obtained under Assumptions (1), (2) and (4) and an additional assumption prohibiting the long term coexistence of two distinct traits, called the “*Invasion Implies Fixation*” (IIF) principle. Assuming furthermore Assumption (3), DIECKMANN and LAW proposed heuristically in [72] an approximation of the TSS in the form of an Ordinary Differential Equation (ODE). This equation, called the *Canonical Equation of Adaptive Dynamics* (abbreviated as CEAD) is an important tool in adaptive dynamics and the focus of this thesis.

**Canonical Equation of Adaptive Dynamics.** The canonical equation allows to describe the long term behaviour of the dominant phenotypic trait in the population as illustrated in the following figures. These numerical illustrations were generated using the *IBMPopSim* package by GIORGI ET AL [116] and tested on a version of the biological model [140] by KISDI.

In Figures 8 and 9, we consider a population characterised by a single quantitative phenotypic trait and whose initial population is monomorphic with a phenotypic trait of value 1.

When  $K$  and  $\sigma$  are not too large and  $p$  not too small, as in Figure 8, the population evolves according to a relatively stable scenario. Initially concentrated around the phenotypic trait of value 1, the population gradually moves towards a phenotypic trait value close to 3.5 corresponding to that of the optimal trait, i.e. the one best adapted to survive and reproduce. Once this point is reached, the population stabilises in its steady state. The canonical equation describes

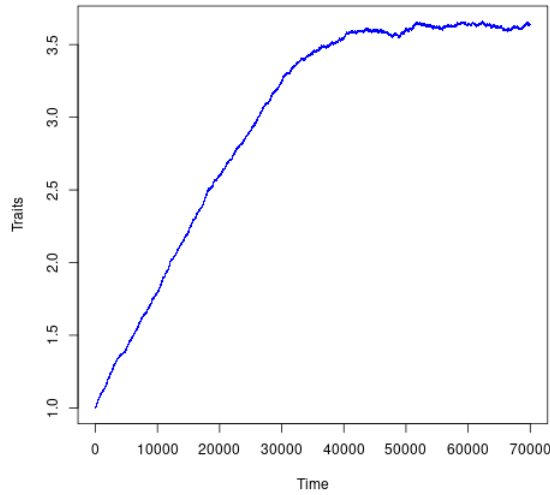


FIGURE 8 – Numerical simulation of the distribution of phenotypic traits of a certain population for the parameters  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.0005$ .

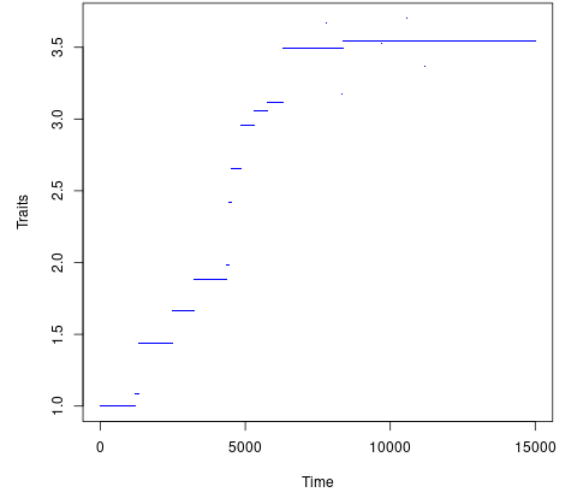


FIGURE 9 – Numerical simulation of the distribution of phenotypic traits of a certain population for the parameters  $K = 300$ ,  $p = 0.00003$ ,  $\sigma = 0.2$ .

the trajectory of the curve between the initial time  $t = 0$  and the time  $t \simeq 40000$  when the population is in equilibrium.

Figure 9 deals with the case of rare mutations, the assumption on which the theory of adaptive dynamics is based. In this figure, it can be seen that evolution proceeds by jumps and describes the successive invasions of mutants replacing the previous resident trait. In other words, Figure 9 illustrates the convergence to an approximation of the TSS where mutations are small, i.e. to the canonical equation. Note that in this situation the number of coexisting traits at the same time is of the order of two or three, whereas in Figure 8 the population is concentrated around a certain trait value with a small phenotypic variance around it.

The goal of this thesis is to describe the evolutionary dynamics of the dominant phenotypic trait in the population as a solution of the canonical equation when mutations are *frequent* and in the asymptotic regime of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ). Figure 8 illustrates precisely this situation.

The canonical equation consists in two terms as summarised in Figure 10.

$$\text{CEAD} = \boxed{\text{Fitness gradient}} \times \boxed{\text{Mutations effects}}$$

FIGURE 10 – Generic form of the canonical equation.

One term is a *fitness* gradient describing the strength of selection that drives the population to increase its fitness locally. The second term describes the effect of mutations.

**Fitness.** The fitness  $\text{Fit}(y, x)$  measures the selective value of a given individual  $y$  in a given environment including the rest of the considered population of trait  $x$ . This individual can be any (fictitious) mutant individual that can be born in the population under consideration. The way a *fitness landscape* is constructed depends on the ecological context [172] and evolves, in the case of adaptive dynamics, along with the evolutionary change of the population. Figure 11 illustrates the variation of fitness as a function of the resident trait.

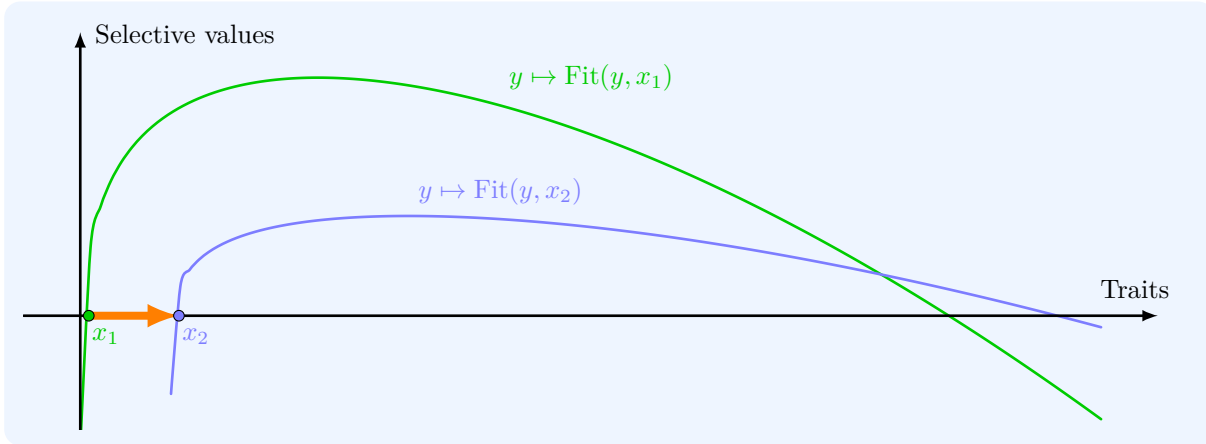


FIGURE 11 – Example of fitness landscape (2D) in adaptive dynamics.

The arrow  $\longrightarrow$  represents the mutation of the resident trait  $x_1$  into the trait  $x_2$  whose value is close to  $x_1$  because of Assumption **(3)**. This mutation will thus change the adaptive landscape. So, Figure 11 captures the dynamic aspect of the fitness landscape in adaptive dynamics.

## 2 Mathematical state of the art in adaptive dynamics

This recent field of adaptive dynamics raises many original mathematical questions. In particular, that of studying the long term behaviour of the dominant phenotypic trait of a given population as a solution of the canonical equation under different asymptotic regimes. Two historical mathematical approaches to the canonical equation are presented below, called respectively “*PDE approach to adaptive dynamics*” and “*stochastic approach to adaptive dynamics*”. Both approaches are based on an *individual-based model* (abbreviated as IBM).

**Individual-based models.** An IBM is a model in which the set of birth, death, competition and mutation events of each individual is described at all times.

We consider a phenotypic traits structured population, of constant size equal to  $K$ , where the survival and reproduction capacities of each individual are characterised by a quantitative trait  $x \in \mathbb{R}$ , i.e. a global characteristic subject to selection, e.g. the capacity to absorb nutrients.

We are interested in approximating the long term dynamics of a large population. The state of the population at time  $t \geq 0$  can be described by the finite point measure on  $\mathbb{R}$ , scaled by  $K$ ,

by

$$\nu_t^{K,p,\sigma} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i(t)}, \quad (2)$$

where  $x_i(t)$  denotes the phenotypic trait of the  $i^{\text{th}}$  individual at time  $t$ . The two historical approaches use different combinations of the biological assumptions **(2)**, **(3)** and **(4)** leading to different versions of the canonical equation. These approaches consider a scaling parameter (in time) under which the population distribution on the trait space concentrates on DIRAC masses, i.e. subpopulations in which all individuals have the same trait. So, the canonical equation corresponds to the movement of DIRAC masses. Note that in 2001 [14] is the first reference that provides a rigorous mathematical formalism to establish the canonical equation heuristically proposed by DIECKMANN and LAW [72].

**PDE approach to adaptive dynamics.** A version of [102] allows us to study the behaviour of this IBM under a large population asymptotic ( $K \rightarrow +\infty$ ). It is proved that this IBM converges in law to a deterministic process obtained as a weak solution of a certain PDE. The PDE approach to adaptive dynamics consists in applying a limit of small mutations ( $\sigma \rightarrow 0$ ) to this deterministic process, to obtain on the  $1/\sigma$  time scale a new version of the DIECKMANN-LAW's canonical equation described by a constrained HAMILTON-JACOBI equation. It is also considered in this approach that mutations frequently occur at the individual level. This original idea was proposed in 2005 in [75] by DIECKMANN, JABIN, MISCHLER and PERTHAME. This approach can be summarised in Figure 12 below and corresponds precisely to the arrow on the right:

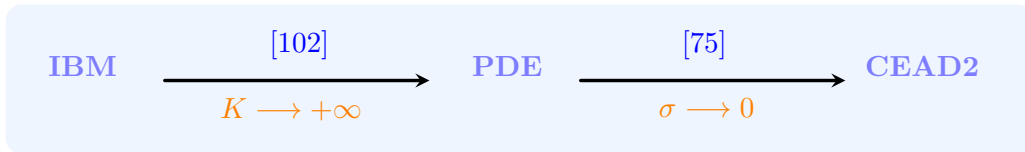


FIGURE 12 – Summary of the PDE approach of adaptive dynamics.

Extensions to this approach are presented in [194, 160, 42, 174].

**Stochastic approach to adaptive dynamics.** This approach leads, in two steps, to the first version of the canonical equation proposed by DIECKMANN-LAW. This stochastic approach can be summarised in Figure 13 below:

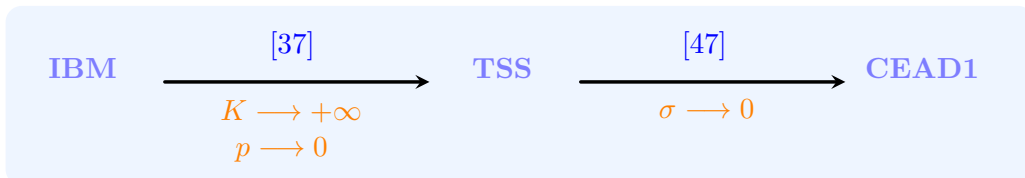


FIGURE 13 – Summary of the stochastic approach of adaptive dynamics.

- **Step 1.** This step consists in proving that the above IBM converges to the TSS under the double asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) satisfying

$Kp \rightarrow 0$ . Note that the limit of rare mutations allows separating mutation events on a long evolutionary scale and the limit of large population allows studying the competition between individuals, between two mutation events on a shorter ecological scale. This original idea of time scale separation between mutation events and ecological interactions was proposed in 2006 by CHAMPAGNAT [37] and corresponds to the left arrow in Figure 13. So, the TSS acts on the time scale of mutations  $1/Kp$ . Note that Figure 9 illustrates the convergence of the IBM to the TSS assuming small mutations.

- **Step 2.** This step consists in proving the convergence of the TSS to the original version of the canonical equation under an asymptotic of small mutations ( $\sigma \rightarrow 0$ ) for the appropriate time scaling of  $1/\sigma^2$  (against  $1/\sigma$  for the PDE approach) because of the particularities of the biological model. This result from 2011 [47], due to CHAMPAGNAT and MÉLÉARD, provides a mathematical justification for the canonical equation of [72] and corresponds to the right arrow in Figure 13. The time scale involved in observing the phenomenon of the canonical equation from the initial IBM is  $1/Kp\sigma^2$ .

Extensions for Step 1 are obtained in [209, 43, 19, 154, 55, 18, 24, 25]. Other extensions for Step 2 are, for example, proposed in [166, 51, 43, 55]. In [6], BAAR ET AL study the convergence – *in one step* – of the the IBM to the canonical equation under the triple asymptotic ( $K \rightarrow +\infty, p \rightarrow 0, \sigma \rightarrow 0$ ).

The PDE and stochastic approaches are interesting because they allow the prediction of the evolutionary fate of ecological communities, but they remain controversial among biologists.

**Biological controversies.** In 2005, a special issue of the *Journal of Evolutionary Biology*, devoted to adaptive dynamics included a critical article describing the assumptions, tools and main controversies of this theory [212].

- **For the PDE approach.** In [195], a biological criticism is formulated against the PDE approach because of the tail problem: exponentially small populations that should in fact become extinct, can have a strong influence on the evolutionary dynamics of the population. In particular, the rate of evolution is too fast. Modifications of the HAMILTON-JACOBI equation have been proposed in [195, 173, 129] to solve this problem.
- **For the stochastic approach.** The main controversy in [212] concerns the stochastic approach since it is based on a non-realistic biological assumption: mutations are supposed to be too rare. As a consequence, the phenomenon in the canonical equation takes place on a time scale  $1/Kp\sigma^2$  that is much too long. The approaches [169, 49] attempt to solve this problem.

**Goal of the thesis.** The goal of this thesis is to correct the biological controversy about rare mutations by proposing more realistic probabilistic models *without* assuming rare mutations. More precisely, our aim is to mathematically evaluate the consequences of a new biological assumption of frequent mutations ( $p = 1$ ) on the canonical equation by placing ourselves in an asymptotic regime of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ) as shown in Figure 14.

Note that both limits  $\sigma \rightarrow 0$  and  $K \rightarrow +\infty$  must be taken simultaneously. Indeed, taking  $\sigma \rightarrow 0$  *without*  $K \rightarrow +\infty$  would lead, as in [37, Remark 3], to the extinction of the population before any significant evolution is observed. Conversely, taking  $K \rightarrow +\infty$  *then*  $\sigma \rightarrow 0$  leads to the PDE approach presented above. On the other hand, note that in a certain sense, our study

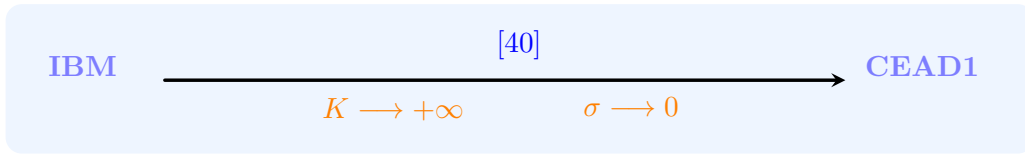


FIGURE 14 – Summary of our approach.

can be seen as an intermediate approach between the stochastic and PDE approaches. The time scale for observing the canonical equation in [47] is  $1/Kp\sigma^2$  and in our case  $1/K\sigma^2$  with  $p = 1$  hence the consistency between these two approaches. We also share with the PDE approach the assumption of frequent mutations at the individual level. Thus, our approach is placed on an intermediate evolutionary time scale satisfying

$$\frac{1}{\sigma} \ll \frac{1}{K\sigma^2} \ll \frac{1}{Kp\sigma^2}$$

where  $f(N) \ll g(N)$  means that  $f(N)/g(N)$  tends to 0 when  $N \rightarrow +\infty$ . Moreover, our approach removes the main drawback of the PDE approach (tail-problem) where evolution is too fast and that of the stochastic approach (rare mutations) where evolution is too slow. Thus, our approach seems to be the best suited to solve these various biological controversies. What is new here is the direct convergence of the IBM to the canonical equation in the proposed regime. A new approach therefore implies a new proof. Here, our proof is based on a *slow-fast asymptotic analysis* that we develop in the next section.

### 3 My contributions

**Our model.** In this thesis, we study a individual-based model describing the Darwinian evolution of an asexual, trait-structured population. We consider a large population of constant size, characterised by a resampling rate modelling selection pressures, and a mutation rate where mutations are assumed to be frequent and to occur over a lifetime. In this model, the state of the population at a fixed time is given by a measure on the trait space and the evolution of the population is described in continuous time by the MARKOV process  $(\nu_t^{K,p,\sigma})_{t \geq 0}$  of the form (2).

**Main result.** The main result of this thesis consists in proving the convergence of this IBM to the canonical DIECKMANN-LAW equation, under the double asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ) and assuming frequent mutations ( $p = 1$ ). This result is only valid for a certain range of parameters  $\sigma$  (as a function of  $K$ ) which must be sufficiently small but not too small either. This result is based on a *slow-fast asymptotic analysis* using an *averaging method* which requires the identification and characterisation of the limit fast component.

**Slow-fast asymptotic analysis.** Our slow component corresponds to the dynamics of the mean phenotypic trait  $(z_t^K)_{t \geq 0} := \left( \int_{\mathbb{R}} x \nu_t^{K,1,\sigma}(dx) \right)_{t \geq 0}$  which operates on the time scale  $1/K\sigma^2$ . Note that this is the mean trait since  $\nu_t^{K,1,\sigma}$  is a probability measure. Our fast component  $(\mu_t^K)_{t \geq 0}$ , on the other hand, operates on the time scale  $K$  and corresponds to the dynamics of the trait distribution, centered with respect to the value of the mean trait and dilated in

$1/\sigma\sqrt{K}$ , on the time scale in  $1/K\sigma^2$ :

$$\mu_t^K := \frac{1}{K} \sum_{i=1}^K \delta_{\frac{1}{\sigma\sqrt{K}}(x_i(t)-z_t^K)}.$$

- **Study of the fast component.** The long time behaviour of the fast component is non-classical and corresponds to that of an original measure-valued diffusion that is interpreted as a *centered FLEMING-VIOT* process. This process is defined as the original FLEMING-VIOT process translated by minus its empirical mean, which we study in detail in this thesis. We characterise the centered FLEMING-VIOT process by a *martingale problem* whose existence proof is based on the martingale problem of the original FLEMING-VIOT process and asymptotic expansions. The uniqueness proof is based on a *weak version* of the *duality method* and on the assumption that the initial condition of the centered FLEMING-VIOT process has all its finite moments. This non-classical version of the duality relation allows to show that two solutions to the centered FLEMING-VIOT martingale problem have the same 1-dimensional finite marginals which is sufficient to conclude that the solution is unique according to ETHIER-KURTZ's result [94, Theorem 4.4.2]. Finally, we prove by coupling arguments that this centered FLEMING-VIOT process benefits from (*strong*) *ergodic properties* and an explicit characterisation of its *invariant measure* by exploiting a variant of the DONNELLY-KURTZ *modified look-down* [78, 80].
- **Averaging method.** The implementation of our averaging method, inspired by KURTZ [144], is based on *compactness-uniqueness* arguments following three main steps. The first step consists in proving the compactness of the laws of the couple  $(z^K, \Gamma^K)_{K \in \mathbb{N}^*}$  where  $\Gamma^K$  denotes the occupation measure of the fast component  $\mu^K$ . The second step consists in establishing a martingale problem for all accumulation points of the family of laws of  $(z^K, \Gamma^K)_{K \in \mathbb{N}^*}$ . The last step identifies these accumulation points in two parts. In the first part, we determine the limit value  $\Gamma$  of  $\Gamma^K$  in terms of the limit of  $z^K$  and by exploiting ergodicity arguments, hence our ergodic study of the centered FLEMING-VIOT process. Note that a DAWSON's argument [60], based on an extension of our weakened duality relation, is also required in the identification of  $\Gamma$ . In a second part, we establish the uniqueness to the martingale problem of the slow component with  $\Gamma$  given. This averaging method requires in our situation the introduction of stopping times to control the moments of the fast component and prove that they converge in probability to  $+\infty$  using results of large deviations on random walks and estimates of exit from an attractive domain. Our method also requires reducing the initial problem on the real line to the torus case to prove the tightness of the slow component. Since the limiting slow component is directed by the canonical equation which is a non-explosive ODE, we choose the torus large enough to conclude the proof to the real line case.

**Scientific contributions.** This thesis has given rise to two main contributions:

N. CHAMPAGNAT and V. HASS. *Existence, uniqueness and ergodicity for the centered Fleming-Viot process*. To appear in *Stochastic Processes and Their Applications* (<https://doi.org/10.1016/j.spa.2023.09.006>).

and



N. CHAMPAGNAT and V. HASS. *Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics*. Preprint (<https://arxiv.org/abs/2303.09936>). 2023. Submitted to *Annals of Applied Probability*.

which correspond respectively to Chapters 2 and 3 of this manuscript.

**Organisation of this thesis.** Chapter 1 of the thesis is a *detailed* version of this introduction and is divided into three sections. The first section is devoted to a detailed development of the biology of adaptive dynamics with a focus on the canonical equations. The DIECKMANN-LAW heuristic approach is explained and compared with other approaches in evolutionary biology. The phenomenon of the canonical equation of adaptive dynamics is also situated on a temporal axis in connection with other eco-evolutionary events. In this first part, an individual-based model is also constructed as a basis for describing the mathematical heuristics of both historical approaches and for our study. Mathematical extensions to the canonical equation are discussed as well as other mathematical approaches with a different time scaling. Finally, the two main biological controversies in the PDE and stochastic approaches are described and our study is justified as a more realistic intermediate approach.

The second section of Chapter 1 presents the mathematical theories used in this thesis in an educational way. In particular, they provide the state of the art and methods for the original FLEMING-VIOT process as well as those for the slow-fast asymptotic analysis with averaging method.

The third section of Chapter 1 is devoted to a more detailed presentation of the contributions of this thesis, part of which aims at extending the original FLEMING-VIOT methods to the case of the centered FLEMING-VIOT and the implementation of our stochastic averaging method inspired by KURTZ. We also present all the difficulties and the solutions implemented to obtain our different results. Different natural perspectives are proposed.

Chapter 2 deals with questions of existence, uniqueness and ergodicity of the centered FLEMING-VIOT process. Chapter 3 is devoted to the convergence of our individual-based model to the canonical equation in the asymptotic regime of large population and small mutations while assuming that mutations occur frequently over the lifetime of individuals. This manuscript ends with two appendices which firstly include the technical lemmas for the existence proof of the centered FLEMING-VIOT process and then the source codes allowing to simulate under different regimes the convergence of the IBM to the canonical equation.

Enjoy reading!



# List of figures

1	Simulation numérique de la distribution des traits phénotypiques d'une certaine population pour les paramètres $K = 300$ , $p = 1$ , $\sigma = 0.0005$ . . . . .	xiii
2	Simulation numérique de la distribution des traits phénotypiques d'une certaine population pour les paramètres $K = 300$ , $p = 0.00003$ , $\sigma = 0.2$ . . . . .	xiii
3	Forme générale de l'équation canonique. . . . .	xiv
4	Exemple de paysage de fitness (2D) en dynamiques adaptatives. . . . .	xiv
5	Résumé de l'approche EDP des dynamiques adaptatives. . . . .	xv
6	Résumé de l'approche stochastique des dynamiques adaptatives. . . . .	xvi
7	Résumé de notre approche. . . . .	xvii
8	Numerical simulation of the distribution of phenotypic traits of a certain population for the parameters $K = 300$ , $p = 1$ , $\sigma = 0.0005$ . . . . .	xxiii
9	Numerical simulation of the distribution of phenotypic traits of a certain population for the parameters $K = 300$ , $p = 0.00003$ , $\sigma = 0.2$ . . . . .	xxiii
10	Generic form of the canonical equation. . . . .	xxiii
11	Example of fitness landscape (2D) in adaptive dynamics. . . . .	xxiv
12	Summary of the PDE approach of adaptive dynamics. . . . .	xxv
13	Summary of the stochastic approach of adaptive dynamics. . . . .	xxv
14	Summary of our approach. . . . .	xxvii
1.1	Diagram of evolutionary mechanisms. . . . .	4
1.2	Example of PIP and possible paths of the TSS monomorphic. . . . .	8
1.3	Generic form of the CEAD. . . . .	11
1.4	Time scales involved in adaptive dynamics in the stochastic approach. . . . .	13
1.5	Case where $\text{Fit}(y, x) > 0$ . . . . .	20
1.6	Case where $\text{Fit}(y, x) < 0$ . . . . .	20
1.7	Invasion fitness landscape of $y \mapsto \text{Fit}(y, x_i) = b(x_i, y) - b(y, x_i)$ with $x_1 := 1.75$ and $x_2 := 2.25$ . . . . .	21
1.8	Profile of the competition function $(x, y) \mapsto c(x, y)$ . . . . .	21
1.9	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 100$ , $p = 0.1$ , $\sigma = 0.03$ . . . . .	22
1.10	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 3000$ , $p = 0.1$ , $\sigma = 0.03$ . . . . .	22
1.11	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 100$ , $p = 0.001$ , $\sigma = 0.01$ . . . . .	23
1.12	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 300$ , $p = 0.00003$ , $\sigma = 0.2$ . . . . .	23

1.13	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 300, p = 1, \sigma = 0.001$ . . . . .	23
1.14	Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and $K = 300, p = 1, \sigma = 0.0005$ . . . . .	23
1.15	Summary of the PDE approach of adaptive dynamics. . . . .	24
1.16	Two possible profiles for $\varphi$ concave where $\varphi_1$ has a higher curvature than $\varphi_2$ . . .	26
1.17	Dispersion of the support around the dominant trait $x^*(t)$ . . . . .	26
1.18	Summary of the stochastic approach of adaptive dynamics. . . . .	27
1.19	The three steps of invasion and fixation of a mutant $y$ in a monomorphic resident population of trait $x$ . . . . .	29
1.20	The case of finite populations: the CDAD. . . . .	36
1.21	The case of three simultaneous limits. . . . .	36
1.22	The case of a system of ODE. . . . .	38
1.23	The case of another time scale to capture a HAMILTON-JACOBI (HJ) equation. .	38
1.24	Graphical summary of all convergences to the CEAD. We consider the triplet of parameters $(\frac{1}{K}, p, \sigma)$ . . . . .	39
1.25	Summary of our approach. . . . .	40
1.26	Graphical representation of the modified look-down process with $N = 5$ on $[0, T]$ . 47	47
1.27	KINGMAN's genealogy at time $T$ in the modified look-down of Figure 1.26. . . . .	47
1.28	Different time scales involved in our IBM expected in the original time scale of the process $(\nu_t^K)_{t \geq 0}$ . . . . .	74
1.29	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 100, p = 0.1, \sigma = 0.03$ . . . . .	80
1.30	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 300, p = 0.1, \sigma = 0.03$ . . . . .	80
1.31	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 3000, p = 0.1, \sigma = 0.03$ . . . . .	81
1.32	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 300, p = 0.00003, \sigma = 0.2$ . . . . .	81
1.33	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 300, p = 1, \sigma = 0.001$ . . . . .	82
1.34	Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and $K = 300, p = 1, \sigma = 0.0005$ . . . . .	82
1.35	Different time scales involved in our IBM expected in the original time scale of the process $(\hat{\nu}_t^{K,1,\sigma})_{t \geq 0}$ . . . . .	83
2.1	Graphical representation of the MORAN model with $N = 5$ . . . . .	111
2.2	KINGMAN's genealogy $(k_{5,T-t})_{0 \leq t \leq T}$ under the MORAN model on the left, tracing back from $T$ to 0. . . . .	111
2.3	Illustration of the centered MORAN process where $ k_{4,T}  = 1$ . . . . .	112
2.4	Illustration of the centered MORAN process where $ k_{4,T}  = 2$ . . . . .	112
2.5	Graphical representation of the modified look-down process with $N = 5$ . . . . .	116
2.6	KINGMAN's genealogy $(\check{k}_{5,t})_{0 \leq t \leq T}$ under the modified look-down model on the left, tracing back from time 0 to time $-T$ . . . . .	116
3.1	Different time scales involved in our individual-based model expected in the original time scale of the process $(\nu_t^K)_{t \geq 0}$ . . . . .	159

---

3.2 For  $K$  large enough, coupling between  $(L_k^K)_{0 \leq k \leq 7}$  and  $(Z_k^K)_{0 \leq k \leq 7}$  up to time  $T$  and before  $\check{\tau}^K$  and where  $M_2(\mu_0^K) \in [0, 2u_1 + 1)$  so that  $L_0^K = 1$  and  $\tilde{Z}_k^K := u_{Z_k^K} = 3^{Z_k^K} K^{\varepsilon/2}$  for all  $k \in \mathbb{N}$ . . . . . 186

*List of figures*

---

# Chapter 1

## A detailed introduction

### Contents

---

<b>1.1 Adaptive dynamics</b>	<b>2</b>
1.1.1 Darwinian evolutionary biology	3
1.1.2 Biological context of adaptive dynamics	5
1.1.2.1 Biological assumptions, fitness and vocabulary	6
1.1.2.2 Successes and application fields	8
1.1.2.3 Canonical equation	10
1.1.2.4 Adaptive dynamics: a multiscale phenomenon and the approach followed in this thesis	12
1.1.3 Mathematical context	14
1.1.3.1 Individual-based models	14
1.1.3.2 Construction of the individual-based model	15
1.1.4 Canonical Equation	22
1.1.4.1 PDE approach of adaptive dynamics	24
1.1.4.2 Stochastic approach of adaptive dynamics	27
1.1.4.3 Mathematical extensions in different contexts	32
1.1.4.4 Other mathematical approaches	35
1.1.4.5 Graphical summary and our approach	38
1.1.4.6 Biological controversies	41
<b>1.2 Mathematical theories</b>	<b>42</b>
1.2.1 FLEMING-VIOT processes	42
1.2.1.1 MORAN's particle process and existence of the martingale problem	43
1.2.1.2 DONNELLY-KURTZ's modified look-down	47
1.2.1.3 Dual process and uniqueness of the martingale problem	49
1.2.1.4 Ergodicity	54
1.2.2 Slow-fast asymptotic analysis	56
1.2.2.1 The perturbed test function and derived extensions	56
1.2.2.2 The KURTZ stochastic averaging method	57
<b>1.3 My contributions</b>	<b>63</b>
1.3.1 From Chapter 2	63
1.3.1.1 Existence for the centered FLEMING-VIOT process	64
1.3.1.2 Uniqueness for the centered FLEMING-VIOT process	66
1.3.1.3 Ergodicity for the centered FLEMING-VIOT process	69

1.3.2	From Chapter 3 . . . . .	70
1.3.2.1	Main result . . . . .	71
1.3.2.2	Slow-fast asymptotic analysis . . . . .	72
1.3.2.3	Notations and difficulties encountered . . . . .	74
1.3.2.4	Our averaging method . . . . .	75
1.3.3	Prospects . . . . .	78
1.3.3.1	Extension to models including fluctuations in population size . . . . .	78
1.3.3.2	Other strategies to resolve the controversy related to the rare mutations assumption . . . . .	83

---

Let us group together some of the abbreviations used throughout this thesis:

- **ODE** Ordinary Differential Equation
- **PDE** Partial Differential Equation
- **SDE** Stochastic Differential Equation
- **IBM** Individual-Based Model
- **TSS** Trait Substitution Sequence
- **PES** Polymorphic Evolution Sequence
- **CEAD** Canonical Equation of Adaptive Dynamics

In the sequel, when referring to a *phenotypic trait* of an individual, the terms “*phenotype*” or “*trait*” will be used independently.

In this thesis we will limit ourselves to the study of the evolution of a single species. The biological study of the adaptive dynamics of several interacting species is notably covered in [161, 72, 14, 66].

## 1.1 Adaptive dynamics

The theory of adaptive dynamics is integrated in the general context of Darwinian evolutionary biology. The ingredients and mechanisms of Darwinian evolution as well as several modelling approaches to study some of these mechanisms are detailed in Section 1.1.1. The adaptive dynamics approach is developed in Section 1.1.2 where its historical background, its fundamental ideas and some differences with population genetics and game theory are specified. The biological assumptions that define its framework and the key notion of fitness are explained in Section 1.1.2.1. The theory of adaptive dynamics has had many successes and two of them are detailed in Sections 1.1.2.2 and 1.1.2.3: the *Trait Substitution Sequence* (TSS) and the Canonical Equation of Adaptive Dynamics (CEAD). The time scale of the CEAD is compared with other eco-evolutionary events in Section 1.1.2.4. We also present in this section the approach followed in this thesis.

While the biological theory of adaptive dynamics is based on partly heuristic derivations, various aspects of this theory have been derived mathematically. Mathematical models of adaptive dynamics describe the population at the level of individuals. This can be done with stochastic *Individual-Based Models* (IBMs) which are described in Section 1.1.3.1. In Section 1.1.3.2, we construct our IBM for a trait-structured population with constant size and we study its behaviour under an asymptotic of large population. The derivations of this result provide the foundations of the so-called *PDE approach of adaptive dynamics* which leads to another version of the DIECKMANN-LAW’s CEAD. The heuristics of this approach will be detailed in Section 1.1.4.1. Another approach, called the *stochastic approach of adaptive dynamics* is developed in Section 1.1.4.2 and numerically illustrated in Section 1.1.3.2. Numerical illustrations of our study are also presented. In Section 1.1.4.3, some mathematical extensions to the CEAD are discussed



in different biological frameworks. In Section 1.1.4.4 we present other approaches which use new sets of combination of biological assumptions and so new time scale changes.

Although successful in predicting the evolutionary fate of ecological communities, both PDE and stochastic approaches are criticised by the biological community. Section 1.1.4.6 is devoted to explain some of these biological controversies. These controversies justify the interest of this thesis as an intermediate approach to correct them. Our approach is detailed in heuristic way in Section 1.1.4.5. We will see in Section 1.2 the mathematical tools for its implementation and Section 1.3 will be devoted to present our contributions.

### 1.1.1 Darwinian evolutionary biology

The modern theory of evolutionary biology has its origin in DARWIN's book [57], where DARWIN stated the famous basic principles of evolution, based on the principle of natural selection, later summarised by the phrase "*Survival as the fittest*". The theory of evolution through natural selection, independently also conceived and described by WALLACE [58], is based on three basic mechanisms:

- **Heredity** which allows the transmission of individual phenotypic characteristics from one generation to another. Two forms are possible:
  - ◆ by *clonal reproduction*, i.e. a mother cell gives two daughter cells with the same genetic inheritance (as long as there was no error in the process of replication). It concerns e.g. bacteria, many plants and fungi.
  - ◆ by *sexual reproduction*, i.e. the genomes of two gametes recombine and form an organism including genetic material from both gametes.
- **A source of variation** on these individual phenotypic characteristics. In clonal reproduction, *genetic mutations*<sup>1</sup> are the main source of variation [68] although *horizontal gene transfer* also exists in micro-organisms. In sexual reproduction, another possible source of variation is *gene recombination*.
- **Natural selection** which can result from inter-species (e.g. predator/prey, hosts/parasites, plant/pollinator) or intra-species (e.g. competition for resources or for reproduction) interactions. Individuals which are better adapted to their natural environment survive and reproduce more likely and thus transmit their characteristics to a larger number of descendants than the ones that are less adapted. This creates the natural selection process and implies, on a large time scale, the disappearance of the disadvantaged phenotypes and the persistence of the adapted phenotypes.

In this thesis, we will focus on clonal reproduction mechanism and consider mutations as the main source of variation.

Evolutionary mechanisms, as understood nowadays, can be represented as in Figure 1.1.

Each individual is characterised by a *genotype*, which contains the full hereditary information and encoded in the DNA. This genotype expresses into an individual *phenotype* during the developpement and through the metabolism of this individual (see  $\xrightarrow{(a)}$  in Figure 1.1). The phenotype or *trait* of an individual describes his observable characteristics, such as morphological or physiological characteristics (e.g. body size, rate of food intake, age at maturity). Individuals of a given population, reduced to their phenotype, interact with each other, they compete e.g.

---

1. (Genetic) mutation: Permanent alteration of the genome of an organism and results e.g. from errors of copy during a cell division.

for resources (space, nutrients, energy, information) or with other species (predator/prey, host/parasite). This competition process affects each individual's survival and reproduction ability and thus induces a *population dynamics* (see  $\xrightarrow{(b)}$ ). From this competition follows a natural selection process (see  $\xrightarrow{(c)}$ ) which acts on the phenotypes (and not on the genotypes) (see  $\xrightarrow{(d)}$ ) and is a consequence of the differences between each individual's reproductive efficiency. Note that the population dynamics depends on the *outer environment* (see  $\xrightarrow{(e)}$ ), e.g. through natural resources or other species, which itself depends of the population interacting with the other species or exploiting these resources. Thus, adaptation of phenotypes depends on the outer environment and on the composition of the rest of the population. Population dynamics also affects the frequencies of genotypes of a population through the mechanisms of reproduction and mutation in the genome of the newborn individual (see  $\xrightarrow{(f)}$  and  $\xrightarrow{(g)}$ ). Lastly, the environment of an individual (including the other individuals of the population) may also have an effect on the translation of its genotype to its phenotype during its development (epigenetics) (see  $\xrightarrow{(h)}$ ).

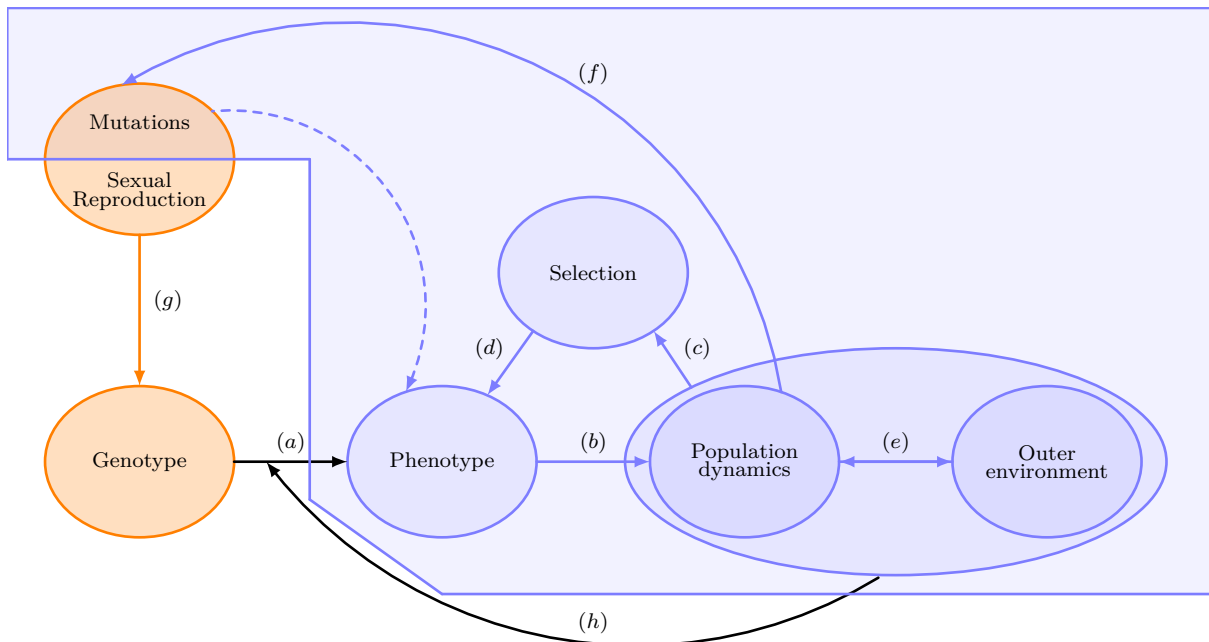


FIGURE 1.1 – Diagram of evolutionary mechanisms.

In view of this diagram of evolutionary mechanisms in Figure 1.1, simplifications and approximations have been proposed according to several approaches to study evolution: population genetics, evolutionary game theory and adaptive dynamics. The last one, providing the biological background of this thesis, is treated in details in Section 1.1.2 and its modelling can be represented by the blue part of Figure 1.1. Note that the theory of adaptive dynamics does not take into account sexual reproduction and the translation of genotype into a phenotype.

- **Population genetics.** The theory of *population genetics* insists on the mechanisms of sexual reproduction. More precisely, it is interested in the genetic differences within a population and studies changes in the allele frequencies from one generation to another but does not take into account the ecological interactions between individuals and with their environment. This simplification makes it possible to model a realistic inheritance law in order to understand the complex patterns of genetic variation and not consider

explicitly phenotypes in the models [95]. It corresponds to the orange part of Figure 1.1. Population genetics was introduced at the beginning of 20<sup>th</sup> century in the fundamental works of FISHER [98], HALDANE [122] and WRIGHT [214] which integrate the inheritance principles of MENDEL in the Darwinian theory of natural selection.

In this approach, the population size is often assumed *constant* and the *evolutionary advantage* of the individual, usually called *fitness*, is often an *a priori* given quantity which depends only on genotypes. This allows to ignore the effects of the outer environment and the rest of the population [95]. In [214], WRIGHT introduced the concept of a fixed *adaptive landscape*, described by the fitness, on which natural selection drives populations to reach a (local) maximum. We will come back to this idea in Section 1.1.2.3 (see also Section 1.1.2.1 in the adaptive dynamics context).

- **Evolutionary game theory.** An alternative approach consists in the *evolutionary game theory*. It ignores genes and sexual reproduction mechanisms to focus on studying phenotypic evolution in some interacting environment. This theory has its source in the works of MAYNARD SMITH and PRICE [165]. In this context of game theory, the players are individuals of a population, the strategies are their heritable phenotypes, the payoff is their evolutionary advantage in this environment and the population state changes according to the rule of a certain game. Here, evolutionary advantage or *fitness* of an individual depends on its phenotype and on the frequencies of the other phenotypes currently present in the population [163]. Unlike population genetics, fitness is not a given constant quantity but depends on the population in which the individuals live and interact with.

This theory studies successive invasions of ecological strategies i.e. analysing whether an alternative mutant strategy or rather phenotype that is initially rare can invade the current population state [126, 163]. An important particular strategy studied is the so-called *Evolutionary Stable Strategy* (ESS). If all individuals in the population adopt the ESS, then no initially rare mutant strategy has a higher selective advantage (payoff) and can thus invade the population [163]. In other words, natural selection alone is sufficient to prevent that mutant strategies successfully invade such a strategy. Evolutionary game theory has been able to provide significant insight, with the idea of ESS, to explain many complex and challenging aspects of biology such as the evolution of altruism in the context of Darwinian evolution [123]. Note that the notion of ESS will be discussed in more detail in Section 1.1.2.3 in the context of adaptive dynamics.

### 1.1.2 Biological context of adaptive dynamics

Since the 1990s, a new theoretical branch of evolutionary biology has been developed significantly for modelling phenotypical evolution in various complex ecological systems. This recent approach, known as the theory of *adaptive dynamics*, grew out of the early work of evolutionary game theorists. It has its origin in the three fundamental papers of HOFBAUER and SIGMUND [126], MARROW ET AL [161] and METZ ET AL [172]. The following works of METZ ET AL [171], DIECKMANN and LAW [72] and GERITZ ET AL [112, 114], in the second half of the 1990s, which integrates and expands the methods of evolutionary game theory [68] form the foundation of the biological framework of adaptive dynamics: biological assumptions, methods, models and graphical tools are introduced. The most detailed presentation of the basis and main consequences of the theory of adaptive dynamics can be found in the fundamental article [171]. Note that the introduction to adaptive dynamics, given in [36, 38, 5], served as basis for parts of this section.

### 1.1.2.1 Biological assumptions, fitness and vocabulary

The goal of this section and the following is to present the main results of the theory of adaptive dynamics. Here, we explain the biological ideas and assumptions needed in this framework, and we introduce the notions of fitness, invasion, and the principle of *invasion implies fixation*. We will try to define as rigorously as possible the key biological notions, while remaining sufficiently general to cover various situations. In particular, the following definitions must not be understood in a formal mathematical way. In the next two sections, we will try to present carefully and rigorously the TSS model and the CEAD and its equilibrium.

**Biological assumptions.** To set the framework for adaptive dynamics, biologists make four simplifying biological assumptions denoting by **(BA)**:

- **(BA 1) Asexual population.** The population considered is asexual in order to simplify the reproduction scheme by assuming it to be clonal to pay more attention to the ecology<sup>2</sup> of the system. Mutations are assumed act directly, at the level of phenotypes, as represented in Figure 1.1 (dashed arrow in the blue part). To be visible, i.e. having a phenotypic effect, mutations have to be non-synonymous<sup>3</sup>, affecting a coding part of the genome and altering the chemical properties of the encoded protein [3], so it is biologically reasonable to consider that the mutations are rare.
- **(BA 2) Rare mutations.** This assumption is fundamental to introduce a time scale separation between an ecological scale, fast, which corresponds to population dynamics and an evolutionary scale, slower one, which corresponds to the arrival of mutations. Since mutations are rare, the selection process has enough time to eliminate disadvantageous phenotypes in the population before a mutation produces a new phenotype. Therefore, the number of coexisting phenotypes, at the same time in the population, is reduced as much as possible, leading to an *adaptive walk*.
- **(BA 3) Small mutations.** Another evolutionary assumption is to suppose that mutations have a small phenotypic effect, i.e. evolution acts slowly on *quantitative* individual characteristics.
- **(BA 4) Large population.** The size of the population considered is large. This ecological assumption allows to consider population densities and assume them to follow deterministic dynamics. However, mutations remain stochastic.

**Fitness and invasion.** Under **(BA 2)**, the only possible evolution for a monomorphic population<sup>4</sup> with trait  $x$  is the arrival of a unique mutant trait  $y$ , followed by the death or the survival of traits  $x$  and  $y$ . As long as the mutant subpopulation is small, competition between individuals with trait  $y$  can be omitted, which corresponds classically in ecology to a Malthusian growth, i.e. an exponential increase (or decrease) of the mutant population size.

For the stochastic Individual-Based Models (IBMs) as studied below (see Section 1.1.3.2), the appropriate notion of exponential growth rate, which measures quantitatively the adaptive value of the mutant trait of the corresponding population, is given by the *fitness*. This leads to the following biological definition (METZ ET AL [172]), that we will illustrate and discuss below.

**Definition 1.1.1.** We call fitness, of the trait  $y \in \mathbb{R}$  with respect to the trait  $x \in \mathbb{R}$  and we denote by  $\text{Fit}(y, x)$  the exponential growth rate of an initially small population with trait  $y$ , in a monomorphic resident population with trait  $x$ , at equilibrium.

---

2. Ecology: Study of the interactions between individuals and their environment including other individuals.

3. Non-synonymous: Change in the amino-acid chain of the protein encoded in the corresponding gene.

4. Monomorphic population: All the individuals of the population hold the same phenotypic trait.

Note that the biological originality of this notion of fitness lies in the fact that it does not depend only on the mutant trait value, as in most of the models in population genetics, but also on the resident trait value. This leads to construct a dynamical *invasion fitness landscape* which evolves along with the evolutionary change of the population (see Figure 1.7) unlike in population genetics where it remains fixed [214]. An invasion fitness landscape, in the case of a 1–dimensional continuously varying traits, is a 3–dimensional plot where the fitness function is plotted as a surface depending on two variables (see [212, Section 2.5] to observe how the population move on it). The way to construct an invasion fitness landscape depends on the ecological context [172]. Moreover, the idea of a fitness landscape co-evolving with the population is initially proposed in [161] then improved in [171, 169].

Thanks to **(BA 4)**, we can approximate the resident population dynamics by an Ordinary Differential Equation (ODE) which admits, in the IBM studied below, a unique non-null globally asymptotically stable steady state. In the previous definition, it is implicitly assumed that there exists a unique equilibrium population state associated with trait. We will see below examples where this is true. Hence, when the current population of trait  $x$  is close to this equilibrium, determined by the underlying ecological dynamics, we can consider that the fitness of a given mutant trait  $y$ , which has just appeared in this population, governs the possibility of invasion of the descendants of this mutant in the population in the following way:

- if  $\text{Fit}(y, x) < 0$ , the (deleterious) mutant trait will disappear with a very high probability, so invasion by the mutant is impossible,
- if  $\text{Fit}(y, x) > 0$ , invasion by the (advantageous) mutant is possible (but not sure since the random fluctuations of the mutant population cannot be neglected, as long as the size of its subpopulation remains small),

with an undetermined case when  $\text{Fit}(y, x) = 0$  for  $y \neq x$ . The invasion is said to be *successful* when the size of the mutant population (initially reduced to one individual) reaches a significant size compared to the size of the resident population (see Section 1.1.4.2). Once the invasion of mutant  $y$  is successful, the competition between  $x$  and  $y$  may drive both, one or none of them to extinction. Thanks to **(BA 4)**, the dynamics of this two traits is determined deterministically. When only the mutant trait survives, we will say that *fixation* has occurred. Furthermore, when the mutant trait disappears we say that the *residents strikes back* [181]. When the mutant and resident traits both persist, we say that there is *coexistence*. If these both disappear, we say that there has been *evolutionary suicide* [121, 71].

In addition, the fitness as defined above, satisfies the following biological property:

$$\forall x \in \mathbb{R}, \quad \text{Fit}(x, x) = 0. \quad (1.1)$$

Indeed, when the resident population with trait  $x$  is at its equilibrium, the growth rate of a single individual with trait  $x$  (seen as a mutant which is identical to the resident) has to be null.

**Invasion-Implies-Fixation principle.** Another common assumption in the context of studying a competitive population is called *Invasion-Implies-Fixation* (IIF) principle [113] given by

- (IIF) If a mutant’s trait invades in the population, it replaces entirely the previous resident trait such that long term coexistence is excluded.

This will allow to define an evolutionary time scale where the population is monomorphic at all times and where evolution proceeds by jumps from one phenotypic trait to another fitter one.

### 1.1.2.2 Successes and application fields

Under these biological assumptions **(BA)**, adaptive dynamics have had some successes notably for its ability to propose tools that can be adapted to analyse and predict the evolution of populations in a wide variety of ecological situations: e.g. to describe competitive or cooperative interactions between different individuals or species, or to describe predator-prey, immune-pathogen, host-parasite or plant-insect relationships. We can mention the study of the evolution of virulence in a host-pathogen population [201], of seed size [115] or of social behaviour [151]. The different tools of adaptive dynamics, used by modellers in ecology and evolution, are described in numerous references [74, 212, 68].

**TSS.** A fundamental tool in adaptive dynamics is the so-called *Trait Substitution Sequence* (TSS). This stochastic process has been introduced in [171, 72] and mathematically studied in [14, 37, 47]. This pure jump MARKOV process, with values in the space of monomorphic trait populations, allows to describe, *at the population level*, the successive invasions of advantageous mutants (also called *adaptive walks*) that invade and settle in the population. Note that Figure 1.2, called *Pairwise Invasibility Plot* (PIP) and introduced in [171], gives an illustration of this evolution phenomenon.

The heuristics leading to the TSS model [171, 72] is based on **(BA 1)**, **(BA 2)**, **(BA 4)** and (IIF) principle. Note that **(BA 2)** induces a separation between the time scale of mutations and those of births and deaths, that **(BA 4)** allows to predict deterministically the outcome of the competition between a resident trait and a mutant trait and that (IIF) principle ensures that two different traits of individuals cannot coexist, on a long time scale, because the competition eliminates one of them.

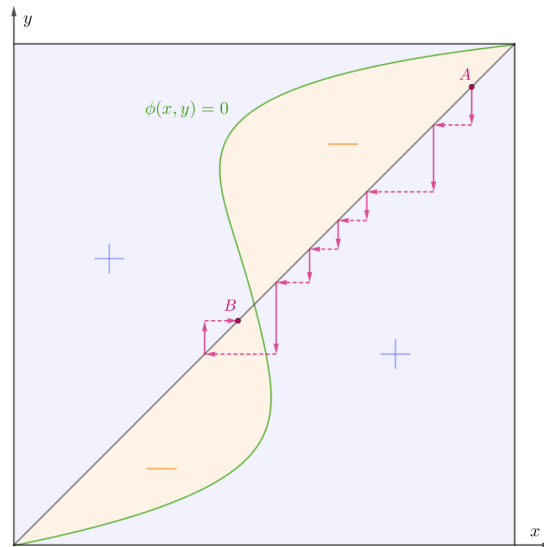


FIGURE 1.2 – Example of PIP and possible paths of the TSS monomorphic.

As seen in Section 1.1.2.1, the possibility of invasion only depends on the sign of the fitness function, so let us decompose the trait space according to the sign of  $\text{Fit}(y, x)$  as in Figure 1.2. From (1.1) and if we assume that the fitness function is of class  $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ , we can write thanks

to the implicit function theorem that

$$\text{Fit}(y, x) = (y - x)\phi(x, y) \quad (1.2)$$

where the implicit function  $\phi$  is continuous on  $\mathbb{R}^2$  and  $\mathcal{C}^1$  on  $\mathbb{R}^2 \setminus \{(x, x) \mid x \in \mathbb{R}\}$ . Hence, the zeros of  $\text{Fit}$  are those on the straight line  $y = x$  and the zeros of  $\phi$ .

Note that  $+$  and  $-$  signs in Figure 1.2 designate the sign of  $\text{Fit}(y, x)$ . We start with a resident monomorphic population with trait  $x$  in point  $A$ . For a mutant invasion with trait  $y$  occurs in this population with trait  $x$ , the point  $(x, y)$  must belong to an area where  $\text{Fit}(y, x) > 0$ . That is the meaning of the vertical arrows. The dashed horizontal arrows indicates that once invasion occurred, there is fixation of the mutant. Once fixed, the population with trait  $y$  is the new resident monomorphic population until it is replaced by a new mutant population and so on. So, the arrows represent a sequence of mutant invasions which is stopped here in  $B$  after 7 iterations.

A key point is the fact that, while the time of mutations, the mutant trait and the invasion of the mutant trait are stochastic, the direction of evolution in the TSS is deterministic. As we see above, it is established by the set of traits  $y$  such that  $\text{Fit}(y, x) > 0$  when the TSS model is in the state of a monomorphic population of trait  $x$ . Since, according to (1.1) and under **(BA 3)**, we see that only one evolution is possible locally depending on the sign of the *gradient fitness*  $\partial_1 \text{Fit}(x, x) := \frac{\partial}{\partial y} \text{Fit}(y, x)|_{y=x}$  except when this derivative cancels. Indeed,

- if  $\partial_1 \text{Fit}(x, x) > 0$ , then  $\text{Fit}(y, x) > 0$  for all  $y > x$  in a neighbourhood of  $x$ , and the evolution (from the mutant point of view) can only proceed towards greater traits which is represented by the single arrow  $\uparrow$  in Figure 1.2,
- if  $\partial_1 \text{Fit}(x, x) < 0$ , then  $\text{Fit}(y, x) < 0$  for all  $y < x$  in a neighbourhood of  $x$ , and the evolution can only proceed towards lesser traits which is represented by all the arrows  $\downarrow$  in Figure 1.2.

Note that the points  $x^*$  such that  $\partial_1 \text{Fit}(x^*, x^*) = 0$  are called *evolutionary singularities* which will review in Section 1.1.2.3. Note that from (1.2), we have  $\phi(x^*, x^*) = 0$ . In other words,  $(x^*, x^*)$  is an intersection point between the curve  $\phi(x, y) = 0$  and the line  $y = x$  as we can see in Figure 1.2.

The approximation of the population dynamics by the TSS under the additional assumption **(BA 3)** constitutes the foundation of another biological concepts: the *Canonical Equation of Adaptive Dynamics* (CEAD). This important advance of adaptive dynamics will be developed biologically in Section 1.1.2.3 and mathematically in Section 1.1.4.

The TSS model and its generalisations have revealed a powerful tool for understanding various evolutionary phenomena, such as *polymorphism* or *evolutionary branching*. By polymorphism we mean the stable coexistence of different phenotypes, which arises if the (IIF) principle is not assumed [171, 72, 47]. Note that the TSS model has been generalised in different contexts, as trait-structured with horizontal gene transfer [18], age-structured population [209], multi-resources chemostat population [43], predator-prey [55], population structured in phenotypic trait under selection and neutral (genetic) marker [19], in a spatial context [154] or populations with a dormancy mechanism [24, 25].

**Evolutionary branching and sympatric speciation**<sup>5</sup>. Evolutionary branching is a di-

5. Speciation: Birth of two daughter species from a single mother species.

versification phenomenon where selection pressures push a population, initially concentrated around a single dominant phenotype, to split into two interacting subpopulations of different dominant phenotypes. This phenomenon is firstly described in detail in [171] (see also [112, 114]). Moreover, METZ ET AL identify the points in the phenotype trait space, where the evolutionary branching is likely to happen, the so-called *evolutionary singular strategies*, and give a criterion for evolutionary branching depending on the derivatives of the fitness function at these points. Note that the concept of evolutionary singular strategies can be seen as a generalisation of the ESS concept introduced in the context of evolutionary game theory [114].

The success of adaptive dynamics is also due to the fact that the mechanism of diversification linked to evolutionary branching provides new insights into the fundamental issue on the origin of species. Indeed, the most widely accepted speciation scenario assumes that a species is divided into two geographically separated subpopulations (e.g. following the appearance of a geographical barrier or the colonisation of a new habitat): this is *allopatric speciation*. The opposite scenario of *sympatric speciation* assumes that reproductive isolation takes place without geographical separation and through various mechanisms related to natural selection (e.g. evolution of reproductive preferences according to sexual traits). Although described in the context of asexual reproduction, evolutionary branching has been proposed as a mechanism that may cause sympatric speciation [171, 70, 76, 47].

**Developments/Extensions.** The tools of adaptive dynamics have received numerous developments and extensions, among which we can mention the study of the dynamics of interacting subpopulations, in order to determine, for instance, whether a mutant type, although invasive, becomes extinct [113, 111] or whether the invasion of the mutant leads to the extinction of the population by *evolutionary suicide* [121, 67, 96]. In [65], it is presented the first application of adaptive dynamics outside the biological framework, by studying an economic system in which traits measure technological characteristics of products, mutations represent product innovations, selection is exerted by customers, and the environment is the market in which products compete. Other extensions of adaptive dynamics standard models are proposed e.g. to study the evolution of cancer under treatment [7]. Note that É. KISDI has compiled a list of references on the developments and applications of adaptive dynamics at <https://www.mv.helsinki.fi/home/kisdi/addyn.htm> whose the application part includes nearly 150 references.

However, these various successes remain to be qualified because the biological assumptions that justify the approximations of the adaptive dynamics are discussed and controversial (see [212]). We will come back to this in Section 1.1.4.6 to motivate the problem we study in this thesis.

### 1.1.2.3 Canonical equation

The goal of this section is to introduce the central object of this thesis: the CEAD. Firstly, we present the historical approach leading to this equation in the context of adaptive dynamics, we explicit its expression and discuss its equilibrium states. Then, we compare the CEAD to other approaches in quantitative evolutionary biology context.

**In adaptive dynamics context.** The CEAD was introduced heuristically by DIECKMANN and LAW in [72] in order to provide a description, in the long term, of the evolutionary dynamics of the dominant trait in a monomorphic competitive population as a solution of an ODE. This



ODE can be obtained as an approximation of the TSS model, in the monomorphic case, under the additional assumption **(BA 3)** (i.e. the trait of the mutant is only slightly different from the trait of its progenitor). So, the heuristic leading to the CEAD is based on the biological assumption **(BA)** and the (IIF) principle.

$$\text{CEAD} = \boxed{\text{Fitness gradient}} \times \boxed{\text{Mutations effects}}$$

FIGURE 1.3 – Generic form of the CEAD.

This equation consists in two terms and is summarised in Figure 1.3. One term is a fitness gradient describing the strength of selection that pushes the population to increase its fitness locally. A second term which describes the effect of mutations. Assume that for any  $x \in \mathbb{R}$ ,  $p(x)$  is the probability that a mutation occurs in a birth, from an individual with trait  $x$ , according to the mutation law  $m(x, \cdot)$  and  $\bar{n}_x$  the equilibrium density of a monomorphic population with trait  $x$  (when there is no mutation). Then, the ODE of [72] writes, in the case of a monomorphic initial population of trait  $x_0$ ,

$$\forall t > 0, \quad \frac{dx}{dt} = \partial_1 \text{Fit}(x, x) \times \frac{\bar{n}_x p(x)}{2} \int_{\mathbb{R}} h^2 m(x, dh), \quad (1.3)$$

with  $x(0) := x_0$  where  $\int_{\mathbb{R}} h^2 m(x, dh)$  is the variance of the mutation probability  $m(x, \cdot)$  which is assumed *symmetrical* on  $\mathbb{R}$  with respect to 0 for any  $x \in \mathbb{R}$ . Let us insist on the fact that, the evolutionary rate  $dx/dt$  does not point in the direction that is the best for the whole population, but in the direction that is the best from the mutant point of view [65]. The factor  $1/2$  is a scaling factor that takes into account that only half of the mutations have an effect on the variation of  $x$  (those in the positive area of the fitness gradient).

DIECKMANN-LAW's approach provided the conceptual background for modelling population dynamics as a MARKOV process that incorporates reproduction, mutation and selection where [14] is the first reference which provides a rigorous mathematical formalism.

For a 1-dimensional trait space, any solution of the canonical equation (1.3) must converge in a long time to an equilibrium state which corresponds to an evolutionary singularity. Once the TSS has reached the neighbourhood of such an evolutionary singularity, biologists have classified the different possible behaviours of the mutation invasion dynamics [171, 74]. These results generalise of the game theoretical approach of evolution as presented above. In particular, they extend the notion of ESS and describe the evolutionary transition leading to these equilibria, so that one can predict in which ESS the population is going to settle. The type of evolutionary singularities (ESS, Convergence Stable Strategy (CSS), Mutual invisibility or Evolutionary Branching) is characterised in terms of the derivatives of the fitness function.

For example,  $x^*$  is an *ESS* if no other trait  $y$  in a neighbourhood of  $x^*$  can invade the equilibrium resident population with trait  $x^*$ , i.e. for any  $y$  close enough to  $x^*$ ,  $\text{Fit}(y, x^*) < 0$ , i.e. the function  $y \mapsto \text{Fit}(y, x^*)$  admit a local maximum at  $x^*$ . Therefore,  $x^*$  is an ESS if  $\partial_{11}^2 \text{Fit}(x^*, x^*) < 0$  (with an undetermined case when  $\partial_{11}^2 \text{Fit}(x^*, x^*) = 0$ ). Conversely, note that if  $\partial_{11}^2 \text{Fit}(x^*, x^*) > 0$  then  $y \mapsto \text{Fit}(y, x^*)$  admits a local minimum and so any mutant trait  $y$  sufficiently close to  $x^*$  can invade the population with trait  $x^*$ .

Note that an ESS corresponds to an *absorbing state* for the monomorphic TSS model, so, in this sense, the expression “evolutionary stable strategy” is somewhat incorrect.

**In quantitative evolutionary biology context.** The canonical equation proposed in [72] is a derivation of those given in quantitative evolutionary biology. In this context, canonical equations are well-known tools used to predict the evolutionary fate of ecological communities.

- **Quantitative genetics.** In quantitative genetics approach, there exists a structurally similar canonical equation for sexually reproducing populations, introduced by LANDE [148] and later extended to frequency-dependent selection [128, 211, 1]. The canonical equation of evolution, obtained by LANDE as a limit of quantitative genetics models, is based on the principle of maximisation of the fitness with respect to its gradient (see WRIGHT [214]). The general form of LANDE’s canonical equation is

$$\frac{dx}{dt} = \nabla W(x) \times k(x), \quad (1.4)$$

where  $W(x)$  is the fitness of trait  $x$  and  $k$  a non-negative coefficient which describes the speed of evolution and is proportional to *genetic variance*<sup>6</sup>. This equation describes the long term Darwinian evolution of the mean phenotypic trait of biological population driven by selection. The main difference between (1.3) and (1.4) lies on the fact that the fitness in (1.3) depends on the mutant and resident traits while in (1.4) it depends only on the mutant trait. Hence a *fixed* fitness landscape in quantitative genetics approach.

- **Evolutionary game theory.** In the context of evolutionary game theory, Darwinian evolution of quantitative traits in several co-evolutionary scenarios can be modeled as a gradual process given by the solution of the following ODE which takes the form in the 1–dimensional case

$$\frac{dx}{dt} = \partial_1 W(x, x) \times k(x), \quad (1.5)$$

where  $W(y, x)$  is the fitness of individual with trait  $y$  in the environment determined by the resident trait  $x$ , and  $k$  a non-negative coefficient which scales the rate of evolutionary change. Dynamics of this kind have been proposed to approximate the evolutionary dynamics of dominant strategies and studied by many authors e.g. as a hill-climbing process on a adaptive landscape (or adaptive walks) [30, 31, 125, 126, 32] (see also [72] and references therein).

Moreover, unlike (1.4) and (1.5) the CEAD (1.3) gives explicit expressions for the fitness and the coefficient  $k(x)$  as functions of individual parameters.

#### 1.1.2.4 Adaptive dynamics: a multiscale phenomenon and the approach followed in this thesis

The goal of this section is first to situate, on a graduated time axis, the phenomenon of the CEAD in relation to other eco-evolutionary events as represented in Figure 1.4. This representation corresponds to the probabilistic approach described in Section 1.1.4.2. In a second part, the approach followed in this thesis will be explained.

---

6. Genetic variance: Part of phenotypic variance (related to a quantitative phenotype). Genetic variance has three major components: the additive genetic variance, dominance variance, and epistatic variance.

**Multiscale phenomenon.** We denote  $\mathbb{N} := \{0, 1, 2, \dots\}$  and  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ . In order to graduate a time axis, let us introduce three *scaling parameters* in relation to the three biological assumptions **(BA 2)**, **(BA 3)** and **(BA 4)** made above:

- $p \in [0, 1]$  is equal to the occurrence probability of a mutation which arises during life at the individual level. Note that  $p \rightarrow 0$  means the mutations are *rare* while  $p = 1$  they are *frequent*.
- $\sigma \in (0, +\infty)$  is linked to the size of a mutation. Note that  $\sigma \rightarrow 0$  means mutations have a *small phenotypic effect*, i.e. evolution acts slowly on the individual phenotypic characteristics.
- $K \in \mathbb{N}^*$  is linked to the size of the population. Note that  $K \rightarrow +\infty$  means a *large population*. We assume that the size of the initial population is proportional to  $K$ .

Hence, the biological framework of adaptive dynamics requires the asymptotics  $K \rightarrow +\infty$ ,  $p \rightarrow 0$  and  $\sigma \rightarrow 0$ .

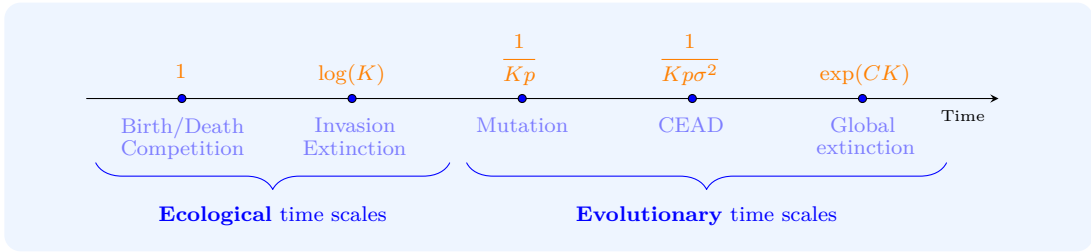


FIGURE 1.4 – Time scales involved in adaptive dynamics in the stochastic approach.

Adaptive dynamics is a multiscale phenomenon involving ecological and evolutionary time scales. The ecological time scale corresponds to interactions between individuals and is visible on relatively short time. In particular, birth, death and competition events of each individual take a time of the order of 1 to occur. Note that thanks to **(BA 4)** and once the invasion of the mutant is successful, the competition between the mutant and resident populations takes a time  $O(1)$ . If we accelerate the time, we can observe on the time scale  $\log(K)$ , the *successful invasion* of an advantageous mutant in the resident monomorphic population but also the *extinction* of this resident population in favour of this mutant population. Therefore, the successive replacement of resident trait by advantageous mutants take a time  $O(\log(K))$ .

Thanks to **(BA 2)**, the ecological time scale is separate from the evolutionary one, which is visible over very long times as mutations events. Indeed, the population size being proportional to  $K$  and mutations arising during life of each individual at rate  $p$ , the overall rate of appearance of a mutant is proportional to  $Kp$  and satisfying  $Kp \rightarrow 0$  when  $K \rightarrow +\infty$  and  $p \rightarrow 0$ . So, the renormalisation of time  $1/Kp$  corresponds to the time scale of mutations. Since mutations are rare, the selection process has enough time to discriminate between phenotypes in the population selecting the best adapted to survive before the arrival of a new mutation. In other words, the phase of replacement of resident trait by a mutant trait takes place on a infinitesimal time scale with respect to the mutation time scale. Hence, the time scale separation assumption  $\log(K) \ll 1/Kp$  is the suitable mathematical formulation of **(BA 2)**. Note that, the TSS model acts on the mutation time scale  $1/Kp$ . Thanks to **(BA 3)**, scaling the size of the TSS's jumps by  $\sigma$ , inducing a rescaling of time by  $1/\sigma^2$ , allows to observe the phenomenon of the CEAD on the evolutionary time scale  $1/Kp\sigma^2$ . Finally, on an exponential time scale, the global extinction of the population can be observed, which corresponds in a language of large deviation and domain

exit [104, 64], to the time scale of the exit from a stable equilibrium neighbourhood.

**Our approach.** As will be seen in more details in Section 1.1.4.6, the previous approach is criticised because of *too rare mutations*. In this thesis, our goal is to correct this biological controversy by studying mathematically, under a double asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ) the consequences of a new biological assumption of *frequent mutations* ( $p = 1$ ) on the canonical equation. More precisely, the goal is to determine from our stochastic IBM, the long term behaviour of the mean phenotypic trait of the population. To do this, we will assume  $\sigma \ll K^{-2}$  and we will prove that the time scale to observe the phenomenon of the CEAD is in  $1/K\sigma^2$ . Recall that, in the approach presented above, the time scale for the CEAD is in  $1/Kp\sigma^2$ . In our case  $p = 1$ , so there is some consistency between this approach and ours. In addition, both canonical equations are the same. However, note that the one we propose is biologically more reasonable because it takes place on a shorter evolutionary timescale.

As we will see in Section 1.1.4.5, our goal can be formulated into a *slow-fast asymptotic analysis* acting on two eco-evolutionary time scales. A *slow scale* corresponding to the dynamics of the mean trait, driven by the CEAD, and acting on  $1/K\sigma^2$ . A *fast scale* corresponding to the evolutionary dynamics of the centered and dilated distribution of trait and acting on  $K$ . This slow-fast asymptotic analysis is based on *averaging techniques* and requires the *identification* and *characterisation* of the long term behaviour of the fast component. This latter is non-classical and corresponds to that of an original measure-valued diffusion which we interpret as a *centered FLEMING-VIOT process*.

Before stating these results in Section 1.3, we will introduce the notion of IBM in Section 1.1.3.1 and construct our model in Section 1.1.3.2. Then in Section 1.1.4 we will present the two main mathematical approaches leading to the CEAD and various extensions. Thanks to the heuristic of the stochastic approach, given in Section 1.1.4.2, we will obtain an explicit expression of the CEAD that we expect in our situation. The proof of this result is based on FLEMING-VIOT processes and averaging techniques. Section 1.2 is dedicated to the introduction of these two notions (FLEMING-VIOT processes in Section 1.2.1 and averaging techniques in Section 1.2.2).

### 1.1.3 Mathematical context

In this thesis, we study the evolution of an asexual reproducing population that is composed of a finite number of individuals, each of them characterised by a 1-dimensional trait and taking into account the interaction between the individuals. This can be done with a stochastic IBM which we introduce in this section. The goal of Section 1.1.3.1 is to present this notion, many examples of its use in the case of populations whose size can fluctuate and then in those of constant size. The role played by Section 1.1.3.2, is to lay the foundations of our study by constructing our stochastic model. This section is also devoted to study the behaviour of our IBM under the asymptotic of large population in Theorem 1.1.2 which allows to compute explicitly the fitness function. Finally, we propose some numerical simulations to illustrate, in particular, the convergence of the IBM to the CEAD by testing several sets of parameters.

#### 1.1.3.1 Individual-based models

While the biological theory of adaptive dynamics is based on partly heuristic derivations, various aspects of the theory have been derived rigorously over the last years in the context of *stochastic IBM*. An IBM is a system of interacting particle modelling Darwinian evolution at the individual level, so including the three basic mechanisms of evolution: *heredity*, *mutation*,

and *selection*. Such models involve a finite *structured population* in which the set of birth, death, competition and mutation events are described for each individual. Note that, *structured populations* are populations in which individuals differ according to variables that affect their survival or reproductive capacities. These variables can be phenotypic, genotypic or behavioral *traits* that are assumed to be hereditarily transmitted from a parent to its offspring unless a mutation occurs. They can also be position, sex or age [143].

Let us insist on the fact that IBMs describe all events at a *microscopic* level and allow to give a realistic description of these phenomena and to obtain *macroscopic* equations, at a large population scale.

IBMs were first introduced in ecology as a tool to describe local interactions or complex phenomena at the individual level [186, 26, 27, 77, 118, 63]. Ecological studies using IBMs are mainly numerical and the models are mostly posed in discrete space as systems of interacting particles [186] and more rarely in continuous space [26, 27, 77]. Many IBMs (with non-constant population size) have been proposed and studied in the context of Darwinian evolutionary by the biology community [171, 73, 96] and the mathematical community [37, 39, 10]. Others are dispersal models in spatially structured populations where the trait is viewed as a spatial location and mutations as dispersal [26, 27, 149, 102]. Other models, structured in age, were developed in [50] and studied mathematically in [187, 209], or structured in age and traits in [97, 166, 167]. Note that adding an age-structure allows to study life history traits, such as age at maturity, or to consider senescence phenomena, e.g. to describe the fixation of deleterious genes or the decrease of reproductive and survival capacities in the course of life [166].

Other IBMs consider population with *fixed* population size, so that births and deaths occur simultaneously in so-called *resampling* (or swap phenomenon) events. The mutation and resampling rate of an individual depends on its phenotype. When a mutation occurs, the new mutant trait is close to its parent's one yielding a slow variation of the trait. In population genetics, the WRIGHT-FISHER model (and its extensions with selection, mutation or immigration), CANNINGS model or the MORAN model are interested in the evolution of allele frequencies according to various mechanisms [86]. In [99, 61, 60, 85], the authors construct the FLEMING-VIOT process as a scaling limit of large population from the MORAN process. Several extensions exist, including frequency-dependent selection, recombination, other reproduction mechanisms [91, 117]. In particular, the last article provides a bridge between population genetics and eco-evolutionary models. Others, based on KERMACK-MCKENDRICK's model, are interested in epidemiological questions [180, 182]. Apart from studying the *forward* going evolutionary process, as the classical theory surrounding the MORAN model does, it is also common to focus on the retrospective analysis, i.e. to look *backwards* in time e.g. the most recent common ancestor using the KINGMAN *coalescent* process [95, 86].

### 1.1.3.2 Construction of the individual-based model

In this thesis, we are interested in modelling, at the individual level, the adaptive evolution of a trait-structured population, describing the capacity of a population to generate and select trait diversity. More precisely, the survival and reproductive abilities of each individual is characterised by a continuous quantitative phenotypic trait  $x \in \mathbb{R}$ , i.e. an overall characteristic subject to selection such as body size at maturity.

**Parameters of the model.** We consider a discrete population of constant size  $K$  in continuous time. The individuals reproduce *asexually* during their lives, i.e. the reproduction scheme is assumed *clonal* and die with *frequency-dependent* rates i.e. depending on the trait. Each birth

of an individual occurs simultaneously with the death of another individual in the population. A mutation occurs during life at the individual level. Each mutation has an amplitude of the order of magnitude  $\sigma \in (0, +\infty)$ .

We are interested in approximating the long term dynamics of a large population. We assume that the number of individuals alive at each time  $t \geq 0$  is always equal to  $K$ . Let us denote by  $x_1(t), \dots, x_K(t)$ , the phenotypic trait values of these individuals at time  $t$ . The state of the population at time  $t$ , can be described by the following finite point measure on  $\mathbb{R}$  rescaled by  $K$

$$\nu_t^{K,p,\sigma} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i(t)}$$

where  $\delta_x$  is the DIRAC measure at  $x$ .

*Biological parameters.* For all  $x, y \in \mathbb{R}$ , let us introduce the following biological parameters:

- $b(x, y) \in \mathbb{R}_+$  is the *resampling* rate, i.e. the rate of simultaneous birth of an individual holding trait  $y$  and death of an individual holding trait  $x$  and it can be interpreted as modelling a competitive pressure driving selection.
- $\theta(x) \in \mathbb{R}_+$  is the mutation rate of an individual holding trait  $x$ .
- $m(x, dh)$  is the mutation law of the scaled mutation step  $h$ , born from an individual with trait  $x$ , where the mutant trait is given by  $y := x + \sigma h \in \mathbb{R}$ . It is a probability measure on  $\mathbb{R}$ .
- $\text{Fit}(y, x) := b(x, y) - b(y, x)$  is the adaptive value or *fitness* of a mutant individual with trait  $y$  in the population of  $(K - 1)$  individuals of trait  $x$ . Recall that the fitness can be interpreted as the initial growth rate of a mutant individual  $y$  in a resident monomorphic population with trait  $x$ . In Section 1.1.3.2 and Remark 1.1.5, two distinct strategies are given to recover the value of this fitness.

Let  $\mathcal{M}_1(\mathbb{R})$  be the set of probability measures on  $\mathbb{R}$ , endowed with the topology of weak convergence making it a Polish space [20]. We define for all  $K \in \mathbb{N}^*$ ,

$$\mathcal{M}_{1,K}(\mathbb{R}) := \left\{ \frac{1}{K} \sum_{i=1}^K \delta_{x_i} \mid (x_i)_{1 \leq i \leq K} \in \mathbb{R}^K \right\} \quad (1.6)$$

the set of probability measures on  $\mathbb{R}$  of  $K$  atoms of mass  $1/K$ . For all  $\ell \in \mathbb{N}$ , we denote by  $\mathcal{C}^\ell(E, \mathbb{R})$  the space of functions of class  $\mathcal{C}^\ell$  from  $E$  to  $\mathbb{R}$  and by  $\mathcal{C}_b^\ell(E, \mathbb{R})$  the space of bounded functions of class  $\mathcal{C}^\ell(E, \mathbb{R})$  with bounded derivatives where  $E$  is a metric space.

Let us now give the infinitesimal generator  $L^{K,p,\sigma}$  of the  $\mathcal{M}_{1,K}(\mathbb{R})$ -valued MARKOV process  $\nu^{K,p,\sigma} := (\nu_t^{K,p,\sigma})_{t \geq 0}$  describing the ecological dynamics of the population with resampling. The generator  $L^{K,p,\sigma}$  is defined for any bounded measurable map  $\phi$  from  $\mathcal{M}_{1,K}(\mathbb{R})$  to  $\mathbb{R}$ , by

$$\begin{aligned} L^{K,p,\sigma} \phi(\nu) &= K \int_{\mathbb{R}} \nu(dx) \int_{\mathbb{R}} \nu(dy) b(x, y) \left[ \phi \left( \nu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right) - \phi(\nu) \right] \\ &+ K \int_{\mathbb{R}} \theta(x) \nu(dx) \int_{\mathbb{R}} [(1-p) \delta_0(dh) + p m(x, h) dh] \left[ \phi \left( \nu - \frac{\delta_x}{K} + \frac{\delta_{x+\sigma h}}{K} \right) - \phi(\nu) \right]. \end{aligned} \quad (1.7)$$

The first term describes the resampling event of one individual by another and the second term describes the effect of mutations over the lifetime of individuals. Note that [102] can be

easily adapted to give the construction of the process  $(\nu_t^{K,p,\sigma})_{t \geq 0}$  with POISSON point measures whose generator is precisely given by (1.7) under the following assumptions, as recalled below.

*Assumptions.* Let us denote by **(A)** the following two assumptions:

**(A1)** The maps  $b : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ , and  $\theta : \mathbb{R} \rightarrow \mathbb{R}_+$  are respectively in  $\mathcal{C}_b^2(\mathbb{R}^2, \mathbb{R})$  and  $\mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  and there exists  $0 < \underline{b}, \bar{b}, \underline{\theta}, \bar{\theta} < \infty$  such that:

$$\underline{b} \leq b(\cdot, \cdot) \leq \bar{b}, \quad \text{and} \quad \underline{\theta} \leq \theta(\cdot) \leq \bar{\theta}.$$

**(A2)** There exists a measurable function  $\bar{m} : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$\forall x \in \mathbb{R}, \forall h \in \mathbb{R}, m(x, h) \leq \bar{m}(h) \quad \text{and} \quad \int_{\mathbb{R}} \bar{m}(h) dh < \infty.$$

*Construction of  $\nu^{K,p,\sigma}$  with POISSON point measures.* Let us recall this construction. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space large enough to contain the following three independent random objects:

**(1)** A  $\mathcal{M}_{1,K}(\mathbb{R})$ -random variable  $\nu_0^{K,p,\sigma}$ : the *initial condition*.

**(2)** A POISSON point measure  $N_1(ds, di, dj, d\alpha)$  on  $[0, +\infty) \times \mathbb{N}^* \times \mathbb{N}^* \times [0, 1]$  with intensity measure

$$q_1(ds, di, dj, d\alpha) = \frac{1}{K} \bar{b} ds \sum_{k=1}^K \delta_k(di) \sum_{\ell=1}^K \delta_\ell(dj) d\alpha :$$

the *resampling POISSON point measure*.

**(3)** A POISSON point measure  $N_2(ds, di, dh, d\alpha)$  on  $[0, +\infty) \times \mathbb{N}^* \times \mathbb{R} \times [0, 1]$  with intensity measure

$$q_2(ds, di, dh, d\alpha) = \frac{1}{K} \bar{\theta} ds \sum_{k=1}^K \delta_k(di) \bar{m}(h) dh d\alpha :$$

the *mutation POISSON point measure*.

In order to solve the enumeration problem of the population traits, let us introduce the function  $H = (H^1, \dots, H^k) : \mathcal{M}_{1,K}(\mathbb{R}) \rightarrow \mathbb{R}^K$  defined by  $H\left(\frac{1}{K} \sum_{k=1}^K \delta_{x_k}\right) = (x_{\zeta(1)}, \dots, x_{\zeta(K)})$ , where  $x_{\zeta(1)} \preceq \dots \preceq x_{\zeta(K)}$  for the lexicographic order on  $\mathbb{R}$ .

Let us introduce the process  $\nu^{K,p,\sigma}$  with values in the SKOROHOD space  $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_{1,K}(\mathbb{R}))$  of càdlàg functions of  $\mathbb{R}_+$  in  $\mathcal{M}_{1,K}(\mathbb{R})$ , defined as

$$\begin{aligned} \nu_t^{K,p,\sigma} = & \nu_0^{K,p,\sigma} + \frac{1}{K} \int_0^t \int_{\{1, \dots, K\}} \int_{\{1, \dots, K\}} \int_0^1 \left[ \delta_{H^j(\nu_{s^-}^{K,p,\sigma})} - \delta_{H^i(\nu_{s^-}^{K,p,\sigma})} \right] \\ & \times \mathbb{1} \left\{ \alpha \leq \frac{b(H^i(\nu_{s^-}^{K,p,\sigma}), H^j(\nu_{s^-}^{K,p,\sigma}))}{\bar{b}} \right\} N_1(ds, di, dj, d\alpha) \\ & + \frac{1}{K} \int_0^t \int_{\{1, \dots, K\}} \int_{\mathbb{R}} \int_0^1 \left[ \delta_{H^i(\nu_{s^-}^{K,p,\sigma}) + \sigma h} - \delta_{H^i(\nu_{s^-}^{K,p,\sigma})} \right] \\ & \times \mathbb{1} \left\{ \alpha \leq p \frac{\theta(H^i(\nu_{s^-}^{K,p,\sigma}))}{\bar{\theta}} \times \frac{m(H^i(\nu_{s^-}^{K,p,\sigma}), h)}{\bar{m}(h)} \right\} N_2(ds, di, dh, d\alpha). \end{aligned} \tag{1.8}$$

The relation (1.8) describes how to simulate exactly the process  $\nu^{K,p,\sigma}$  (see [10, Section 6.4] and Section 1.1.3.2): the POISSON point measure associated with each type of jumps more

often than the process  $\nu^{K,p,\sigma}$  should. We then decide whether a jump occurs by comparing  $\alpha$  to a quantity related to the rate of the corresponding event.

In similar way to [102, Proposition 2.6], using ITÔ's formula for jump process, we can prove under **(A)** that a process solution of (1.8) has (1.7) for infinitesimal generator. Moreover, in similar way to [102, Theorem 3.1], using GRONWALL's lemma and denoting by  $\mathbf{1}$  the constant function equal to 1, we can prove that if  $\mathbb{E}(\langle \mathbf{1}, \nu_0^{K,p,\sigma} \rangle) < \infty$ , a such process satisfies for all  $T > 0$ ,  $\mathbb{E}(\sup_{t \in [0, T]} \langle \mathbf{1}, \nu_t^{K,p,\sigma} \rangle) < \infty$ . We can prove in this case that there is no explosion of the jump process, which makes it possible to deduce the existence and the uniqueness of the solution  $\nu^{K,p,\sigma}$  of (1.8).

**Limit of large population.** For a measurable real bounded function  $f$ , and a measure  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote  $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x)\nu(dx)$ . We fix  $K \in \mathbb{N}^*$ ,  $p \in [0, 1]$  and  $\sigma \in (0, +\infty)$ .

First of all, we begin to give, for all  $f : \mathbb{R} \rightarrow \mathbb{R}$ , the semi-martingale decomposition of  $\langle f, \nu_t^{K,p,\sigma} \rangle$  and a strong solution to a integro-differential equation satisfied by the deterministic limit process of  $(\nu_t^{K,p,\sigma})_{t \geq 0}$  when  $K \rightarrow +\infty$  with  $p$  and  $\sigma$  fixed.

Adapting [102, Lemma 5.2] and applying (1.7) to the function  $\Phi(\nu) := \langle f, \nu \rangle$  with  $f : \mathbb{R} \rightarrow \mathbb{R}$  bounded measurable, we obtain that

$$\begin{aligned} M_t^{K,p,\sigma,f} := & \langle f, \nu_t^{K,p,\sigma} \rangle - \langle f, \nu_0^{K,p,\sigma} \rangle - \int_0^t \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dx) \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dy) b(x, y) [f(y) - f(x)] ds \\ & - p \int_0^t \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dx) \theta(x) \int_{\mathbb{R}} m(x, h) dh [f(x + \sigma h) - f(x)] ds \end{aligned}$$

is a square integrable martingale started from 0 with quadratic variation:

$$\begin{aligned} \langle M^{K,p,\sigma,f} \rangle_t = & \frac{1}{K} \int_0^t \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dx) \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dy) b(x, y) [f(y) - f(x)]^2 ds \\ & + \frac{p}{K} \int_0^t \int_{\mathbb{R}} \nu_s^{K,p,\sigma}(dx) \theta(x) \int_{\mathbb{R}} m(x, h) dh [f(x + \sigma h) - f(x)]^2 ds. \end{aligned}$$

It is standard to adapt the proof of [102, Theorem 5.3], to establish the following result:

**Theorem 1.1.2** (Version of FOURNIER-MÉLÉARD's result [102]). *Under **(A)**, let us consider a process sequence  $(\nu_{K \in \mathbb{N}^*}^{K,p,\sigma})$  with  $p \in [0, 1]$ ,  $\sigma > 0$  fixed, of infinitesimal generators (1.7) such that the initial condition  $\nu_0^{K,p,\sigma}$  converges in law for the weak topology on  $\mathcal{M}_1(\mathbb{R})$  when  $K \rightarrow +\infty$  to some deterministic finite measure  $\xi_0 \in \mathcal{M}_1(\mathbb{R})$  and  $\sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle \mathbf{1}, \nu_0^{K,p,\sigma} \rangle^3) < \infty$ .*

*Then, for all  $T > 0$ , the sequence  $(\nu_{K \in \mathbb{N}^*}^{K,p,\sigma})$  converges in law in  $\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R}))$ , when  $K \rightarrow +\infty$ , to a deterministic continuous function  $(\xi_t)_{t \geq 0} \in \mathcal{C}^0([0, T], \mathcal{M}_1(\mathbb{R}))$ . This measure-valued function  $\xi$  is the unique solution of the integro-differential equation written in its weak form: for all bounded and measurable functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , for all  $t \geq 0$ ,*

$$\begin{aligned} \langle f, \xi_t \rangle = & \langle f, \xi_0 \rangle + \int_0^t \int_{\mathbb{R}} \xi_s(dx) \int_{\mathbb{R}} \xi_s(dy) b(x, y) [f(y) - f(x)] ds \\ & + p \int_0^t \int_{\mathbb{R}} \xi_s(dx) \theta(x) \int_{\mathbb{R}} m(x, h) dh [f(x + \sigma h) - f(x)] ds, \end{aligned} \tag{1.9}$$



If for all  $t \geq 0$ ,  $\xi_t$  has density  $x \mapsto u(t, x)$  with respect to LEBESGUE measure,  $u(t, x)$  is a weak solution of the following (strong) equation: for all  $x \in \mathbb{R}$ , for all  $t \geq 0$ ,

$$\begin{aligned} \partial_t u(t, x) = & u(t, x) \left( \int_{\mathbb{R}} u(t, y) [b(y, x) - b(x, y)] dy \right) \\ & + \frac{p}{\sigma} \int_{\mathbb{R}} u(t, y) \theta(y) m \left( y, \frac{x-y}{\sigma} \right) dy - pu(t, x) \theta(x). \end{aligned} \quad (1.10)$$

In addition, we can prove [102, Proposition 5.4], under the assumption that  $\xi_0$  has  $x \mapsto u(0, x)$  as density with respect to the LEBESGUE measure, then the solution  $(\xi_t)_{t \geq 0}$  to (1.9) satisfies for all  $t \geq 0$  that the measure  $\xi_t$  has a density  $x \mapsto u(t, x)$  with respect to LEBESGUE measure. Note that the deterministic model given by (1.10) is the basis of the PDE approach of adaptive dynamics (see Section 1.1.4.1) leading to a new version of the CEAD given in (1.3).

**Fitness computations.** The aim of this section is to recover the expression of the fitness  $\text{Fit}(y, x)$  using deterministic approach and the study of equilibria, in two steps. According to Definition 1.1.1 and **(BA 2)**, it is sufficient to consider the case where there is no mutation, i.e.  $p = 0$  to compute this fitness function.

*Step 1. Preservation of mass and dimorphic state.* Note that for all  $K \in \mathbb{N}^*$ ,  $p \in [0, 1]$  and  $\sigma \in (0, +\infty)$ , for all  $t \geq 0$ ,  $\langle \mathbf{1}, \nu_t^{K,p,\sigma} \rangle = 1$ . From Theorem 1.1.2, we have for all  $p \in [0, 1]$  and  $\sigma > 0$  that

$$\forall t \geq 0, \quad \langle \mathbf{1}, \xi_t \rangle = \lim_{K \rightarrow +\infty} \langle \mathbf{1}, \nu_t^{K,p,\sigma} \rangle = 1,$$

in other words, there is mass preservation at the limit.

The following result ensures, in the case where there is no mutation, that if the process  $(\xi_t)_{t \geq 0}$  is initially dimorphic<sup>7</sup> then it remains so for all time.

**Proposition 1.1.3.** *Assume  $p = 0$  and **(A)**. Let  $\alpha_0 \in [0, 1]$  and  $x, y \in \mathbb{R}$ . If  $\xi_0 := \alpha_0 \delta_y + (1 - \alpha_0) \delta_x$ , then there exists  $\alpha : \mathbb{R}_+ \rightarrow [0, 1]$  such that  $\alpha(0) = \alpha_0$  and satisfying the following logistic ODE*

$$\forall t > 0, \quad \alpha'(t) = (b(x, y) - b(y, x)) \alpha(t) (1 - \alpha(t)), \quad (1.11)$$

in other words, for all  $t \geq 0$ ,  $\alpha(t) = \frac{\alpha_0}{\alpha_0 + (1 - \alpha_0) \exp(-[b(x, y) - b(y, x)]t)}$  and we have

$$\forall t \geq 0, \quad \xi_t = \alpha(t) \delta_y + (1 - \alpha(t)) \delta_x. \quad (1.12)$$

*Proof.* Let us consider a bounded measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $f(x) = f(y) = 0$ . From (1.9), we deduce that  $\langle f, \xi_t \rangle \leq \langle f, \xi_0 \rangle + (\bar{b} - \underline{b} + \bar{\theta} - \underline{\theta}) \int_0^t \langle f, \xi_s \rangle ds$ . By GRONWALL's lemma, we obtain for all  $t \geq 0$ ,

$$\langle f, \xi_t \rangle \leq \langle f, \xi_0 \rangle \exp([\bar{b} - \underline{b} + \bar{\theta} - \underline{\theta}] t).$$

Now,  $\langle f, \xi_0 \rangle = \alpha_0 f(y) + (1 - \alpha_0) f(x) = 0$ . Hence, for all  $t \geq 0$ ,  $\langle f, \xi_t \rangle = 0$  and by continuity of  $(\xi_t)_{t \geq 0}$ , we deduce that there exists  $\beta : \mathbb{R}_+ \rightarrow [0, 1]$  such that  $\beta(0) = 1 - \alpha_0$  and

$$\forall t \geq 0, \quad \xi_t = \alpha(t) \delta_y + \beta(t) \delta_x.$$

7. Dimorphic population: A population in which the individuals have only two distinct phenotypes.

As for all  $t \geq 0$ ,  $\alpha(t) + \beta(t) = \langle \mathbf{1}, \xi_t \rangle = 1$ , the announced relation (1.12) follows. Applying (1.9) for all bounded measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $f(x) \neq f(y)$ , we have that the process  $(\alpha(t))_{t \geq 0}$  is solution to (1.11) which completes the proof.  $\square$

*Step 2. Dynamic systems approach.* Equation (1.11) admits two equilibria:  $\alpha^* = 0$  and  $\alpha^* = 1$ . Denoting by

$$V_{x,y}(\alpha) := (b(x,y) - b(y,x)) \alpha (1 - \alpha),$$

the eigenvalue associated to the Jacobian matrix  $\text{Jac}(V_{x,y})(\alpha^*)$  is

$$\lambda^* = (1 - 2\alpha^*) (b(x,y) - b(y,x)).$$

As the fitness function  $\text{Fit}(y,x)$  represents the initial growth rate of a mutant individual of trait  $y$  in a monomorphic resident population of trait  $x$ , it is obtained from  $\lambda^*$  for the case  $\alpha^* = 0$ . Hence the fitness function is given by

$$\text{Fit}(y,x) = b(x,y) - b(y,x).$$

We have in Figures 1.5 and 1.6 two phase portraits according to the sign of the fitness  $\text{Fit}(y,x)$ . Note that, in the case where  $\text{Fit}(y,x) > 0$ , the equilibrium  $\alpha^* = 1$  is *attractive* and  $\alpha^* = 0$  is *repulsive* which means that the invasion by the mutant is possible. However, in the case where  $\text{Fit}(y,x) < 0$  the equilibrium  $\alpha^* = 0$  is *attractive* which means that the mutant invasion is impossible. Note that in the case where the fitness function is zero, each point of the phase portrait is an equilibrium for the autonomous vector field (1.11).

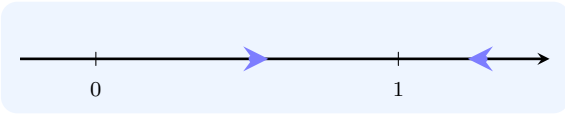


FIGURE 1.5 – Case where  $\text{Fit}(y,x) > 0$ .

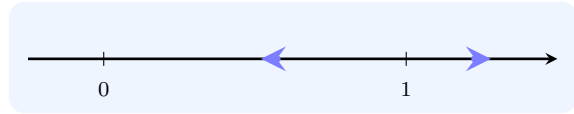


FIGURE 1.6 – Case where  $\text{Fit}(y,x) < 0$ .

Moreover, in our case the parameters of the microscopic model prevent the coexistence of any two traits in the population on a long time scale. Indeed, they satisfy the *Invasion-Implies-Fixation* (IIF) principle [113] (see section 1.1.2.1) which implies:

$$\text{(IIF)} \quad \text{For all } x \in \mathbb{R}, \text{ almost all } y \in \mathbb{R} \text{ such that } \text{Fit}(y,x) > 0 \text{ satisfy } \text{Fit}(x,y) < 0.$$


Note that Figure 1.5 illustrates the (IIF) principle.

**Numerical simulations.** The goal of this section is to illustrate numerically the phenomenon of the CEAD. We have performed numerical simulations of the process  $\nu^{K,p,\sigma}$  for a classical and simple model from the biological literature. The simulation is based on the acceptance-rejection method of the construct given in (1.8). The R-source of the program, provided in Appendix B, is based on the `IBMPopSim` package developed by GIORGI ET AL in [116]. Note that, other numerical simulations are proposed in Section 1.3.3, in the case of a more general model in which the population size can fluctuate.

The biological model is adapted from KISDI [140] for which the trait space is  $\mathcal{X} := [0, 4]$  (and not  $\mathbb{R}$  as previously) and the parameters are:

$$\theta(x) := 1, \quad c(x,y) := 1 - \frac{1}{1 + 1.2 \exp(-4(x-y))}, \quad b(x,y) := (4-y) \times c(x,y) \quad (1.13)$$

and  $m(x, dh)$  is a Gaussian distribution  $\mathcal{N}(x, \sigma^2)$  conditional on the mutant trait remaining in  $\mathcal{X}$ .

Note that, in Figure 1.7, is represented the invasion fitness landscape associated to the invasion fitness  $\text{Fit}(y, x) = b(x, y) - b(y, x)$ . The fitness landscape  $y \mapsto \text{Fit}(y, x)$ , that evolution seeks to maximise (locally), depends on the resident trait  $x$ , unlike in classical population genetics models, where the fitness landscape is fixed once and forever. The arrow  represents the effect of a small mutation of the resident trait  $x_1 := 1.75$  into the trait  $x_2 := 2.25$ . Hence the change of landscape when there is a change of resident trait.

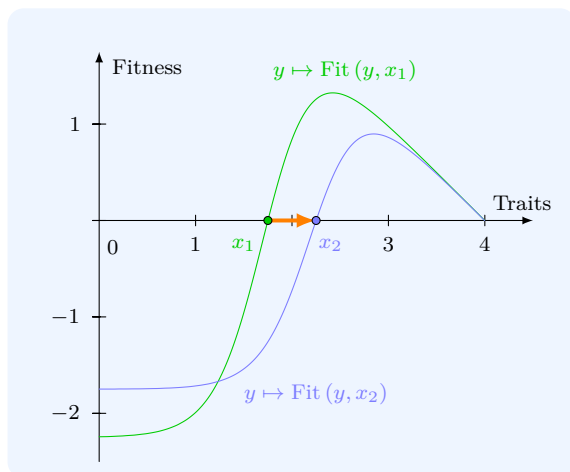


FIGURE 1.7 – Invasion fitness landscape of  $y \mapsto \text{Fit}(y, x_i) = b(x_i, y) - b(y, x_i)$  with  $x_1 := 1.75$  and  $x_2 := 2.25$ .

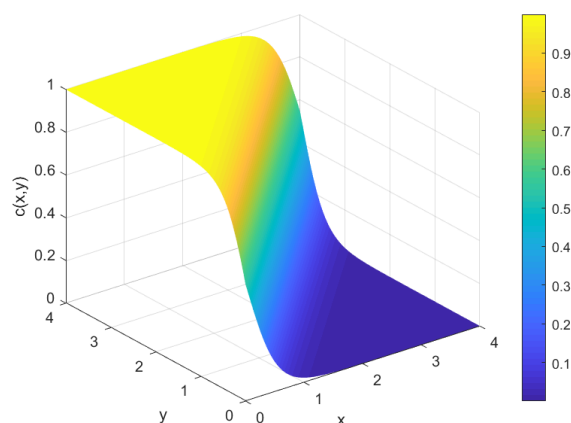


FIGURE 1.8 – Profile of the competition function  $(x, y) \mapsto c(x, y)$ .

Let us recall that  $b(x, y)$  is a resampling rate, i.e. the rate of simultaneous death of an individual with trait  $x$  and birth of an individual with trait  $y$  and it can be interpreted as modelling a competitive pressure driving selection. Here, it is composed of two terms which both have a biological interpretation. Note that the first term  $c(x, y)$  depends only on  $x - y$  and tends to 0 when  $x - y \rightarrow +\infty$  and to 1 when  $x - y \rightarrow -\infty$ . So, this function models an *asymmetric competition* in favour of high traits as we can see in Figure 1.8:  $x$  feels little competition from a smaller trait  $y$ , but strong competition from a higher trait  $y$ . Since the second term  $4 - y$  only depends on the reproducing trait and is favourable to small traits, evolution is expected to select an optimal trait.

We performed simulations of KISDI's model, starting with a monomorphic initial condition of trait 1 and varying the parameters  $K, p$  and  $\sigma$ . Some of these simulations are presented in the following figures and show a wide variety of evolutionary behaviours.

Note that, when  $K$  and  $\sigma$  are not too large and  $p$  not too small, the population evolves according to a relatively stable general scenario as in Figures 1.9, 1.10, 1.11, 1.13 and 1.14: it initially remains concentrated around the trait value equal to 1 that progressively moves towards a trait  $x^*$  close to 3.5 corresponding to the optimal trait value of this population, in the sense that it is best adapted to survive and reproduce. Once this point is reached, the population stabilises in its steady state.

In order to better understand these phenomena, biologists have been interested in the case where mutations are rare (very small  $p$ ), on which the theory of adaptive dynamics is based. Simulations show that the number of traits coexisting at the same time is restricted to a small

number, of the order of 2 or 3. As illustrated in Figure 1.12, the evolution proceeds first by jumps with appearances and then successive invasions of mutant traits replacing the previous resident traits. In other words, Figure 1.12 illustrates the convergence of our IBM to the TSS approximation where the mutational effect is small, i.e. to the CEAD.

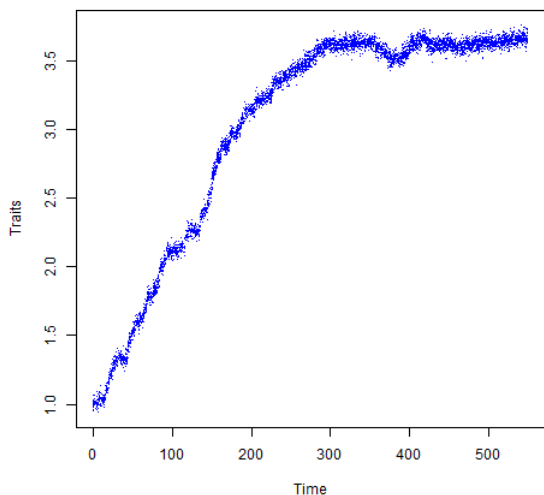


FIGURE 1.9 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 100$ ,  $p = 0.1$ ,  $\sigma = 0.03$ .

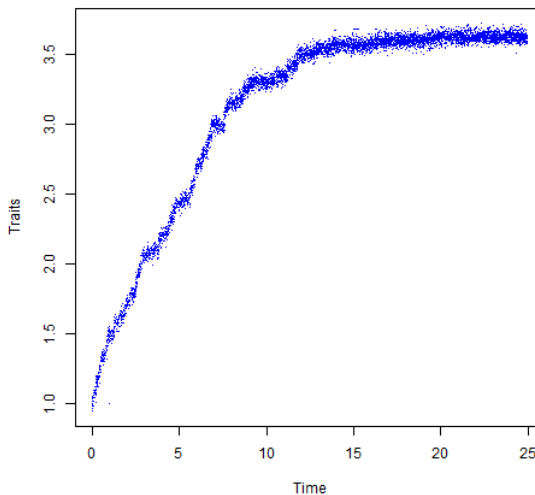


FIGURE 1.10 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 3000$ ,  $p = 0.1$ ,  $\sigma = 0.03$ .

Note that in Section 1.1.4.2, we will establish, in a heuristic way, the convergence result illustrated in Figure 1.12. The assumption of rare mutations leads to a population where the phenotypic variability is *de facto* reduced to a small number of phenotypes, which practically never happens in real populations. The phenotypic variability is a fundamental feature of evolution since it gives to the population the opportunity to react to a sudden environment change. We will come back to these biological criticisms in Section 1.1.4.6.

This motivates our approach, developed in Section 1.1.4.5, which consists in replacing the rare mutations assumption of Figure 1.12, with an assumption of *frequent* mutations. More precisely, the goal of this thesis is to evaluate mathematically, under a simultaneous asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ), the consequences of this new biological assumption. Figures 1.13 and 1.14 illustrates the expected convergence result, in this regime, from the IBM to the CEAD.

### 1.1.4 Canonical Equation

This section is devoted to the presentation of the state of the art about the mathematical derivation of the canonical equation. Two main approaches have been developed so far and are the subject of Sections 1.1.4.1 and 1.1.4.2. We will present, these two approaches, based on the IBM of Section 1.1.3.2 using different combinations of the biological assumptions (**BA**), leading to different versions of the CEAD. In Section 1.1.4.3, we discuss some mathematical

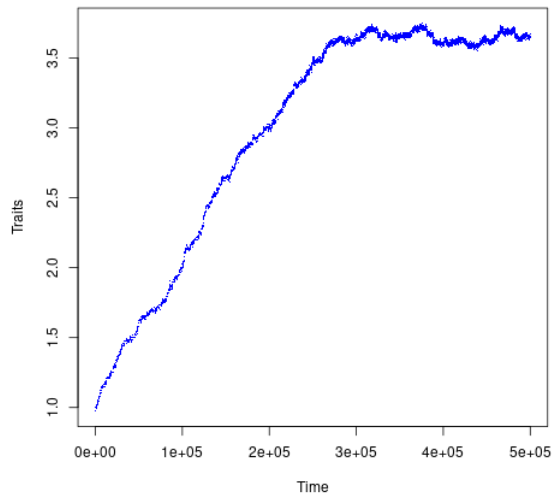


FIGURE 1.11 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 100$ ,  $p = 0.001$ ,  $\sigma = 0.01$ .

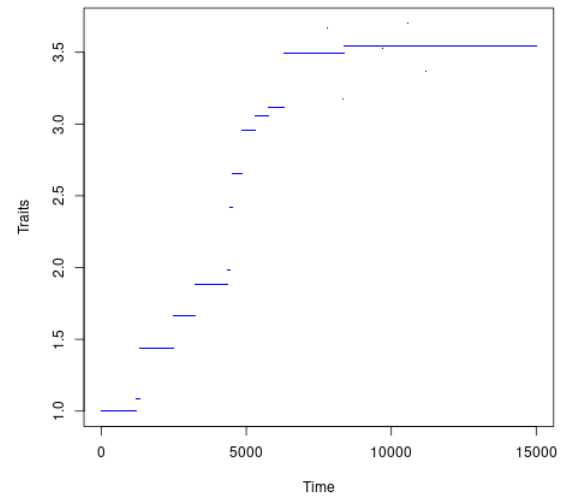


FIGURE 1.12 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 300$ ,  $p = 0.00003$ ,  $\sigma = 0.2$ .

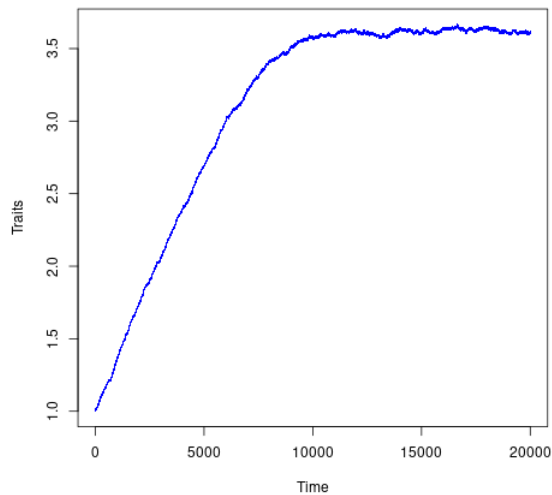


FIGURE 1.13 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.001$ .

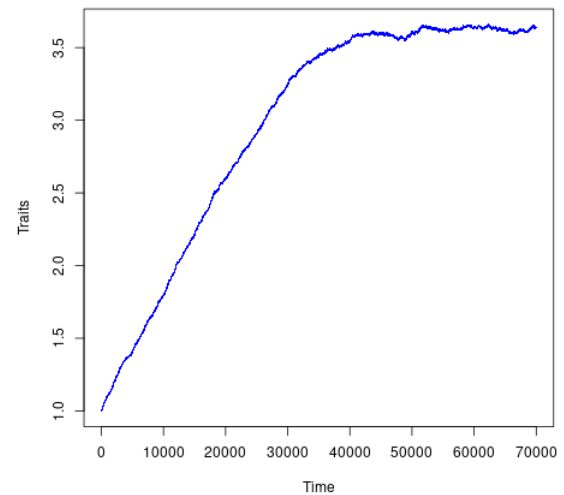


FIGURE 1.14 – Numerical simulation of the trait distribution of the microscopic model with parameters (1.13) and  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.0005$ .

extensions of the CEAD in different biological contexts. Finally, in Section 1.1.4.4, we present other mathematical approaches using new sets of combination of biological assumptions (**BA**).

### 1.1.4.1 PDE approach of adaptive dynamics

The goal of this section is to present the PDE approach of adaptive dynamics which consists in applying a limit of small mutations and strong selection to the deterministic process, solution of PDE (1.10) that was obtained as a large population limit of the IBM of Section 1.1.3.2. This original idea which was proposed in 2005 by DIEKMANN, JABIN, MISHLER and PERTHAME in [75] allows to obtain, on the time scale  $1/\sigma$ , a new version of the DIECKMANN-LAW's CEAD, described by a HAMILTON-JACOBI equation with constraint. This approach can be summarised by Figure 1.15 bellow (right arrow).

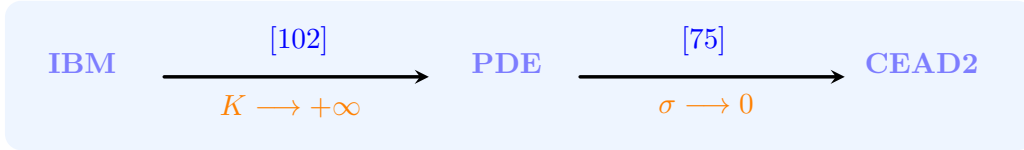


FIGURE 1.15 – Summary of the PDE approach of adaptive dynamics.

To obtain a simpler form of the PDE (1.10), we introduce the following additional assumptions.

*Assumptions (A3)* (PDE approach of adaptive dynamics)

- For all  $x \in \mathbb{R}$ ,  $\theta(x) \equiv \theta_0$  where  $\theta_0$  is some real constant.
- For all  $x, h \in \mathbb{R}$ ,  $m(x, h) = \bar{m}(h)$  where  $\bar{m} \in \mathcal{C}_K^1(\mathbb{R}, \mathbb{R})$  with compact support.
- For all  $x \in \mathbb{R}$ ,  $h \mapsto m(x, h)$  is symmetric on  $\mathbb{R}$  with respect to 0.

**Heuristics of the approach.** We remain, in this paragraph, at an informal level and refer to [75] and other cited references for details. Note that these approaches, based on assumptions of small and *frequent* mutations, employ asymptotic analysis to derive a HAMILTON-JACOBI equation. The scaling proposed in [75] consists in a time scaling in  $1/\sigma$ . We then obtain the following PDE: for all  $x \in \mathbb{R}$ , for all  $t \geq 0$ ,

$$\begin{aligned} \partial_t u_\sigma(t, x) &= \frac{1}{\sigma} u_\sigma(t, x) \left( \int_{\mathbb{R}} u_\sigma(t, y) [b(y, x) - b(x, y)] dy \right) \\ &+ \frac{p\theta_0}{\sigma} \int_{\mathbb{R}} \bar{m}(h) [u_\sigma(t, x + \sigma h) - u_\sigma(t, x)] dh. \end{aligned} \quad (1.14)$$

The change of unknown function

$$u_\sigma(t, x) = \exp\left(\frac{\varphi_\sigma(t, x)}{\sigma}\right) \quad (1.15)$$

is known in quantum mechanics as the *WKB Ansatz* and has already been used for reaction-diffusion equations with a scaling similar to the above in [103, 11]. We obtain for all  $x \in \mathbb{R}$ , for all  $t \geq 0$

$$\begin{aligned} \partial_t \varphi_\sigma(t, x) &= \int_{\mathbb{R}} u_\sigma(t, y) [b(y, x) - b(x, y)] dy \\ &+ p\theta_0 \int_{\mathbb{R}} \bar{m}(h) \left[ \exp\left(\frac{\varphi_\sigma(t, x + \sigma h) - \varphi_\sigma(t, x)}{\sigma}\right) - 1 \right] dh, \end{aligned}$$

which suggests the convergence, when  $\sigma \rightarrow 0$ , of  $\varphi_\sigma$  to a solution  $\varphi$  of the HAMILTON-JACOBI equation

$$\partial_t \varphi(t, x) = \int_{\mathbb{R}} [b(y, x) - b(x, y)] \mu_t(dy) + p\theta_0 H(\partial_x \varphi(t, x)), \quad (1.16)$$

where for all  $t \geq 0$ ,  $\mu_t$  is a measure on  $\mathbb{R}$ , obtained as a weak limit of  $u_\sigma(t, \cdot)$  when  $\sigma \rightarrow 0$ : for all bounded function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\int_{\mathbb{R}} f(x) u_\sigma(t, x) dx \xrightarrow{\sigma \rightarrow 0} \int_{\mathbb{R}} f(x) \mu_t(dx)$  and

$$H(\phi) := \int_{\mathbb{R}} \bar{m}(h) (\exp(\phi h) - 1) dh$$

is an Hamiltonian operator.

This approach has been used in many works for variants of this model (with non-constant population size) [196, 69, 130, 202, 160]. Assumptions **(A3)** come from [160] and allow to control uniformly in  $\sigma$  the total density  $\int_{\mathbb{R}} u_\sigma(t, x) dx$  of the population.

Note that equation (1.15) shows that, when we have uniform controls on the total population density:  $\int_{\mathbb{R}} u_\sigma(t, x) dx \leq C$  for some constant  $C > 0$ , we must have for all  $t \geq 0$  that  $\varphi(t, \cdot) \leq 0$ . Indeed, if we had  $\max_{x \in \mathbb{R}} \varphi(t, x) > 0$  for some time  $t \geq 0$ , then also  $\varphi_\sigma(t, x) > 0$  uniformly in  $\sigma$  for  $t, x$  in some open set (by uniform continuity). But this is impossible since  $\int_{\mathbb{R}} \exp\left(\frac{\varphi_\sigma(t, x)}{\sigma}\right) ds \leq C$ . Thus, we obtain a HAMILTON-JACOBI equation (1.16) with the constraint  $\max_{x \in \mathbb{R}} \varphi(t, x) \leq 0$  for all  $t \geq 0$ . Moreover, note that from (1.15), for all  $t \geq 0$  there exists  $x \in \mathbb{R}$  such that  $\varphi(t, x) = 0$  which means the population at the limit is concentrated on

$$Z_\varphi := \left\{ y \in \mathbb{R} \mid \varphi(t, y) = 0 \right\}$$

that is to say,  $\mu_t$  can only give mass to the zeros of  $\varphi$  at time  $t$ .

Note that  $H(0) = 0$  and that for all  $t \geq 0$ ,  $\partial_x \varphi(t, x) = 0$  at a maximum point for  $x \mapsto \varphi(t, x)$ , hence we expect to have  $\int_{\mathbb{R}} [b(y, x) - b(x, y)] \mu_t(dy) \leq 0$  for all  $x \in Z_\varphi$  since, from (1.16), otherwise the condition that  $\max_{x \in \mathbb{R}} \varphi(t, x) = 0$  would be violated just after time  $t$ . Moreover, since for all  $t \geq 0$ ,  $\max_{x \in \mathbb{R}} \varphi(t, x) = 0$ , we expect to have  $\int_{\mathbb{R}} [b(y, x) - b(x, y)] \mu_t(dy) = 0$  for at least one  $x \in Z_\varphi$ . Unfortunately, these conditions are in general not sufficient to characterise  $\mu_t$ . However, under the additional assumption that for all  $t \geq 0$ , there exists a unique  $x^*(t)$  such that  $\varphi(t, x^*(t)) = 0$ , that is to say  $x \mapsto \varphi(x, t)$  admits only one maximum at  $x^*(t)$  (for instance if for all  $t \geq 0$ ,  $x \mapsto \varphi(x, t)$  is concave [160]), then  $\mu_t$  must be equal to  $\delta_{x^*(t)}$  (since it must have mass 1) the solution of the HAMILTON-JACOBI equation (1.16) can be uniquely characterised [13, 194, 42, 174].

**CEAD of the PDE approach.** Let us assume that for all  $t \geq 0$ , there exists a unique  $x^*(t)$  such that  $x \mapsto \varphi(x, t)$  admits only one maximum at  $x^*(t)$ . So, we deduce in the one hand that  $\partial_x \varphi(t, x^*(t)) = 0$  and the other hand that

$$u_\sigma(t, y) dy \xrightarrow{\sigma \rightarrow 0} \delta_{x^*(t)}(dy). \quad (1.17)$$

In particular,

$$\partial_t \varphi(t, x) = b(x^*(t), x) - b(x, x^*(t)) + p\theta_0 H(\partial_x \varphi(t, x)). \quad (1.18)$$

It is now assumed that  $t \mapsto x^*(t)$  is differentiable. Then, we obtain from  $\partial_x \varphi(t, x^*(t)) = 0$  and the chain rule that

$$\partial_{tx}^2 \varphi(t, x^*(t)) + \dot{x}^*(t) \partial_{xx}^2 \varphi(t, x^*(t)) = 0$$

and from (1.18) that

$$\begin{aligned} \partial_{xt}^2 \varphi(t, x) &= \partial_2 b(x^*(t), x) - \partial_1 b(x, x^*(t)) + \partial_{xx}^2 \varphi(t, x) \int_{\mathbb{R}} h \bar{m}(h) \exp(h \partial_x \varphi(t, x)) dh \\ &= \partial_1 \text{Fit}(x^*(t), x) + \partial_{xx}^2 \varphi(t, x) \int_{\mathbb{R}} h \bar{m}(h) \exp(h \partial_x \varphi(t, x)) dh. \end{aligned}$$

Hence, using the symmetry around 0 of  $h \mapsto \bar{m}(h)$  we have

$$\partial_{xt}^2 \varphi(t, x^*(t)) = \partial_1 \text{Fit}(x^*(t), x^*(t)).$$

Therefore, we deduce a new form of the CEAD presented in (1.3):

$$\forall t > 0, \quad \dot{x}^*(t) = [-\partial_{xx}^2 \varphi(t, x^*(t))]^{-1} \times \partial_1 \text{Fit}(x^*(t), x^*(t)), \quad x^*(0) = x_0^*. \quad (1.19)$$

This equation consists in two terms: the first one scaling the rate of evolutionary change, proportional to mutational variance in adaptive dynamics (or genetic variance in quantitative genetics); the second one the gradient of the fitness. In this way, (1.19) may be seen as a generalisation of the CEAD or LANDE's equation (1.4) to a case where the mutations are not assumed to be very rare (on the contrary to adaptive dynamics) and such that the genetic variance is explicitly expressed in terms of the second-order derivative of the function  $\varphi$ .

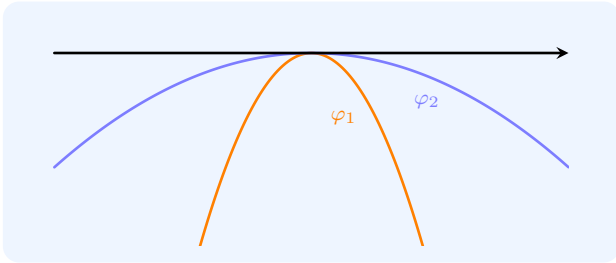


FIGURE 1.16 – Two possible profiles for  $\varphi$  concave where  $\varphi_1$  has a higher curvature than  $\varphi_2$ .

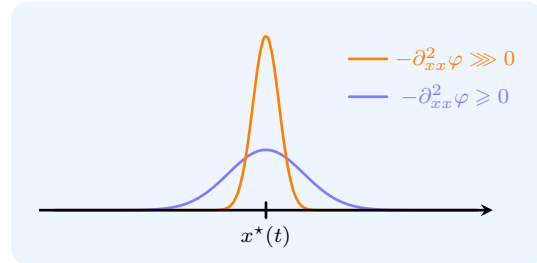


FIGURE 1.17 – Dispersion of the support around the dominant trait  $x^*(t)$ .

Note that  $\partial_{xx}^2 \varphi(t, x^*(t))$  is a non-positive coefficient because  $x^*(t)$  corresponds to a maximum of  $x \mapsto \varphi(t, x)$ . Therefore one can read the direction in which  $x^*(t)$  moves from the selection gradient  $\partial_1 \text{Fit}(x^*(t), x^*(t))$  and the speed of motion is governed by  $\partial_{xx}^2 \varphi(t, x^*(t))$ . Note that, if for some  $t \geq 0$ ,  $x \mapsto \varphi(t, x)$  is very concave, as the orange curve in Figure 1.16, near its maximum  $x^*(t)$ , then  $-\partial_{xx}^2 \varphi(t, x^*(t)) \gg 0$  i.e. it is very non-negative and so  $[-\partial_{xx}^2 \varphi(t, x^*(t))]^{-1}$  is very small in (1.19). Thus, in this case the speed of evolution of the optimal trait is slow. In this case, we see that the phenotypic distribution of the population is closer to the dominant/mean trait  $x^*(t)$  than for larger values of  $-\partial_{xx}^2 \varphi(t, x^*(t))$  as we can see in Figure 1.17 with the orange curve.

Some derivations of this approach have been used in various fields as population systems with adults and juveniles are considered in [33] or evolution of mutualism in [96].



### 1.1.4.2 Stochastic approach of adaptive dynamics

The goal of this section is to present heuristically the stochastic approach of adaptive dynamics leading, in two steps, to the first version of the CEAD initially proposed by DIECKMANN-LAW in [72]. This approach can be summarised by Figure 1.18 bellow.

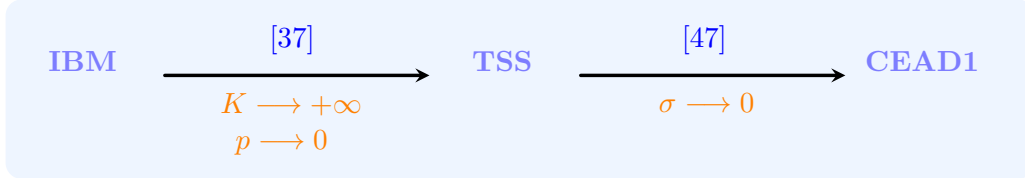


FIGURE 1.18 – Summary of the stochastic approach of adaptive dynamics.

Our first goal is to present a support concentration result on the IBM of Section 1.1.3.2, i.e. to guarantee the convergence of this process, on the time scale of (rare) mutations, to the TSS, with values in finite sums of DIRAC masses. For this purpose, we will combine simultaneously a limit of rare mutations ( $p \rightarrow 0$ ), allowing to separate mutation events on a long evolutionary time scale, and a limit of large population ( $K \rightarrow +\infty$ ), allowing to study the competition between two mutations on a shorter ecological time scale satisfying  $Kp \rightarrow 0$ . This original idea of time scale separation, between mutations and birth and death events, was proposed by CHAMPAGNAT in [37] and corresponds to the left arrow in Figure 1.18.

Our second goal consists in proving the convergence of the TSS to the original version of the CEAD under an asymptotic of small mutations ( $\sigma \rightarrow 0$ ) and when the time is rescaled in an unusual way by  $1/\sigma^2$  (versus  $1/\sigma$  for the PDE approach), because of the particularities of the biological model. This result, due to CHAMPAGNAT ET AL [14], provides a mathematical justification of the canonical equation of [72]. Note that, in [37, 47], births and deaths are not simultaneous and mutations do not occur during life, but one can be convinced that their proposed approach will work in our situation.

**Step 1. Convergence to the TSS.** The following result proves that, under a double limit of rare mutations ( $p \rightarrow 0$ ) and large population ( $K \rightarrow +\infty$ ), the IBM converges, on the time scale of mutations (in  $1/Kp$ ), to a pure jump MARKOV process with values in the space of monomorphic trait populations, called TSS.

**Theorem 1.1.4** (Version of CHAMPAGNAT’s result [37]). *Let  $\sigma \in (0, 1]$  be fixed. Let us consider (A), that  $\nu_0^{K,p,\sigma} := \delta_x$  for some  $x \in \mathbb{R}$  be fixed and assume that  $K \rightarrow +\infty$  and  $p \rightarrow 0$  such that*

$$\forall C > 0, \quad \log(K) \ll \frac{1}{Kp} \ll \exp(CK). \quad (1.20)$$

*Then, the process  $(\nu_{t/Kp}^{K,p,\sigma})_{t \geq 0}$  converges, in the sense of finite-dimensional marginals on  $\mathcal{M}_1(\mathbb{R})$ <sup>8</sup>, to the pure jump homogeneous MARKOV process  $(\Lambda_t^\sigma)_{t \geq 0}$  with values the set of DIRAC masses such that for all  $t \geq 0$ ,*

$$\Lambda_t^\sigma = \delta_{X_t^\sigma},$$

8. More precisely, for the topology of  $\mathcal{M}_1(\mathbb{R})$  induced by the applications  $\nu \mapsto \langle f, \nu \rangle$  where  $f : \mathbb{R} \rightarrow \mathbb{R}$  bounded measurable function.

where the  $\mathbb{R}$ -valued MARKOV process  $(X_t^\sigma)_{t \geq 0}$  is such that  $X_0^\sigma = x$  has for infinitesimal generator, for all bounded measurable function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$L^\sigma \varphi(x) := \frac{1}{\sigma^2} \int_{\mathbb{R}} (\varphi(x + \sigma h) - \varphi(x)) \theta(x) \frac{[\text{Fit}(x + \sigma h, x)]_+}{b(x + \sigma h, x + \sigma h)} m(x, h) dh \quad (1.21)$$

where  $[a]_+$  denotes the positive part of  $a \in \mathbb{R}$ .

The process  $\Lambda^\sigma$  is called TSS and was introduced in [171], first worked out in [72] and rigorously studied in [37]. This process was extended to the case of *Polymorphic Evolution Sequence* (PES) in [171, 47] where the coexistence of several distinct traits in the population is allowed.

Note that, Assumption (1.20) ensures the separation of three time scales: the invasion/extinction time scale ( $\log(K)$ ), the mutation time scale ( $1/Kp$ ), and the overall population extinction time scale ( $\exp(CK)$ ) (see Figure 1.4). Theorem 1.1.4 means that on the time scale of mutations, the population is at all times monomorphic, and that the phase of replacement of a resident trait by a mutant trait occurs (the so-called *fixation* of the mutant trait) on a time scale infinitesimal compared to the mutations.

The rate of jump  $\theta(x) \frac{[\text{Fit}(x + \sigma h, x)]_+}{b(x + \sigma h, x + \sigma h)} m(x, h)$  from a resident individual of trait  $x$  to a mutant individual of trait  $y$  of the form  $y := x + \sigma h$  (the trait of the mutant is only slightly different from the trait of its progenitor) is factorised as follows:  $\theta(x)$  is the mutation rate of an individual of trait  $x$  on the time scale  $1/Kp$ ,  $m(x, h)$  is the density of the mutation law and  $\frac{[\text{Fit}(x + \sigma h, x)]_+}{b(x + \sigma h, x + \sigma h)}$  is the *invasion probability* of a mutant of trait  $y = x + \sigma h$ .

*Proof intuitions of Theorem 1.1.4.* It is divided into two main phases: firstly the arrival of a mutation, secondly the invasion and fixation or extinction of the mutant.

- **Phase 1. Arrival of a mutation.** We start with a monomorphic population of trait  $x$ :  $\nu_0^{K,p,\sigma} = \delta_x$ . On the initial time scale, the arrival rate of a mutant is  $K \times \theta(x) \times p$  where the first factor is the number of individuals having the trait  $x$ , the second factor is the mutation rate of an individual of trait  $x$  and the third factor is the mutation probability. On the  $1/Kp$  time scale, we thus obtain a mutation rate of  $\theta(x)$  in the population of trait  $x$ , in other words if we note  $\tau^K$  the arrival of the first mutation, then  $Kp\tau^K$  converges to a random variable with an exponential law of parameter  $\theta(x)$ . The mutant trait is then given by  $y = x + \sigma H$  where  $H$  is a random variable following the law  $m(x, h)dh$ .
- **Phase 2. Invasion and fixation or extinction of a mutant.** The study of this second phase can be summarised in Figure 1.19 below. We will distinguish three steps in the invasion and fixation of a mutant  $y$  in a monomorphic population of trait  $x$ .
  - ◆ **Invasion step.** At the moment when the mutant trait appears (at time  $t = 0$  in Figure 1.19), its density in the measure  $\nu_0^{K,p,\sigma}$  is  $1/K$  (a single individual). In other words, at the arrival of the mutant, the population  $\nu_0^{K,p,\sigma} = \frac{K-1}{K}\delta_x + \frac{1}{K}\delta_y$  is *dimorphic*. We will call the *density* of trait  $u \in \{x, y\}$  in the population  $\nu_t^{K,p,\sigma}$  the number  $N_t^{K,u} := \langle \mathbb{1}_{\{u\}}, \nu_t^{K,p,\sigma} \rangle$ . Note that, before the next mutation,  $(N_t^{K,y})_{t \geq 0}$  is a  $\{0, \frac{1}{K}, \frac{2}{K}, \dots, 1\}$ -valued MARKOV jump process with initial condition  $N_0^{K,y} = 1/K$  and with transition rates as follows:
    - $\frac{i}{K} \rightarrow \frac{i+1}{K}$  at rate  $Kb(x, y) \times \frac{i}{K} \times \frac{K-i}{K}$
    - $\frac{i}{K} \rightarrow \frac{i-1}{K}$  at rate  $Kb(y, x) \times \frac{K-i}{K} \times \frac{i}{K}$

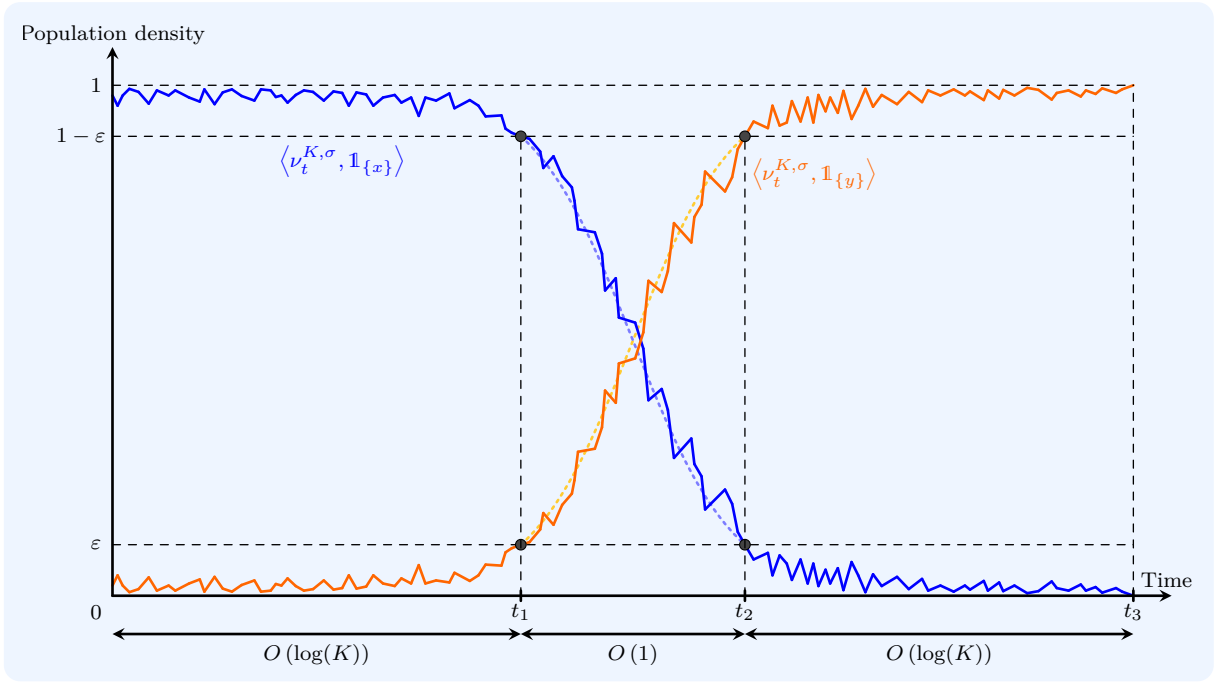


FIGURE 1.19 – The three steps of invasion and fixation of a mutant  $y$  in a monomorphic resident population of trait  $x$ .

for all  $i \in \mathbb{N}^*$  where, in the first case, the first factor is the global resampling rate of a resident individual of trait  $x$  by a mutant individual of trait  $y$ , the second factor is the probability of selecting a resampling individual of trait  $x$  and the third is the probability of selecting a mutant individual of trait  $y$  in the population. As long as  $i \ll K$  (i.e. the mutant population is negligible), these rates can be approximated by

- $i \rightarrow i + 1$  at rate  $i \times b(x, y)$
- $i \rightarrow i - 1$  at rate  $i \times b(y, x)$

If we denote by  $B_P(b, d)$  the transitions rates of a branching process with birth rate per individual  $b$  and death rate per individual  $d$ , the transition rates of a  $B_P(b(x, y), b(y, x))$  branching process are recognised above. This branching process, starting from a single individual at the beginning, and denoted by  $(Z_t)_{t \geq 0}$ , has the *survival probability*

$$\mathbb{P}(Z_t > 0, \forall t \geq 0) = \frac{[b(x, y) - b(y, x)]_+}{b(x, y)}.$$

This suggests that the probability of mutant invasion is  $[b(x, y) - b(y, x)]_+ / b(x, y)$ . This argument can be made rigorous by proving, by comparison (couplings) with two branching processes

$$Z^1 \sim B_P((1 - \varepsilon)b(x, y), b(y, x)) \quad \text{and} \quad Z^2 \sim B_P(b(x, y), (1 - \varepsilon)b(y, x))$$

such that

$$Z_t^1 \leq N_t^{K,y} \leq Z_t^2 \tag{1.22}$$

up to the stopping time  $t_1 := \inf \left\{ t \geq 0 \mid N_t^{K,y} > \varepsilon \text{ or the next mutation occurs} \right\}$ , that the probability of the mutant population  $N_t^{K,y}$  reaching a fixed  $\varepsilon > 0$  density (see Figure 1.19), is close to  $[b(x, y) - b(y, x)]_+ / b(x, y)$  and the probability of the mutant population extinguishing before reaching the density  $\varepsilon$  is close to  $1 - [b(x, y) - b(y, x)]_+ / b(x, y)$ . Moreover, both events take place in  $O(\log(K))$  time with high probability. Note that as long as the mutant density is low, the dynamics of the resident population is close to what it would be in the absence of mutants, so  $N_t^{K,x}$  remains close to 1. So, the *invasion step*, between 0 and  $t_1$ , can be interpreted as the transition of the mutant density from  $1/K$  to a small fixed  $\varepsilon > 0$  level (which should tend towards 0 to complete the proof).

If there is *extinction* of the mutant  $y$ , we have finished this phase and we wait for the arrival of a new mutation. Note that invasion is only possible if  $\text{Fit} > 0$ , i.e. if the branching process  $(Z_t)_{t \geq 0}$  is super-critical.

◆ **Competitive step.** Once the mutant invasion is successful, i.e. once  $N_t^{K,y}$  reaches the density  $\varepsilon$ , for large  $K$ ,  $N_t^{K,y}$  is close to the solution of the logistic equation (1.11) of Proposition 1.1.3 with  $\alpha_0 = \varepsilon$ . This allows us to approximate the joint dynamics of  $(N_t^{K,x}, N_t^{K,y})$  with initial condition  $(1 - \varepsilon, \varepsilon)$ , represented in dotted line in Figure 1.19 between  $t_1$  and  $t_2$ . We then note that this ODE converges to the attractive equilibrium  $(0, 1)$ . Thus, the density of mutants reaches the  $1 - \varepsilon$  level and the density of residents reaches the  $\varepsilon$  level at some time  $t_2$  (see Figure 1.19) such that  $t_2 - t_1 = O(1)$  (where the  $O(1)$  depends on  $\varepsilon$ ) with high probability.

◆ **Fixation of the mutant population/Extinction of the resident population.** This last step consists in using, in similar way as in the first step, comparisons with branching processes to bound the resident population from time  $t_3$  (see Figure 1.19) as follows and prove that the latter dies out in  $O(\log(K))$  with high probability:

$$N_t^{K,x} \leq Z_t^3$$

where  $Z^3 \sim B_P(b(y, x), (1 - \varepsilon)b(x, y))$ . This branching process being sub-critical, it becomes extinct almost surely even if the initial number of individuals of trait  $x$  is  $\varepsilon K$  at time  $t_3$ . We then obtain a new monomorphic population  $\nu_{t_3}^{K,p,\sigma} = \delta_y$  and we start again Phase 1 by waiting for the arrival of a new mutation.

This discussion explains the infinitesimal generator of the TSS given by (1.21).

**Remark 1.1.5.** (1) *As stated above, the dynamics of the mutant population  $y$ , with size  $N_t^{K,y}$  is close to a continuous time binary branching process  $B_P(b(x, y), b(y, x))$ . The population being initially rare means here that  $N_0^{K,y} = 1$ . In this case, the fitness of  $y$  with respect to  $x$  can be defined by as the initial growth rate*

$$\frac{K-1}{K} \text{Fit}(y, x) := \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E} \left( \frac{N_t^{K,y} - N_0^{K,y}}{N_0^{K,y}} \mid N_0^{K,y} = 1 \right) = \lim_{t \rightarrow 0} \frac{\mathbb{E} \left( N_t^{K,y} \mid N_0^{K,y} = 1 \right) - 1}{t}.$$

*Another definition is possible, as the long term growth rate of the approximating branching process  $(Z_t)_{t \geq 0}$  since, from [4, Chapter 3, Section 7], the process  $(\exp(-[b(x, y) - b(y, x)]t) Z_t)_{t \geq 0}$  is a martingale.*

(2) From (1.22), we deduce that the survival probability satisfies

$$\frac{(1-\varepsilon)b(x,y) - b(y,x)}{(1-\varepsilon)b(x,y)} \leq \mathbb{P}(N_t^{K,y} > 0, \forall t \geq 0) \leq \frac{b(x,y) - (1-\varepsilon)b(y,x)}{b(x,y)}$$

and thus, when  $\varepsilon \rightarrow 0$ , we obtain that

$$\mathbb{P}(N_t^{K,y} > 0, \forall t \geq 0) = \begin{cases} 1 - \frac{b(y,x)}{b(x,y)} = \frac{b(x,y) - b(y,x)}{b(x,y)} & \text{if } b(x,y) > b(y,x) \\ 0 & \text{otherwise} \end{cases}$$

and we find again the value of the fitness  $\text{Fit}(y,x) = b(x,y) - b(y,x)$ . Here again, we see that the possibility of invasion is determined by the sign of the fitness.

**Step 2. Convergence to the CEAD.** The following result proves, under a limit of small mutations ( $\sigma \rightarrow 0$ ), the convergence of the TSS of Theorem 1.1.4 on the time scale  $1/\sigma^2$ , which leads to a global time scale  $1/Kp\sigma^2$  from the initial IBM, to the CEAD.

**Theorem 1.1.6** (Version of CHAMPAGNAT-MÉLÉARD's result [47]). *Let us consider (A) and for all  $\sigma \in (0, 1]$  a  $\mathbb{R}$ -valued MARKOV process  $(X_t^\sigma)_{t \geq 0}$  with infinitesimal generator (1.21) and of initial condition  $X_0^\sigma := x_0 \in \mathbb{R}$ . Then, the process  $(X_{t/\sigma^2}^\sigma)_{t \geq 0}$  converges in law, when  $\sigma \rightarrow 0$ , for the SKOROHOD topology in  $\mathbb{D}([0, +\infty), \mathbb{R})$  to the deterministic process  $(x(t))_{t \geq 0}$  with  $\mathcal{C}^0([0, +\infty), \mathbb{R})$  paths, solution to the CEAD:*

$$\forall t > 0, \quad x'(t) = \frac{\theta(x(t))}{b(x(t), x(t))} \int_{\mathbb{R}} h [h \partial_1 \text{Fit}(x(t), x(t))]_+ m(x(t), h) dh \quad (1.23)$$

with  $x(0) := x_0$ . Moreover, in the case where  $m(x, \cdot)$  is a symmetrical measure on  $\mathbb{R}$  for all  $x \in \mathbb{R}$ , (1.23) gets the simpler form

$$\forall t > 0, \quad x'(t) = \partial_1 \text{Fit}(x(t), x(t)) \times \frac{\theta(x(t))}{2b(x(t), x(t))} \int_{\mathbb{R}} h^2 m(x(t), h) dh, \quad (1.24)$$

Note that (1.24) gets the classical form of DIECKMANN-LAW's CEAD. Moreover, this CEAD will correspond to the asymptotic behaviour of the IBM under our scaling of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ) (but frequent mutations) developed in this thesis (see Section 1.1.4.5)

*Proof intuitions of Theorem 1.1.6.* A limited expansion, to the first order, of the expression (1.21) leads to the deterministic operator  $A$ , defined for all  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  by

$$A\varphi(x) := \varphi'(x) \times \frac{\theta(x)}{b(x,x)} \int_{\mathbb{R}} h [h \partial_1 \text{Fit}(x,x)]_+ m(x,h) dh$$

combined with classical arguments of *uniform tightness* and *martingale property* allow to conclude that the TSS converges to the CEAD.

**Remark 1.1.7.** *Note that the PDE and stochastic approaches obtain a simplified model by applying a concentration limit aiming at replacing the population densities by DIRAC masses concentrated on a small number of traits (i.e. to subpopulations in which all individuals have the same trait). So, the CEAD corresponds to the motion of DIRAC masses. While these two approaches provide approximate models with similar qualitative behaviours, they exploit different ideas and have significant differences in terms of time scale and dynamics. The canonical equations obtained by these two methods are therefore structurally different.*

### 1.1.4.3 Mathematical extensions in different contexts

The aim of this section is to discuss different mathematical extensions to the CEAD in various biological contexts: trait and age-structured populations, Mendelian diploid populations, multi-resources chemostat models with a PDE and a stochastic point of view, co-evolution of predator-prey populations and population structured in space and traits. For these different developments, methods and time scales for observing the CEAD phenomenon remain unchanged with respect to the two main approaches.

Note that, when the trait space is multidimensional and/or the number of coexisting traits in the population is arbitrary provides another extension in the study of the canonical equation [72, 162, 14, 152], [36, Chapter 3], [82, 153].

**Trait and age-structured populations.** In [166], MÉLÉARD and TRAN give a description, in continuous time, of a discrete stochastic IBM of asexual population structured by trait and age undergoing mutation and selection. On the one hand, they consider approximations of the microscopic process under a large population asymptotic ( $K \rightarrow +\infty$ ) and obtain, in the limit, a deterministic PDE involving trait and age, that generalises classical demographic PDEs [120, 50, 208, 180]. This gives them information on the long time behaviour of the population as in the approach of [102] (without age-structure).

On the other hand, taking simultaneously the asymptotics of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) under the well-chosen time-scale separation condition, of the form (1.20), they obtain, on the time scale  $1/Kp$ , the convergence of their microscopic process to the so-called *TSS process for age-structured populations* that jumps from a monomorphic equilibrium (where all the individuals carry the same trait) to another. This process extends to populations with age-structure the TSS rigorously studied in [37]. The age is taken into account as a structuration variable and not as a trait in the sense of [171].

Moreover, under the additional assumption of small mutations ( $\sigma \rightarrow 0$ ), they prove, on the time scale  $1/\sigma^2$ , the convergence of the TSS process for age-structured populations to the solution of an age-dependent ODE that extends the CEAD introduced by DIECKMANN-LAW [72]. Here again, the time scale leading to the canonical equation is  $1/Kp\sigma^2$ .

**Mendelian diploid populations.** *In a population of sexually reproduced diploid individuals, the rules of transmission of genetic material from a pair of parents to their children are called Mendel's laws [168]. These rules can be summarised as follows. We consider a given locus (location) of the genome where a gene is located. Mendel's laws state that each of the two parents of a child transmits one of its two alleles, chosen uniformly at random, to that child. Unlike asexual populations reproducing clonally, the average number of individuals born with a given genotype is not proportional to the number of individuals in the population carrying that same genotype, which leads mathematical difficulties.*

In [51], COLLET ET AL consider a diploid hermaphroditic population in which the survival and Mendelian reproduction ability of each individual is subject to selection and mutation and depends only on a quantitative phenotypic trait. This hereditary trait is determined by its genotype, represented by the types of two alleles  $a$  and  $A$  on a single locus. When a mutation occurs at reproduction, it changes one of the two allelic traits of the newborn. They assume that the genotype to phenotype map  $\phi$  is smooth and there are no *parental effects*, so that  $\phi(u_A, u_a) = \phi(u_a, u_A)$  for some two allelic trait values  $(u_a, u_A)$  leading to, genotype to genotype

map is locally additive.

Making the combined large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) asymptotics and mutation steps are not too large ( $\sigma > 0$ ) under the slightly new time-scale separation

$$\forall C > 0, \quad \frac{\log(K)}{\sigma} \ll \frac{1}{Kp} \ll \exp(CK),$$

they study the long time behaviour of the population process, in the mutation time scale  $1/Kp$ , and prove it converges to a pure jump process moving between *homozygous states*. This process is a generalisation of the simple TSS that for the *haploid* case were derived in [171, 37, 47].

Once again, under the additional assumption of small mutations ( $\sigma \rightarrow 0$ ), they study the convergence of the diploid TSS, on the adequate time scale  $1/\sigma^2$ , to the canonical equation. The time scale leading to the canonical equation is again  $1/Kp\sigma^2$ .

The main difference of Mendelian models compared to the clonal ones in [72, 171] and successors, is a difference in the invasion probability of a mutant due to the additional noise inherent in the Mendelian mechanism, and the occurrence of an additional factor 2 in the canonical equation due to the fact the mutants invade as heterozygotes but establish as homozygotes. See also [170] for another extension of the CEAD for Mendelian diploids and haplo-diploids.

**Multi-resources chemostat models.** *Since the first works of MONOD [176] and NOVICK and SZILAR [184, 183], biologists have developed experimental procedures which allow to maintain a population of micro-organism (bacteria, micro-algae) at a stationary finite size while, at the same time, the micro-organisms have a positive individual growth rate. The procedure is based on a chemostat: micro-organisms live in a growth container of constant volume in which liquid is injected continuously. This liquid contains nutrients which are consumed by the micro-organisms. They assume that the chemostat is well stirred, so that the distribution of micro-organisms and nutrients are spatially uniform. Moreover, since the container has a finite volume and fresh liquid enters permanently, an equal amount of liquid pours out containing both unconsumed nutrients and micro-organisms. This pouring out helps regulating the micro-organisms population size. These chemostats are extremely useful to make bacteria adapt to the nutrients environment in such away that they will increase their growth rate. This experimental device is for instance used on an industrial scale in water treatment stations [15].*

*PDE approach.* In the various deterministic approaches proposed below, the authors start from the intermediate “PDE” model, as illustrated in Figure 1.15, and apply a small mutations limit ( $\sigma \rightarrow 0$ ) to it.

In [75, 160, 42, 174] they consider an integro-differential nonlinear model, more sophisticated than (1.10), that describes the evolution of a population structured by a 1-dimensional quantitative trait and consuming resources whose dynamics depend on the current state of the population. Note that the case of chemostat for a single or two resource(s) is described in [75]. Note also that the case of a finite and infinite number of resources are treated in [42] and [174] respectively. Other models with a single resource are considered in [160].

Following the formalism proposed in [75], these references study a concentration phenomenon arising in the limit of strong selection (competition for resources) and small mutations as in (1.17). They prove that the population density converges to a sum of DIRAC masses characterised by the solution  $\varphi$  of a HAMILTON-JACOBI equation similar to (1.16) which depends on resources concentrations that they fully characterise in terms of the function  $\varphi$ . In [160], and [174], the authors obtain, on the time scale  $1/\sigma$ , an extended form of the CEAD similar to (1.19). Note that, by Theorem 1.1.2, the CEAD is also observable in the time scale  $1/\sigma$ . In a model

with a single resource but multidimensional traits, [160] establishes rigorously the CEAD result which was only *formally* given in [194].

*Stochastic approach.* In [43], CHAMPAGNAT ET AL consider an asexual population (bacteria here) in a stochastic multi-resource chemostat model which couples deterministic and stochastic dynamics. The bacteria dynamics follows a birth and death process with reproduction due to resources consumption and mutation. The resource process is deterministic between birth or death events in the bacterial population, with coefficients depending on the composition of the population.

Considering a double asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ), their stochastic IBM, on the mutational time scale  $1/Kp$ , converges to a pure jump process describing the successive invasions of advantageous mutants. This process extends the TSS and PES processes (see Theorem 1.1.4 and [37] for the TSS and [47] for the PES) to chemostat system.

In the small mutational steps limit ( $\sigma \rightarrow 0$ ), the *TSS/PES process for chemostat system* in turn gives rise, on the time scale  $1/\sigma^2$ , to a differential equation in phenotypic space describing the dynamics of coexisting traits and extending the CEAD. So, the global time scale leading to the CEAD observation is  $1/Kp\sigma^2$ .

**Co-evolution of predator-prey populations.** *When different species interact within an ecological network, each species induces selection pressure on the species with which it interacts, linked to the individual characteristics involved in the interaction. For example, in a predator-prey interaction, prey defence and predator preference mechanisms are subject to co-evolutionary pressures and therefore generate co-evolutionary dynamics [131]. By considering interactions between species, MARROW ET AL have heuristically extended the theory of adaptive dynamics into the framework of predator-prey dynamics in [161]. More precisely, they introduce a two-dimensional version of the canonical equation that describes, using a coupled system of two differential equations, the co-evolutionary dynamics of prey and predator traits. (see also [141])*

In [55], COSTA ET AL are interested in the impact in a predator-prey community. They introduce an IBM of the community that takes into account both prey and predator phenotypes in order to describe the phenotypic co-evolution of preys and predators. The community evolves as multi-type birth and death process with mutations.

They first consider the infinite particle approximation of the process without mutation to study the long term behaviour of a predator-prey community. In this limit of large population ( $K \rightarrow +\infty$ ), the process can be approximated by a system of differential equations. They prove the existence of a unique globally asymptotically stable equilibrium under specific conditions on the interaction among prey individuals.

Taking simultaneously the asymptotics of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) under the separation between the demographic and mutational time scales of the form (1.20), the community evolves, on the mutational time scale  $1/Kp$ , according to a Markovian jump process. This process describes the successive equilibria of the predator-prey community and extends the TSS/PES to a co-evolutionary framework.

The limit of small mutations ( $\sigma \rightarrow 0$ ) leads, on the time scale  $1/\sigma^2$ , to a system of two differential equations which is a version of the CEAD for the predator-prey co-evolution.

**Spatially and phenotypically structured populations.** *Spatial aspects are fundamental in ecology [179], both for the study of the colonisation of new habitats by dispersing species and for the propagation of epidemics [175, 179] or population agglomeration phenomena [215]. The*



*spatial aspect of the environment plays an important role in the appearance and maintenance of species diversity and stability of polymorphism in particular for speciation [70, 77]. It is recognised that a heterogeneous environment facilitates the coexistence of species [156, 83].*

In [46, 154], the authors consider an IBM which describes spatially and phenotypically structured asexual populations. More precisely, the birth, death, competitive pressure, mutation and spatial motion of each individual depend both on its position and its phenotypic trait. The spatial motion is driven by a reflected diffusion in a bounded domain.

In [46], CHAMPAGNAT and MÉLÉARD prove, under a large population asymptotic ( $K \rightarrow +\infty$ ), the convergence of the microscopic model to a deterministic continuous process, solution to a non-linear reaction diffusion defined on the of traits and locations.

Moreover, in [154], LEMAN establishes the convergence of the microscopic process under the double asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) satisfying the separation of time scales of the form (1.20), towards a jump MARKOV process. This limiting jump process describes, on the mutational time scale  $1/Kp$ , an evolutionary phenomenon using a sequence of monomorphic equilibria characterised by their spatial pattern and their trait and extends the TSS to a spatial context.

A natural perspective that has been proposed in [155] but not yet conducted would be to establish a spatial canonical equation for this spatially structured model.

#### 1.1.4.4 Other mathematical approaches

In this section, we present other mathematical strategies leading to the CEAD, whose our study, or to other mathematical objects, using new sets of combination of biological assumptions. So, new time scales are proposed. Note that, all these approaches can be summarised graphically in Figure 1.24.

**Populations of finite sizes.** *The canonical equation completely omits an important aspect of evolution called genetic drift (this terms refers to allelic fluctuations which have stochastic causes) or weak selection, which captures the stochasticity in evolutionary dynamics due to the preponderance of near-neutral mutations [138], i.e. with slightly advantageous or deleterious effects, that invade and fix in the population by chance. In contrast, evolution governed by strong selection produces selective sweeps [164, 206], where mutations rapidly settle in the population, leading to deterministic behaviour as in the CEAD. Controversy exists around this issue of neutrality, as biologists debate the preponderance of neutral mutations over strong selection. The classical theory of population genetics is largely based on this assumption of neutrality or weak selection [138, 95].*

*Although a limit of small mutations (i.e. weak selection) is used in Theorem 1.1.6, genetic drift is absent from the canonical equation because of the order of the limits, since an infinite population limit was first used to prove the convergence of the IBM to the TSS in Theorem 1.1.4. It is then natural to ask what happens to the CEAD when the large population assumption ( $K \rightarrow +\infty$ ) is removed, that is, when we apply a limit of rare mutations ( $p \rightarrow 0$ ) and then a limit of small mutations ( $\sigma \rightarrow 0$ ) to the IBM.*

In [44, 45], CHAMPAGNAT and LAMBERT consider a stochastic IBM of population structured by trait undergoing mutation and selection. In order to include the effect of stochasticity (genetic drift), they consider *finite populations* whose size fluctuates randomly through time, in contrast to classical works in both population genetics (where the population size traditionally remains

constant) and adaptive dynamics (where population sizes are infinite). Applying a limit of rare mutations ( $p \rightarrow 0$ ) to these populations, while keeping the population size finite, they obtain the convergence, on the time scale  $1/p$ , of their IBM to the *equivalent of the TSS* in finite population, denoted by “TSS” in Figure 1.20. Note that for the limit of rare mutations to give this non-trivial result, it is necessary to prevent the extinction of the IBM, which would otherwise occur in the limit even before the first mutation.

Then, they apply a limit of small mutations ( $\sigma \rightarrow 0$ ) to the equivalent of the TSS, that leads, on the time scale  $1/\sigma^2$ , to a diffusion process, unique solution to a (Stochastic Differential Equation) SDE that they call the *Canonical Diffusion of Adaptive Dynamics* (CDAD). The diffusion term of this SDE captures genetic drift while the drift term embodies *directional selection* driven by the gradient of the fixation probability, also interpreted as an invasion fitness. Therefore, the time scale in observing the CDAD phenomenon is  $1/p\sigma^2$ .

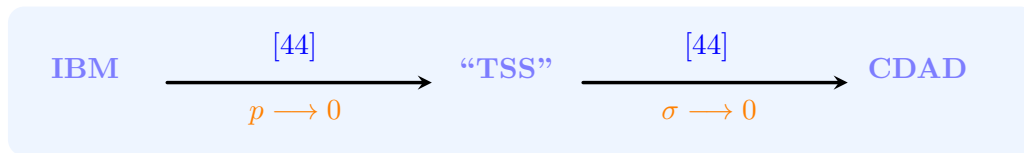


FIGURE 1.20 – The case of finite populations: the CDAD.

**The three simultaneous limits.** The order of the different scales of the TSS/PES approach leads to some remarks. First, the large population limit alone leads to the PDE (1.10), the basis of the PDE approach. For the TSS/PES approach, it is essential to combine simultaneously an asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ), since the rare mutations limit alone would lead to population extinction even before the first mutation [37, Remark 3]. Moreover, the TSS/PES and then the CEAD are obtained successively by applying the two limits ( $K \rightarrow +\infty, p \rightarrow 0$ ) then  $\sigma \rightarrow 0$  separately and on different time scales. This implies that they give no clue about how the biological parameters  $K, p$  and  $\sigma$  should compare to ensure that the CEAD approximation of the IBM is correct. So, in [6], BAAR ET AL analyse the situation when the limits of large population ( $K \rightarrow +\infty$ ), rare mutations ( $p \rightarrow 0$ ) and small mutation steps ( $\sigma \rightarrow 0$ ) are taken *simultaneously* (see Figure 1.21) and it gives a range of parameters leading to the CEAD.

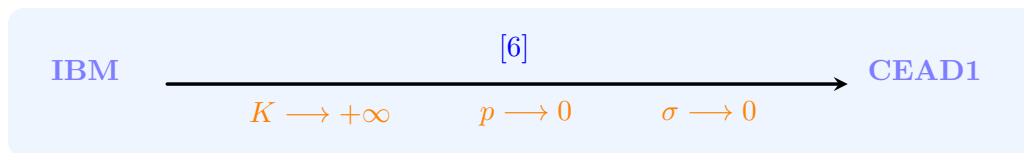


FIGURE 1.21 – The case of three simultaneous limits.

To be a bit more precise, they use in [6], the following conditions on the scaling parameters. There exists a small  $\alpha > 0$  such that

$$K^{-\frac{1}{2}+\alpha} \ll \sigma \ll 1 \quad \text{and} \quad \exp(-K^\alpha) \ll p \ll \frac{\sigma^{1+\alpha}}{K \log(K)}, \quad \text{as } K \rightarrow +\infty, p, \sigma \rightarrow 0.$$

The time scale, on which they control the population process driven by the CEAD, is  $1/Kp\sigma^2$  and can be explained as follows. The expected time for a mutation event to happen is of order  $1/Kp$ , the probability that a mutant invades is of the order  $\sigma$ , and an order of  $1/\sigma$  mutant invasion is required to observe a macroscopic change in the trait value of the essentially monomorphic population in the limit.

The conditions on the scaling parameters can be explained as follows. First, they prove that the duration of one successful invasion phase is of order  $\log(K)/\sigma$ . Therefore, the condition  $p \ll \sigma^{1+\alpha}/(K \log(K))$  allows mutation events during some successful invasions phases but ensures that there are never the appearance of two invading (successful) mutants during a such a phase and thus guarantees the separation of the ecological and evolutionary time scales. Second, the random fluctuations of the population process are of the order  $K^{-1/2}$ , therefore the condition  $K^{-\frac{1}{2}+\alpha} \ll \sigma$  ensures that the sign of mutant's fitness is not influenced by the fluctuations of the population size. Finally, they prove that  $\exp(K^\alpha)$  is the time the population stays within an  $(\varepsilon\sigma)$ -neighbourhood, with small  $\varepsilon > 0$ , when a mutation occurs. Therefore, the condition  $\exp(-K^\alpha) \ll p$  ensures that the population is still in this neighbourhood when a mutation occurs.

The two following approaches do not give a convergence result towards a version of the CEAD but they provide a convergence of the IBM to a solution of, either an ODE system for exponents of population sizes or a HAMILTON-JACOBI equation, each proposing an original time scale change.

**A system of ODE for exponents of population sizes.** In [49], CHAMPAGNAT ET AL consider a stochastic model for the evolution of a discrete population of bacteria structured by a 1-dimensional phenotypic trait taking into account a mechanism of Horizontal Gene Transfer (HGT) and *finitely* many values with density-dependent competition. More precisely, the trait space is the grid of mesh  $\delta > 0$  of  $[0, 4]$ , i.e.  $[0, 4] \cap \delta\mathbb{N} = \{0, \delta, \dots, L\delta\}$  with  $L = \lfloor 4/\delta \rfloor$  where  $\lfloor a \rfloor$  designates the lower integer part of  $a \in \mathbb{R}$ . The population is described by the vector  $(N_0^K(t), \dots, N_\ell^K(t), \dots, N_L^K(t))$  where  $N_\ell^K(t)$  is the number of individual of trait  $x := \ell\delta$  at time  $t$ .

They focus on a parameter scaling where individual mutations are rare but the global mutation rate tends to infinity. More precisely, they suppose that a mutation occurs, at the individual level, with a probability  $p := K^{-\alpha}$  with  $\alpha \in (0, 1)$ , which tends to 0 when  $K \rightarrow +\infty$ . So the total mutation rate in a population with size of order  $K$  is equal to  $Kp = K^{1-\alpha}$  which goes to infinity with  $K$ . Note that this specific mutation scale is far from the stochastic approach presented above in Section 1.1.4.2 where the total mutation rate  $Kp \rightarrow 0$  by (1.20) leading to a time scale separation between demographic and mutational events. In [49], this specific mutational scale implies that negligible subpopulations of size order  $K^\beta$ ,  $\beta \in (0, 1)$  may have a strong contribution to evolution.

Their main result quantifies the asymptotic dynamics of subpopulation sizes on a logarithmic scale. Defining for all  $\ell \in \{0, \dots, L\}$ ,  $\beta_\ell^K(t)$  such that  $N_\ell^K(t \log(K)) := K^{\beta_\ell^K(t)} - 1$  which evolves on the time scale  $\log(K)$ , they prove on the double asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p := K^{-\alpha} \rightarrow 0$ ) that the limit of the sequence  $(\beta_0^K, \dots, \beta_L^K)_{K \in \mathbb{N}}$  is a system of ODE (see Figure 1.22). This limit function is a deterministic piecewise affine continuous function which is characterised along successive phases determined by their resident or dominant trait. Note that local extinction is possible when the exponent  $\beta$  hits 0. A similar scaling was also studied in [29, 52, 23].

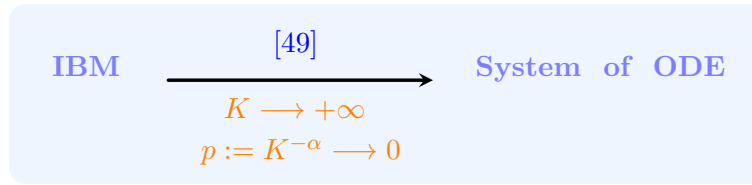


FIGURE 1.22 – The case of a system of ODE.

**Another time scaling to capture a Hamilton-Jacobi equation.** As in the approach described in the previous paragraph, CHAMPAGNAT ET AL consider in [48] a stochastic IBM for the evolution of a discrete population structured by a 1-dimensional trait with values on a finite grid  $[0, 1] \cap \delta_K \mathbb{N}$  with mesh grid  $\delta_K > 0$ , and subject to mutation and selection. Their main result quantifies the asymptotic dynamics of subpopulation sizes, of the order  $K^{\beta_{\ell}^K(t)}$ . They prove on the double asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma := 1/\log(K) \rightarrow 0$ ) but *frequent mutations* ( $p = 1$ ) and such that the mesh grid satisfies  $\delta_K \ll \sigma$ , that the limit of the sequence  $(\beta_0^K, \dots, \beta_{1-\delta_K}^K)_{K \in \mathbb{N}}$  is the unique solution to HAMILTON-JACOBI equation. This convergence is summed up in Figure 1.23.

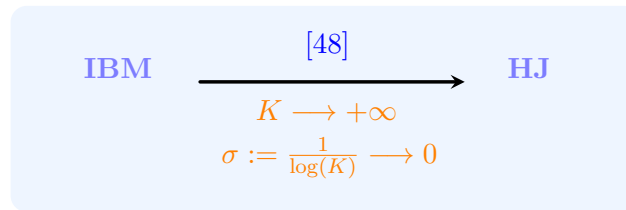


FIGURE 1.23 – The case of another time scale to capture a HAMILTON-JACOBI (HJ) equation.

In [23], BLATH ET AL extend the result of [49] by considering an additional mechanism to HGT, mutation and competition: the dormancy mechanism.

#### 1.1.4.5 Graphical summary and our approach

The mathematical approaches developed previously can be summarised in Figure 1.24, where our mathematical contribution is also represented. All these approaches are based on IBM.

- The large population limit ( $K \rightarrow +\infty$ ) alone (see Theorem 1.1.2 and [102]), represented by  $\xrightarrow{\text{orange arrow}}$ , leads to the PDE (1.16) (  $\bullet$  ).
- The PDE approach (see Section 1.1.4.1 and [75]), represented by  $\xrightarrow{\text{dotted orange arrow}}$ , consists in applying a small mutations limit ( $\sigma \rightarrow 0$ ) and a time change in  $1/\sigma$  to (1.16) where the occurrence of mutations is assumed *frequent*, to obtain the CEAD (1.19) (  $\bullet$  ).
- The stochastic approach (see Section 1.1.4.2), is done in two successive steps. The first one (see Theorem 1.1.4 and [37]), represented by  $\xrightarrow{\text{green arrow}}$ , consists in taking a simultaneous asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) such that  $Kp \rightarrow 0$  by (1.20) to prove the convergence from the IBM to the TSS (  $\bullet$  ). The second one (see

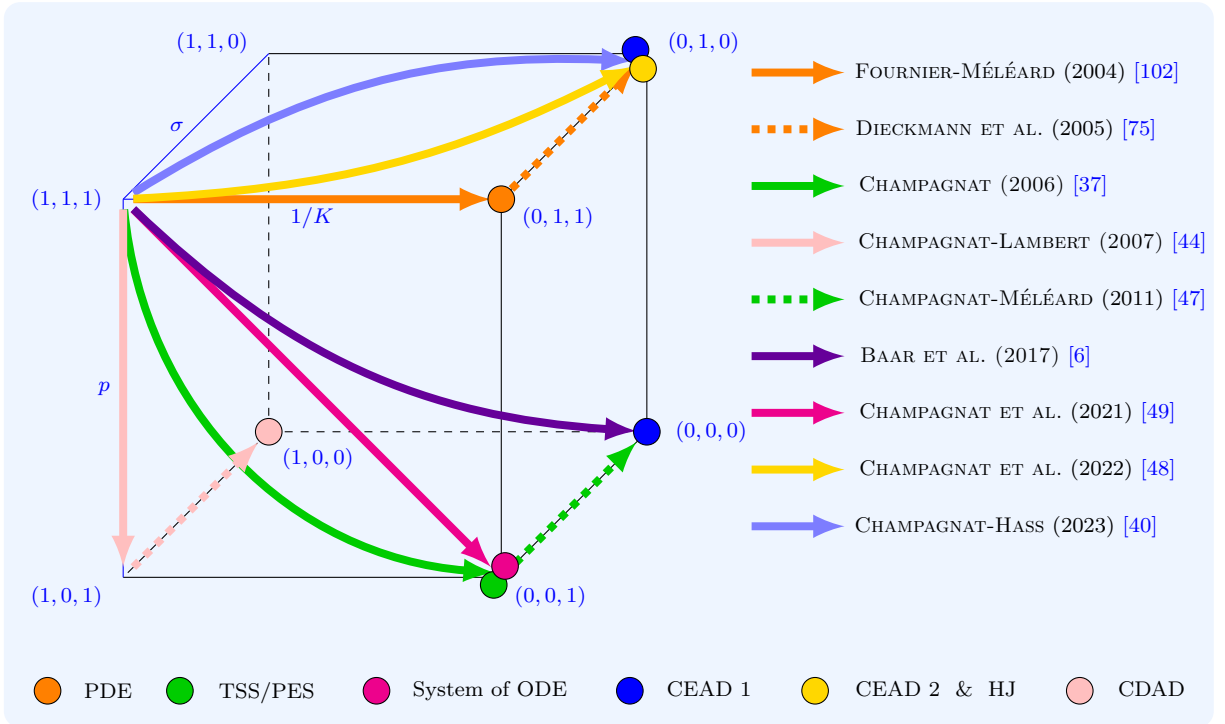




FIGURE 1.24 – Graphical summary of all convergences to the CEAD. We consider the triplet of parameters  $(\frac{1}{K}, p, \sigma)$ .

Theorem 1.1.6 and [47]), represented by  $\text{---}\blacktriangleright$ , consists in applying a small mutations limit ( $\sigma \rightarrow 0$ ) and a time change in  $1/\sigma^2$  to the TSS to obtain the original version of the CEAD ( $\bullet$ ) visible on  $1/Kp\sigma^2$  from the initial IBM.

- The approach with populations of finite sizes (see Section 1.1.4.4 and [44]) is represented by  $\text{---}\blacktriangleright$ . It requires two successive asymptotics, first of rare mutations ( $p \rightarrow 0$ ) and then of small mutations ( $\sigma \rightarrow 0$ ) to obtain convergence of the IBM to the CDAD ( $\circ$ ) visible on the time scale  $1/p\sigma^2$ .
- The approach with the three simultaneous limits ( $K \rightarrow +\infty, p \rightarrow 0, \sigma \rightarrow 0$ ), (see Section 1.1.4.4 and [6]), represented by  $\text{---}\blacktriangleright$ , is also visible on  $1/Kp\sigma^2$ .
- The approach with a system of ODE limit (see Section 1.1.4.4 and [49]) is represented by  $\text{---}\blacktriangleright$ . It consists in establishing the convergence of the IBM to a system of ODE ( $\bullet$ ) under a double asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p := K^{-\alpha} \rightarrow 0$ ) such that  $Kp \rightarrow +\infty$ . Evolution acts on the time scale  $\log(K)$ .
- The approach with a HAMILTON-JACOBI equation (see Section 1.1.4.4 and [48]) is represented by  $\text{---}\blacktriangleright$ . It consists in establishing the convergence of the IBM under the assumption of frequent mutations ( $p = 1$ ) and the double asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma := 1/\log(K) \rightarrow 0$ ) to a solution of the HAMILTON-JACOBI equation ( $\bullet$ ).

**Our approach.** Our approach considers the IBM of Section 1.1.3.2 (asexual population with a large but constant population size, structured by a 1-dimensional quantitative trait subject to selection and mutation occurring during life). In our study, we investigate under the assumption

of frequent mutations ( $p = 1$ ) (as defended in Section 1.1.4.6), the long term behaviour of the system in the simultaneous limit of large population ( $K \rightarrow +\infty$ ) and small mutational effects ( $\sigma \rightarrow 0$ ) proving convergence to the CEAD (1.24) on the time scale  $1/K\sigma^2$ . Figures 1.13 and 1.14 illustrate this convergence result. After conveniently scaling the population state, the result relies on a *slow-fast dynamics* (see Section 1.2.2). As the mutation size parameter is small, the population distribution tends to be close to a DIRAC mass. Therefore, the mean trait appears as the natural slow component. We will prove that it acts on the time scale  $1/K\sigma^2$  and that the fast component acts on the time scale  $K$ . The fast component is given by the dynamics of the centered and dilated distribution of traits following a discrete version of the centered FLEMING-VIOT process. We prove that the centered FLEMING-VIOT process is ergodic, and thus we expect the centered and dilated distribution of traits to stabilise on the slow time scale. Hence the distribution of traits should stabilise to a DIRAC mass and the dynamics of the dominant trait, driven by the CEAD, corresponds to that of the mean trait.

In contrast to earlier works, e.g. by CHAMPAGNAT and MÉLÉARD in [47], we establish this convergence – *in one step* – as represented in Figure 1.25 and in Figure 1.24 by  (pointing to ).

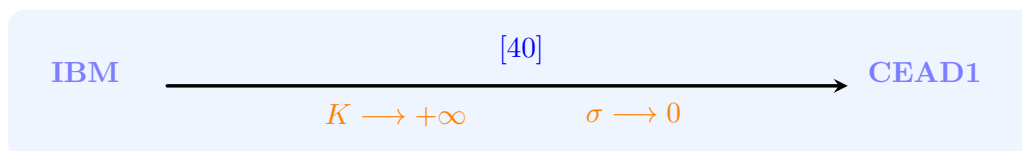


FIGURE 1.25 – Summary of our approach.

As with the approach of [48], we consider a similar configuration where  $p = 1$ ,  $K \rightarrow +\infty$  and  $\sigma \rightarrow 0$ . However, in [48], their framework which allows to describe negligible but non-extinct population (present everywhere in area) and to characterise the dynamics of exponents  $\beta$ , is close to the PDE setting of [12], so that their stochastic population size process, conveniently normalised, converges to a solution of a HAMILTON-JACOBI equation. In contrast, in our case population concentrates around a single trait value.

As with the approach of BAAR ET AL in [6], our result holds only for a certain range of parameters  $\sigma$  (as a function of  $K$ ) which must be small enough but not too small either:

$$\forall C > 0, \quad K^{-C \log(K)} \ll \sigma \ll K^{-2}.$$

The reason for these two bounds will be discussed in Section 1.3.

The reason for not considering the same IBM as in [37, 47, 6] is because, in this case, the slow-fast analysis involves three time scales: a slow one corresponding to the dynamics of the mean trait acting on the evolutionary time scale  $1/K\sigma^2$ , a fast one corresponding to the dynamics of the centered, dilated distribution of traits acting on the evolutionary time scale  $K$  and a very fast one corresponding to the population size dynamics acting on the ecological time scale 1. For simplicity, we focus in this thesis on a model with constant population size to reduce the number of time scales to two. We expect our results to extend to general IBMs as in [37, 47, 6] and we leave this for future works (see Section 1.3.3).

As our approach is new compared to the above references, it implies a new proof which is based on a slow-fast asymptotic analysis using the stochastic averaging method of KURTZ [144].

Section 1.2.2 gives an overview of the averaging methods, while Section 1.2.1 provides state of the art and methods for the FLEMING-VIOT process, which plays the role of the limit fast component in our slow-fast analysis.

#### 1.1.4.6 Biological controversies

The PDE and stochastic approaches leading to the CEAD are interesting because they allow us to predict the evolutionary fate of an ecological community, but they remain controversial in the biological community. In 2005, a special issue of *Journal of Evolutionary Biology* devoted to adaptive dynamics included a number of articles by researchers from this community, notably a central article describing the assumptions, tools and main criticisms of this theory [212]. These controversies suggest several mathematical issues that may help to identify the relevance of various biological objections and are summarised in [38].

In [195], a biological criticism is raised for the PDE approach because of the so-called *tail-problem*: in PDEs like (1.14) and under the scaling (1.15), exponentially small populations still have positive density everywhere in the trait space. This can have a strong influence on the future evolutionary dynamics of the population compared to the IBM where exponentially small density should correspond to extinct traits. In particular, the evolutionary speed is too fast. Modification of the HAMILTON-JACOBI equation were proposed in [195, 173, 129] to solve this problem. The idea is to introduce population extinction when the exponent  $\beta$  is too small.

The main controversy in [212] concerns the stochastic approach since it is based on a non-realistic biological assumption: mutations are assumed to be *too rare*. We recall that in adaptive dynamics, only mutations that have a non-deleterious phenotypic influence and produce viable individuals able to reproduce should be taken into account. Since proteins are encoded by a small fraction of DNA and many mutations produce non-functional proteins by altering their chemical properties, thus producing non-viable individuals, it is not unreasonable to assume that mutations are rare but certainly not as rare as in (1.20) in order to apply Theorem 1.1.4 in the TSS/PES approach. As a consequence of the assumption of rare mutations is that the CEAD takes place on  $1/Kp\sigma^2$  which is a too long time scale. Moreover, this assumption also implies that populations are at all times monomorphic, which is not the case for real populations. Hence the approaches [169, 49] to try to solve this problem.

**Our solution.** The goal of this thesis is to provide another approach whose aim is to correct the biological controversy on rare mutations by proposing more realistic probabilistic models in which the rare mutations assumption no longer applies. More precisely, our aim is to investigate mathematically, under a simultaneous asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ), the consequences of a new biological assumption of *frequent* mutations ( $p = 1$ ) on the CEAD. Note that the two limits  $\sigma \rightarrow 0$  and  $K \rightarrow +\infty$  need to be combined because making  $\sigma \rightarrow 0$  *without*  $K \rightarrow +\infty$  would lead, in similar way to [37, Remark 3], to the fast extinction of the population, and the converse problem ( $K \rightarrow +\infty$  *then*  $\sigma \rightarrow 0$ ) has already been studied (see Section 1.1.4.1).

In some sense, our study can be seen as an intermediate approach between the PDE and stochastic approaches of adaptive dynamics. Indeed, there is some consistency between our approach and the stochastic ones [47, 6] in the sense that, the time scales to observe the CEAD have the same expression. Also, we share with the PDE approach the assumption that mutations, at the individual level, are *frequent*. Hence, our approach takes places on a *intermediate*

evolutionary time scale satisfying

$$\frac{1}{\sigma} \ll \frac{1}{K\sigma^2} \ll \frac{1}{Kp\sigma^2}$$

and removes the main biological drawback of the PDE approach (tail-problem) where the evolutionary speed is too fast and that of the stochastic approach (rare mutations) where the evolutionary speed is too slow. Thus, our approach therefore seems to be the most suitable solution to these various biological controversies

Note that another strategy can be considered to address biological criticism on rare mutations, which will be presented in the prospect of this thesis in Section 1.3.3.

## 1.2 Mathematical theories

Multiscale models are usually represented by two components: fast and slow variables, which are characterised by different time scales. The entire model becomes intricate due to the interaction between the fast and slow variables, making it difficult to control the whole model. The averaging principle provides a way to study the long time behaviour of slow variables individually, thereby simplifying the model. The goal of Section 1.2.1 is to provide the state of the art and methods for the original FLEMING-VIOT process which is the basis of the fast component of our slow-fast analysis. Part of our contribution in this thesis aim at extending some of these methods to the centered version of the FLEMING-VIOT process. The goal of Section 1.2.2 is to present several averaging techniques and especially the stochastic averaging method of KURTZ, which will be used in this thesis.

### 1.2.1 Fleming-Viot processes

FLEMING and VIOT have introduced in [99] a probability-measure-valued MARKOV stochastic process modelling the dynamics of the distribution of allelic frequencies in a *selectively neutral* genetic population as influenced by mutation and random genetic drift: the original FLEMING-VIOT process. The initial model of [99] was progressively enriched with further mechanisms of Darwinian evolution: selection [99, 89, 91, 60, 86], recombination [93, 91] or the effect of an environment [124]. FLEMING and VIOT characterise in [99] the law of their process as a solution of a STROOCK-VARADHAN measure-valued martingale problem [207] both in the selective neutral case and the case with selection. To obtain the existence of such solution on a compact metric space, their method is based on discretisation of the mutation operator and tightness arguments. An alternative approach for *existence* is used, in the studies [177, 178, 139, 135, 205] based on the OTHA-KIMURA model [188, 189] and in the references [99, 61, 91, 85] based on its continuous-time version: the MORAN model (also called *continuous-state stepwise mutation model*). If we denote by  $N$  the population size, these authors construct a particle process whose limiting behaviour is analysed under the assumptions that the mutation step is proportional to  $1/\sqrt{N}$  and on the time scale  $(Nt)_{t \geq 0}$  which is presented below. In [124], another particle process, based on the look-down construction [78] is used to show the existence of the FLEMING-VIOT process in a random environment. This look-down construction also allows to analyse sample path properties of the process and has been used in numerous references since then, such as [80] [85, Chapter 5]. Below, we will discuss in detail the look-down construction that will be useful in the thesis.

In [99], uniqueness of the solution to the FLEMING-VIOT martingale problem in the selectively neutral case, is proved using uniqueness of moments of certain finite-dimensional distributions and arguments on semigroup. However, in the case where natural selection acts, the previous



method fails, but the result can be obtained from a version of the CAMERON-MARTIN-GIRSANOV formula [60, Chapter 10] [59, Theorem 5.1]. See also [92] for an application of this method in the case of unbounded selection function. In most references such as [85, 61, 88, 89, 91, 92], under a variety of assumptions, the *duality method* [94, Proposition 4.4.7] is used to prove the uniqueness of the FLEMING-VIOT process. The idea is to relate the distribution of the FLEMING-VIOT process with a process which is easier to understand, called *dual process*. It consists in establishing a duality identity which guarantees uniqueness of the solution to the FLEMING-VIOT martingale problem. Below, we recall the main ideas of this approach, that will be useful in the thesis. Other methods are used in some references: [62] makes use of resolvent estimates; [191, 190] prove existence and uniqueness of FLEMING-VIOT processes with *unbounded selection* intensity functions by using DIRICHLET's forms.

Questions of ergodicity of the FLEMING-VIOT processes were also the subject of many works discussed in Section 1.2.1.4 below.

This first mathematical section is organised as follows. In Section 1.2.1.1, we construct the MORAN's particle process leading to *existence* of a solution to the FLEMING-VIOT martingale problem. In Section 1.2.1.2, we give the construction of a version of the look-down. In Section 1.2.1.3, we prove uniqueness of this solution by developing the *duality method*. Finally, in Section 1.2.1.4, we present various *ergodicity* results on FLEMING-VIOT processes.

### 1.2.1.1 Moran's particle process and existence of the martingale problem

Following [85, 61, 99], we present the continuous-time OHTA-KIMURA model on  $\mathbb{Z}$  for a randomly mating population of MORAN type with stepwise mutation that we will call MORAN process in the sequel. This model comes from the population genetics literature. We define  $\mathcal{M}_{1,N}(\mathbb{Z})$  the set of probability measures on  $\mathbb{Z}$  consisting of  $N$  atoms of mass  $1/N$ .

Let us consider a fixed finite population of  $N$  individuals on  $\mathbb{Z}$ , where each individual is representing by an atom of mass  $1/N$ . An individual's type is represented by its position  $k \in \mathbb{Z}$  (which can also be interpreted as a trait or a site). So, the state of the population at time  $t$ , can be described by the following empirical measure of  $\mathcal{M}_{1,N}(\mathbb{Z})$

$$p_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)}$$

where  $X_i$  designates the type of the individual  $i$  on  $\mathbb{Z}$ . For simplicity, we denote  $p_t^N(\{k\}) := p_t^N(k)$  for all  $k \in \mathbb{Z}$ .

This model consider two mechanisms:

- **A mutation mechanism.** Each individual, independently of all other individuals, mutates from type  $i$  to type  $j$ , according to a (continuous time) simple random walk at rate

$$\theta_{ij} := \begin{cases} \frac{1}{2} & \text{if } j \in \{i-1, i+1\} \\ 0 & \text{if } j \notin \{i-1, i+1\} \end{cases}, \quad i \in \mathbb{Z}.$$

Note that the numbers of individuals of type  $i \in \mathbb{Z}$  is  $Np(i)$  if the distribution of population is  $p$ . So, the total rate of appearance of a mutant individual of type  $j$ , born of individuals of type  $i$  is  $Np(i)\theta_{ij}$ .

- **A sampling mechanism.** Each of the  $Np(i)$  individuals of type  $i \in \mathbb{Z}$ , independently of all other individuals, has a resampling rate  $\lambda > 0$ , and when they resample, the chance that they sample an individual of type  $j$  is  $p(j)$ . Hence, the total rate of resampling an

individual of type  $i$  into an individual of type  $j$  is  $\lambda N p(i)p(j)$ .

Therefore, the rate at which an individual of type  $i$  is replaced by an individual of type  $j$ , given that the distribution of the population is  $p$ , is  $Np(i)\theta_{ij} + \lambda N p(i)p(j)$ .

Let  $E$  be a metric space. Recall that for all  $\ell \in \mathbb{N}$ , we denote by  $\mathcal{C}^\ell(E, \mathbb{R})$  (respectively  $\mathcal{C}_b^\ell(\mathbb{R}, \mathbb{R})$ ) the space of functions of class  $\mathcal{C}^\ell$  from  $E$  to  $\mathbb{R}$  (respectively the space of real bounded functions of class  $\mathcal{C}^\ell(\mathbb{R}, \mathbb{R})$  with bounded derivatives). The generator of the  $\mathcal{M}_{1,N}(\mathbb{Z})$ -valued continuous-time MARKOV process  $(p_t^N)_{t \geq 0}$  is given for all  $\psi \in \mathcal{C}^0(\mathcal{M}_{1,N}(\mathbb{Z}), \mathbb{R})$  by

$$L_1 \psi(p) := \sum_{i \in \mathbb{Z}} \sum_{\substack{j \in \mathbb{Z} \\ j \neq i}} (Np(i)\theta_{ij} + N\lambda p(i)p(j)) [\psi(p^{ij}) - \psi(p)],$$

where  $p^{ij} = \sum_{k \in \mathbb{Z}} p^{ij}(k) \delta_k \in \mathcal{M}_{1,N}(\mathbb{Z})$  with for all  $k \in \mathbb{Z}$ ,

$$p^{ij}(k) := \begin{cases} p(k) - \frac{1}{N} & \text{if } k = i \\ p(k) + \frac{1}{N} & \text{if } k = j \\ p(k) & \text{if } k \notin \{i, j\} \end{cases}.$$

The original idea of FLEMING and VIOT [99, Appendix] consists in reducing the step size of the mutation process to  $1/\sqrt{N}$ , and so in order to obtain a non-trivial limit as  $N \rightarrow +\infty$ , this is matched by a speeding up of time by a factor of  $N$ . So, we define the process  $(Y_t^N)_{t \geq 0}$  on  $\mathbb{Z}/\sqrt{N}$  for all  $t \geq 0$ , for all  $A \subset \mathbb{Z}/\sqrt{N}$  by

$$Y_t^N(A) := p_{Nt}^N(\sqrt{N}A) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i(Nt)/\sqrt{N}}(A). \quad (1.25)$$

The tightness of the rescaled processes can be proved using the ALDOUS-REBOLLEDO criterion [2, 132]. In order to identify the limit as  $N \rightarrow +\infty$ , we consider the generator of the rescaled processes  $(Y_t^N)_{t \geq 0}$ , given for all  $\psi \in \mathcal{C}_b^0(\mathcal{M}_{1,N}(\mathbb{Z}/\sqrt{N}), \mathbb{R})$  by

$$L_N \psi(p) := \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} N (Np(i)\theta_{ij}^{(N)} + N\lambda p(i)p(j)) [\psi(p^{ij}) - \psi(p)], \quad (1.26)$$

where for all  $i, j \in \mathbb{Z}/\sqrt{N}$ ,  $\theta_{ij}^{(N)} = \frac{1}{2} \mathbb{1}_{j \in \{i - \frac{1}{\sqrt{N}}, i + \frac{1}{\sqrt{N}}\}}$ . Note that the factor  $N$  in (1.26) reflects

the time change, the sum is taken over the rescaled lattice  $\mathbb{Z}/\sqrt{N}$  and  $\theta_{ij}$  is modified in  $\theta_{ij}^{(N)}$ .

For a measurable real bounded function  $f$  and a measure  $p \in \mathcal{M}_{1,N}(\mathbb{Z})$ , we denote  $\langle f, p \rangle := \sum_{k \in \mathbb{Z}} f(k)p(k)$ . Prescribing a convenient set of test functions of the form  $\psi(p) := F_g(p) := F(\langle g, p \rangle)$ , with  $F \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$ ,  $g \in \mathcal{C}_b^3(\mathbb{R}, \mathbb{R})$ , we have from (1.26) that

$$L_N F_g(p) = \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} N^2 (p(i)\theta_{ij}^{(N)} + \lambda p(i)p(j)) \left[ F \left( \langle g, p \rangle - \frac{g(i)}{N} + \frac{g(j)}{N} \right) - F(\langle g, p \rangle) \right].$$

We will see below that this type of test function makes it possible to characterise the FLEMING-VIOT process. We will use a similar approach to prove existence of the *centered FLEMING-VIOT*

process. Applying TAYLOR's formula, we have that

$$L_N F_g(p) = \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} \left[ (\mathbf{A})_1^{ij} + (\mathbf{A})_2^{ij} + (\mathbf{A})_3^{ij} + (\mathbf{A})_4^{ij} + O\left(\frac{1}{N}\right) \right],$$

where

$$\begin{aligned} (\mathbf{A})_1^{ij} &:= N^2 F'(\langle g, p \rangle) p(i) \theta_{ij}^{(N)} \left[ \frac{g(j)}{N} - \frac{g(i)}{N} \right], & (\mathbf{A})_2^{ij} &:= N^2 F'(\langle g, p \rangle) \lambda p(i) p(j) \left[ \frac{g(j)}{N} - \frac{g(i)}{N} \right], \\ (\mathbf{A})_3^{ij} &:= N^2 \frac{F''(\langle g, p \rangle)}{2} p(i) \theta_{ij}^{(N)} \left[ \frac{g(j)}{N} - \frac{g(i)}{N} \right]^2, & (\mathbf{A})_4^{ij} &:= N^2 \frac{F''(\langle g, p \rangle)}{2} \lambda p(i) p(j) \left[ \frac{g(j)}{N} - \frac{g(i)}{N} \right]^2. \end{aligned}$$

Applying again TAYLOR's formula to  $g$ , we have

$$\begin{aligned} \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} (\mathbf{A})_1^{ij} &= F'(\langle g, p \rangle) \frac{N}{2} \sum_{i \in \mathbb{Z}/\sqrt{N}} p(i) \left( g\left(i - \frac{1}{\sqrt{N}}\right) - 2g(i) + g\left(i + \frac{1}{\sqrt{N}}\right) \right) \\ &= F'(\langle g, p \rangle) \sum_{i \in \mathbb{Z}/\sqrt{N}} p(i) \left[ \frac{g''(i)}{2} + O\left(\frac{1}{\sqrt{N}}\right) \right] \\ &= F'(\langle g, p \rangle) \left\langle \frac{g''}{2}, p \right\rangle + O\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

By symmetry,  $\sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} p(i) p(j) [g(j) - g(i)] = 0$ , hence

$$\sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} (\mathbf{A})_2^{ij} = 0.$$

Applying again TAYLOR's formula to  $g$ , we have

$$\begin{aligned} \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} (\mathbf{A})_3^{ij} &= \frac{F''(\langle g, p \rangle)}{4} \sum_{i \in \mathbb{Z}/\sqrt{N}} p(i) \left( \left[ g\left(i - \frac{1}{\sqrt{N}}\right) - g(i) \right]^2 + \left[ g\left(i + \frac{1}{\sqrt{N}}\right) - g(i) \right]^2 \right) \\ &= \frac{F''(\langle g, p \rangle)}{2N} \sum_{i \in \mathbb{Z}/\sqrt{N}} p(i) [g'(i)]^2 + O\left(\frac{1}{N}\right) \\ &= O\left(\frac{1}{N}\right). \end{aligned}$$

Finally, we have

$$\begin{aligned} \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{\substack{j \in \mathbb{Z}/\sqrt{N} \\ j \neq i}} (\mathbf{A})_4^{ij} &= \lambda \frac{F''(\langle g, p \rangle)}{2} \sum_{i \in \mathbb{Z}/\sqrt{N}} \sum_{j \in \mathbb{Z}/\sqrt{N}} p(i) p(j) [g^2(i) + g^2(j) - 2g(i)g(j)] \\ &= \lambda F''(\langle g, p \rangle) (\langle g^2, p \rangle - \langle g, p \rangle^2). \end{aligned}$$

The expression for the generator now reduces to

$$L_N F_g(p) = F'(\langle g, p \rangle) \left\langle \frac{g''}{2}, p \right\rangle + \lambda F''(\langle g, p \rangle) (\langle g^2, p \rangle - \langle g, p \rangle^2) + O\left(\frac{1}{\sqrt{N}}\right).$$

We are now in a position to identify the martingale problem that characterises the original FLEMING-VIOT process. Granted tightness of our sequence of rescaled processes and some technical work to check that the MARKOV property is preserved under passage to the limit [61, 91, 85], we obtain Theorem 1.2.2 below. Let us introduce some notations.

The original FLEMING-VIOT process is a measure-valued diffusion in  $\mathcal{M}_1(\mathbb{R})$ , the set of probability measures on  $\mathbb{R}$  with respect to the BOREL  $\sigma$ -field  $\mathcal{B}(\mathbb{R})$ , which is endowed with the topology of weak convergence making it a Polish space [20]. We consider the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$  where

$$\Omega := \mathcal{C}^0([0, +\infty), \mathcal{M}_1(\mathbb{R}))$$

is endowed with the SKOROHOD topology,  $\mathcal{F}$  is the associated BOREL  $\sigma$ -field and  $(\mathcal{F}_t)_{t \geq 0}$  is the canonical filtration. For a measurable real bounded function  $f$  and a measure  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote  $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x) \nu(dx)$ . Recall that for all  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  we denote for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ ,  $F_g(\nu) := F(\langle g, \nu \rangle)$ .

**Definition 1.2.1.** *The probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the original FLEMING-VIOT martingale problem with resampling rate  $\lambda \in (0, +\infty)$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$  if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$  and for each  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,*

$$M_t^F(g) := F(\langle g, Y_t \rangle) - F(\langle g, Y_0 \rangle) - \int_0^t \mathcal{L}_{\text{FV}} F_g(Y_s) ds \quad (1.27)$$

with for all  $\mu \in \mathcal{M}_1(\mathbb{R})$ ,

$$\mathcal{L}_{\text{FV}} F_g(\mu) := F'(\langle g, \mu \rangle) \left\langle \frac{g''}{2}, \mu \right\rangle + \lambda F''(\langle g, \mu \rangle) [\langle g^2, \mu \rangle - \langle g, \mu \rangle^2]$$

is a continuous square integrable  $\mathbb{P}_\nu^{\text{FV}}$ -martingale whose martingale bracket satisfies for all  $G, H \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and for all  $g, h \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\langle M^G(g), M^H(h) \rangle_t = 2\lambda \int_0^t G'(\langle g, Y_s \rangle) H'(\langle h, Y_s \rangle) [\langle gh, Y_s \rangle - \langle g, Y_s \rangle \langle h, Y_s \rangle] ds.$$

**Theorem 1.2.2** (FLEMING-VIOT [99]). **(1)** *For all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , there exists a probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  satisfying the martingale problem of Definition 1.2.1.*

**(2)** *Let  $Y^N$  be given as in (1.25). Then  $(Y_t^N)_{t \geq 0}$  converges in law, when  $N \rightarrow +\infty$ , to  $(Y_t)_{t \geq 0}$  whose law is solution to the martingale problem (1.27)*

As the original FLEMING-VIOT process is obtained as a scaling limit of MORAN's model, we can interpret the terms involving the first order derivative  $F'$  as describing the effect of the mutation whereas the ones involving the second order derivative  $F''$  as describing the effect of the random genetic drift. As we will see in Section 1.3.1, the centered version of the FLEMING-VIOT process is also solution of a measure-valued martingale problem.

We shall check that there is a unique solution to the FLEMING-VIOT martingale problem in Section 1.2.1.3. In the next section, we propose another particle construction of the FLEMING-VIOT: the so-called *DONNELLY-KURTZ's modified look-down*. An adaptation of this explicit construction will be exploited, in this thesis, to prove ergodicity results for the centered version of the FLEMING-VIOT process.

### 1.2.1.2 Donnelly-Kurtz's modified look-down

In [78, 80], DONNELLY and KURTZ have given a countable representation of the FLEMING-VIOT measure-valued process, in a selectively neutral case, as the empirical measure carried by a certain infinite interactive particle system. In contrast to the MORAN approach (see Section 1.2.1.1), here, the type of individual is interpreted as a *level* in  $\mathbb{N}^*$ . The DONNELLY-KURTZ's construction provides an explicit description of the genealogy of the individuals in the population. The role of neutrality is to ensure that, conditional on all the information up to any time, each individual has the same chance of having the longest line of descent, second longest line of descent and so on. Therefore, this ordering is just a random permutation of the labels of the original system and so the empirical measure remains unchanged. In other words, their approach exploits the notion of *exchangeability* which we recall the definition.

**Definition 1.2.3.** We say that the random variables  $Z_1, \dots, Z_n$ ,  $n \in \mathbb{N}^*$ , are exchangeable if, for any permutation  $\sigma$  of  $\{1, \dots, n\}$ , the vectors  $(Z_1, \dots, Z_n)$  and  $(Z_{\sigma(1)}, \dots, Z_{\sigma(n)})$  have the same distribution.

In [79], DONNELLY and KURTZ incorporate, to their representation, selection and recombination mechanisms (see also [85, Section 5.4]).

In this section, we present an adaptation of DONNELLY-KURTZ's modified look-down construction in the selectively neutral case.

We consider a probability space where we define the modified look-down process on  $\mathbb{R}_+$  as a population dynamics on the set  $\mathbb{N}^*$  of levels where one individual is assigned to each level. To each pair of levels  $(i, j) \in \mathbb{N}^2$  with  $1 \leq i < j$ , we assign an independent POISSON process  $(N_{ij}(t))_{t \geq 0}$  with intensity  $2\lambda > 0$  (corresponding to a resampling rate  $\lambda$  in the MORAN model) and to each level  $i \in \mathbb{N}^*$ , we assign an independent standard Brownian motion  $(B_i(t))_{t \geq 0}$  on  $\mathbb{R}_+$ . Jointly with the modified look-down is constructed for all  $N \in \mathbb{N}^*$ , the so-called  $N$ -look-down process whose evolution is given as follows:

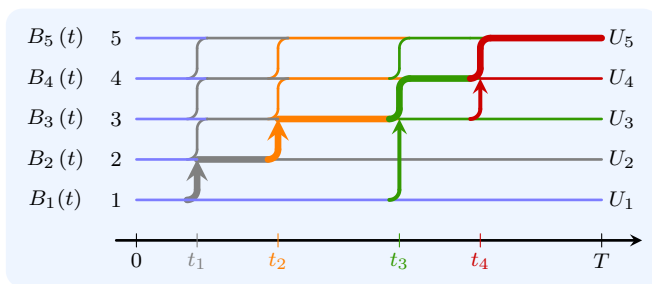


FIGURE 1.26 – Graphical representation of the modified look-down process with  $N = 5$  on  $[0, T]$ .

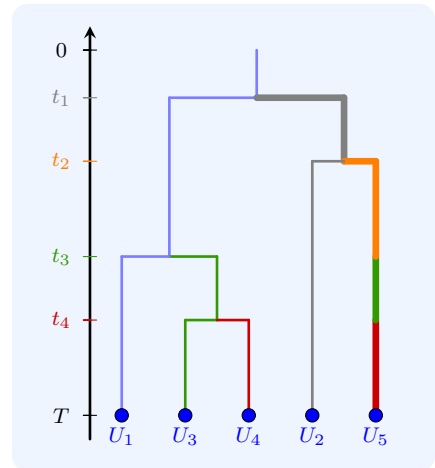


FIGURE 1.27 – KINGMAN's genealogy at time  $T$  in the modified look-down of Figure 1.26.

- (0) **Initial positions of particles.** An *exchangeable* initial condition  $(U_1(0), \dots, U_N(0))$  is chosen at time 0.

- (1) **Birth/Death rule.** Each jump time  $t_k$  of one of the POISSON process  $(N_{ij})_{1 \leq i < j \leq N}$  corresponds to a reproduction event at time  $t_k$ . When the time  $t_k$  is the jump time of the POISSON process  $N_{ij}$ , we put an arrow from  $i$  to  $j$ , as illustrated in Figure 1.26, which means that the individual at level  $i$  puts a child at level  $j$ . The offspring at level  $j$  adopts the current spatial position of its parent at level  $i$ . The parent level and position do not change. Individuals previously at level  $\ell \in \{j, \dots, N-1\}$  are shifted one level up to  $\ell+1$  and the individual at level  $N$  dies. Note that in this construction, the individual at level 1 is immortal, and so it is said that it has a descent that never dies out.
- (2) **Spatial motion.** Between reproduction events, individuals' spatial positions at each level  $i$  evolve according to the increments of the standard Brownian motion  $B_i(t)$  (which can be interpreted as a mutation in the MORAN model) and positions are transmitted to offsprings at birth.

Note that the  $N$ -modified look-down process is simply the first  $N$  levels of the  $(N+k)$ -modified look-down for any  $k \in \mathbb{N}^*$ . In other words, the modified look-down construction can be done with an infinite population as a projective limit of the so-called (infinite) modified look-down. Note also that this construction with infinite population is possible thanks to the mechanism (1) above. This would not be the case with the MORAN model with infinite population. Indeed, in the MORAN model with infinitely many individuals, there would be infinitely many arrows towards any individual at any level, in any time interval of positive length. This is a great advantage of the modified look-down construction.

For  $T > 0$  fixed, from [78, 80], the genealogy  $(k_{N,t})_{t \in [0,T]}$  in backward time since time  $T$ , of the population evolving according to the  $N$ -modified look-down  $(U_1(t), \dots, U_N(t))_{t \in [0,T]}$ , is exactly determined by KINGMAN's  $N$ -coalescent with coalescence rate  $2\lambda$ . In Figure 1.27 we give the KINGMAN genealogy associated to the 5-modified look-down of the Figure 1.26. Note that the mechanism (1) allows to order the individuals by decreasing persistence of their descent as we can see in Figure 1.27. We refer to [86, Chapter 2] for a brief overview of KINGMAN's genealogies and its main properties.

The previous construction of the  $N$ -look-down provides positions  $U_i(t)$  for the particle at level  $i \in \mathbb{N}^*$  and at time  $t \geq 0$  and it leads to Theorem 1.2.4 below. Let us introduce some notations.

We consider the  $N$ -MORAN particle process  $Y^N$  of Section 1.2.1.1 defined by

$$Y_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)} \tag{1.28}$$

where mutations are replaced by a Laplacian operator. Its state space is  $\mathcal{M}_{1,N}(\mathbb{R})$ , the set of probability measures on  $\mathbb{R}$  consisting of  $N$  atoms of mass  $1/N$ . In addition, ETHIER and KURTZ prove in [91, Theorem 6.1] that if  $(X_i(0))_{i \in \mathbb{N}^*}$  is exchangeable, then for all  $t > 0$ ,  $(X_i(t))_{i \in \mathbb{N}^*}$  is exchangeable. The infinitesimal generator of the  $\mathbb{R}$ -measure-valued process  $Y^N$  is given for all  $n \in \mathbb{N}^*$ ,  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,  $\mu_N \in \mathcal{M}_{1,N}(\mathbb{R})$  by

$$\begin{aligned} & L_N P_{f,n}(\mu_N) \\ &= P_{\frac{\Delta f}{2},n}(\mu_N) + \lambda N(N-1) \int_{\mathbb{R}} \int_{\mathbb{R}} \left[ P_{f,n} \left( \mu_N - \frac{\delta_x}{N} + \frac{\delta_y}{N} \right) - P_{f,n}(\mu_N) \right] \mu_N(dx) \mu_N(dy) \end{aligned}$$

where  $P_{f,n}(\mu_N) := \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu_N(dx_1) \dots \mu_N(dx_n)$ . The first term of the generator describes the effect of mutations according to the Laplacian operator. The second term describes

the sampling replacement mechanism: at rate  $\lambda$  (the sampling rate) an individual of type  $x$  is immediately replaced by one of type  $y$ . Note that the population size remains constant over time.

**Theorem 1.2.4** (DONNELLY-KURTZ [78, 80]). *For all  $t \geq 0$ ,  $N \in \mathbb{N}^*$ ,  $(U_i(t))_{1 \leq i \leq N}$  is exchangeable and*

$$\frac{1}{N} \sum_{i=1}^N \delta_{U_i(t)} \stackrel{\text{law}}{=} \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)}$$

where  $(X_i(t))_{1 \leq i \leq N}$  is the vector of positions of particles in the  $N$ -MORAN process in (1.28). In particular,  $\left(\frac{1}{N} \sum_{i=1}^N \delta_{U_i(t)}\right)_{t \geq 0}$  converges, when  $N \rightarrow +\infty$ , to the original FLEMING-VIOT process.

To study the invariant probability measure  $\pi$  of the centered FLEMING-VIOT process, we will use a variant of the previous modified look-down construction for time running backward from 0 to  $-\infty$  up to the coalescence time of the population. This will provide an explicit construction of  $\pi$ .

### 1.2.1.3 Dual process and uniqueness of the martingale problem

**Theorem 1.2.5** (FLEMING-VIOT [99]). *The FLEMING-VIOT martingale problem (1.27) has a unique solution.*

The proof of Theorem 1.2.5 relies on the *duality method* [85, 60, 89, 91] and [94, Proposition 4.4.7]. The idea is to relate the distribution of the original process with that of a simpler process, called *dual process*. This leads to a duality relation which ensures that two solutions to the martingale problem have the same 1-dimensional marginal laws, from which classical arguments, based on MARKOV's property [94, Theorem 4.4.2], allow to deduce uniqueness.

The first goal of this section is to construct a dual process for the original FLEMING-VIOT process. For this, it is required to introduce two other versions of the original FLEMING-VIOT martingale problem for test functions with multiple variables and for test functions given by polynomials in  $\nu \in \mathcal{M}_1(\mathbb{R})$ . These martingale problems are equivalent to the one of Definition 1.2.1 and it is usual to switch from one to the other [91]. The second goal of this section is to establish the duality relation and prove uniqueness of the solutions of (1.27). We propose, at the end of this section, more general FLEMING-VIOT processes, in order to present some ergodicity results in Section 1.2.1.4.

To prove that the centered version of the FLEMING-VIOT process admits a unique solution to its martingale problem, we will follow the same approach as proposed in the next three paragraphs and detail the arguments for this purpose.

**Extension to multiple variables.** We firstly introduce the version of the original FLEMING-VIOT martingale problem with  $p \in \mathbb{N}^*$  variables.

**Definition 1.2.6.** *A probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the original FLEMING-VIOT martingale problem with  $p$  variables with resampling rate  $\lambda$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$ , if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$  and for each*

$F \in \mathcal{C}^2(\mathbb{R}^p, \mathbb{R})$  and  $g = (g_1, \dots, g_p) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R}^p)$ ,

$$\begin{aligned} M_t^F(g) &:= F(\langle g_1, Y_t \rangle, \dots, \langle g_p, Y_t \rangle) - F(\langle g_1, Y_0 \rangle, \dots, \langle g_p, Y_0 \rangle) \\ &\quad - \int_0^t \sum_{k=1}^p \partial_k F(\langle g_1, Y_s \rangle, \dots, \langle g_p, Y_s \rangle) \left\langle \frac{g_k''}{2}, Y_s \right\rangle ds \\ &\quad - \lambda \int_0^t \sum_{i,j=1}^p \partial_{ij}^2 F(\langle g_1, Y_s \rangle, \dots, \langle g_p, Y_s \rangle) [\langle g_i g_j, Y_s \rangle - \langle g_i, Y_s \rangle \langle g_j, Y_s \rangle] ds \end{aligned} \quad (1.29)$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale.

As mentioned in [61, Theorem 5.1], the solution  $\mathbb{P}_\nu^{\text{FV}}$  of the martingale problem (1.27) is the *unique* solution to the previous martingale problem (1.29).

**Extension to polynomials in  $\nu$ .** Our goal here is to study the DOOB semi-martingale decomposition of polynomial functions of the original FLEMING-VIOT process of the form

$$P_{f,n}(\nu) := \langle f, \nu^n \rangle := \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) \nu(dx_1) \dots \nu(dx_n) \quad (1.30)$$

with  $n \in \mathbb{N}^*$ ,  $\nu \in \mathcal{M}_1(\mathbb{R})$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  and  $\nu^n$  is the  $n$ -fold product measure of  $\nu$ . To obtain the expression of  $\mathcal{L}_{\text{FV}}$  for polynomial functions  $P_{f,n}$  with  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ , we first look for this expression when  $f$  has product form. The martingale problem (1.29) gives heuristics for this issue: for the choice of  $F(x_1, \dots, x_n) := \prod_{i=1}^n x_i$ , noting that  $F(\langle g_1, \mu \rangle, \dots, \langle g_n, \nu \rangle) = P_{f,n}(\nu)$  with  $f(x_1, \dots, x_n) := \prod_{i=1}^n g_i(x_i)$  and  $g_i \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,  $i \in \{1, \dots, n\}$ , we deduce that for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ ,

$$\mathcal{L}_{\text{FV}} P_{f,n}(\nu) = \sum_{i=1}^n \left\langle \frac{g_i''}{2}, \nu \right\rangle \prod_{\substack{j=1 \\ j \neq i}}^n \langle g_j, \nu \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle g_i g_j, \nu \rangle - \langle g_i, \nu \rangle \langle g_j, \nu \rangle] \prod_{\substack{k=1 \\ k \neq i,j}}^n \langle g_k, \nu \rangle.$$

This leads us to introduce our second extension of the martingale problem (1.27).

**Definition 1.2.7.** *The probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the FLEMING-VIOT martingale problem for polynomials with resampling rate  $\lambda$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$ , if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$ , for all  $n \in \mathbb{N}^*$ , and for each  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,*

$$M_t^{(n)}(f) := \langle f, Y_t^n \rangle - \langle f, Y_0^n \rangle - \int_0^t \mathcal{L}_{\text{FV}} P_{f,n}(Y_s) ds \quad (1.31)$$

with for all  $\mu \in \mathcal{M}_1(\mathbb{R})$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathcal{L}_{\text{FV}} P_{f,n}(\mu) := \left\langle \frac{\Delta f}{2}, \mu^n \right\rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle] \quad (1.32)$$

is a continuous  $\mathbb{P}_\nu^{\text{FV}}$ -martingale in  $L^2(\Omega)$  where, for all  $i, j \in \{1, \dots, n\}$ ,  $\Phi_{i,j} : \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^0(\mathbb{R}^{n-1}, \mathbb{R})$ , with  $i \neq j$ , is the function obtained from  $f$  by inserting the variable  $x_i$  between  $x_{j-1}$  and  $x_j$  when  $i < j$  and by inserting the variable  $x_{i-1}$  between  $x_{j-1}$  and  $x_j$  when  $i > j$ :

$$\begin{aligned} \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_i, x_j, x_{j+1}, \dots, x_{n-1}) & i < j \\ \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_{i-1}, x_j, x_{j+1}, \dots, x_{n-1}) & i > j \end{aligned} \quad (1.33)$$



By [91, Theorem 3.2], the solution  $\mathbb{P}_\nu^{\text{FV}}$  of the martingale problem (1.27) is the unique solution to the previous martingale problem (1.31). As martingale problems of Definitions 1.2.1, 1.2.6 and 1.2.7 are equivalent, we deduce the following result.

**Corollary 1.2.8** (ETHIER-KURTZ [91]). *Existence and uniqueness hold for all martingale problems of Definitions 1.2.1, 1.2.6 and 1.2.7 and they all admit the same solution.*

The operator  $\mathcal{L}_{\text{FV}}$  acts on the function  $\nu \mapsto P_{f,n}(\nu)$  with fixed  $f$ . In view of (1.32), we can define an operator  $\mathcal{L}_{\text{FV}}^*$  acting on the function  $f \mapsto P_{f,n}(\nu)$  with fixed  $\nu$  such that  $\mathcal{L}_{\text{FV}}^* P_{f,n}(\nu) = \mathcal{L}_{\text{FV}} P_{f,n}(\nu)$ , which can be interpreted as the generator of a new stochastic process – the *dual process* – on the state space  $\mathcal{C} := \bigcup_{n \in \mathbb{N}^*} \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R})$ . The next paragraph is devoted to construct the dual process and establish the duality relation.

**Dual process and duality relation.** The relation (1.32) suggests that the  $\mathcal{C}$ -valued dual process  $(\xi_t)_{t \geq 0}$  only jumps, for all  $i, j \in \{1, \dots, n\}, i \neq j$  and  $n \geq 2$  from  $f \in \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R})$  to  $\Phi_{i,j} f \in \mathcal{C}_b^0(\mathbb{R}^{n-1}, \mathbb{R})$  at rate  $\lambda$ . Note that if  $n = 1$ , the dual process is in  $\mathcal{C}_b^0(\mathbb{R}, \mathbb{R})$  and so no further jumps occur. Between jumps, the dual process evolves according to the semi-group of operator  $(T_{\text{FV}}^{(n)}(t))_{t \geq 0}$  associated to the Laplacian operator, defined for all  $f \in L^\infty(\mathbb{R}^n)$  by

$$\forall t \geq 0, \forall x \in \mathbb{R}^n, \quad T_{\text{FV}}^{(n)}(t)f(x) := \int_{\mathbb{R}^n} f(u)g_{t,x}(u)du, \quad (1.34)$$

where  $g_{t,x}$  is a density of the Gaussian distribution  $\mathcal{N}^{(n)}(x, t^2 I_n)$  where  $I_n$  is the identity matrix of order  $n$ .

In view of the previous discussion, let us consider the MARKOV's a birth and death process  $M := (M(t))_{t \geq 0}$  in  $\mathbb{N}$  whose transition rates  $q_{i,j}$  from  $i$  to  $j$  are given by:

$$(1) \quad q_{n,n-1} = \lambda n(n-1) \quad (2) \quad q_{i,j} = 0 \text{ otherwise.}$$

Note that this process is a *pure death process* absorbing in 1. We define the dual process as follows:

**Definition 1.2.9.** *For all  $M(0) \in \mathbb{N}^*$ , for all  $\xi_0 \in \mathcal{C}_b^0(\mathbb{R}^{M(0)}, \mathbb{R})$ , for all  $k \leq M(0) - 1$ ,*

$$\xi_t := T_{\text{FV}}^{(M(\tau_k))}(t - \tau_k) \Lambda_k T_{\text{FV}}^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1}) \Lambda_{k-1} \cdots \Lambda_1 T_{\text{FV}}^{(M(0))}(\tau_1) \xi_0, \quad \tau_k \leq t < \tau_{k+1},$$

where  $(\tau_n)_{0 \leq n \leq M(0)-1}$  is the sequence of jump times of the birth-death process  $M$  with  $\tau_0 = 0$  and the convention  $\tau_{M(0)} = +\infty$  and where  $(\Lambda_n)_{0 \leq n \leq M(0)-1}$  is a sequence of random operators of type  $\Phi_{i,j}$ .

Following ETHIER-KURTZ's works [89, 91] (see also [85, Sections 1.6 and 1.12] and [61, Section 5.2]), we obtain the following *duality relation*.

**Theorem 1.2.10.** *We consider fixed  $M(0) \in \mathbb{N}^*$ ,  $\xi_0 \in \mathcal{C}_b^0(\mathbb{R}^{M(0)}, \mathbb{R})$  and  $(Y_t)_{t \geq 0}$  a stochastic process whose law  $\mathbb{P}_\nu^{\text{FV}}$  is a solution of the martingale problem (1.27) with  $\nu \in \mathcal{M}_1(\mathbb{R})$ . We introduce a dual process  $(\xi_t)_{t \geq 0}$  independent of  $(Y_t)_{t \geq 0}$  built on the same probability space (enlarging it if necessary). We shall denote by  $\mathbb{P}_{(\nu, \xi_0)}^{\text{FV}}$ , the law of  $((Y_t, \xi_t))_{t \geq 0}$  on this probability space. The duality identity is given by:*

$$\forall t \geq 0, \quad \mathbb{E}_{(\nu, \xi_0)} \left( \left\langle \xi_0, Y_t^{M(0)} \right\rangle \right) = \mathbb{E}_{(\nu, \xi_0)} \left( \left\langle \xi_t, Y_0^{M(t)} \right\rangle \right). \quad (1.35)$$

We rely on ETHIER-KURTZ's result [94, Theorem 4.4.2] to deduce uniqueness of the martingale problem (1.27) from the last result. This ETHIER-KURTZ's theorem ensures that it is sufficient to verify that if we give ourselves two solutions to the martingale problem (1.27), they have the same 1-dimensional marginal laws.

*Proof of Theorem 1.2.5.* Let  $(Y_t)_{t \geq 0}$  and  $(\widetilde{Y}_t)_{t \geq 0}$  be two solutions of the martingale problem (1.27) with the same initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$ . Let  $(\xi_t)_{t \geq 0}$  be the dual process with initial condition  $\xi_0 \in \mathcal{C}_b^0(\mathbb{R}^{M(0)}, \mathbb{R})$  with  $M(0) \in \mathbb{N}^*$ . We suppose that these three processes are built on the same probability space and independent of each other. These processes satisfy, the duality identity (1.35):

$$\forall t \geq 0, \quad \mathbb{E}_{(\nu, \xi_0)} \left( \langle \xi_0, Y_t^{M(0)} \rangle \right) = \mathbb{E}_{(\nu, \xi_0)} \left( \langle \xi_t, \nu^{M(t)} \rangle \right) = \mathbb{E}_{(\nu, \xi_0)} \left( \langle \xi_0, \widetilde{Y}_t^{M(0)} \rangle \right)$$

i.e.

$$\forall t \geq 0, \quad \mathbb{E}_{(\nu, \xi_0)} \left( P_{\xi_0, M(0)}(Y_t) \right) = \mathbb{E}_{(\nu, \xi_0)} \left( P_{\xi_0, M(0)}(\widetilde{Y}_t) \right).$$

As the set of test functions  $\left\{ P_{f, n} \mid f \in \mathcal{C}_b^0(\mathbb{R}^n, \mathbb{R}), n \in \mathbb{N}^* \right\}$  is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -convergence determining [60, Lemma 2.1.2], it is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -separating [94, Chapter 3, Section 4, p.112], so it follows that for any  $t \geq 0$ ,  $Y_t^{M(0)}$  and  $\widetilde{Y}_t^{M(0)}$  have the same law. In particular for the choice  $M(0) := 1$ , [94, Theorem 4.4.2] ensures uniqueness to the martingale problem (1.27).  $\square$

**Extension to more general Fleming-Viot processes.** There are many more general forms of the FLEMING-VIOT process that are of considerable importance in mathematical population genetics taking into account mechanisms of mutation, selection and recombination. An excellent survey is [91] (see also [60, Section 5.7 and Chapter 10]). Here, we focus on a class of FLEMING-VIOT processes with a *generic* mutation operator  $A$  and where the type space  $E$  is a Polish space endowed with the topology of weak convergence. Following [60, 91], the aim of this paragraph is to present a general framework and a general identity duality in order to enounce and prove some ergodicity results in Section 1.2.1.4.

We denote by  $\mathcal{M}_1(E)$  the set of probability measures on  $E$ , endowed with the topology of weak convergence making it a Polish space [20]. The function  $\varphi \in \mathcal{C}^0(\mathcal{M}_1(E), \mathbb{R})$  is said to be differentiable [60, p 18-19] if the limit

$$\frac{\delta \varphi(\mu)}{\delta \mu(x)} := \lim_{\varepsilon \rightarrow 0} \frac{\varphi(\mu + \varepsilon \delta_x) - \varphi(\mu)}{\varepsilon} = \frac{\partial}{\partial \varepsilon} \varphi(\mu + \varepsilon \delta_x) \Big|_{\varepsilon=0}$$

exists for each  $x \in E$  and belongs to  $\mathcal{C}^0(E, \mathbb{R})$  for all  $\mu \in \mathcal{M}_1(E)$ . The second derivative is defined by

$$\frac{\delta^2 \varphi(\mu)}{\delta \mu(x) \delta \mu(y)} = \frac{\partial^2}{\partial \varepsilon_1 \partial \varepsilon_2} \varphi(\mu + \varepsilon_1 \delta_x + \varepsilon_2 \delta_y) \Big|_{\varepsilon_1 = \varepsilon_2 = 0}$$

if it exists for each  $x, y \in E$  and belongs to  $\mathcal{C}^0(E \times E, \mathbb{R})$  for all  $\mu \in \mathcal{M}_1(E)$ . The generator  $\mathcal{L}_{\text{FV}}^{E, A}$  of such FLEMING-VIOT processes (see [60, Section 2.6] and [91]) is given for all test functions  $\varphi \in \mathcal{C}_b^2(\mathcal{M}_1(E), \mathbb{R})$  by

$$\mathcal{L}_{\text{FV}}^{E, A} \varphi(\mu) := \int_E \left( A \frac{\delta \varphi(\mu)}{\delta \mu(x)} \right) \mu(dx) + \lambda \int_E \int_E \frac{\delta^2 \varphi(\mu)}{\delta \mu(x) \delta \mu(y)} [\mu(dx) \delta_x(dy) - \mu(dx) \mu(dy)]. \quad (1.36)$$

Note that the model originally studied by FLEMING and VIOT [99] is the one where  $A$  is the Laplacian operator on  $E := \mathbb{R}$  for a class of test functions  $\varphi(\mu) := F_g(\mu) := F(\langle g, \mu \rangle)$ ,

$\mu \in \mathcal{M}_1(\mathbb{R})$  with  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . Another example studied in [21] is where  $A$  is the fractional Laplacian operator on  $E := \mathbb{R}$ .

For all  $n \in \mathbb{N}^*$ , let us denote  $A^{(n)} := \sum_{i=1}^n A_i$  an operator on  $\mathcal{D}(A^{(n)}) \subset \mathcal{C}_b^0(E^n, \mathbb{R})$  where  $A_i$  denotes the action of  $A$  on the  $i^{\text{th}}$  variable. Assume that for all  $n \in \mathbb{N}^*$ ,  $A^{(n)}$  generates a strongly continuous contraction semi-group of operator  $\left(T_{\text{FV}}^{E, A^{(n)}}(t)\right)_{t \geq 0}$ . Let us consider the space  $\Omega_E := \mathcal{C}^0([0, +\infty), \mathcal{M}_1(E))$  and the following martingale problem:

**Definition 1.2.11.** *A probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\Omega_E)$  is said to solve the general original FLEMING-VIOT martingale problem for polynomials with resampling rate  $\lambda > 0$  and with initial condition  $\mu \in \mathcal{M}_1(E)$ , if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega_E$  satisfies  $\mathbb{P}_\mu(Y_0 = \mu) = 1$  and for all  $n \in \mathbb{N}^*$  and  $f \in \mathcal{D}(A^{(n)})$*

$$P_{f,n}(Y_t) - P_{f,n}(Y_0) - \int_0^t \mathcal{L}_{\text{FV}}^{E,A} P_{f,n}(Y_s) ds \quad (1.37)$$

with for all  $\mu \in \mathcal{M}_1(E)$ , ((1.36) reduces informally to)

$$\mathcal{L}_{\text{FV}}^{E,A} P_{f,n}(\mu) := \langle A^{(n)} f, \mu^n \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle],$$

is a martingale, where  $\Phi_{i,j} : \mathcal{C}_b^0(E^n, \mathbb{R}) \rightarrow \mathcal{C}_b^0(E^{n-1}, \mathbb{R})$  is defined as in (1.33) and  $P_{f,n}$  in (1.30).

This leads to define the so-called *dual process*  $(\xi_t)_{t \geq 0}$ :

**Definition 1.2.12.** *For all  $M(0) \in \mathbb{N}^*$ , for all  $\xi_0 \in \mathcal{C}_b^0(E^{M(0)}, \mathbb{R})$ , for all  $k \leq M(0) - 1$ , for all  $t \in [\tau_k, \tau_{k+1})$*

$$\xi_t := T_{\text{FV}}^{E,A^{(M(\tau_k))}}(t - \tau_k) \Lambda_k T_{\text{FV}}^{E,A^{(M(\tau_{k-1}))}}(\tau_k - \tau_{k-1}) \Lambda_{k-1} \cdots \Lambda_1 T_{\text{FV}}^{E,A^{(M(0))}}(\tau_1) \xi_0, \quad (1.38)$$

where  $(\tau_n)_{0 \leq n \leq M(0)-1}$  is the sequence of jump times of the birth-death process  $M$  with  $\tau_0 := 0$  and the convention  $\tau_{M(0)} = +\infty$  and where  $(\Lambda_n)_{0 \leq n \leq M(0)-1}$  is a sequence of random operators of type  $\Phi_{i,j}$ .

We are now in position to write the *duality identity*.

**Theorem 1.2.13.** *We consider fixed  $M(0) \in \mathbb{N}^*$ ,  $\xi_0 \in \mathcal{C}_b^0(E^{M(0)}, \mathbb{R})$  and  $(Y_t)_{t \geq 0}$  a stochastic process whose law  $\mathbb{P}_\mu$  is a solution of the martingale problem (1.37) with  $\mu \in \mathcal{M}_1(E)$ . We introduce a dual process  $(\xi_t)_{t \geq 0}$  independent of  $(Y_t)_{t \geq 0}$  built on the same probability space (enlarging it if necessary). We shall denote by  $\mathbb{P}_{(\mu, \xi_0)}$  the law of  $((Y_t, \xi_t))_{t \geq 0}$  on this probability space. The duality identity is given by*

$$\forall t \geq 0, \quad \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_0, Y_t^{M(0)} \rangle \right) = \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_t, Y_0^{M(t)} \rangle \right). \quad (1.39)$$

In similar way as above, we can prove that the duality identity (1.39) determines the 1-dimensional distributions of a solution of the martingale problem (1.37). Uniqueness of the 1-dimensional distributions in turn implies uniqueness of the finite-dimensional distributions and the MARKOV property of the solution (see [94, Theorem 4.4.2]). In the next section, we will see that duality identity gives a simple proof of *weak ergodicity*.

### 1.2.1.4 Ergodicity

The goal of this section is to present the state of the art on ergodicity results for FLEMING-VIOT processes. Let us begin this section by recalling from [91, Section 5] some definitions on ergodicity.

In the first paragraph, we present some ergodicity results. In the second paragraph, we focus on particular cases where the mutation operator is shift invariant.

**Definition 1.2.14.** *Let  $E$  be a Polish space and  $\mathcal{B}(E)$  the BOREL  $\sigma$ -field on  $E$ . An  $E$ -valued MARKOV process  $(Z_t)_{t \geq 0}$  with a unique stationary distribution  $\pi \in \mathcal{M}_1(E)$  is:*

- weakly ergodic if for all  $f \in \mathcal{C}_b^0(E, \mathbb{R})$ , for all initial condition  $x_0 \in E$ ,

$$\lim_{t \rightarrow +\infty} \mathbb{E}_{x_0} (f(Z_t)) = \int_E f(x) \pi(dx) =: \langle f, \pi \rangle \quad (1.40)$$

- strongly ergodic if

$$\lim_{t \rightarrow +\infty} \sup_{B \in \mathcal{B}(E)} |\mathbb{P}_{x_0}(Z_t \in B) - \pi(B)| = 0, \quad x_0 \in E. \quad (1.41)$$

**State of the art.** Recall that the duality identity (1.39) is a powerful tool implying uniqueness of solutions of the martingale problem (1.37). This duality identity also gives a simple proof of the following weak ergodic theorem below.

**Theorem 1.2.15** (ETHIER-KURTZ [91]). *Let  $E$  be compact and suppose that the closure of  $A$  generates a FELLER semi-group  $(T_{\text{FV}}^{E,A}(t))_{t \geq 0}$  on  $\mathcal{C}_b^0(E, \mathbb{R})$ , such that there exists  $\pi \in \mathcal{M}_1(E)$  satisfying  $\lim_{t \rightarrow +\infty} T_{\text{FV}}^{E,A}(t)f(x) = \langle f, \pi \rangle$  for all  $f \in \mathcal{C}_b^0(E, \mathbb{R})$  and  $x \in E$ . Then, the  $\mathcal{M}_1(E)$ -valued FLEMING-VIOT process  $(Y_t)_{t \geq 0}$  with mutation operator  $A$  is weakly ergodic.*

*Sketch of proof.* Recall that the process  $(M(t))_{t \geq 0}$  is a pure death process, so, then after  $M(0) - 1$  jumps,  $M$  absorbs at 1 at an a.s. finite time  $\tau_{M(0)-1}$ . Therefore, from (1.38), (1.39) and the dominated convergence theorem, for all  $\xi_0 \in \mathcal{C}_b^0(E^{M(0)}, \mathbb{R})$  and  $Y_0 \in \mathcal{M}_1(E)$ ,

$$\begin{aligned} \lim_{t \rightarrow +\infty} \mathbb{E}_{(Y_0, \xi_0)} (\langle \xi_0, Y_t^{M(0)} \rangle) &= \lim_{t \rightarrow +\infty} \mathbb{E}_{(Y_0, \xi_0)} (\langle T_{\text{FV}}^{E,A}(t - \tau_{M(0)-1}) \xi_{\tau_{M(0)-1}}, Y_0 \rangle \mathbf{1}_{\tau_{M(0)-1} \leq t}) \\ &= \mathbb{E}_{(Y_0, \xi_0)} (\langle \xi_{\tau_{M(0)-1}}, \pi \rangle). \end{aligned}$$

i.e.

$$\lim_{t \rightarrow +\infty} \mathbb{E}_{(Y_0, \xi_0)} (P_{\xi_0, M(0)}(Y_t)) = \mathbb{E}_{(Y_0, \xi_0)} (P_{\xi_{\tau_{M(0)-1}, 1}(\pi)}). \quad (1.42)$$

Since the right-hand side of (1.42) does not depend on  $Y_0$  (and  $\xi_0$ ), if there exists a stationary distribution  $\Pi \in \mathcal{M}_1(\mathcal{M}_1(E))$ , then the RIESZ representation theorem ensures us that for all  $\xi_0 \in \mathcal{C}_b^0(E^{M(0)}, \mathbb{R})$  and  $Y_0 \in \mathcal{M}_1(E)$ ,

$$\lim_{t \rightarrow +\infty} \mathbb{E}_{Y_0} (P_{\xi_0, M(0)}(Y_t)) = \langle P_{\xi_0, M(0)}, \Pi \rangle.$$

Since the set of test functions

$$\left\{ P_{f,n} \mid f \in \mathcal{C}_b^0(E^n, \mathbb{R}), n \in \mathbb{N}^* \right\}$$

is  $\mathcal{M}_1(\mathcal{M}_1(E))$ –convergence determining [60, Lemma 2.1.2], it is  $\mathcal{M}_1(\mathcal{M}_1(E))$ –separating [94, Chapter 3, Section 4, p.112], so we have for all  $\varphi \in \mathcal{C}_b^0(\mathcal{M}_1(E), \mathbb{R})$ ,  $Y_0 \in \mathcal{M}_1(E)$  that

$$\lim_{t \rightarrow +\infty} \mathbb{E}_{Y_0}(\varphi(Y_t)) = \int_{\mathcal{M}_1(E)} \varphi(\mu) \Pi(d\mu)$$

and the announced result follows from (1.40). The existence of  $\Pi$  is deduced from the assumption that  $A$  generates a FELLER semi-group [91, p.359].  $\square$

Other cases of weak ergodicity are treated with a similar method in [91].

*Coupling arguments* provide one approach to *strong ergodicity*. The basic idea is to construct  $E$ –valued MARKOV processes  $(Z_t^{x_1})_{t \geq 0}$  and  $(Z_t^{x_2})_{t \geq 0}$ , with the same transition function but different initial condition  $x_1$  and  $x_2$  in  $E$ , on the same probability space, in such a way that there is a random time  $\tau$  such that  $Z_t^{x_1} = Z_t^{x_2}$  for all  $t \geq \tau$ . We then have the basic coupling inequality [158]

$$\sup_{B \in \mathcal{B}(E)} |\mathbb{P}(Z_t^{x_1} \in B) - \mathbb{P}(Z_t^{x_2} \in B)| \leq \mathbb{P}(\tau > t), \quad t \geq 0. \quad (1.43)$$

In the special case where the mutation operator of the FLEMING-VIOT process has the form

$$Af(x) = \frac{\theta}{2} \int_E (f(y) - f(x)) P(x, dy), \quad \theta \in (0, +\infty), f \in \mathcal{D}(A), \quad (1.44)$$

it is proved in [91, Theorem 5.3], using a coupling argument, that the FLEMING-VIOT process is strongly ergodic if  $P(x, dy) = \nu(dy)$  for some  $\nu \in \mathcal{M}_1(E)$ . In addition, in this context, it is proved in [91, Theorem 8.1] that the FLEMING-VIOT process has a reversible stationary distribution.

Some of the previous results were extended in [90] to models with recombination and in [92] to models with unbounded selection, with the additional tool of DAWSON’S GIRSANOV-type formula for strong ergodicity.

For the mutation operator (1.44), it is proved in [88, 89] and [60, Theorem 8.2.1] that the FLEMING-VIOT process is purely atomic for every time, in other words the solutions of the martingale problem take values in the set of purely atomic probability measures. In [93], the ergodicity result of [90] was extended to the weak atomic topology.

**When the mutation operator is the (fractional) Laplacian.** If we consider the case where the mutation operator is the (fractional) Laplacian on  $\mathbb{R}$ , there exists no stationary distribution [91, 157], [94, Problem 11 p.450]. Instead the process exhibits a wandering phenomenon (see [61] in the Laplacian case and [21] in the fractional Laplacian case). In [61], the wandering coherence of FLEMING-VIOT process is obtained exploiting a centering effect by the “centre of mass process”  $z(t) := \int_{\mathbb{R}} x Y_t(dx)$ . However, in [21] the authors consider for  $z(t)$  the lowest-level-particle in the modified look-down construction.

Nevertheless, [205, 91] considered the FLEMING-VIOT process shifted by minus its empirical mean – that we call *centered* FLEMING-VIOT process – and established existence of a unique invariant measure and *weak* ergodicity for this process using moment and duality arguments. More precisely in [91, Theorem 9.1.2], thanks to some estimates of the original FLEMING-VIOT dual process and the finiteness of all moments of the FLEMING-VIOT process shifted by minus its empirical mean for any time  $t$ , the authors obtain an expression for these in the asymptotic  $t \rightarrow +\infty$ . Then, by tightness arguments and characterisation of the limit, the result follows. Note

that [91, Theorem 9.1.2] extends [61, Theorem 7.2]. In [205], an analogous approach is used for the MORAN model. In both references, only *weak* ergodic results are obtained. Note that none of these references consider the questions of existence and uniqueness for the martingale problem for the centered FLEMING-VIOT process. Up to our knowledge, our study of these questions (see Section 1.3.1) is new.

The study of ergodic properties of the centered version of the FLEMING-VIOT was motivated, in our case, by biological questions in adaptive dynamics. Recall that to solve some controversies in this field (see Section 1.1.4.6), it was proposed to apply a double asymptotic of small mutations and large population, but *frequent mutations*. After conveniently scaling the population state, this leads to a slow-fast dynamics where the fast dynamics appears to be given by a discrete version of the centered FLEMING-VIOT process. One of the ingredients of this asymptotic analysis (see Section 1.2.2) is the ergodicity of the limit fast component, which explains why we are interested in the ergodicity properties of such processes. In addition, using a coupling argument based on the DONNELLY-KURTZ modified look-down (see Section 1.2.1.2) it is possible to obtain strong ergodicity bounds (see (1.41)) and an explicit construction of the invariant measure.

## 1.2.2 Slow-fast asymptotic analysis

To prove convergence in the framework of slow-fast dynamics (also called *stochastic singular perturbations*), different techniques can be used. All these approaches are based on *compactness-uniqueness* arguments and consist in identifying and characterising the limit of the fast component, which requires ergodic properties, in order to establish the convergence of the slow component and identify its limit. In other words, the fast component is only a tool to study the behaviour, of the slow component which is the main study. In Section 1.2.2.1 we present the *perturbed test function method* and some derived extensions. In Section 1.2.2.2, the *KURTZ stochastic averaging method* is presented. This approach, which we take care to detail, will be exploited in this thesis.

### 1.2.2.1 The perturbed test function and derived extensions

Firstly, the method of the *perturbed test function*, initially proposed by PAPANICOLAOU, STROOCK and VARADHAN in [192], identifies the generator of the limit process with a martingale approach whose idea is the following. If we consider a family of stochastic processes  $((X^\varepsilon, Y^\varepsilon))_{\varepsilon>0}$  of generator  $L^\varepsilon$  where  $X^\varepsilon$  is the *slow component* and  $Y^\varepsilon$  is the *fast component* in the form  $Y^\varepsilon(t) = Y(t/\varepsilon)$  where  $Y$  is a MARKOV process, the slow-fast problem consists in identifying the limit process of  $X^\varepsilon$  using ergodicity properties of the fast dynamics. Assuming that the family  $(X^\varepsilon)_{\varepsilon>0}$  is tight, we consider  $X$  a limiting value. We can expect to characterise  $X$  with a martingale problem derived from the martingale problem of  $(X^\varepsilon, Y^\varepsilon)$  provided that the solution to this problem is unique. We would then obtain the convergence of  $X^\varepsilon$  to  $X$  in law. However, for multiscale singular problems, the convergence of  $L^\varepsilon \varphi((X_t^\varepsilon, Y_t^\varepsilon))$  when  $\varepsilon \rightarrow 0$ , for a test function  $\varphi$  depending only on the  $X$  component, does not (in general) take place because it contains diverging terms in  $\varepsilon$ . To overcome this difficulty, firstly the idea is to decompose the generator  $L^\varepsilon$  in the following form  $L^\varepsilon = \frac{1}{\varepsilon}L_1 + L_0$  where  $L_1$  is the infinitesimal generator of the MARKOV process  $Y$  in the variable  $y$  and  $L_0$  the operator of the slow component depending on slow and fast variables. Secondly, the idea is to perturb the initial test function  $\varphi(x)$  into a test function  $\varphi^\varepsilon(x, y) := \varphi(x) + \varepsilon\varphi_1(x, y)$  such that

$$L^\varepsilon \varphi^\varepsilon = \left[ \frac{1}{\varepsilon}L_1 + L_0 \right] (\varphi + \varepsilon\varphi_1) = (L_1\varphi_1 + L_0\varphi - \bar{L}_0\varphi) + \bar{L}_0\varphi + \varepsilon L_0\varphi_1 \quad (1.45)$$

because of  $L_1\varphi = 0$  and where  $\bar{L}_0$  is the operator of the limit slow component *averaged* by the unique invariant probability measure of the limit fast component  $Y$ . Provided that  $\varphi_1$  solves the POISSON equation (with respect to  $L_1$  and the variable  $y$ )

$$L_1\varphi_1 + L_0\varphi - \bar{L}_0\varphi = 0, \quad (1.46)$$

we deduce that  $L^\varepsilon\varphi^\varepsilon = \bar{L}_0\varphi + O(\varepsilon)$ . Then,  $\varphi^\varepsilon(x, y) \rightarrow \varphi(x)$  and  $L^\varepsilon\varphi^\varepsilon(x, y) \rightarrow \bar{L}_0\varphi(x)$  when  $\varepsilon \rightarrow 0$  as expected.

The perturbed test function method has been extended in [22, 145, 146] and in the books [100, Section 6.3], [147]. Many other references apply the perturbed test function strategy in various settings: in finance [101] in transport problems as in [192] or in [56] where the tightness of the fast component is established using its occupation measure. Similar methods are used in homogenisation [192, 17, 193] (see also [16] which exploits spectral and semi-group properties in addition), and in stochastic stability and control [147].

The perturbed test function method has also been extended to analyse the behaviour of random perturbed systems of the form

$$\begin{cases} dX_t^\varepsilon &= A(X_t^\varepsilon, Y_t^\varepsilon) dt + \sigma(X_t^\varepsilon, Y_t^\varepsilon) dW_t, & X_0^\varepsilon = x_0, \\ dY_t^\varepsilon &= \frac{1}{\varepsilon} B(X_t^\varepsilon, Y_t^\varepsilon) dt + \frac{1}{\sqrt{\varepsilon}} \varphi(X_t^\varepsilon, Y_t^\varepsilon) dW_t, & Y_0^\varepsilon = y_0, \end{cases}$$

for some smooth functions  $A, B, \sigma, \varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(W_t)_{t \geq 0}$  a 1-dimensional Brownian motion and  $x_0, y_0 \in \mathbb{R}$ . Assuming that the fast equation with frozen slow component  $x$

$$dY_t^{x,y} = B(x, Y_t^{x,y}) dt + \varphi(x, Y_t^{x,y}) dW_t, \quad Y_0^{x,y} = y,$$

admits a unique solution  $Y^{x,y}$  for all  $x, y \in \mathbb{R}$ , and that there exists two smooth mappings  $\bar{A}, \bar{\sigma} : \mathbb{R} \rightarrow \mathbb{R}$  such that for all  $T > 0$ , for all  $x, y \in \mathbb{R}$ ,

$$\begin{aligned} \left| \frac{1}{T} \int_0^T \mathbb{E}(A(x, Y_t^{x,y})) dt - \bar{A}(x) \right| &\leq \alpha(T) (1 + |x|^2) \\ \left| \frac{1}{T} \int_0^T \mathbb{E}(\sigma^2(x, Y_t^{x,y})) dt - \bar{\sigma}^2(x) \right| &\leq \alpha(T) (1 + |x|^2) \end{aligned} \quad (1.47)$$

where  $\lim_{T \rightarrow +\infty} \alpha(T) = 0$ , KHASHINSKII proves in [136] that  $(X^\varepsilon)_{\varepsilon \in (0,1]}$  converges in *probability*, as  $\varepsilon \rightarrow 0$ , to the solution  $\bar{X}$  of the *averaged equation*

$$d\bar{X}_t = \bar{A}(\bar{X}_t) dt + \bar{\sigma}(\bar{X}_t) dW_t, \quad \bar{X}_0 = x_0.$$

The weak convergence proof from  $(X^\varepsilon)_{\varepsilon \in (0,1]}$  to  $\bar{X}$  consists in proving the tightness of the family  $(X^\varepsilon)_{\varepsilon \in (0,1]}$  and then identifying the limit using the martingale problem associated with the averaged system. The convergence in probability relies on the ergodic conditions (1.47). Others extensions of KHASHINSKII's approach are proposed in [192, 210, 137, 193, 35, 34, 109, 106, 110, 105, 53, 107].

### 1.2.2.2 The Kurtz stochastic averaging method

In the first two paragraphs of this section, the arguments and multiple applications of the KURTZ method are presented. Finally, in a last paragraph, this original method is implemented in relation to the case studied in this thesis.

**Arguments of Kurtz’s method.** An important method of the stochastic averaging, in a generic framework, was proposed by KURTZ in [144]. The main idea consists in exploiting the *occupation measure*  $\Gamma^\varepsilon$  of the fast component  $Y^\varepsilon$  which is formally defined for all  $t \geq 0$ , for any Borelian  $B$  by

$$\Gamma^\varepsilon([0, t] \times B) = \int_0^t \mathbf{1}_B(Y_s^\varepsilon) ds,$$

and to establish the convergence of the couple  $(X^\varepsilon, \Gamma^\varepsilon)$  when  $\varepsilon \rightarrow 0$  using *compactness-uniqueness* arguments. Proceeding in this way, it allows us to escape the difficulties created by the fluctuations of the fast component and avoids to obtain tightness result of the “slow-fast” couple  $(X^\varepsilon, Y^\varepsilon)$ . The proof can be divided into four steps:

- The first one consists in establishing uniform tightness of the sequence of laws of the couple  $(X^\varepsilon, \Gamma^\varepsilon)$ .
- The second one consists in establishing a martingale problem for any limiting value of the family of laws of the couple.
- The third one is based on establishing the uniqueness of the limiting value of  $\Gamma^\varepsilon$  expressed in terms of the limit  $X$  of  $X^\varepsilon$  which is given. The characterisation of the limiting value  $\Gamma$  of  $\Gamma^\varepsilon$  is usually based on ergodicity arguments.
- Finally, the last step is to establish the uniqueness of the martingale problem for the slow component limit  $X$  when  $\Gamma$  is given as above.

As will appear in Section 1.3.2, the KURTZ approach seems to be better adapted to our situation than the perturbed test function method.

**Applications of Kurtz’s method.** Many approaches are based on it: for trait and age-structured populations in [167], in a medical setting with hematopoiesis in [28] or the development of ovarian follicles in [9], for Piecewise Deterministic MARKOV Processes (PDMPs) in the predator-prey setting in [54], for some simple constrained MARKOV processes in [108] or in adaptive dynamics context with [119, 19]. Note that in [119], GUPTA ET AL, give a new proof of the convergence of the IBM to the TSS (see Theorem 1.1.6) that skips many technicalities linked with fine approximations of the IBM. In [19], they consider a finite population where each individual is characterised by a trait under selection and completely linked neutral (genetic) marker. They consider a large population limit with rare mutations under the assumption that the neutral marker mutates faster than the trait under selection. They prove the convergence of the IBM to a new measure diffusive process with jumps, called *Substitution FLEMING-VIOT Process* (SFVP) generalising the TSS. KURTZ’s averaging method is also used for chemical reaction network in [8, 199, 133, 197, 198]. In [8], an application of a cell’s viral infection is proposed. Using a compartmental model, [199] situates in a neuroscience setting and describes fast axonal transport, whereas [197, 198] study the effects of fast spatial movements of molecules on the dynamics of chemical species. In the setting of chemical reaction network modelised by MARKOV chain models, they authors proved in [133] a central limit theorem to capture the fluctuations of the original model around the averaged model.

**Implementation of Kurtz’s method.** In order to be self-contained, we present three technical lemmas as well as the main theorem of [144] in the particular case of real-valued slow component and real probability measure-valued fast component, which is the case studied in this thesis. We propose to give a proof of this theorem, assuming the technical lemmas are granted, whose approach will be the one followed in this thesis. Since these results are based on



compactness-uniqueness arguments, we recall some definitions.

**Definition 1.2.16.** Let  $E$  be a Polish space and  $\mathcal{B}(E)$  the BOREL  $\sigma$ -field on  $E$ . A family  $\Pi$  of probability measures on  $(E, \mathcal{B}(E))$  is relatively compact for the topology of the weak convergence if

$$\forall (P_n)_{n \in \mathbb{N}^*} \in \Pi^{\mathbb{N}^*}, \exists P \in \Pi, \exists \chi : \mathbb{N}^* \nearrow \mathbb{N}^*, \quad P_{\chi(n)} \xrightarrow[n \rightarrow +\infty]{\text{law}} P.$$

A family  $\Pi$  of probability measures on  $(E, \mathcal{B}(E))$  is tight if

$$\forall \eta > 0, \exists D_\eta \subset E \text{ compact}, \forall P \in \Pi, \quad P(D_\eta) > 1 - \eta.$$

PROHOROV's theorem [20, Theorems 5.1 and 5.2] ensures that on a Polish space, the notions of relative compactness and tightness are equivalent, i.e. the family of probability measures  $\Pi$  on  $(E, \mathcal{B}(E))$  is relatively compact if and only if  $\Pi$  is tight. Recall that  $\mathcal{M}_1(\mathbb{R})$ , the set of probability measures on  $\mathbb{R}$  with respect to the BOREL  $\sigma$ -field  $\mathcal{B}(\mathbb{R})$  which is endowed with the topology of weak convergence making it a Polish space [20].

We denote by  $\mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$  the set of measures  $\Gamma$  on  $\mathbb{R}_+ \times \mathcal{M}_1(\mathbb{R})$  such that for all  $t \geq 0$ ,  $\Gamma([0, t] \times \mathcal{M}_1(\mathbb{R})) = t$ . For any  $t \geq 0$ , we denote by  $\mathcal{M}_m^t(\mathcal{M}_1(\mathbb{R}))$  the set of measures  $\Gamma \in \mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$  restricted to  $[0, t] \times \mathcal{M}_1(\mathbb{R})$ . For all  $T > 0$ , we denote by  $\mathbb{D}([0, T], E)$  the space of  $E$ -valued functions that are right-continuous on  $[0, T]$  and have left-hand limits.

**Lemma 1.2.17.** Let  $T > 0$ . A sequence of  $\mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ -valued random variable  $(\Gamma_n)_{n \in \mathbb{N}^*}$  is relatively compact if and only if for each  $\varepsilon > 0$  and each  $t \in [0, T]$  there exists a compact  $D \subset \mathcal{M}_1(\mathbb{R})$  such that  $\inf_{n \in \mathbb{N}^*} \mathbb{E}(\Gamma_n([0, t] \times D)) \geq (1 - \varepsilon)t$ .

**Lemma 1.2.18.** Let  $T > 0$  and  $((x_n, \Gamma_n))_{n \in \mathbb{N}^*} \subset \mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  such that  $(x_n, \Gamma_n) \xrightarrow[n \rightarrow +\infty]{} (x, \Gamma)$  in the SKOROHOD topology. Let  $\psi \in \mathcal{C}_b^0([0, T] \times \mathcal{M}_1(\mathbb{R}), \mathbb{R})$ . Define

$$u_n(t) := \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(x_n(s), \mu) \Gamma_n(ds, d\mu), \quad u(t) := \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(x(s), \mu) \Gamma(ds, d\mu)$$

$z_n(t) := \Gamma_n([0, t] \times \mathcal{M}_1(\mathbb{R}))$  and  $z(t) := \Gamma([0, t] \times \mathcal{M}_1(\mathbb{R}))$ .

- (a) If  $x$  is continuous on  $[0, t]$  and  $\lim_{n \rightarrow +\infty} z_n(t) = z(t)$ , then  $\lim_{n \rightarrow +\infty} u_n(t) = u(t)$ .
- (b) If  $(x_n, z_n, \Gamma_n) \xrightarrow[n \rightarrow +\infty]{} (x, z, \Gamma)$  in  $\mathbb{D}([0, T], \mathbb{R} \times \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ , then

$$(x_n, z_n, u_n, \Gamma_n) \xrightarrow[n \rightarrow +\infty]{} (x, z, u, \Gamma)$$

in  $\mathbb{D}([0, T], \mathbb{R} \times \mathbb{R} \times \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ . In particular,  $\lim_{n \rightarrow +\infty} u_n(t) = u(t)$  at all points of continuity of  $z$ .

- (c) The continuity assumption on  $\psi$  can be replaced by the assumption that  $\psi$  is continuous a.e.  $\nu_t$  for each  $t$  where  $\nu_t \in \mathcal{M}_F([0, T] \times \mathcal{M}_1(\mathbb{R}))$  (the set of finite measures on  $[0, T] \times \mathcal{M}_1(\mathbb{R})$ ) is the measure determined by

$$\nu_t(A \times B) = \Gamma(\{(s, y) | x(s) \in A, s \leq t, y \in B\}).$$

- (d) In both (a) and (b), the boundedness assumption on  $\psi$  can be replaced by the assumption that there exists a non-negative convex function  $\Psi$  on  $\mathbb{R}_+$  satisfying  $\lim_{r \rightarrow +\infty} \frac{\Psi(r)}{r} = +\infty$  such that

$$\sup_{n \in \mathbb{N}^*} \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \Psi(|\psi(x_n(s), \mu)|) \Gamma_n(ds, d\mu) < \infty$$

for each  $t > 0$ .

**Theorem 1.2.19** (KURTZ [144]). *Let  $T > 0$  and  $((X_n, Y_n))_{n \in \mathbb{N}^*}$  be a stochastic process with sample paths in  $\mathbb{D}([0, T], \mathbb{R} \times \mathcal{M}_1(\mathbb{R}))$  adapted to a filtration  $(\mathcal{F}_n(t))_{t \in [0, T]}$ . Assume that:*

(1)  $(X_n)_{n \in \mathbb{N}^*}$  satisfies the compact containment condition:

$$\forall \eta > 0, \exists D \subset \mathbb{R} \text{ compact, } \forall t \in [0, T], \quad \inf_{n \in \mathbb{N}^*} \mathbb{P}(X_n(t) \in D) \geq 1 - \eta,$$

(2) the family of laws of the marginal random variables  $(Y_n(t))_{t \in [0, T], n \in \mathbb{N}^*}$  is tight on  $\mathcal{M}_1(\mathbb{R})$ ,

(3) there is an operator  $A : \mathcal{D}(A) \subset \mathcal{C}_b^0(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{C}^0(\mathbb{R} \times \mathcal{M}_1(\mathbb{R}), \mathbb{R})$  such that for all  $f \in \mathcal{D}(A)$  there is a process  $\mathcal{E}_n^f$  for which

$$M_n^f(t) := f(X_n(t)) - f(X_n(0)) - \int_0^t Af(X_n(s), Y_n(s)) ds + \mathcal{E}_n^f(t) \quad (1.48)$$

is an  $(\mathcal{F}_n(t))_{t \in [0, T]}$ -martingale,

(4)  $\mathcal{D}(A)$  is dense in  $\mathcal{C}_b^0(\mathbb{R}, \mathbb{R})$  in the topology of uniform convergence on compact sets,

(5) for all  $f \in \mathcal{D}(A)$ , there exists  $p_0 > 1$  such that

$$\sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \int_0^T |Af(X_n(t), Y_n(t))|^{p_0} dt \right) < \infty \quad (1.49)$$

and

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left( \sup_{t \in [0, T]} |\mathcal{E}_n^f(t)| \right) = 0. \quad (1.50)$$

Let  $\Gamma_n$  be the  $\mathcal{M}_m^t(\mathcal{M}_1(\mathbb{R}))$ -valued random variable defined for all  $t \in [0, T]$ , for any borelian  $B$  by

$$\Gamma_n([0, t] \times B) = \int_0^t \mathbf{1}_B(Y_n(s)) ds.$$

Then, the sequence of laws of  $((X_n, \Gamma_n))_{n \in \mathbb{N}^*}$  is tight in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$  and for any limiting value  $\mathbb{Q}$  of this sequence, the canonical process  $(X, \Gamma)$  on  $\mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  satisfies that for all  $f \in \mathcal{D}(A)$ ,

$$N^f(t) := f(X(t)) - f(X(0)) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(X(s), \mu) \Gamma(ds, d\mu) \quad (1.51)$$

is a  $\mathbb{Q}$ -martingale.

In order to characterise the occupation measure limit  $\Gamma$ , KURTZ gives, in the following lemma, a desintegration result for  $\Gamma$ .

**Lemma 1.2.20.** *Let  $T > 0$  and  $\Gamma \in \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ . There exists a random probability measure-valued process  $(\gamma_t)_{t \in [0, T]}$  that is predictable in  $(\omega, t)$  and such that for all bounded measurable function  $\psi : [0, T] \times \mathcal{M}_1(\mathbb{R}) \rightarrow \mathbb{R}$ ,*

$$\int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s, u) \Gamma(ds, d\mu) = \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s, \mu) \gamma_s(d\mu) ds.$$

*Proof of Theorem 1.2.19.* The proof is divided in six steps. In Steps 1 and 2, we establish respectively the tightness of the family of laws of  $(X_n)_{n \in \mathbb{N}^*}$  and  $(\Gamma_n)_{n \in \mathbb{N}^*}$ . In Step 3, we deduce the tightness of the family of laws of  $((X_n, \Gamma_n))_{n \in \mathbb{N}^*}$ . In Step 4, we prove uniform integrability

results for a family  $(N_n^f(t))_{t \in [0, T], n \in \mathbb{N}^*}$  constructed from  $X_n$  and  $\Gamma_n$  similarly as  $N^f(t)$  in (1.51). Step 5 is devoted to establish the convergence in distribution from  $(N_n^f)_{n \in \mathbb{N}^*}$  to  $N^f$ . In Step 6, we prove that the limit  $N^f$  is a martingale.

**Step 1. Tightness of the family of laws of  $(X_n)_{n \in \mathbb{N}^*}$  on  $\mathbb{D}([0, T], \mathbb{R})$ .** Let  $f \in \mathcal{D}(A)$ . For all  $n \in \mathbb{N}^*$  and for all  $t \in [0, T]$ , let us consider the two processes  $\mathcal{Y}_n(t) := f(X_n(t))$  and  $\mathcal{Z}_n(t)$  defined by the relation

$$\mathcal{Y}_n(t) = \mathcal{Y}_n(0) + \int_0^t \mathcal{Z}_n(s) ds + M_n^f(t)$$

where  $M_n^f$  is the martingale given by (1.48). Note that from (1.49) and (1.50), there exists a constant  $C > 0$  such that

$$\begin{aligned} \sup_{n \in \mathbb{N}^*} \mathbb{E} \left[ \left( \int_0^T |\mathcal{Z}_n(s)|^{p_0} ds \right)^{1/p_0} \right] &\leq C \left[ 1 + \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \int_0^T |Af(X_n(s), Y_n(s))|^{p_0} ds \right) \right. \\ &\quad \left. + \sup_{n \in \mathbb{N}^*} \mathbb{E} \left( \sup_{t \in [0, T]} |\mathcal{E}_n^f(t)| \right) \right] \\ &< \infty. \end{aligned}$$

Hence, from [94, Theorem 3.9.4], the family of  $(\mathcal{Y}_n)_{n \in \mathbb{N}^*}$ , on  $\mathbb{D}([0, T], \mathbb{R})$  is tight. From Assumptions (1) and (4), we deduce from [94, Theorem 3.9.1] that the family of laws of  $(X_n)_{n \in \mathbb{N}^*}$  is tight on  $\mathbb{D}([0, T], \mathbb{R})$ .

**Step 2. Tightness of the family of laws of  $(\Gamma_n)_{n \in \mathbb{N}^*}$  on  $\mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ .** Let  $\eta > 0$  be fixed. From Assumption (2), there exists  $D_\eta^T \subset \mathcal{M}_1(\mathbb{R})$  compact such that for all  $t \in [0, T], n \in \mathbb{N}^*$ ,  $\mathbb{P}(Y_n(t) \in D_\eta^T) \geq 1 - \eta$ . It follows that for all  $t \in [0, T], n \in \mathbb{N}^*$ ,

$$\mathbb{E}(\Gamma_n([0, t] \times D_\eta^T)) = \int_0^t \mathbb{P}(Y_n(s) \in D_\eta^T) ds \geq t(1 - \eta).$$

Therefore the tightness of the family of laws of  $(\Gamma_n)_{n \in \mathbb{N}^*}$  follows from Lemma 1.2.17.

**Step 3. Tightness of the family of laws of  $((X_n, \Gamma_n))_{n \in \mathbb{N}^*}$  on  $\mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ .** From Steps 1 and 2 and PROHOROV's theorem, we deduce that the family of laws of the couple  $((X_n, \Gamma_n))_{n \in \mathbb{N}^*}$  is relatively compact in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$ .

**Step 4. Uniform integrability.** Let us consider  $(N_n^f(t))_{t \in [0, T]}$ , the stochastic process defined for all  $f \in \mathcal{D}(A)$  by

$$\begin{aligned} N_n^f(t) &:= f(X_n(t)) - f(X_n(0)) - \int_0^t Af(X_n(s), Y_n(s)) ds \\ &= f(X_n(t)) - f(X_n(0)) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(X_n(s), \mu) \Gamma_n(ds, d\mu). \end{aligned}$$

From (1.49) and [213, Chapter 13],  $(N_n^f)_{n \in \mathbb{N}^*}$  is uniformly integrable.

**Step 5. Almost sure convergence of  $(\tilde{N}_n^f)_{n \in \mathbb{N}^*}$  to  $\tilde{N}^f$ .** From Step 3, there exists a probability measure  $\mathbb{Q}$  on the canonical space  $\mathbb{D}([0, T], \mathbb{R}) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  and an increasing

function  $\chi : \mathbb{N}^* \rightarrow \mathbb{N}^*$  such that the subsequence of laws of  $\left( (X_{\chi(n)}, \Gamma_{\chi(n)}) \right)_{n \in \mathbb{N}^*}$  converges weakly to the limiting value  $\mathbb{Q}$  when  $n \rightarrow +\infty$ . From SKOROHOD's representation theorem, there exists a probability space on which we define the families  $\left( \tilde{X}_{\chi(n)} \right)_{n \in \mathbb{N}^*}$  and  $\left( \tilde{\Gamma}_{\chi(n)} \right)_{n \in \mathbb{N}^*}$  and  $\tilde{X}, \tilde{\Gamma}$  copies of  $(X_{\chi(n)})_{n \in \mathbb{N}^*}, (\Gamma_{\chi(n)})_{n \in \mathbb{N}^*}, X$  and  $\Gamma$  under  $\mathbb{Q}$  such that the sequence  $\left( (\tilde{X}_{\chi(n)}, \tilde{\Gamma}_{\chi(n)}) \right)_{n \in \mathbb{N}^*}$  converges a.s. to  $(\tilde{X}, \tilde{\Gamma})$  when  $n \rightarrow +\infty$ . Note that for all  $t \in [0, T]$ ,

$$\tilde{\Gamma}_{\chi(n)}([0, t] \times \mathcal{M}_1(\mathbb{R})) = t \xrightarrow[n \rightarrow +\infty]{\text{a.s.}} t = \tilde{\Gamma}([0, t] \times \mathcal{M}_1(\mathbb{R})).$$

Applying Lemma 1.2.18 **(b)(c)(d)** with the function  $\Psi(x) := x^{p_0}$ , we deduce from (1.49) that for all  $f \in \mathcal{D}(A)$ , for all  $t \in [0, T]$ , a.s.

$$\lim_{n \rightarrow +\infty} \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(\tilde{X}_{\chi(n)}, \mu) \tilde{\Gamma}_{\chi(n)}(ds, d\mu) = \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(\tilde{X}(s), \mu) \tilde{\Gamma}(ds, d\mu)$$

and thus, the sequence  $\left( \tilde{N}_{\chi(n)}^f \right)_{n \in \mathbb{N}^*}$  converges a.s. to the process  $\tilde{N}^f$  where  $\tilde{N}_{\chi(n)}^f(t)$  and  $\tilde{N}^f(t)$  are respectively defined for all  $n \in \mathbb{N}^*, t \in [0, T]$  and  $f \in \mathcal{D}(A)$  by

$$\begin{aligned} \tilde{N}_n^f(t) &:= f(\tilde{X}_n(t)) - f(\tilde{X}_n(0)) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(\tilde{X}_n(s), \mu) \tilde{\Gamma}_n(ds, d\mu), \\ \tilde{N}^f(t) &:= f(\tilde{X}(t)) - f(\tilde{X}(0)) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} Af(\tilde{X}(s), \mu) \tilde{\Gamma}(ds, d\mu). \end{aligned}$$

In particular,  $\tilde{N}^f$  and  $N^f$  have the same distribution under  $\mathbb{Q}$  and  $\left( N_{\chi(n)}^f \right)_{n \in \mathbb{N}^*}$  converges in distribution to  $N^f$ .

**Step 6.  $N^f$  is a  $\mathbb{Q}$ -martingale.** Let us consider the filtration  $\left( \tilde{\mathcal{F}}_n(t) \right)_{t \in [0, T]}$  defined by  $\tilde{\mathcal{F}}_n(t) := \sigma \left( \tilde{X}_n(s), \tilde{\Gamma}_n([0, s] \times H) \mid s \leq t, H \in \mathcal{B}(\mathcal{M}_1(\mathbb{R})) \right)$  and  $(f_i)_{1 \leq i \leq q}, q \in \mathbb{N}^*$  bounded LIPSCHITZ functions from  $\mathbb{R}$  to  $\mathbb{R}$  and  $0 < t_1 \leq \dots \leq t_q \leq s < t$ . For all  $n \in \mathbb{N}^*, t \in [0, T]$ , note that  $\tilde{M}_n^f(t) := \tilde{N}_n^f(t) + \tilde{\mathcal{E}}_n^f(t)$  is a  $\left( \tilde{\mathcal{F}}_n(t) \right)_{t \in [0, T]}$ -martingale from (1.48). Hence,

$$\mathbb{E} \left( \tilde{M}_{\chi(n)}^f(t) \prod_{i=1}^q f_i(\tilde{M}_{\chi(n)}^f(t_i)) \right) = \mathbb{E} \left( \tilde{M}_{\chi(n)}^f(s) \prod_{i=1}^q f_i(\tilde{M}_{\chi(n)}^f(t_i)) \right).$$

As for all  $i \in \{1, \dots, q\}$ ,  $f_i$  is LIPSCHITZ and since the error term  $\tilde{\mathcal{E}}_n^f(t)$  satisfies (1.50), there exists a constant  $C > 0$  such that

$$\begin{aligned} \mathbb{E} \left( \left| f_i(\tilde{N}_{\chi(n)}^f(t_i)) - f_i(\tilde{M}_{\chi(n)}^f(t_i)) \right| \right) &= \mathbb{E} \left( \left| f_i(\tilde{M}_{\chi(n)}^f(t_i) - \tilde{\mathcal{E}}_{\chi(n)}^f(t_i)) - f_i(\tilde{M}_{\chi(n)}^f(t_i)) \right| \right) \\ &\leq C \mathbb{E} \left( \left| \tilde{\mathcal{E}}_{\chi(n)}^f(t_i) \right| \right) \end{aligned}$$

which goes to 0 when  $n \rightarrow +\infty$ . We deduce that for all  $u \in \{s, t\}$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left( \tilde{M}_{\chi(n)}^f(u) \prod_{i=1}^q f_i(\tilde{M}_{\chi(n)}^f(t_i)) \right) = \lim_{n \rightarrow +\infty} \mathbb{E} \left( \tilde{N}_{\chi(n)}^f(u) \prod_{i=1}^q f_i(\tilde{N}_{\chi(n)}^f(t_i)) \right).$$

From Steps 4 and 5, we deduce that

$$\mathbb{E} \left( \tilde{N}^f(t) \prod_{i=1}^q f_i(\tilde{N}^f(t_i)) \right) = \mathbb{E} \left( \tilde{N}^f(s) \prod_{i=1}^q f_i(\tilde{N}^f(t_i)) \right).$$

Hence,

$$\mathbb{E} \left( N^f(t) \prod_{i=1}^q f_i(N^f(t_i)) \right) = \mathbb{E} \left( N^f(s) \prod_{i=1}^q f_i(N^f(t_i)) \right).$$

Since the last property is true for all  $q \in \mathbb{N}^*$ ,  $0 < t_1 \leq \dots \leq t_q \leq s < t$  and for all bounded LIPSCHITZ functions  $(f_i)_{1 \leq i \leq q}$  from  $\mathbb{R}$  to  $\mathbb{R}$ , the monotone class theorem ensures us that

$$\mathbb{E} \left( N^f(t) \middle| \sigma \left( N^f(u) \middle| u \leq s \right) \right) = N^f(s).$$

Hence,  $N^f$  is a  $\mathbb{Q}$ -martingale which ends the proof.  $\square$

### 1.3 My contributions

This thesis is divided in two parts, which are the content of Chapters 2 and 3. The topics are related, however, the chapters can be read independently from each other. In the first part, we study the *centered* FLEMING-VIOT process defined as the original FLEMING-VIOT process shifted by minus its empirical mean. Our goals are to characterise it with a martingale problem, to establish (strong) ergodicity properties and to characterise its invariant probability measure. This part corresponds to the publication [41].

N. CHAMPAGNAT and V. HASS. *Existence, uniqueness and ergodicity for the centered Fleming-Viot process*. To appear in *Stochastic Processes and Their Applications* (<https://doi.org/10.1016/j.spa.2023.09.006>).

In the second part of this thesis, we study the long term behaviour of the stochastic IBM, defined in Section 1.1.3.2, where mutations are *frequent*, in the double limit of *large population* with *small mutational effects* proving convergence to the CEAD. This result is based on a slow-fast asymptotic analysis where we use an averaging method inspired by KURTZ (see Section 1.2.2.2). The contribution of the fast component, which converges to the centered FLEMING-VIOT process, is obtained by averaging according to its invariant measure, characterised in Chapter 2. This part corresponds to the preprint [40]

N. CHAMPAGNAT and V. HASS. *Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics*. Preprint (<https://arxiv.org/abs/2303.09936>). 2023. Submitted to *Annals of Applied Probability*.

This section is organised as follows. In Sections 1.3.1 and 1.3.2, we present respectively the main results of Chapter 2 and 3 as well as our approach. Finally in Section 1.3.3, we propose several natural prospects to this thesis.

#### 1.3.1 From Chapter 2

In Chapter 2, we are interested in the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$  defined by  $Z_t := \tau_{-(\text{id}, Y_t)} \# Y_t$ , where  $(Y_t)_{t \geq 0}$  is the original FLEMING-VIOT proces. As the previous works on the original FLEMING-VIOT process (see Section 1.2.1), it is natural to ask question of existence, uniqueness and ergodicity when the mutation operator is the Laplacian on  $\mathbb{R}$ . We establish existence of the *centered FLEMING-VIOT martingale problem* exploiting the martingale problem

of  $(Y_t)_{t \geq 0}$  (see Section 1.2.1.1). We prove uniqueness for initial conditions admitting finite moments using a weakened version of the duality method (see Section 1.2.1.3), allowing us to prove uniqueness for  $\cdot$ . Finally, we establish strong ergodicity properties of  $(Z_t)_{t \geq 0}$  that extend the weak ergodicity results obtained in Section 1.2.1.4. Based on a variant of the DONNELLY-KURTZ modified look-down (see Section 1.2.1.2), a characterisation of the unique invariant probability measure  $\pi$  of  $(Z_t)_{t \geq 0}$  is also given.

This section is organised as follows. In Section 1.3.1.1 we define the martingale problem for the centered FLEMING-VIOT process and establish our existence result. In Section 1.3.1.2 we give the necessary ingredients to prove uniqueness to the centered FLEMING-VIOT martingale problem for initial conditions admitting finite moments and we discuss this assumption. Finally, ergodicity properties with exponential convergence in total variation for the centered FLEMING-VIOT process and a characterisation of  $\pi$  are established in Section 1.3.1.3.

### 1.3.1.1 Existence for the centered Fleming-Viot process

**Centered Fleming-Viot martingale problem and main result.** Recall several notations that we use repeatedly in the sequel. For a measurable real bounded function  $f$  and a measure  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote  $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x)\nu(dx)$ . We denote by  $\text{id}$  the identity function. We denote  $\mathbb{N} := \{0, 1, 2, \dots\}$  and  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ . If  $E_1$  and  $E_2$  are two metric spaces, then for all  $\ell \in \mathbb{N}$ , we denote by  $\mathcal{C}^\ell(E_1, E_2)$  the space of functions of class  $\mathcal{C}^\ell$  from  $E_1$  to  $E_2$ . We denote  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$  the filtered probability space where

$$\Omega := \mathcal{C}^0([0, +\infty), \mathcal{M}_1(\mathbb{R}))$$

is endowed with the SKOROHOD topology,  $\mathcal{F}$  is the associated BOREL  $\sigma$ -field and  $(\mathcal{F}_t)_{t \geq 0}$  is the canonical filtration.

The centered FLEMING-VIOT process is a measure-valued diffusion in

$$\mathcal{M}_1^{c,2}(\mathbb{R}) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \mid \langle \text{id}^2, \mu \rangle < \infty, \langle \text{id}, \mu \rangle = 0 \right\}$$

which is endowed with the trace of the topology of weak convergence on  $\mathcal{M}_1(\mathbb{R})$ . We consider the filtered probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0})$  where

$$\tilde{\Omega} := \left\{ X \in \mathcal{C}^0([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R})) \mid \forall T > 0, \sup_{0 \leq t \leq T} \langle \text{id}^2, X_t \rangle < \infty \right\}$$

is endowed with the trace of the SKOROHOD topology on  $\Omega$ ,  $\tilde{\mathcal{F}}$  is the trace of the  $\sigma$ -field  $\mathcal{F}$  and  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  is the trace of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . For all  $k \in \mathbb{N}$  and for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , we also denote

$$M_k(\mu) = \langle |\text{id}|^k, \mu \rangle \tag{1.52}$$

the  $k^{\text{th}}$  moment of  $\mu$ . For all  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  we denote for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ ,  $F_g(\nu) := F(\langle g, \nu \rangle)$ .

**Definition 1.3.1.** A probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  is said to solve the centered FLEMING-VIOT martingale problem with resampling rate  $\lambda \in (0, +\infty)$  and with initial condition  $\mu \in$

$\mathcal{M}_1^{c,2}(\mathbb{R})$  if the canonical process  $(X_t)_{t \geq 0}$  on  $\tilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$  and for each  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\widehat{M}_t^F(g) := F_g(X_t) - F_g(X_0) - \int_0^t \mathcal{L}_{\text{FVc}} F_g(X_s) ds \quad (1.53)$$

with for all  $\varpi \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,

$$\begin{aligned} \mathcal{L}_{\text{FVc}} F_g(\varpi) &:= F'(\langle g, \varpi \rangle) \left( \left\langle \frac{g''}{2}, \varpi \right\rangle + \lambda [\langle g'', \varpi \rangle M_2(\varpi) - 2 \langle g' \times \text{id}, \varpi \rangle] \right) \\ &+ \lambda F''(\langle g, \varpi \rangle) \left( \langle g^2, \varpi \rangle - \langle g, \varpi \rangle^2 + \langle g', \varpi \rangle^2 M_2(\varpi) - 2 \langle g', \varpi \rangle \langle g \times \text{id}, \varpi \rangle \right) \end{aligned} \quad (1.54)$$

is a continuous  $\mathbb{P}_\mu$ -martingale in  $L^2(\tilde{\Omega})$  with quadratic variation process

$$\begin{aligned} \langle \widehat{M}^F(g) \rangle_t &= 2\lambda \int_0^t [F'(\langle g, X_s \rangle)]^2 [\langle g^2, X_s \rangle - \langle g, X_s \rangle^2 \\ &+ \langle g', X_s \rangle^2 M_2(X_s) - 2 \langle g', X_s \rangle \langle g \times \text{id}, X_s \rangle] ds. \end{aligned}$$

Note that a major difficulty comes from the presence of terms in  $M_2(\mu)$  in (1.54). We will see in particular that this leads to the creation of particles in the dual process. Note also that the operator given by (1.54) is the generator (in the sense of DYNKIN) of the  $\mathcal{M}_1^{c,2}(\mathbb{R})$ -valued centered FLEMING-VIOT process.

**Remark 1.3.2.** *The additional terms in the martingale problem (1.53) with respect to the martingale problem (1.27) describe the impact of centering and ensure that at all times the centered FLEMING-VIOT process remains  $\mathcal{M}_1^{c,2}(\mathbb{R})$ -valued.*

Let us define by  $\tau_\alpha$  the translation operator of vector  $\alpha \in \mathbb{R}$ . For all  $\mu \in \mathcal{M}_1(\mathbb{R})$ , for all  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\tau_\alpha \# \mu(A) := \mu(\tau_\alpha^{-1}(A)) = \mu\left(\left\{x - \alpha \mid x \in A\right\}\right),$$

where  $\#$  is the pushforward operator ( $\tau_\alpha \# \mu$  is the pushforward measure of  $\mu$  by  $\tau_\alpha$ ).

The main result of this section is the following:

**Theorem 1.3.3.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , there exists a probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  satisfying the martingale problem of Definition 1.3.1, given by the law of the process  $(Z_t)_{t \geq 0}$  defined by*

$$Z_t := \tau_{-\langle \text{id}, Y_t \rangle} \# Y_t := Y_t(\cdot + \langle \text{id}, Y_t \rangle), \quad t \geq 0 \quad (1.55)$$

where  $(Y_t)_{t \geq 0}$  is the original FLEMING-VIOT process.

Note that the characterisation of the centered FLEMING-VIOT process with our martingale problem approach is new. The definition (1.55) justifies the term ‘‘centered FLEMING-VIOT process’’.

*Sketch of proof of Theorem 1.3.3.* Our method is to exploit the original FLEMING-VIOT martingale problem (1.27) and asymptotic expansions of  $F_g(Z_{t+h}) - F_g(Z_t)$  when  $h \rightarrow 0$ . An

additional difficulty occurs in our case since we need to apply the original FLEMING-VIOT martingale problem to *predictable test functions* which we do by classical methods based on *regular conditional probabilities*.

Following the approach of Section 1.2.1.3, we then extend the martingale problem of Definition 1.3.1 to polynomial test functions.

**Definition 1.3.4.** *The probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  is said to solve the centered FLEMING-VIOT martingale problem for polynomials with resampling rate  $\lambda$  and with initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , if the canonical process  $(X_t)_{t \geq 0}$  on  $\tilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$ , for all  $n \in \mathbb{N}^*$ , and for each  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,*

$$\widehat{M}_t^{(n)}(f) := \langle f, X_t^n \rangle - \langle f, X_0^n \rangle - \int_0^t \mathcal{L}_{\text{FVc}} P_{f,n}(X_s) \, ds \quad (1.56)$$

with for all  $\varpi \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathcal{L}_{\text{FVc}} P_{f,n}(\varpi) := \langle B^{(n)} f, \varpi^n \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \varpi^{n-1} \rangle - \langle f, \varpi^n \rangle] + \lambda \sum_{i=1}^n \sum_{j=1}^n \langle K_{i,j} f, \varpi^{n+1} \rangle \quad (1.57)$$

is a continuous  $\mathbb{P}_\mu$ -martingale in  $L^2(\tilde{\Omega})$  where, for all  $i, j \in \{1, \dots, n\}$ ,

- $\Phi_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$ , with  $i \neq j$ , is the function obtained from  $f$  by inserting the variable  $x_i$  between  $x_{j-1}$  and  $x_j$  when  $i < j$  and by inserting the variable  $x_{i-1}$  between  $x_{j-1}$  and  $x_j$  when  $i > j$ :

$$\begin{aligned} \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_i, x_j, x_{j+1}, \dots, x_{n-1}) & i < j \\ \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_{i-1}, x_j, x_{j+1}, \dots, x_{n-1}) & i > j \end{aligned}$$

- $K_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}^2(\mathbb{R}^{n+1}, \mathbb{R})$  is defined as

$$K_{i,j} f(x_1, \dots, x_n, x_{n+1}) := \partial_{ij}^2 f(x_1, \dots, x_n) x_{n+1}^2.$$

Compared to the original FLEMING-VIOT process of Definition 1.2.7, we require more regularity here by taking functions  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ . This is due to the operator  $K_{i,j}$ .

**Theorem 1.3.5.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , the probability measure  $\mathbb{P}_\mu$  constructed in Theorem 1.3.3, satisfies the martingale problem of Definition 1.3.4.*

### 1.3.1.2 Uniqueness for the centered Fleming-Viot process

As for the original FLEMING-VIOT martingale problem, we will prove uniqueness to the martingale problem (1.53) by relying on the *duality method* (see Section 1.2.1.3). However, additional difficulties occur in our case since bounds on the dual process are much harder to obtain and the duality identity cannot be proved in its usual form. As a consequence, our uniqueness result only holds for initial conditions admitting finite moments contrary to the original approach.

#### Main result.

**Theorem 1.3.6.** *The centered FLEMING-VIOT martingale problem (1.53) has a unique solution if the initial condition has all its moments finite.*



We comment on the assumption on the moments of the initial condition in the sketch of the proof given below.

**Duality and sketch of the uniqueness proof.** Our proof of Theorem 1.3.6 is based on the *duality method* as proposed in Section 1.2.1.3. From (1.57), the operator  $\mathcal{L}_{\text{FVc}}$  applied on the function  $P_{f,n}$  defined in (1.30) with fixed  $f$  and  $n \in \mathbb{N}^*$ , satisfies the following identity:

$$\begin{aligned} \mathcal{L}_{\text{FVc}}P_{f,n}(\mu) &= \langle B^{(n)}f, \mu^n \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j}f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle] \\ &\quad + \lambda \sum_{i=1}^n \sum_{j=1}^n [\langle K_{i,j}f, \mu^{n+1} \rangle - \langle f, \mu^n \rangle] + \lambda n^2 \langle f, \mu^n \rangle \\ &=: \tilde{\mathcal{L}}_f^* P_{f,n}(\mu) + \lambda n^2 \langle f, \mu^n \rangle \end{aligned} \tag{1.58}$$

We note that  $\tilde{\mathcal{L}}_f^*$  can be seen as an operator acting on the function  $f \mapsto P_{f,n}(\mu)$  with fixed  $\mu$ . The operator  $\tilde{\mathcal{L}}_f^*$  can be interpreted as the generator of a stochastic process on the state space  $\tilde{\mathcal{C}} := \bigcup_{n \in \mathbb{N}^*} \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ . Following Section 1.2.1.3 and ETHIER-KURTZ's works [89, 91], this suggests to introduce a dual process  $(\xi_t)_{t \geq 0}$ , of generator  $\tilde{\mathcal{L}}_f^*$  and to prove a duality relation of the form:

$$\forall t \geq 0, \quad \mathbb{E} \left( \langle \xi_0, X_t^{M(0)} \rangle \right) = \mathbb{E} \left( \langle \xi_t, X_0^{M(t)} \rangle \exp \left( \lambda \int_0^t M^2(u) du \right) \right) \tag{1.59}$$

where  $M := (M(t))_{t \geq 0}$  is a MARKOV's birth and death process in  $\mathbb{N}$  whose transition rates  $q_{i,j}$  from  $i$  to  $j$  are given by:

$$\textbf{(1)} \quad q_{n,n+1} = \lambda n^2 \quad \textbf{(2)} \quad q_{n,n-1} = \lambda n(n-1) \quad \textbf{(3)} \quad q_{i,j} = 0 \text{ otherwise.}$$

It is standard to prove that this birth and death process is non-explosive. It is known that the relation (1.59) implies uniqueness [94, Theorem 4.4.2]. However in our situation, it is difficult to obtain the strong version (1.59). For technical reasons, we will obtain only a weakened version. The main difference that explains this difficulty, compared to Section 1.2.1.3, is that the process  $(M(t))_{t \geq 0}$  is not a pure death process. Indeed, because of transition rate **(1)** of  $M$ , there is creation of particles in  $(\xi_t)_{t \geq 0}$  which corresponds to the operator  $K_{i,j}$ . Therefore, the proof will be divided in two large steps.

*Step 1. Construction of the dual process  $(\xi_t)_{t \geq 0}$ .* The relation (1.58) suggests that the  $\tilde{\mathcal{C}}$ -valued dual process  $(\xi_t)_{t \geq 0}$  jumps, for all  $i, j \in \{1, \dots, n\}$  from  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  to  $K_{i,j}f \in \mathcal{C}^2(\mathbb{R}^{n+1}, \mathbb{R})$  at rate  $\lambda$  and if  $i \neq j$ , from  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  to  $\Phi_{i,j}f \in \mathcal{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$  at rate  $\lambda$ . Moreover, note that if  $n = 1$ , the dual process can only jump from  $f \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  to  $K_{i,j}f$ . Between jumps, this dual process evolves according to the semi-group of operator  $(T^{(n)}(t))_{t \geq 0}$  associated to the generator  $B^{(n)}$  defined by

$$B^{(n)}f(x) := \frac{1}{2} \Delta f(x) - 2\lambda (\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}), \quad x \in \mathbb{R}^n$$

where we denote by  $\mathbf{1} \in \mathbb{R}^n$ , the vector whose coordinates are all 1 and by  $\Delta$  the Laplacian operator on  $\mathbb{R}^n$ . In this thesis, we will give an explicit expression of the semi-group  $(T^{(n)}(t))_{t \geq 0}$  defined as an integral against Gaussian kernels whose expression is more sophisticated than (1.34). This representation will be derived from a probabilistic interpretation of the semi-group using a FEYNMAN-KAC's formula. We then define the dual process as follows:

**Definition 1.3.7.** For all  $M(0) \in \mathbb{N}^*$ , for all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ ,

$$\xi_t := T^{(M(\tau_n))} (t - \tau_n) \Lambda_n T^{(M(\tau_{n-1}))} (\tau_n - \tau_{n-1}) \Lambda_{n-1} \cdots \Lambda_1 T^{(M(0))} (\tau_1) \xi_0, \\ \tau_n \leq t < \tau_{n+1}, \quad n \in \mathbb{N},$$

where  $(\tau_n)_{n \in \mathbb{N}}$  is the sequence of jump times of the birth-death process  $M$  with  $\tau_0 = 0$  and where  $(\Lambda_n)_{n \in \mathbb{N}}$  is a sequence of random operators. These are conditionally independent given  $M$  and satisfy for all  $k \in \mathbb{N}$ ,  $n \geq 1$  and  $1 \leq i \neq j \leq n$ ,

$$\mathbb{P} \left( \Lambda_k = \Phi_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n - 1\} \right) = \frac{1}{n(n-1)}$$

and for all  $n \geq 1$  and  $1 \leq i, j \leq n$ ,

$$\mathbb{P} \left( \Lambda_k = K_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n + 1\} \right) = \frac{1}{n^2}.$$

*Step 2. Weakened duality relation.* We consider fixed  $M(0) \in \mathbb{N}^*$ ,  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  and  $(X_t)_{t \geq 0}$  a stochastic process whose law  $\mathbb{P}_\mu$  is a solution of the martingale problem (1.53) with  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ . We introduce a dual process  $(\xi_t)_{t \geq 0}$  independent of  $(X_t)_{t \geq 0}$  built on the same probability space (enlarging it if necessary). We shall denote by  $\mathbb{P}_{(\mu, \xi_0)}$ , the law of  $((X_t, \xi_t))_{t \geq 0}$  on this probability space. For any  $k \in \mathbb{N}$ , we introduce the stopping time

$$\theta_k := \inf \left\{ t \geq 0 \mid M(t) \geq k \quad \text{or} \quad \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq k \right\}. \quad (1.60)$$

**Theorem 1.3.8.** Given any  $(X_t)_{t \geq 0}$ ,  $(\xi_t)_{t \geq 0}$  as above, we have the weakened duality identity: for all  $k \in \mathbb{N}$  and any stopping time  $\theta$  such that  $\theta \leq \theta_k$ ,

$$\forall t \geq 0, \quad \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_0, X_{t \wedge \theta}^{M(0)} \rangle \right) = \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_{t \wedge \theta}, X_0^{M(t \wedge \theta)} \rangle \exp \left( \lambda \int_0^{t \wedge \theta} M^2(u) du \right) \right). \quad (1.61)$$

Note that this result holds true for any initial measure  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ . The stopping time  $\theta_k$  ensures that each of the quantities involved in (1.61) are bounded and thus that their expectations are finite. Afterwards, we want to prove that if two solutions of the martingale problem satisfy the weakened duality identity, then their 1-dimensional marginals coincide. This is where we need stronger assumptions on the initial condition  $\mu$ .

**Lemma 1.3.9.** Assume that  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  has all its moments finite. Then, the stopping time  $\theta_k$  defined by (1.60) satisfies  $\lim_{k \rightarrow +\infty} \theta_k = +\infty$ ,  $\mathbb{P}_{(\mu, \xi_0)}$  - a.s.

We will see in Remark 2.6.5 of Chapter 2 that the assumption on  $\mu$  is optimal in the following sense: even if  $\xi_0$  is bounded,  $\xi_t$  may have polynomial growth of any exponent  $k$  in some direction of  $\mathbb{R}^{M(t)}$  so that  $\langle |\xi_t|, \mu^{M(t)} \rangle$  is infinite if  $\mu$  has infinite  $k^{\text{th}}$  moment. This shows that we cannot expect to have  $\theta_k \rightarrow +\infty$  when  $k \rightarrow +\infty$  if  $\mu$  has not all its moments finite. This means that we cannot expect that the duality method could give uniqueness for weaker assumptions on the initial condition.

Once Theorem 1.3.8 and Lemma 1.3.9 are proved, the proof of uniqueness of the centered FLEMING-VIOT martingale problem (1.53) can be completed in similar way to the one of Theorem 1.2.5.

### 1.3.1.3 Ergodicity for the centered Fleming-Viot process

In this section, we establish ergodicity properties with exponential convergence in total variation for the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$ . Recall in the case of the original FLEMING-VIOT process, standard duality arguments provide weak ergodicity estimates (see (1.40)). Here, we extend the weak ergodicity results obtained in [205, 91] using a coupling argument to obtain strong ergodicity properties (see (1.41)) which is an original contribution to the current literature. We also prove an explicit construction of the invariant probability measure from a variant of the modified look-down.

**Exponential ergodicity in total variation.** For all  $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$  the total variation distance between  $\mu$  and  $\nu$ , is defined by

$$\|\mu - \nu\|_{TV} := \sup_{B \in \mathcal{B}(\mathbb{R})} |\mu(B) - \nu(B)| = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} |\langle f, \mu \rangle - \langle f, \nu \rangle|.$$

The main result of this section is the following:

**Theorem 1.3.10.** *There exists a unique invariant probability measure  $\pi \in \mathcal{M}_1(\mathcal{M}_1^{c,2}(\mathbb{R}))$  for  $(Z_t)_{t \geq 0}$  and constants  $\alpha, \beta \in (0, +\infty)$  such that for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,*

$$\|\mathbb{P}_\mu(Z_T \in \cdot) - \pi\|_{TV} \leq \alpha \exp(-\beta T).$$

To prove this result we proceed according to the following sketch of proof. Let us begin fixing  $N \in \mathbb{N}^*$  and denoting  $\mathcal{M}_{1,N}^{c,2}(\mathbb{R}) := \left\{ \mu_N \in \mathcal{M}_{1,N}(\mathbb{R}) \mid \langle \text{id}^2, \mu_N \rangle < \infty, \langle \text{id}, \mu_N \rangle = 0 \right\}$ .

*Sketch of proof of Theorem 1.3.10.*

- In Step 1, we construct the *centered*  $N$ -MORAN process  $(Z_t^N)_{t \geq 0}$  from the (original)  $N$ -MORAN process  $(Y_t^N)_{t \geq 0}$  defined by (1.28) as:

$$Z_t^N := \tau_{-\langle \text{id}, Y_t^N \rangle} \# Y_t^N, \quad t \geq 0$$

and we prove the following result:

**Proposition 1.3.11.** *For all initial condition  $Z_0^N := \frac{1}{N} \sum_{i=1}^N \delta_{V_i} \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$  with  $(V_i)_{1 \leq i \leq N}$  exchangeable  $\mathbb{R}$ -valued random variables such that  $Z_0^N$  converges in law to  $Z_0 \in \mathcal{M}_1^{c,2}(\mathbb{R})$  as  $N \rightarrow +\infty$  and satisfying  $\sup_{N \in \mathbb{N}^*} \mathbb{E}(\langle \text{id}^2, Z_0^N \rangle) < \infty$ , the centered MORAN process  $(Z_t^N)_{t \geq 0}$  converges in law on  $\mathcal{C}^0([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R}))$ , as  $N \rightarrow +\infty$ , to the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$  solution of the martingale problem (1.56) with initial condition  $Z_0$ .*

Note that this proposition extends the convergence result [60, Theorem 2.7.1] from the  $N$ -MORAN process to the original FLEMING-VIOT process.

- In Step 2, we construct backward in time, the original and centered  $N$ -MORAN processes whose their genealogies are exactly determined by KINGMAN's coalescent with coalescence rate  $2\lambda$  (see Section 1.2.1.2). The state of the population at the final time is constructed from the ancestral positions by following the genealogy and adding mutations on the

genealogical tree of the population according to standard Brownian motions. Therefore, at the final time, the position of each individual of the population is equal to the sum of the position of its ancestor at time 0 and the Brownian mutation steps that occurred during its ancestral branches in the coalescent.

- In Step 3, we couple (see (1.43)) two centered  $N$ -MORAN processes from different initial conditions but with the same KINGMAN genealogy and the same mutations in order to establish a.s. coupling of the two processes on the event where the KINGMAN's coalescent describing the genealogy has a single ancestor. The following exponential ergodicity result follows.

**Proposition 1.3.12.** *For all  $N \in \mathbb{N}^*$  there exists a unique invariant probability measure  $\pi_N \in \mathcal{M}_1(\mathcal{M}_{1,N}^{c,2}(\mathbb{R}))$  for  $(Z_t^N)_{t \geq 0}$  and constants  $\alpha, \beta \in (0, +\infty)$  independent of  $N$  such that for all  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,*

$$\left\| \mathbb{P}_{\mu_N} \left( Z_T^N \in \cdot \right) - \pi_N \right\|_{TV} \leq \alpha \exp(-\beta T).$$

Finally, we deduce Theorem 1.3.10 by letting  $N \rightarrow +\infty$  and using Propositions 1.3.11 and 1.3.12. In addition, these propositions allow to prove the convergence in law of  $(\pi_N)_{N \in \mathbb{N}^*}$  to  $\pi$  when  $N \rightarrow +\infty$ .

**Characterisation of  $\pi_N$  and  $\pi$ .** Another new contribution of our part consists in providing an explicit construction of  $\pi_N$  and  $\pi$  using an adaptation of the modified look-down of Section 1.2.1.2 for time running backward from 0 to  $-\infty$  up to the (negative) coalescence time of the population  $T_{coal}^\infty$ . Taking the same notations of Section 1.2.1.2, this leads to define the vector of positions of individuals  $(U_1, \dots, U_N) := (U_1(0), \dots, U_N(0))$  from the vector  $(U_1(T_{coal}^\infty), \dots, U_N(T_{coal}^\infty))$ . We then construct the empirical distributions  $Y_{coal}^N := \frac{1}{N} \sum_{i=1}^N \delta_{U_i}$  and its centered version  $Z_{coal}^N := \frac{1}{N} \sum_{i=1}^N \delta_{V_i}$  where  $V_i := U_i - \frac{1}{N} \sum_{j=1}^N U_j$ ,  $i \in \mathbb{N}^*$  and deduce from Proposition 1.3.12 that  $Z_{coal}^N$  is distributed as  $\pi_N$ .

We then wish to let  $N \rightarrow +\infty$ . The difficulty consists in establishing the exchangeability property of the family  $(U_i)_{i \in \mathbb{N}^*}$  which will allow to apply the DE FINETTI representation theorem [142, Theorem 12.26 and Remark 12.27] to prove existence of a random variable  $Y_{coal}^\infty$  such that  $(Y_{coal}^N)_{N \in \mathbb{N}^*}$  converges a.s. when  $N \rightarrow +\infty$  to  $Y_{coal}^\infty$ . Then, defining the random variable  $Z_{coal}^\infty \in \mathcal{M}_1(\mathbb{R})$  as

$$Z_{coal}^\infty := \tau_{-\langle \text{id}, Y_{coal}^\infty \rangle} \# Y_{coal}^\infty,$$

we prove that  $(Z_{coal}^N)_{N \in \mathbb{N}^*}$  converges a.s. when  $N \rightarrow +\infty$  to  $(Z_{coal}^\infty)$  where  $(Z_{coal}^\infty)$  has the law  $\pi$ .

This characterisation of  $\pi$  makes explicit the expression of the second moment under  $\pi$  given in the next result. This point will be useful in our slow-fast analysis (see Section 1.3.2) when the limit slow component is averaged over the unique invariant probability measure of the limit fast component which is precisely  $\pi$ .

**Corollary 1.3.13.** *We have  $\int_{\mathcal{M}_1^{c,2}(\mathbb{R})} M_2(\mu) \pi(d\mu) = 1/2\lambda$ .*

### 1.3.2 From Chapter 3

In Chapter 3, the stochastic IBM describing Darwinian evolution of asexual, phenotypic trait-structured population of Section 1.1.3.2 is considered. Recall that its study is motivated

by biological questions on adaptive dynamics as presented in Section 1.1 and more specically in Section 1.1.4.6. The aim of this chapter is to describe, on the time scale  $1/K\sigma^2$ , the evolutionary dynamics of the mean trait (at the population level) as solution of the CEAD when mutations (at the individual level) are *frequent* ( $p = 1$ ) and under the simultaneous asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ). This convergence result from the IBM to the CEAD can be reformulated into a slow-fast asymptotic analysis acting on two evolutionary time scales. A slow scale corresponding to the dynamics of the mean trait and a fast scale corresponding to the centered and dilated distribution of traits seen as a discrete version of the centered FLEMING-VIOT process whose limit is characterised in Section 1.3.1. The slow-fast asymptotic analysis is based on KURTZ's averaging method developed in Section 1.2.2.2.

This section is organised as follows. In Section 1.3.2.1 we present the main convergence result to the CEAD. In Section 1.3.2.2, we introduce the slow and fast dynamics involved in our IBM and we give the heuristic leading to the CEAD. In Section 1.3.2.3, we present some difficulties encountered to prove our main result. Finally, in Section 1.3.2.4, we develop our stochastic averaging method inspired by the KURTZ averaging approach.

### 1.3.2.1 Main result

We consider the stochastic IBM of Section 1.1.3.2 taking the probability of mutation  $p = 1$  such that, now, we use the simplified notation  $\nu^{K,\sigma}$  for  $\nu^{K,1,\sigma}$ . Recall that  $\nu^{K,\sigma}$  is a MARKOV process taking values in  $\mathcal{M}_{1,K}(\mathbb{R})$  defined in (1.6). For all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , let us denote by  $\text{Diam}(\text{Supp } \nu)$  the diameter of the support of  $\nu$ . Let us denote by **(A)'** the following two assumptions:

**(A1)'** The maps  $b : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ , and  $\theta : \mathbb{R} \rightarrow \mathbb{R}_+$  are respectively in  $\mathcal{C}_b^2(\mathbb{R}^2, \mathbb{R})$  and  $\mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  and there exists  $0 < \underline{b}, \bar{b}, \underline{\theta}, \bar{\theta} < \infty$  such that:

$$\underline{b} \leq b(\cdot, \cdot) \leq \bar{b}, \quad \text{and} \quad \underline{\theta} \leq \theta(\cdot) \leq \bar{\theta}.$$

**(A2)'** (a) There exists  $A_m \in (0, +\infty)$  such that the law  $m(x, dh)$  is absolutely continuous with respect to the LEBESGUE measure on  $\mathbb{R}$  with density  $m(x, h)$  centered and satisfies

$$\forall x \in \mathbb{R}, \forall |h| \geq A_m, \quad m(x, h) = 0.$$

(b) For all  $\alpha \in \mathbb{R}$ , for all  $\ell \in \mathbb{N}$ , we denote by

$$m_\ell(\alpha) := \int_{\mathbb{R}} |h|^\ell m(\alpha, h) dh$$

the  $\ell^{\text{th}}$  moment of the mutation law. We assume for  $\ell \in \{2, 4, 6\}$  that  $m_\ell$  is LIPSCHITZ and there exists  $\underline{m}_\ell, \bar{m}_\ell \in (0, +\infty)$  such that for all  $\alpha \in \mathbb{R}$ ,  $\underline{m}_\ell \leq m_\ell(\alpha) \leq \bar{m}_\ell$ .

Note that **(A1)'** is exactly **(A1)** from Section 1.1.3.2.

The main result of this thesis is the following:

**Theorem 1.3.14.** *Assume **(A)'** and for all  $K \in \mathbb{N}^*$ ,  $\sigma \in (0, +\infty)$ , for some  $x_0 \in \mathbb{R}$ , for all  $\ell \in \{1, 2, 3\}$ ,*

$$\langle \text{id}, \nu_0^{K,\sigma} \rangle := x_0, \quad \text{and} \quad \langle (\text{id} - x_0)^{2\ell}, \nu_0^{K,\sigma} \rangle \leq C_{2\ell}^* K^\ell \sigma^{2\ell}$$

for some  $C_{2\ell}^* > 0$ . Assume that  $K \rightarrow +\infty$ ,  $\sigma \rightarrow 0$  and  $\sigma = \sigma(K)$  such that

$$\exists \varepsilon > 0, \quad \forall C > 0, \quad K^{-C \log(K)} \ll \sigma \ll K^{-(2+\varepsilon)}. \quad (1.62)$$

Then, for the SKOROHOD topology, the sequence of the mean trait processes  $\left( \left( \langle \text{id}, \nu_{t/K\sigma^2}^{K,\sigma} \rangle \right)_{t \geq 0} \right)_{K \in \mathbb{N}^*}$  converges in law in  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$  when  $K \rightarrow +\infty$ ,  $\sigma \rightarrow 0$ , to the unique deterministic process  $(\zeta_t)_{t \geq 0} \in \mathcal{C}^0(\mathbb{R}_+, \mathbb{R})$  with initial condition  $\zeta_0 := x_0$ , which is solution to the CEAD:

$$\dot{y} = \partial_1 \text{Fit}(y, y) \times \frac{\theta(y)}{2b(y, y)} m_2(y). \quad (1.63)$$

In addition, we have the following support concentration property: for all  $T > 0$ ,

$$\sup_{t \in [0, T]} \text{Diam} \left( \text{Supp } \nu_{t/K\sigma^2}^{K,\sigma} \right) \leq \frac{1}{K} \quad (1.64)$$

holds with probability which tends to 1 when  $K \rightarrow +\infty$ .

We recover the classical form of canonical equation (see Section 1.1.2.3) and, more precisely, the same canonical equation as in Section 1.1.4.2 (see Theorem 1.1.6). Note that (1.64) describes, in some sense, when  $K \rightarrow +\infty$ , the convergence of the population distribution  $(\nu_{t/K\sigma^2}^{K,\sigma})$  to a DIRAC mass. Note also that Theorem 1.3.14 holds only for a certain range of  $\sigma$  parameters (as a function of  $K$ ) given by (1.62) which must be small enough but not too small either.

Finally, note that Figures 1.13 and 1.14 illustrate the convergence result to the CEAD of Theorem 1.3.14. In these figures, the CEAD describes the phase between the initial time  $t = 0$  and the time when the population stabilises in its equilibrium (at time  $t \simeq 10000$  in Figure 1.13 and  $t \simeq 40000$  in Figure 1.14).

**Remark 1.3.15.** We conjecture that the assumption  $\sigma \ll K^{-(2+\varepsilon)}$  given by (1.62) is too restrictive and that it can be weakened by  $\sigma \ll K^{-(\frac{3}{2}+\varepsilon)}$ . This conjecture is supported by the discussion in Section 3.2.3.4 of Chapter 3 and constitutes a future prospect. This is motivated by the asymptotic expansion of the generator (1.67) given below, where the leading error term is  $\sigma K^{3/2} M_2(\mu)$ .

The aim of the rest of the section is to give a sketch of proof of Theorem 1.3.14. Since  $\sigma = \sigma(K)$ , we will use the notation  $\nu^K$  to designate  $\nu^{K,\sigma}$ .

### 1.3.2.2 Slow-fast asymptotic analysis

In this section, we introduce the slow and fast dynamics involved in our IBM and we give the heuristic leading to the CEAD (1.63).

Let  $a, \alpha \in \mathbb{R}$  and  $\mathcal{B}(\mathbb{R})$  the BOREL  $\sigma$ -field on  $\mathbb{R}$ . Let us define respectively by  $\tau_\alpha$  and  $h_a$ , the translation operator of vector  $\alpha$  and the homothety of ratio  $a$ . For all  $x \in \mathbb{R}$ , for all  $A \in \mathcal{B}(\mathbb{R})$ ,

$$(h_a \circ \tau_\alpha) \# \delta_x(A) = \delta_{a(x+\alpha)}(A) \quad \text{and} \quad (\tau_\alpha \circ h_a) \# \delta_x(A) = \delta_{ax+\alpha}(A),$$

where  $\#$  is still the pushforward operator. Note that for all  $K \in \mathbb{N}^*$ , any  $\nu \in \mathcal{M}_{1,K}(\mathbb{R})$  has all its moments finite. We define for all  $K \in \mathbb{N}^*$ ,  $\mathcal{M}_{1,K}^c(\mathbb{R})$  the set of centered probability measures of  $\mathcal{M}_{1,K}(\mathbb{R})$ . From the population process  $(\nu_{t/K\sigma^2}^K)_{t \geq 0}$ , we introduce two evolutionary dynamics:

- The slow dynamics, with value in  $\mathbb{R}$ , corresponds to that of the mean trait defined by

$$z_t^K := \langle \text{id}, \nu_{t/K\sigma^2}^K \rangle.$$

- The fast dynamics, with value in  $\mathcal{M}_{1,K}^c(\mathbb{R})$ , corresponds to that of the centered and dilated distribution of traits defined by

$$\mu_t^K := \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-z_t^K} \right) \# \nu_{t/K\sigma^2}^K = \frac{1}{K} \sum_{i=1}^K \delta_{\frac{1}{\sigma\sqrt{K}}(x_i(t) - z_t^K)}.$$

$$\text{Note that } \nu_t^K = \left( \tau_{z_t^K} \circ h_{\sigma\sqrt{K}} \right) \# \mu_{tK\sigma^2}^K.$$

Let us consider for all  $\ell \in \mathbb{N}^*$ ,

$$\mathcal{M}_1^{c,\ell}(\mathbb{R}) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \mid M_\ell(\mu) < \infty, \langle \text{id}, \mu \rangle = 0 \right\}$$

where  $M_\ell(\nu)$  is the  $\ell^{\text{th}}$  moment of  $\nu$  defined in (1.52) and with the convention  $M_\ell(\mu) = +\infty$  if  $\mu \notin \mathcal{M}_1^{c,\ell}(\mathbb{R})$ . Let us denote by  $\mathcal{B}(E, \mathbb{R})$  the set of BOREL functions on the metric space  $E$ .

In the following, for all  $K \in \mathbb{N}^*$ , we study the couple of processes  $((z_t^K, \mu_t^K))_{t \geq 0}$ . From (1.7) and an appropriate change of variable, the infinitesimal generator of the process  $((z_t^K, \mu_t^K))$ , denoted by  $\mathcal{L}^K$  is defined for all bounded measurable function  $\Phi$  from  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$  to  $\mathbb{R}$ , by

$$\begin{aligned} \mathcal{L}^K \Phi(z, \mu) &= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) \\ &\quad \times \left[ \Phi \left( z + \frac{\sigma\sqrt{K}}{K}(y-x), \tau_{-\frac{y-x}{K}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right) - \Phi(z, \mu) \right] \\ &+ \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \theta(\sigma\sqrt{K}x + z) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z, h) \\ &\quad \times \left[ \Phi \left( z + \frac{\sigma\sqrt{K}h}{K^{\frac{3}{2}}}, \tau_{-\frac{h}{K^{\frac{3}{2}}}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right] \right) - \Phi(z, \mu) \right] dh. \end{aligned}$$

This leads to the following results.

**Proposition 1.3.16.** (1) *If we consider a function  $f : (z, \mu) \mapsto f(z) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ , the generator  $\mathcal{L}^K$  satisfies the following decomposition*

$$\mathcal{L}^K f(z, \mu) = \mathcal{L}_{\text{SLOW}} f(z, \mu) + O\left(\frac{1}{K^2} + \frac{M_2(\mu)}{K} + \sigma\sqrt{K}M_3(\mu)\right) \quad (1.65)$$

where the operator  $\mathcal{L}_{\text{SLOW}}$ , of the limit slow component, is defined from  $\mathcal{C}_b^1(\mathbb{R}, \mathbb{R})$  to  $\mathcal{B}(\mathbb{R} \times \mathcal{M}_1^{c,2}(\mathbb{R}), \mathbb{R})$  by

$$\mathcal{L}_{\text{SLOW}} f(z, \mu) = f'(z)M_2(\mu) \partial_1 \text{Fit}(z, z). \quad (1.66)$$

(2) *If we consider a function  $F_\varphi : (z, \mu) \mapsto F_\varphi(\mu) := F(\langle \varphi, \mu \rangle)$  with  $F \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$  and  $\varphi \in \mathcal{C}_b^3(\mathbb{R}, \mathbb{R})$ , the generator  $\mathcal{L}^K$  satisfies the following decomposition*

$$\mathcal{L}^K F_\varphi(z, \mu) = \frac{\theta(z)m_2(z)}{K^2\sigma^2} \left[ \mathcal{L}_{\text{FVc}}^{\lambda(z)} F_\varphi(z, \mu) + O\left(\frac{1}{\sqrt{K}} + \sigma K^{\frac{3}{2}}M_2(\mu) + \frac{M_3(\mu)}{K}\right) \right] \quad (1.67)$$

where  $\mathcal{L}_{\text{FVc}}^{\lambda(z)}$  is the generator of the centered FLEMING-VIOT process with resampling rate  $\lambda(z) := \frac{b(z,z)}{\theta(z)m_2(z)}$ , given by (1.54).

From (1.67) and (1.65), note that the fast component  $(\mu_{K\sigma^2 t}^K)_{t \geq 0} = \left( \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-\langle \text{id}, \nu_t^K \rangle} \right) \# \nu_t^K \right)_{t \geq 0}$  moves on the evolutionary time scale  $K$  whereas the slow component  $(z_{K\sigma^2 t}^K)_{t \geq 0} = (\langle \text{id}, \nu_t^K \rangle)_{t \geq 0}$  moves over a much longer evolutionary time scale  $1/K\sigma^2$ . The different time scales involved in this model can be represented as in Figure 1.28. It follows that the fast component will be the first to stabilise in its equilibrium state.

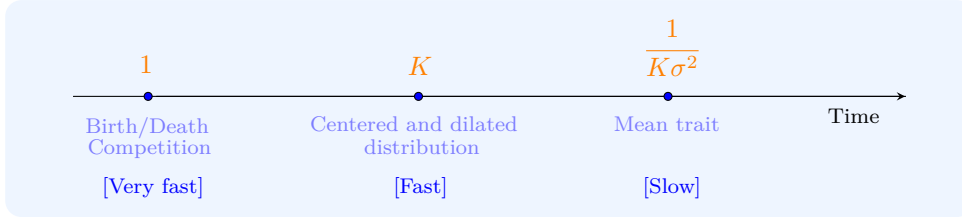


FIGURE 1.28 – Different time scales involved in our IBM expected in the original time scale of the process  $(\nu_t^K)_{t \geq 0}$ .

As the limit fast component is driven by a centered FLEMING-VIOT process, we expect that it will inherit ergodicity properties as stated in Section 1.3.1.3, where we established the existence of a unique invariant probability measure  $\pi^\lambda$  for the centered FLEMING-VIOT process. In particular, we expect that the averaging principle applied to the drift coefficient in (1.66) will lead to the averaged drift

$$\partial_1 \text{Fit}(z_t, z_t) \int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^{\lambda(z_t)}(d\mu) = \frac{\partial_1 \text{Fit}(z_t, z_t)}{2\lambda(z_t)} = \partial_1 \text{Fit}(z_t, z_t) \frac{\theta(z_t)}{2b(z_t, z_t)} m_2(z_t)$$

where  $(z_t)_{t \geq 0}$  is the limit slow component and the computation of the integral is given by Corollary 1.3.13. Formally, by replacing  $M_2(\mu)$  by its mean value  $\int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^{\lambda(z_t)}(d\mu)$  in (1.66), we obtain that

$$\mathcal{L}_{\text{CEAD}} f(z, \mu) = f'(z) \partial_1 \text{Fit}(z, z) \frac{\theta(z)}{2b(z, z)} m_2(z)$$

and we recognise the generator of the announced CEAD.

Before explaining in Section 1.3.2.4 our implementation of the averaging principle we introduce some notations and some difficulties encountered to prove Theorem 1.3.14.

### 1.3.2.3 Notations and difficulties encountered

**(1) For tightness arguments and moments.** • The slow-fast analysis relies on a stochastic averaging result that exploits tightness arguments. The classical approach to prove tightness of sequences of laws [10, 102] requires to have uniform pathwise estimates on the moments of the process. In our situation, we note the presence of moments of order 2 and 3 in (1.65) and (1.67). However, we could not establish control in expectation of  $\sup_{0 \leq t \leq T} M_2(\mu_t^K)$  for all  $T \geq 0$  because of the long time scale  $1/K\sigma^2$  which does not allow to exploit the martingale problem



associated to the decomposition of  $M_2(\mu_t^K)$ :

$$\begin{aligned} M_2(\mu_t^K) &= M_2(\mu_0^K) - \frac{1}{K^2\sigma^2} \int_0^t \{2b(z_s^K, z_s^K) M_2(\mu_s^K) - \theta(z_s^K) m_2(z_s^K)\} ds \\ &\quad + \frac{1}{K^2\sigma^2} \int_0^t O\left(\frac{1}{K} + \sigma K M_2(\mu_s^K) + \sigma K^{\frac{3}{2}} M_3(\mu_s^K)\right) ds + M_t^{K, P_{\text{id}^2, 1}} \end{aligned} \quad (1.68)$$

where  $(M_t^{K, P_{\text{id}^2, 1}})_{t \geq 0}$  is a local-martingale. This is an important difference, for example with [167] which studies similar forms of individual-based processes.

*Our solution.* Instead, we will first use fine pathwise estimates and expectation bounds on  $M_2(\mu_t^K)$  up to the stopping time  $\tau^K$  defined for all  $K \in \mathbb{N}^*$  by  $\tau^K := \check{\tau}^K \wedge \hat{\tau}^K$  where

$$\hat{\tau}^K := \inf \left\{ t \geq 0 \mid M_2(\mu_t^K) \geq K^\varepsilon \right\} \quad \text{and} \quad \check{\tau}^K := \inf \left\{ t \geq 0 \mid \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}} \right\}$$

with  $\varepsilon > 0$  given in Theorem 1.3.14. Note that  $\check{\tau}^K$  allows to control the diameter of the support of the centered and dilated distribution of the traits  $\mu^K$ .

- Another important difficulty arises from the presence of third order moment in the DOOB semi-martingale decomposition of  $M_2(\mu_t^K)$  given by (1.68).

*Our solution.* This difficulty is overcome by introducing the stopping time  $\check{\tau}^K$  which guarantees that the error term in  $M_3(\mu_t^K)$  remains under control:

$$M_3(\mu_{t \wedge \check{\tau}^K}^K) \leq \text{Diam}(\text{Supp } \mu_{t \wedge \check{\tau}^K}^K) M_2(\mu_{t \wedge \check{\tau}^K}^K) \leq \frac{M_2(\mu_{t \wedge \check{\tau}^K}^K)}{\sigma K^{\frac{3+\varepsilon}{2}}}.$$

**(2) For stopping times.** The introduction of the previous stopping time leads to a new difficulty: establishing that  $\tau^K$  tends to  $+\infty$  in probability when  $K \rightarrow +\infty$ .

*Our solution.* To prove  $\check{\tau}^K$  that tends to  $+\infty$ , we use pathwise estimates on  $M_2(\mu_t^K)$  up to time  $\check{\tau}^K$  to estimate the different transitions of  $M_2(\mu_t^K)$  between thresholds of the form  $3^{\ell-1} K^{\frac{\varepsilon}{2}}$  and  $3^{\ell+1} K^{\frac{\varepsilon}{2}}$  and to construct a coupling between these transitions and biased random walks. This allows us to use large deviations results on random walks and estimates on the exit from an attracting domain (see e.g. [104, 64]) to prove that the stopping time  $\hat{\tau}^K$  converges to  $+\infty$  in probability when  $K \rightarrow +\infty$ . This in turn implies that  $\check{\tau}^K \rightarrow +\infty$  in probability when  $K \rightarrow +\infty$ . Note that the left bound of (1.62), where  $\sigma$  must not be too small, is crucial to obtain estimates of exit from an attractive domain for random walks.

**(3) For improvement of Assumption (1.62) (see Remark 1.3.15).** The difficulty consists in having estimates for moments of order  $2\ell$  for all  $\ell \in \mathbb{N}^*$ . But the approach developed in this thesis seems not to be successful for moments of order 8 or more. This is an open issue for a further work (see Section 3.2.3.4).

Other difficulties encountered are due to our implementation of the KURTZ averaging method detailed below.

### 1.3.2.4 Our averaging method

First of all, note that the KURTZ approach seems to be better adapted to our situation than the perturbed test function method (see Section 1.2.2.1). Indeed, the generator  $L_1$  being

that of a measure-valued diffusion, it is thus delicate to find the good class of test function  $\varphi_1$  satisfying (1.46) and the computations are difficult because of the moment terms. In addition, the remainder terms generated by (1.45) are difficult to control and the fast component  $\mu^K$  does not take the form  $\mu(tK)$ . However, KURTZ's averaging technique can neither be implemented directly because of difficulties inherent to our model described above. Therefore, our result given below required a complete reworking of the classical arguments.

To write our stochastic averaging method and to compare it with the original KURTZ's approach of Section 1.2.2.2, we introduce some notations in the first paragraph of this section. The second paragraph of this section is devoted to the tightness result. Finally, the third paragraph is dedicated to identify the limit of this slow-fast analysis.

**Notations.** Let us introduce the torus  $\mathbb{T}_R := [x_0 - 2R, x_0 + 2R]$  of length  $4R$  with  $R > 0$  fixed and  $x_0$  is the value of the mean trait of  $\nu_0^K$  as in Theorem 1.3.14.

We define all the previous notations, in similar way on  $\mathbb{T}_R$ . Let  $b_R \in \mathcal{C}_b^2(\mathbb{T}_R^2, \mathbb{R})$  and  $\theta_R \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{R})$  be two functions satisfying  $b_R = b$  on  $[x_0 - R, x_0 + R]^2$  and  $\theta_R = \theta$  on  $[x_0 - R, x_0 + R]$ . Let  $T > 0$  be fixed. We define the processes  $\nu^{K,R} := (\nu_t^{K,R})_{t \in [0, T]}$ ,  $z^{K,R} := (z_t^{K,R})_{t \in [0, T]}$  and  $\mu^{K,R} = (\mu_t^{K,R})_{t \in [0, T]}$  as follows

$$\nu_t^{K,R} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i^R(t)}, \quad z_t^{K,R} := \langle \text{id}, \nu_{t/K\sigma^2}^{K,R} \rangle \quad \text{and} \quad \mu_t^{K,R} := \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-z_t^K} \right) \# \nu_{t/K\sigma^2}^K$$

where for all  $i \in \{1, \dots, K\}$ ,  $\nu_t^{K,R}$  is defined on  $\mathcal{M}_1(\mathbb{T}_R)$  from  $b_R$  and  $\theta_R$  as  $\nu_t^K$  was defined from  $b$  and  $\theta$  and  $z_t^{K,R} \in \mathbb{T}_R$ . Note that  $\mu_t^{K,R}$  takes values in  $\mathcal{M}_1(\mathbb{T}_R^{K,\sigma})$  where  $\mathbb{T}_R^{K,\sigma}$  is the torus corresponding to the interval  $\left[ \frac{1}{\sigma\sqrt{K}}(x_0 - 2R - z_t^K), \frac{1}{\sigma\sqrt{K}}(x_0 + 2R - z_t^K) \right]$ . However, we will identify in the sequel  $\mathcal{M}_1(\mathbb{T}_R^{K,\sigma})$  as a subset of  $\mathcal{M}_1(\mathbb{R})$  using the natural embedding of  $\mathbb{T}_R^{K,\sigma}$  in  $\mathbb{R}$ . So,  $\mu_t^{K,R} \in \mathcal{M}_1(\mathbb{R})$ . We also define the stopping time  $\tau^{K,R} := \check{\tau}^{K,R} \wedge \widehat{\tau}^{K,R}$  where

$$\widehat{\tau}^{K,R} := \inf \left\{ t \geq 0 \mid M_2(\mu_t^{K,R}) \geq K^\varepsilon \right\} \quad \text{and} \quad \check{\tau}^{K,R} := \inf \left\{ t \geq 0 \mid \text{Diam}(\text{Supp } \nu_t^{K,R}) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}} \right\}.$$

We define on  $\mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  the sequence of random measures  $(\Gamma^{K,R})_{K \in \mathbb{N}^*}$  as follows

$$\Gamma^{K,R}(dt, d\mu) = \delta_{\mu_t^{K,R}}(d\mu) dt.$$

In the sequel, we study the limit as  $K$  tends to  $+\infty$  of the sequence of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$  according two steps. The first one is based on a tightness argument. The second one consists in characterising each accumulation point of the family of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$ .

**Tightness result.** The main result of this section is the following and corresponds to the one of Section 1.2.2.2 and follows a similar method.

**Theorem 1.3.17.** *Let  $T, R > 0$ . The sequence of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  is tight in the set of probability measures on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  and for any limiting value  $\mathbb{Q}$  of this*

sequence, the canonical process  $(\zeta^R, \Gamma^R)$  on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  satisfies that for all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ ,

$$N_t^{f,R} := f(\zeta_t^R) - f(\zeta_0^R) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\zeta_s^R, \mu) \Gamma^R(ds, d\mu) = 0 \quad (1.69)$$

$\mathbb{Q}$ -a.s. Moreover, for all  $t \in [0, T]$ ,

$$\mathbb{E}_{\mathbb{Q}} \left( \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} M_5(\mu) \Gamma^R(ds, d\mu) \right) < \infty, \quad (1.70)$$

so that the definition (1.69) makes sense.

The main modification of KURTZ's setting of [144, Theorem 2.1] is that we have to work with stopped times and need to be careful with moment estimates and uniform integrability properties. This led us to rewrite the proof (in 8 steps) according to the following sketch of proof where we have two additional steps compared to the one of Theorem 1.2.19.

*Sketch of proof of Theorem 1.3.17.*

- In Step 1, we establish the tightness of the sequence of laws of the stopped slow component  $(z_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$ . We exploit again criteria developed by [94, Theorems 3.9.1 and 3.9.4]. It's here that our restriction to the torus case  $\mathbb{T}_R$  is fundamental. This strategy allows to overcome the difficulty related to the verification of the compact containment condition, on the real line, of the slow component stopped at the stopping time  $\tau^{K,R}$ .
- In Step 2, we establish a compact containment condition for the stopped fast component  $(\mu_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$ . This step is new compared to the original proof because it is given by Assumption **(2)** of Theorem 1.2.19. In our situation, we use that  $\tau^{K,R}$  tends to  $+\infty$  in probability when  $K \rightarrow +\infty$  and uniform upper bound of  $\mathbb{E}(M_2(\mu_{t \wedge \tau^{K,R}}^{K,R}))$ .
- In Step 3, the proof of tightness of the family of laws of  $(\Gamma^{K,R})_{K \in \mathbb{N}^*}$  is obtained in similar way as the one of Theorem 1.2.19.
- In Step 4, we deduce the tightness of the family of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$ . In our case, an additional argument is required due to the stopping time  $\tau^{K,R}$ .
- In Step 5, we prove similar uniform integrability result as in the original proof. Note that (1.49) is replaced here by the assertion that the family  $(\int_0^{t \wedge \tau^{K,R}} M_2(\mu_s^{K,R}) ds)_{t \in [0, T], K \in \mathbb{N}^*}$ , is uniformly integrable. Note that we prove little more establishing that the family  $(\int_0^{t \wedge \tau^{K,R}} M_5(\mu_s^{K,R}) ds)_{t \in [0, T], K \in \mathbb{N}^*}$ , is uniformly integrable. This additional result is useful to prove (1.70).
- In Step 6, using also the SKOROHOD representation theorem, we prove an almost sure convergence result.
- In Step 7, we prove with the same method that  $N^{f,R}$  is a  $\mathbb{Q}$ -martingale. Note that, here, the error term  $\mathcal{E}_{t \wedge \tau^{K,R}}^{f,K,R}$  satisfies  $\mathcal{E}_{t \wedge \tau^{K,R}}^{f,K,R} = \frac{1}{K} O(\int_0^t [1 + M_2(\mu_s^{K,R})] \mathbb{1}_{s \leq \tau^{K,R}} ds)$  and (1.50) holds true thanks to uniform integrability of  $(\int_0^{t \wedge \tau^{K,R}} M_2(\mu_s^{K,R}) ds)_{t \in [0, T], K \in \mathbb{N}^*}$ .
- In Step 8, we prove that  $N^{f,R}$  is null  $\mathbb{Q}$ -a.s. This step is new and possible because we have an explicit form of the operator  $A$ . Indeed, here  $A := \mathcal{L}_{\text{SLOW}}$ .

**Identification of the limit.** Using the desintegration result of Lemma 1.2.20 replacing  $\gamma$  by  $\gamma^R$  we have that  $\Gamma^R(dt, d\mu) = \gamma_t^R(d\mu)dt$  under  $\mathbb{Q}$ . To characterise it, we have the following result:

**Proposition 1.3.18.** *For a.e.  $t \in [0, T]$ ,  $\mathbb{Q}$ -a.s.,  $\gamma_t^R = \pi^\lambda(\zeta^R)$  where  $\pi^\lambda$  is the unique invariant probability measure of the centered FLEMING-VIOT process with resampling rate  $\lambda$ .*

For this, we need to check that any measure  $\gamma$  in  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$  satisfying

$$\int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^\lambda \phi(\mu) \gamma(d\mu) = 0 \quad (1.71)$$

for a certain class  $\mathcal{F}$  of functions  $\phi$  must be  $\pi^\lambda$ . Equation (1.67) suggests to take  $\mathcal{F}$  as the set of functions of the form  $F_\varphi(\mu) := F(\langle \varphi, \mu \rangle)$ . However, the last property seems hard to prove for this choice of  $\mathcal{F}$ . Instead, we adapt a result of DAWSON [60, Theorem 2.7.1] that applies to the set  $\mathcal{F}$  of so-called *polynomial functions* of the form

$$P_{f,n}(\mu) := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu(dx_1) \cdots \mu(dx_n).$$

However the duality property used by DAWSON in [60] does not hold in our case. In Theorem 1.3.8 it is proved only a weak duality relation involving stopping times. This difficulty will be solved by proving that the measure  $\gamma$  in (1.71) gives mass only to measures having their first five moments finite.

This allows us to characterise the limit fast component and hence the limit slow component on the torus, as solution to an ODE. Since this ODE is non-explosive, we choose the torus large enough to conclude the proof on the real line.

### 1.3.3 Prospects

In this section, we propose several prospects to this work. As discussed in Remark 1.3.15 and Section 1.3.2.3 the range of  $\sigma$  parameters (1.62) may be improved into the assumption  $\sigma \ll K^{-(\frac{3}{2}+\epsilon)}$  by using estimates for moments of order  $2\ell$ ,  $\ell \in \mathbb{N}^*$ .

In the sequel, we present some other future works. In Section 1.3.3.1 we present an extension to models including fluctuations in population size. In Section 1.3.3.2, we propose other strategies to resolve the biological controversy related to the rare mutations assumption.

#### 1.3.3.1 Extension to models including fluctuations in population size

Recall that our approach consider the IBM of Section 1.1.3.2 where the population size is constant in order to reduce the number of time scales in our slow-fast analysis. Now, we consider the same IBM as in [37, 47, 6] where population size can fluctuate. In this context, we will see that the slow-fast asymptotic analysis involves three eco-evolutionary time scale.

We consider an asexual population model, structured by a 1-dimensional quantitative phenotypic trait. The evolution of the population is driven by two effects: mutations of the traits occurring during life of individuals, and selection as the reproductive and survival abilities of an individual depend of its trait. Each mutation has an amplitude of the order of magnitude  $\sigma \in (0, +\infty)$ .

We consider, at any time  $t \geq 0$ , a finite number  $N_t^{K,p,\sigma}$  of individuals, each of them holding a trait value in  $\mathbb{R}$ . Let us denote by  $x_1, \dots, x_{N_t^{K,p,\sigma}}$  the trait values of these individuals. The state

of the population at time  $t \geq 0$ , rescaled by  $K$ , can be described by the finite point measure on  $\mathbb{R}$

$$\widehat{\nu}_t^{K,p,\sigma} := \frac{1}{K} \sum_{i=1}^{N_t^{K,p,\sigma}} \delta_{x_i(t)}.$$

**Biological parameters.** For all  $x, y \in \mathbb{R}$ , let us introduce the following biological parameters:

- $b(x) \in \mathbb{R}_+$  is the rate of birth from an individual holding trait  $x$ .
- $d(x) \in \mathbb{R}_+$  is the rate of “natural” death for an individual holding trait  $x$ .
- $c(x, y) \in \mathbb{R}_+$  is a competition kernel representing the pressure felt by an individual holding trait  $x$  from an individual holding trait  $y$ .
- $m(x, dh)$  is the mutation law of the scaled mutation step  $h$ , born from an individual with trait  $x$ , where the mutant trait is given by  $y := x + \sigma h \in \mathbb{R}$ . It is a probability measure on  $\mathbb{R}$ .
- $\bar{n}_x := \frac{b(x) - d(x)}{c(x, x)}$  can be interpreted as the equilibrium density of a monomorphic population with trait  $x$  when there is no mutation.
- $\text{Fit}(y, x) := b(y) - d(y) - c(y, x)\bar{n}_x$  can be interpreted as the fitness function of a mutant individual with trait  $y$  in a monomorphic population with trait  $x$ .

Note that this model has two scaling parameters  $K$  and  $\sigma$ . The limit  $K \rightarrow +\infty$  corresponds to a large population scaling as in Theorem 1.1.2 (see [102]). The limit  $\sigma \rightarrow 0$  corresponds to a limit of small mutations. We define for all  $K \in \mathbb{N}^*$ ,  $\mathcal{M}^K(\mathbb{R}) := \left\{ \frac{1}{K} \sum_{i=1}^n \delta_{x_i} \mid n \in \mathbb{N}^*, (x_i)_{1 \leq i \leq n} \in \mathbb{R}^n \right\}$ . The process  $(\widehat{\nu}_t^{K,p,\sigma})_{t \geq 0}$  is an  $\mathcal{M}^K$ -valued MARKOV process with infinitesimal generator defined for any bounded measurable function  $\phi$  from  $\mathcal{M}^K$  to  $\mathbb{R}$ , by

$$\begin{aligned} \widehat{L}^{K,p,\sigma}(\widehat{\nu}) &:= K(1-p) \int_{\mathbb{R}} \widehat{\nu}(dx) b(x) \left[ \phi\left(\widehat{\nu} + \frac{\delta_x}{K}\right) - \phi(\widehat{\nu}) \right] \\ &\quad + Kp \int_{\mathbb{R}} \widehat{\nu}(dx) b(x) \int_{\mathbb{R}} m(x, h) dh \left[ \phi\left(\widehat{\nu} + \frac{\delta_{x+\sigma h}}{K}\right) - \phi(\widehat{\nu}) \right] \\ &\quad + K \int_{\mathbb{R}} \widehat{\nu}(dx) \left( d(x) + \int_{\mathbb{R}} \widehat{\nu}(dy) c(x, y) \right) \left[ \phi\left(\widehat{\nu} - \frac{\delta_x}{K}\right) - \phi(\widehat{\nu}) \right] \end{aligned}$$

The first term describes the births *without* mutation, the second term describes the births *with* mutation, and the third term describes the deaths by age or competition.

As in Section 1.1.3.2, we have performed numerical simulations of the process  $\widehat{\nu}^{K,p,\sigma}$  for a biological model adapted from KISDI [140]. For this biological model, the trait space is  $\mathcal{X} := [0, 4]$  (and not  $\mathbb{R}$  as previously) and the parameters are:

$$b(x) := 4 - x, \quad d(x) := 0, \quad c(x, y) := \frac{2}{K} \left( 1 - \frac{1}{1 + 1.2 \exp(-4(x - y))} \right) \quad (1.72)$$

and  $m(x, dh)$  is a Gaussian distribution  $\mathcal{N}(x, \sigma^2)$  conditional on the mutant trait remaining in  $\mathcal{X}$ .

The function  $c$  models an asymmetric competition in favour of high traits whereas the birth rate  $b$  favours the small traits. Therefore, we expect that evolution will select an optimal trait.

We performed simulations of KISDI’s model, starting with a monomorphic initial condition

of trait 1 and varying the parameters  $K, p$  and  $\sigma$ . Some of these simulations are presented in the following figures and show a wide variety of evolutionary behaviours. The upper part of these simulations describes the evolution of the distribution of the population traits while the lower part describes the evolution of the population size. The R-source of these illustrations, provided in Appendix B, is also based on the `IBMPopSim` package developed by GIORGI ET AL in [116].

In Figures 1.29 to 1.34 the same sets of parameters as in Figures 1.9 to 1.14 are tested. Some of these simulations are similar to the ones of Section 1.1.3.2. Note that, in Figures 1.30 and 1.31, a *diversification phenomenon* is highlighted: in Figure 1.30, once the population reaches the neighbourhood of the optimal trait close to 3.5, the population stabilises and then splits into two subpopulations concentrated around two distinct values of traits that co-evolve: this phenomenon is called *evolutionary branching*.

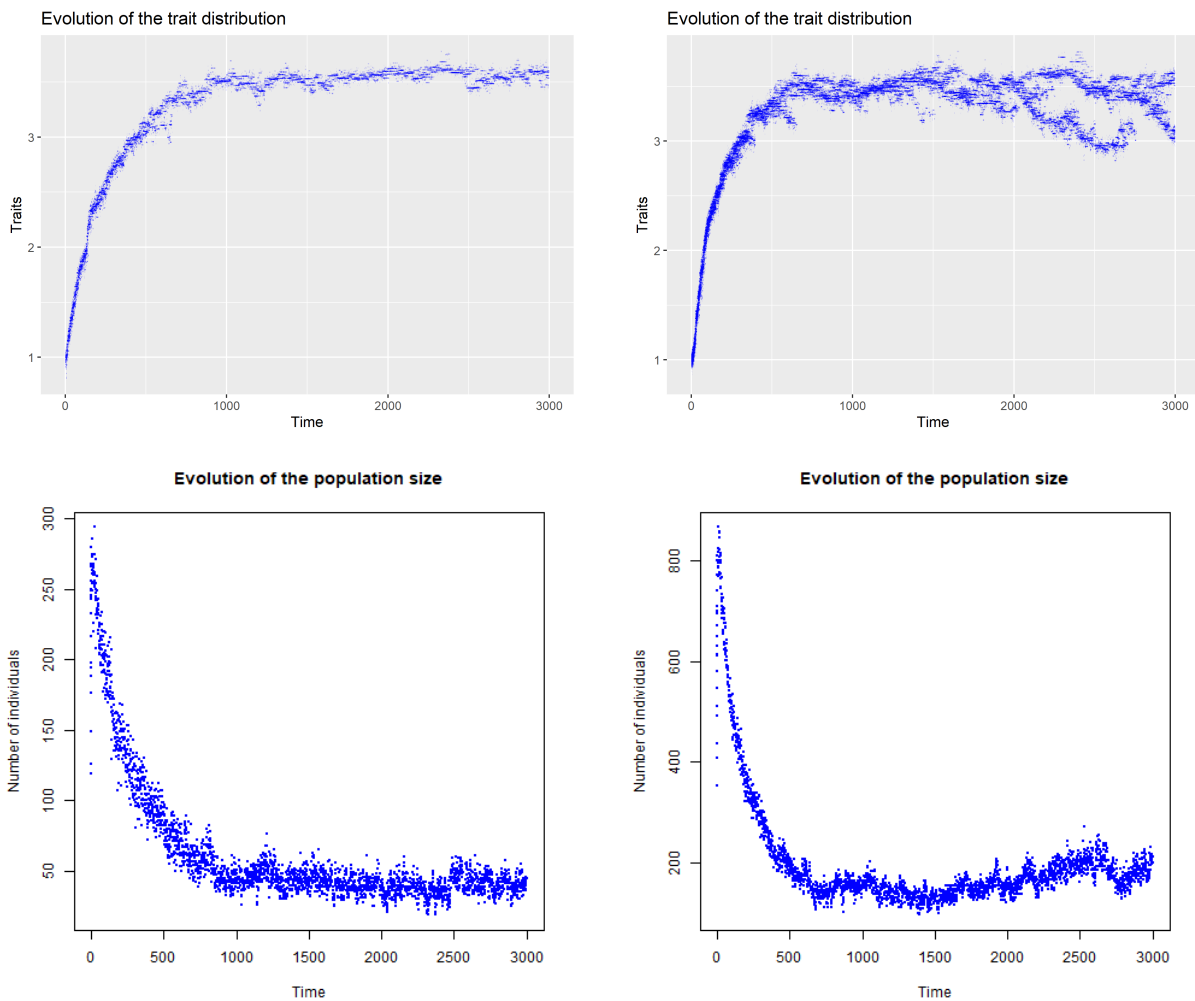


FIGURE 1.29 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 100, p = 0.1, \sigma = 0.03$ .

FIGURE 1.30 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 300, p = 0.1, \sigma = 0.03$ .

Figures 1.33 and 1.34 illustrate the following expected convergence result, from the IBM to

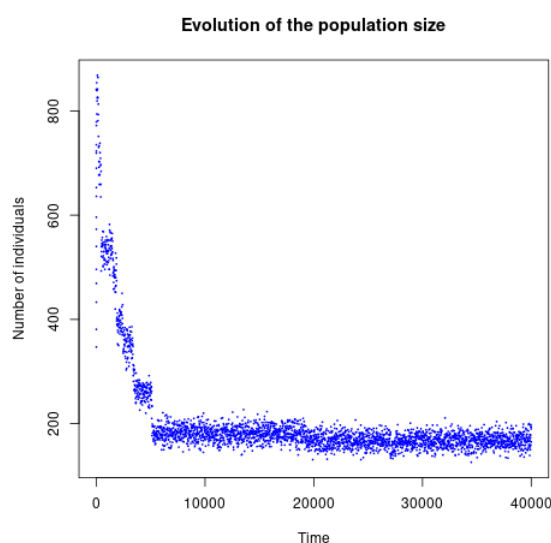
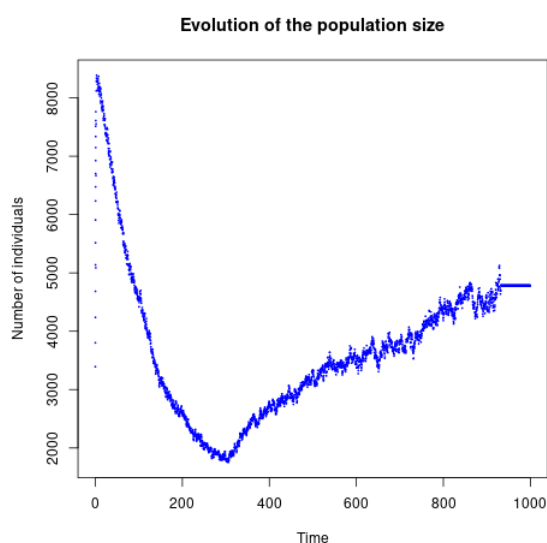
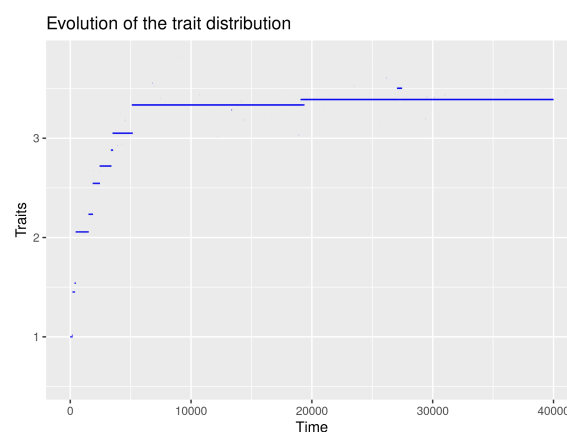
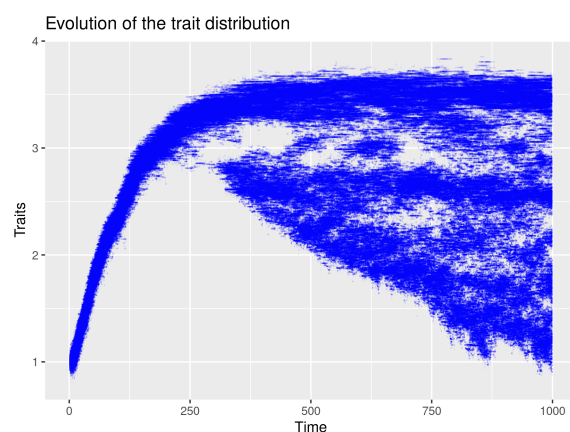


FIGURE 1.31 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 3000$ ,  $p = 0.1$ ,  $\sigma = 0.03$ .

FIGURE 1.32 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 300$ ,  $p = 0.00003$ ,  $\sigma = 0.2$ .

the CEAD, under an assumption of frequent mutations ( $p = 1$ ) and a double asymptotic of large population ( $K \rightarrow +\infty$ ) and small mutations ( $\sigma \rightarrow 0$ ).

**Conjecture 1.3.19.** *Assume that  $p = 1$ ,  $K \rightarrow +\infty$ ,  $\sigma \rightarrow 0$  and  $\sigma = \sigma(K)$  such that*

$$\exists \varepsilon > 0, \quad \forall C > 0, \quad K^{-C \log(K)} \ll \sigma \ll K^{-(2+\varepsilon)}.$$

*Under ad-hoc assumptions on  $b, d, c, m, \hat{\nu}_0$  and denoting  $\mathbf{1}$  the constant function equal to 1, we expect to the sequence of the mean trait process  $\left( \left( \langle \text{id}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle / \langle \mathbf{1}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle \right)_{t \geq 0} \right)_{K \in \mathbb{N}^*}$  converges when  $K \rightarrow +\infty$  and  $\sigma \rightarrow 0$  to the unique deterministic process solution of the CEAD*

$$\frac{dx(t)}{dt} = \partial_1 \text{Fit}(x(t), x(t)) \times \bar{n}_{x(t)} \int_{\mathbb{R}} h^2 m(x(t), h) dh.$$

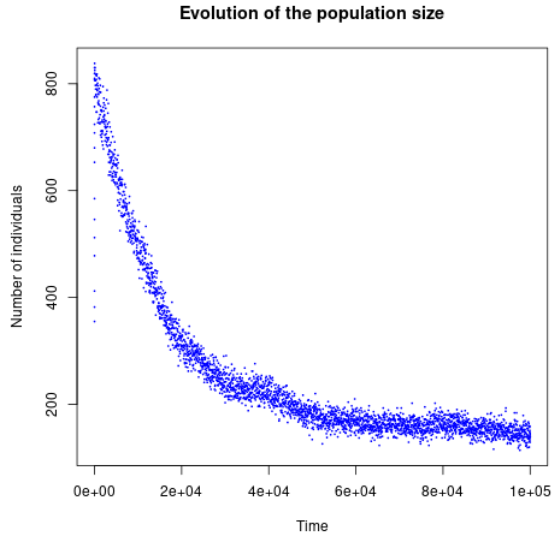
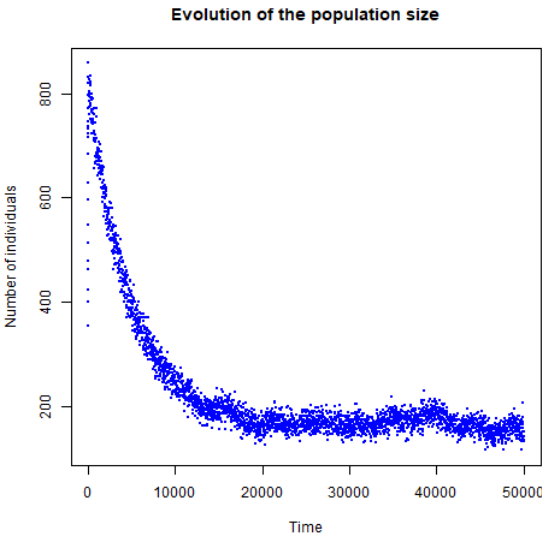
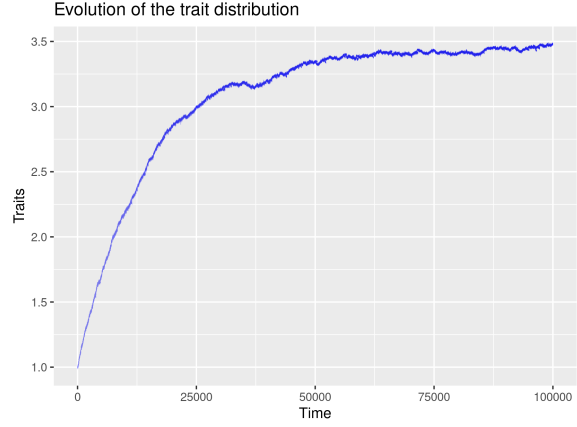
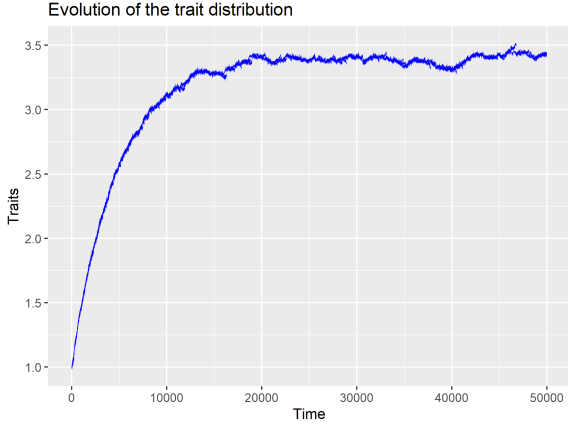


FIGURE 1.33 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.001$ .

FIGURE 1.34 – Numerical simulation of the trait distribution and population size of the microscopic model with parameters (1.72) and  $K = 300$ ,  $p = 1$ ,  $\sigma = 0.0005$ .

We expected that this result can be proved using a slow-fast analysis involving three evolutionary time scales as represented in Figure 1.35:

- The slow dynamics, with value in  $\mathbb{R}$ , corresponds to that of the mean trait defined by

$$\hat{z}_t^K := \frac{\langle \text{id}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle}{\langle \mathbf{1}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle},$$

- The fast dynamics, with value in  $\mathcal{M}_1^{c,K}(\mathbb{R}) := \left\{ \frac{\nu}{\langle \mathbf{1}, \nu \rangle} \mid \nu \in \mathcal{M}^K(\mathbb{R}), \langle \text{id}, \nu \rangle = 0 \right\}$ , corresponds to that of the centered and dilated distribution of traits defined by

$$\hat{\mu}_t^K := \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-\hat{z}_t^K} \right) \# \frac{\hat{\nu}_{t/K\sigma^2}^{K,1,\sigma}}{\langle \mathbf{1}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle} = \frac{1}{K \langle \mathbf{1}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle} \sum_{i=1}^K \langle \mathbf{1}, \hat{\nu}_{t/K\sigma^2}^{K,1,\sigma} \rangle \delta_{\frac{1}{\sigma\sqrt{K}}(x_i(t/K\sigma^2) - \hat{z}_t^K)}.$$



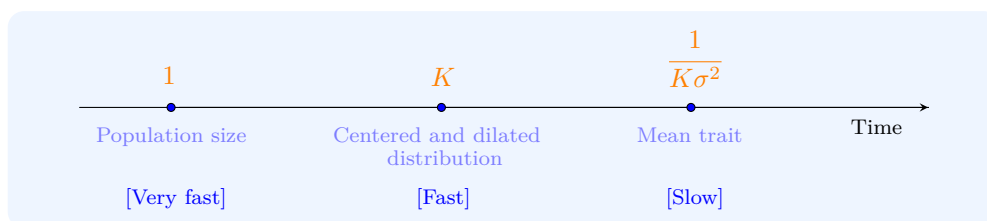


FIGURE 1.35 – Different time scales involved in our IBM expected in the original time scale of the process  $(\widehat{\nu}_t^{K,1,\sigma})_{t \geq 0}$ .

- The very fast dynamics, with value in  $\mathbb{R}_+$ , corresponds to that of the population size defined by

$$\langle \mathbf{1}, \widehat{\nu}_t^{K,1,\sigma} \rangle = \frac{N_t^{K,1,\sigma}}{K}.$$

Note that  $\widehat{\nu}_t^{K,1,\sigma} = \langle \mathbf{1}, \widehat{\nu}_t^{K,1,\sigma} \rangle \left( \tau_{z_{tK\sigma^2}^K} \circ h_{\sigma\sqrt{K}} \right) \# \widehat{\mu}_{tK\sigma^2}^K$ . Note also when the population size is assumed to be fixed at its equilibrium value, the slow-fast asymptotic analysis involves only two time scales. This analysis corresponds precisely to the approach developed in this thesis. In this sense, the model of Section 1.1.3.2 can be seen as a toy model for this one. Our strategy for proving Conjecture 1.3.19 will therefore be to control the deviations of the population size from its equilibrium so that we can reduce to the slow-fast analysis developed in this thesis.

### 1.3.3.2 Other strategies to resolve the controversy related to the rare mutations assumption

As seen in Section 1.1.4.6, the stochastic approach initiated by CHAMPAGNAT [37] and completed by CHAMPAGNAT ET AL [47] is criticised because it is based on an unrealistic biological assumption of too rare mutations.

We propose an alternative approach to solve this controversy based on the observation that most of the mutations have deleterious effects and increase the mortality of individuals, or have nearly neutral effects. It would be more realistic to assume that only *advantageous* mutations are rare. Now, the TSS precisely describes an evolutionary dynamic governed only by advantageous mutations. One can therefore seek to extend the TSS to situations where a large number of neutral or disadvantageous mutations dictate the DIRAC masses of the TSS to be replaced by spread out distributions, and where these distributions evolve by jumps over a long time scale of advantageous mutations. Two possibilities can be considered:

- Adapt the TSS model by imposing that most of the mutations occur in a fixed direction (say  $< 0$ ) of a 1-dimensional type space, and choosing ecological parameters such that selection drives the population in the opposite direction (say  $> 0$ ). This approach requires to prescribe directions of positive fitness. Since fitness is a complex function of ecological parameters, this cannot be done easily in general, but we can restrict to a particular toy models when existing tools should formally apply replacing monomorphic populations at equilibrium with equilibrium population distribution on trait sets of the form  $]-\infty, x^*]$  where  $x^*$  is the dominant trait and all other traits are disadvantaged.

- The distinction between advantageous and disadvantageous mutations is however delicate at the phenotypic level, so it is natural to consider an extension of the IBM of Section 1.1.3.2 including genetics and genotype to phenotype map (see Figure 1.1). For example, we can consider a countable sequence of genes on a chromosome with *additive effects* on phenotypes with weight converging to 0. Only mutations on the first genes will then have a great effect on the phenotypes. Note that, if we assume that most of the mutations affect genes with small effects on phenotype, then most of the mutations are neutral and a “significant” mutation occurs on a long time scale. Here, the study of adaptive dynamics is harder, but might give significant biological insights on the time scale of evolution and how diversity in a species is related to genetic parameters.

The study of the CEAD in both cases is a first step, for which existing methods are expected to apply. The study of evolutionary branching is more complicated since it both involves modelling and methodological difficulties, so this problem will be addressed in a longer term.

## Chapter 2

# Existence, uniqueness and ergodicity for the centered Fleming-Viot process

This chapter corresponds to the publication [41] “Existence, uniqueness and ergodicity for the centered FLEMING-VIOT process”, written in collaboration with NICOLAS CHAMPAGNAT and to appear in *Stochastic Processes and Their Applications*.

### Contents

---

<b>2.1</b>	<b>Introduction</b>	<b>86</b>
<b>2.2</b>	<b>Existence for the centered Fleming-Viot process</b>	<b>89</b>
2.2.1	Centered FLEMING-VIOT martingale problem and main result	89
2.2.2	Sketch of proof of Theorem 2.2.3	91
2.2.2.1	Framework and objective of the proof	91
2.2.2.2	Outline of the proof	92
2.2.3	Some extensions to the centered FLEMING-VIOT martingale problem	93
2.2.3.1	Extensions to multiple variables	93
2.2.3.2	Extension to polynomials in $\mu$	95
2.2.4	Some properties of the centered FLEMING-VIOT process	98
2.2.4.1	MARKOV’s property	98
2.2.4.2	Moments and some martingales	98
2.2.4.3	Compact support	101
<b>2.3</b>	<b>Uniqueness for the centered Fleming-Viot process</b>	<b>102</b>
2.3.1	Main result	102
2.3.2	Notations and outline of the uniqueness proof	102
2.3.3	Proof of Theorem 2.3.1 from Theorem 2.3.4 and Lemma 2.3.5	105
<b>2.4</b>	<b>Ergodicity for the centered Fleming-Viot process</b>	<b>106</b>
2.4.1	MORAN’s models and FLEMING-VIOT’s processes	106
2.4.2	Backward construction of MORAN’s process and KINGMAN’s coalescent	110
2.4.3	Centered variables and centering effects	112
2.4.4	Coupling arguments with two distinct initial conditions	113
2.4.5	Proof of Theorem 2.4.1	115
2.4.6	Characterisation of the invariant probability measure	115
2.4.6.1	Construction of exchangeable random variables allowing to characterise $\pi_N$ and $\pi$	116

2.4.6.2	Characterisation of the invariant probability measure $\pi$	120
<b>2.5</b>	<b>Proof of Theorem 2.2.3</b>	<b>122</b>
2.5.1	DOOB's semi-martingale decomposition of $\langle \text{id}, Y_t - Y_s \rangle, s \leq t$	122
2.5.2	DOOB's semi-martingale decomposition of $\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_s \rangle}, Y_t - Y_s \rangle, s \leq t, j \in \{0, 1, 2\}$	123
2.5.3	Expressions of the terms of (2.11) involving $F'$	124
2.5.3.1	Decomposition and study of $(\mathbf{A})_i^1$	125
2.5.3.2	Martingale contribution of $(\mathbf{A})_i^2$ and $(\mathbf{A})_i^3$	125
2.5.3.3	Contributions of $(\mathbf{A})_i^4$ and $(\mathbf{A})_i^5$	126
2.5.3.4	Study of the error term $(\mathbf{A})_i^6$	127
2.5.4	Expressions of terms of (2.11) involving $F''$	127
2.5.5	Error terms	128
2.5.6	Extension to test functions $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ and $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$	129
2.5.7	$L^2$ -martingale and quadratic variation	129
2.5.8	Technical result for Lemma 2.5.2	130
2.5.9	Proof of Lemma 2.5.2	132
<b>2.6</b>	<b>Proof of the results of Section 2.3</b>	<b>133</b>
2.6.1	Study of a semi-group	133
2.6.1.1	Construction of the semi-group	134
2.6.1.2	MILD formulation	140
2.6.2	Proof of Lemma 2.3.5	141
2.6.2.1	Proof that Lemma 2.6.4 implies Lemma 2.3.5	143
2.6.2.2	Proof of Lemma 2.6.4	143
2.6.3	Proof of Theorem 2.3.4	144

---

**Abstract.** Motivated by questions of ergodicity for shift invariant FLEMING-VIOT process, we consider the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$  defined by  $Z_t := \tau_{-\langle \text{id}, Y_t \rangle} \# Y_t$ , where  $(Y_t)_{t \geq 0}$  is the original FLEMING-VIOT process. Our goal is to characterise the centered FLEMING-VIOT process with a martingale problem. To establish the existence of a solution to this martingale problem, we exploit the original FLEMING-VIOT martingale problem and asymptotic expansions. The proof of uniqueness is based on a weakened version of the duality method, allowing us to prove uniqueness for initial conditions admitting finite moments. We also provide counter examples showing that our approach based on the duality method cannot be expected to give uniqueness for more general initial conditions. Finally, we establish ergodicity properties with exponential convergence in total variation for the centered FLEMING-VIOT process and characterise the invariant measure.

## 2.1 Introduction

FLEMING and VIOT have introduced in [99] a probability-measure-valued stochastic process modelling the dynamics of the distribution of allelic frequencies in a *selectively neutral* genetic population as influenced by mutation and random genetic drift: the original FLEMING-VIOT process. The initial model of [99] was progressively enriched with further mechanisms of Darwinian evolution: selection [99, 89, 91, 60, 86], recombination [93, 91] or the effect of an environment [124]. FLEMING and VIOT characterise in [99] the law of their process as a solution of a STROOCK-VARADHAN measure-valued martingale problem [207] both in the selective neutral case and the case with selection. To obtain the existence of such solution on a compact metric space, their

method is based on discretisation of the mutation operator and tightness arguments. An alternative approach is used, in the studies [177, 178, 139, 135, 205] based on the OTHA-KIMURA model [188, 189] and in the references [99, 61, 91, 85] based on its continuous-time version: the MORAN model (also called *continuous-state stepwise mutation model*). If we denote by  $N$  the population size, these authors construct a particle process whose limiting behaviour is analysed under the assumptions that the mutation step is proportional to  $1/\sqrt{N}$  and on the time scale  $(Nt)_{t \geq 0}$ . In [124], another particle process, based on the look-down construction [78] is used to show the existence of the FLEMING-VIOT process in a random environment. This look-down construction also allows to analyse sample path properties of the process and has been used in numerous references since then, such as [80] [85, Chapter 5].

In [99], uniqueness of the solution to the FLEMING-VIOT martingale problem in the selectively neutral case, is proved using uniqueness of moments of certain finite-dimensional distributions and arguments on semigroup. However, in the case where natural selection acts, the previous method fails, but the result can be obtained from a version of the CAMERON-MARTIN-GIRSANOV formula [60, Chapter 10] [59, Theorem 5.1]. See also [92] for an application of this method in the case of unbounded selection function. In most references such as [85, 61, 88, 89, 91, 92], under a variety of assumptions, the *duality method* [94, Proposition 4.4.7] is used to prove the uniqueness of the FLEMING-VIOT process. The idea is to relate the distribution of the original process with that of a simple process, called *dual process*. This leads to a duality relation which ensures that two solutions to the martingale problem have the same 1-dimensional marginal laws. Uniqueness of the solution to the martingale problem then follows from MARKOV's property [94, Theorem 4.4.2]. Other methods are used in some references: [62] makes use of resolvent estimates; [191, 190] prove existence and uniqueness of FLEMING-VIOT processes with *unbounded selection* intensity functions by using DIRICHLET's forms.

Questions of ergodicity of the FLEMING-VIOT process were also the subject of many works. Let  $E$  be a Polish space and  $\mathcal{B}(E)$  the BOREL  $\sigma$ -field on  $E$ . Let us recall from [91, Section 5] that an  $E$ -valued MARKOV process  $(Z_t)_{t \geq 0}$  with a unique stationary distribution  $\pi$  is *weakly ergodic* if for all bounded continuous functions  $f$  on  $E$ , for all initial condition  $x_0 \in E$ ,

$$\lim_{t \rightarrow +\infty} \mathbb{E}_{x_0} (f(Z_t)) = \int_E f(x) \pi(dx) \quad (2.1)$$

and *strongly ergodic* if

$$\lim_{t \rightarrow +\infty} \sup_{B \in \mathcal{B}(E)} |\mathbb{P}_{x_0}(Z_t \in B) - \pi(B)| = 0, \quad x_0 \in E. \quad (2.2)$$

If  $E$  is compact, for mutation operators  $A$  whose closure generates a FELLER semigroup on the space of continuous functions and such that there is a unique probability measure  $\nu_0$  on  $E$  satisfying  $\int_E Af(x) \nu_0(dx) = 0$ , some ergodicity results for the FLEMING-VIOT process are obtained in [91]. More precisely, in the selectively neutral case and without recombination, a simple proof of weak ergodicity of the FLEMING-VIOT process is given using duality arguments whereas coupling arguments provide an approach to strong ergodicity. These results were extended in [90] to models with recombination and in [92] to models with unbounded selection, with the additional tool of DAWSON's GIRSANOV-type formula for strong ergodicity. In the special case where the mutation operator of the FLEMING-VIOT process has the form

$$Af(x) = \frac{\theta}{2} \int_E (f(y) - f(x)) P(x, dy), \quad \theta \in (0, +\infty), \quad f \in \mathcal{D}(A), \quad (2.3)$$

it is proved in [91] that the FLEMING-VIOT process has a reversible stationary distribution if  $P(x, dy) = \nu(dy)$  for some probability measure  $\nu$  on  $E$  (see [157] for a converse result). For the

mutation operator (2.3), it is proved in [88, 89] and [60, Theorem 8.2.1] that the FLEMING-VIOT process is purely atomic for every time, in other words the solutions of the martingale problem take values in the set of purely atomic probability measures. In [93], the ergodicity result of [90] was extended to the weak atomic topology.

However, if we consider the case where the mutation operator is the Laplacian on  $\mathbb{R}^d$ , there exists no stationary distribution [91, 157], [94, Problem 11 p.450]. Instead the process exhibits a wandering phenomenon [61, 21]. Nevertheless, [205, 91] considered the FLEMING-VIOT process shifted by minus its empirical mean and established existence of a unique invariant measure and weak ergodicity for this process using moment and duality arguments. More precisely in [91], thanks to some estimates of the original FLEMING-VIOT dual process and the finiteness of all moments of the FLEMING-VIOT process shifted by minus its empirical mean for any time  $t$ , the authors obtain an expression for these in the asymptotic  $t \rightarrow +\infty$ . Then, by tightness arguments and characterisation of the limit, the result follows. In [205], an analogous approach is used for the continuous-state stepwise mutation model.

In this chapter we are interested in the FLEMING-VIOT process shifted by minus its empirical mean, which we call *centered FLEMING-VIOT process*. As in previous works it is natural to ask questions of existence, uniqueness and ergodicity when the mutation operator is the Laplacian on  $\mathbb{R}$ . Moreover, the study of this process was motivated by biological questions in adaptive dynamics. The theory of adaptative dynamics [171] is based on biological assumptions of *rare* and *small mutations* and of *large population* under which an ODE approximating the population evolutionary dynamics, the *Canonical Equation of Adaptive Dynamics* (CEAD) was proposed [72]. Two mathematical approaches were developed to give a proper mathematical justification of this theory: a *deterministic* one [75, 194, 160], and a *stochastic* one [37, 47, 43]. Despite their success, the proposed approaches are criticised by biologists [212, 195]. Among the biological assumptions of adaptive dynamics, the assumption of rare mutations is the most criticised as unrealistic. In order to solve this problem, we propose to apply an asymptotic of small mutations and large population, but *frequent mutations*. After conveniently scaling the population state, this leads to a *slow-fast dynamics* [192, 144], where the fast dynamics appears to be given by a discrete version of the centered FLEMING-VIOT process [40]. This explains why we are interested in ergodicity properties of such processes.

To establish the existence of the centered FLEMING-VIOT process, we characterise it as a solution of a measure-valued martingale problem that we called the *centered FLEMING-VIOT martingale problem*. Our method is to exploit the original FLEMING-VIOT martingale problem and asymptotic expansions. An additional difficulty occurs in our case since we need to apply the original FLEMING-VIOT martingale problem to *predictable test functions*. This requires to extend the martingale problem to such test functions using *regular conditional probabilities*. The proof of uniqueness of the solution of the centered FLEMING-VIOT martingale problem is based on duality methods as in the previous works. However, additional difficulties occur in our case since bounds on the dual process are much harder to obtain and the duality identity can only be proved in a weakened version. In particular, our uniqueness result only holds for initial conditions admitting finite moments. We also provide a counter example showing that our uniqueness result is optimal in the sense that we cannot expect to obtain uniqueness for more general initial conditions using the duality approach. Finally, we obtain strong ergodicity properties of the centered FLEMING-VIOT process that extend the weak ergodicity results obtained in [205, 91]. To this aim, we construct the centered version of the MORAN process and we prove that it converges in law to the centered FLEMING-VIOT process. Exploiting the relationship between the MORAN model and the KINGMAN coalescent, we obtain a result of exponential ergodicity in total variation for the centered MORAN model uniformly in the number of particles. This result

is propagated to the centered FLEMING-VIOT process by coupling arguments. Using another strategy proposed by [205, 91], based on the DONNELLY-KURTZ modified look-down [80] we give a characterisation of the unique invariant measure of the centered FLEMING-VIOT process.

This chapter is organised as follows. In Section 2.2 we define the martingale problem for the centered FLEMING-VIOT process and establish an existence result. We give also some equivalent extensions to the centered FLEMING-VIOT martingale problem and some properties of the centered FLEMING-VIOT process. In Section 2.3 we prove uniqueness to the centered FLEMING-VIOT martingale problem for initial conditions admitting finite moments and we discuss this assumption. In Section 2.4, we establish exponential convergence in total variation for the centered FLEMING-VIOT process to its unique invariant measure and provide a characterisation of this measure based on the DONNELLY-KURTZ modified look-down. Finally in Sections 2.5 and 2.6, we prove respectively the main results of existence and uniqueness of the solution of the centered FLEMING-VIOT martingale problem. Appendix A gathers technical lemmas for the existence proof.

## 2.2 Existence for the centered Fleming-Viot process

In this section, our aim is to define the martingale problem for the centered FLEMING-VIOT process and to establish an existence result. This result is stated in Section 2.2.1. Then, we give in Section 2.2.2, the framework and ideas of the proof. In Section 2.2.3, we give some equivalent extensions to the centered FLEMING-VIOT martingale problem with different sets of test functions. We end this section by giving some interesting results about the centered FLEMING-VIOT process: it satisfies the MARKOV property (Section 2.2.4.1), admits finite moments (Section 2.2.4.2) and has compact support (Section 2.2.4.3).

### 2.2.1 Centered Fleming-Viot martingale problem and main result

The original FLEMING-VIOT process is a measure-valued diffusion in  $\mathcal{M}_1(\mathbb{R})$ , the set of probability measures on  $\mathbb{R}$  with respect to the BOREL  $\sigma$ -field  $\mathcal{B}(\mathbb{R})$ , which is endowed with the topology of weak convergence making it a Polish space [20]. If  $\mathcal{I}$  is an interval of  $\mathbb{R}$ , then for all  $\ell \in \mathbb{N}$ , we denote by  $\mathcal{C}^\ell(\mathcal{I}, \mathbb{R})$  the space of functions of class  $\mathcal{C}^\ell$  from  $\mathcal{I}$  to  $\mathbb{R}$ . For  $\ell \in \mathbb{N}$ , we denote by  $\mathcal{C}_b^\ell(\mathbb{R}, \mathbb{R})$  the space of real bounded functions of class  $\mathcal{C}^\ell(\mathbb{R}, \mathbb{R})$  with bounded derivatives. We consider the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$  where

$$\Omega := \mathcal{C}^0([0, +\infty), \mathcal{M}_1(\mathbb{R}))$$

is endowed with the SKOROHOD topology,  $\mathcal{F}$  is the associated BOREL  $\sigma$ -field and  $(\mathcal{F}_t)_{t \geq 0}$  is the canonical filtration. The centered FLEMING-VIOT process is a measure-valued diffusion in

$$\mathcal{M}_1^{c,2}(\mathbb{R}) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \mid \int_{\mathbb{R}} |x|^2 \mu(dx) < \infty, \int_{\mathbb{R}} x \mu(dx) = 0 \right\}$$

which is endowed with the trace of the topology of weak convergence on  $\mathcal{M}_1(\mathbb{R})$ . We consider the filtered probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0})$  where

$$\tilde{\Omega} := \left\{ X \in \mathcal{C}^0([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R})) \mid \forall T > 0, \sup_{0 \leq t \leq T} \int_{\mathbb{R}} |x|^2 X_t(dx) < \infty \right\}$$

is endowed with the trace of the SKOROHOD topology on  $\Omega$ ,  $\tilde{\mathcal{F}}$  is the trace of the  $\sigma$ -field  $\mathcal{F}$  and  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  is the trace of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . We introduce several notations that we use

repeatedly in the sequel. For a measurable real bounded function  $f$  and a measure  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote  $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x) \nu(dx)$ . We denote by  $\text{id}$  the identity function. We denote  $\mathbb{N} := \{0, 1, 2, \dots\}$  and  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ . For all  $k \in \mathbb{N}$  and for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we also denote

$$M_k(\nu) := \int_{\mathbb{R}} |x - \langle \text{id}, \nu \rangle|^k \nu(dx).$$

In particular, for all  $k \in \mathbb{N}$  and for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,  $M_k(\mu) = \langle |\text{id}|^k, \mu \rangle$ . For all  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  we denote for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ ,  $F_g(\nu) := F(\langle g, \nu \rangle)$ .

**Definition 2.2.1.** A probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  is said to solve the centered FLEMING-VIOT martingale problem with resampling rate  $\lambda \in (0, +\infty)$  and with initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  if the canonical process  $(X_t)_{t \geq 0}$  on  $\tilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$  and for each  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\widehat{M}_t^F(g) := F_g(X_t) - F_g(X_0) - \int_0^t \mathcal{L}_{\text{FVc}} F_g(X_s) ds \quad (2.4)$$

with for all  $\varpi \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,

$$\begin{aligned} \mathcal{L}_{\text{FVc}} F_g(\varpi) &:= F'(\langle g, \varpi \rangle) \left( \left\langle \frac{g''}{2}, \varpi \right\rangle + \lambda [\langle g'', \varpi \rangle M_2(\varpi) - 2 \langle g' \times \text{id}, \varpi \rangle] \right) \\ &+ \lambda F''(\langle g, \varpi \rangle) \left( \langle g^2, \varpi \rangle - \langle g, \varpi \rangle^2 + \langle g', \varpi \rangle^2 M_2(\varpi) - 2 \langle g', \varpi \rangle \langle g \times \text{id}, \varpi \rangle \right) \end{aligned} \quad (2.5)$$

is a continuous  $\mathbb{P}_\mu$ -martingale in  $L^2(\tilde{\Omega})$  with quadratic variation process

$$\begin{aligned} \langle \widehat{M}^F(g) \rangle_t &= 2\lambda \int_0^t [F'(\langle g, X_s \rangle)]^2 [\langle g^2, X_s \rangle - \langle g, X_s \rangle^2 \\ &+ \langle g', X_s \rangle^2 M_2(X_s) - 2 \langle g', X_s \rangle \langle g \times \text{id}, X_s \rangle] ds. \end{aligned} \quad (2.6)$$

Note that a major difficulty comes from the presence of terms in  $M_2(\mu)$  in (2.5). We will see in particular that this leads to the creation of particles in the dual process. Note also that the operator given by (2.5) is the generator (in the sense of DYNKIN) of the  $\mathcal{M}_1^{c,2}(\mathbb{R})$ -valued centered FLEMING-VIOT process. We recall that the probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the *original FLEMING-VIOT martingale problem* with resampling rate  $\lambda$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$  if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$  and for each  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\begin{aligned} M_t^F(g) &:= F(\langle g, Y_t \rangle) - F(\langle g, Y_0 \rangle) - \int_0^t F'(\langle g, Y_s \rangle) \left\langle \frac{g''}{2}, Y_s \right\rangle ds \\ &- \lambda \int_0^t F''(\langle g, Y_s \rangle) [\langle g^2, Y_s \rangle - \langle g, Y_s \rangle^2] ds \end{aligned} \quad (2.7)$$

is a continuous square integrable  $\mathbb{P}_\nu^{\text{FV}}$ -martingale whose martingale bracket satisfies for all  $G, H \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and for all  $g, h \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\langle M^G(g), M^H(h) \rangle_t = 2\lambda \int_0^t G'(\langle g, Y_s \rangle) H'(\langle h, Y_s \rangle) [\langle gh, Y_s \rangle - \langle g, Y_s \rangle \langle h, Y_s \rangle] ds. \quad (2.8)$$



In the population genetics literature, the terms involving the first order derivative  $F'$  describe the effect of mutations whereas the ones involving the second order derivative  $F''$  describe the effect of the random genetic drift. As the original FLEMING-VIOT process is obtained as a scaling limit of a population genetics model, called MORAN's model (see Section 2.4.1), the different terms in (2.7) can be interpreted as in the population genetics context. It is well-known that, for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , there exists a unique probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  satisfying the previous martingale problem (2.7) [99, Theorem 3].

**Remark 2.2.2.** *The additional terms in the martingale problem (2.4) with respect to the martingale problem (2.7) describe the impact of centering and ensure that at all times the centered FLEMING-VIOT process remains  $\mathcal{M}_1^{c,2}(\mathbb{R})$ -valued.*

Let us define by  $\tau_\alpha$  the translation operator of vector  $\alpha \in \mathbb{R}$ . For all  $\mu \in \mathcal{M}_1(\mathbb{R})$ , for all  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\tau_\alpha \# \mu(A) := \mu(\tau_\alpha^{-1}(A)) = \mu\left(\left\{x - \alpha \mid x \in A\right\}\right),$$

where  $\#$  is the *pushforward operator* ( $\tau_\alpha \# \mu$  is the *pushforward measure* of  $\mu$  by  $\tau_\alpha$ ). For all  $k \in \mathbb{N}$ , we denote

$$\mathcal{M}_1^k(\mathbb{R}) := \left\{ \nu \in \mathcal{M}_1(\mathbb{R}) \mid \langle |\text{id}|^k, \nu \rangle < \infty \right\}.$$

The main result of this section is the following:

**Theorem 2.2.3.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , there exists a probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  satisfying the martingale problem of Definition 2.2.1, given by the law of the process  $(Z_t)_{t \geq 0}$  defined by*

$$Z_t := \tau_{-\langle \text{id}, Y_t \rangle} \# Y_t := Y_t(\cdot + \langle \text{id}, Y_t \rangle), \quad t \geq 0 \quad (2.9)$$

where  $(Y_t)_{t \geq 0}$  is the original FLEMING-VIOT process.

Note that for all  $\mu \in \mathcal{M}_1^1(\mathbb{R})$ ,  $\langle \text{id}, \mu \rangle$  is the mean value of  $\mu$  and so  $\langle \text{id}, Z_t \rangle = 0$  for all  $t \geq 0$ . As a result, the process  $(Z_t)_{t \geq 0}$  corresponds to the original FLEMING-VIOT process centered by its mean value, hence its name of *centered FLEMING-VIOT process*.

## 2.2.2 Sketch of proof of Theorem 2.2.3

The proof is based on the original FLEMING-VIOT martingale problem (2.7).

### 2.2.2.1 Framework and objective of the proof

Let  $\nu \in \mathcal{M}_1^2(\mathbb{R})$  and  $\mathbb{P}_\nu^{\text{FV}}$  the unique solution to the original FLEMING-VIOT martingale problem (2.7). As the support of the original FLEMING-VIOT process is compact at all positive times  $\mathbb{P}_\nu^{\text{FV}}$ -a.s. [159],  $\mathbb{P}_\nu^{\text{FV}}(\mathcal{C}^0([0, +\infty), \mathcal{M}_1^2(\mathbb{R}))) = 1$ . Moreover, as  $t \mapsto \langle \text{id}, Y_t \rangle$  is continuous  $\mathbb{P}_\nu^{\text{FV}}$ -a.s. (see Lemma 2.5.1 (2)), we deduce that, for all  $t \geq 0$ ,  $Z_t$  given by (2.9), is well defined and is a random variable on  $\Omega$ .

When the dependency of  $Y$  and  $Z$  on the initial condition  $\nu$  of  $Y$  is important, we shall use the notation  $(Y_t^\nu)_{t \geq 0}$  and  $(Z_t^\nu)_{t \geq 0}$  instead of  $(Y_t)_{t \geq 0}$  and  $(Z_t)_{t \geq 0}$ . Our goal is to prove that the law of the process  $(Z_t)_{t \geq 0}$  denoted by  $\mathbb{P}_{\tau_{-\langle \text{id}, \nu \rangle} \# \nu}^{\text{FVc}}$  solves the martingale problem of Definition 2.2.1 with initial condition  $\tau_{-\langle \text{id}, \nu \rangle} \# \nu$ . Note that the notation  $\mathbb{P}_{\tau_{-\langle \text{id}, \nu \rangle} \# \nu}^{\text{FVc}}$  is justified because the original FLEMING-VIOT process is invariant by translation:

**Proposition 2.2.4.** *Let  $\nu \in \mathcal{M}_1^1(\mathbb{R})$  and  $a \in \mathbb{R}$ . Then, the law of  $Z_t^\nu$  is the same as the law of  $Z_t^{\tau_a \# \nu}$ .*

*Proof.* By translation invariance of the original FLEMING-VIOT process, the process  $(\tau_{-a} \# Y_t^{\tau_a \# \nu})_{t \geq 0}$  has the same law as the process  $(Y_t^\nu)_{t \geq 0}$ . Now,

$$Z_t^{\tau_a \# \nu} = \tau_{-\langle \text{id}, Y_t^{\tau_a \# \nu} \rangle} \# Y_t^{\tau_a \# \nu} = \tau_{-\langle \text{id}, \tau_{-a} \# Y_t^{\tau_a \# \nu} \rangle} \# (\tau_{-a} \# Y_t^{\tau_a \# \nu}).$$

Thus,  $(Z_t^{\tau_a \# \nu})_{t \geq 0}$  has the same law as  $(\tau_{-\langle \text{id}, Y_t^\nu \rangle} \# Y_t^\nu)_{t \geq 0} = (Z_t^\nu)_{t \geq 0}$ .  $\square$

### 2.2.2.2 Outline of the proof

We restrict to the time interval  $[0, T]$  for  $T > 0$  arbitrary. By standard arguments, it is sufficient to prove that, for all  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$F_g(Z_t) - F_g(Z_0) - \int_0^t \mathcal{L}_{\text{FVc}} F_g(Z_s) ds \quad (2.10)$$

is a continuous  $\mathbb{P}_\nu^{\text{FV}}$ -martingale,  $\nu \in \mathcal{M}_1^2(\mathbb{R})$  where  $\mathcal{L}_{\text{FVc}}$  is given by (2.5). We start by assuming  $F, g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$  and we seek for the DOOB's semi-martingale decomposition of

$$F_g(Z_t) := F(\langle g, Z_t \rangle) = F(\langle g \circ \tau_{-\langle \text{id}, Y_t \rangle}, Y_t \rangle),$$

using the original FLEMING-VIOT martingale problem (2.7). However,  $F(\langle g \circ \tau_{-\langle \text{id}, Y_t \rangle}, Y_t \rangle)$  does not take the form  $H(\langle h, Y_t \rangle)$  with *deterministic*  $h$ . Therefore, we cannot apply (2.7) directly. To get over this difficulty, we consider for  $t \in [0, T]$ , an increasing sequence  $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = T$  of subdivisions of  $[0, T]$  whose mesh tends to 0. We can observe that

$$\begin{aligned} F_g(Z_t) - F_g(Z_0) &= \sum_{i=0}^{p_n-1} \{F_g(Z_{t_{i+1}^n \wedge t}) - F_g(Z_{t_i^n \wedge t})\} \\ &= \sum_{i=1}^{p_n-1} \left\{ F \left( \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_{i+1}^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} \right\rangle \right) - F \left( \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle \right) \right\}. \end{aligned}$$

Using asymptotic expansions (see Lemma A.1.1 of Appendix A with  $p = 1$ ) of the terms in the previous sum, we prove that

$$\begin{aligned} F_g(Z_t) - F_g(Z_0) &= \sum_{i=0}^{p_n-1} \left\{ (\mathbf{A})_i + (\mathbf{B})_i + O \left( \left| \langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle \right|^3 \right) \right. \\ &\quad \left. + O \left( \sum_{k=0}^2 \left| \langle g^{(k)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle \right|^3 \right) \right\}, \end{aligned} \quad (2.11)$$

where

$$\begin{aligned}
 (\mathbf{A})_i &= F' \left( \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle \right) \left\{ \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_{i+1}^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \right. \\
 &\quad \left. - \langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle \left[ \left\langle g' \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle + \left\langle g' \circ \tau_{-\langle \text{id}, Y_{t_{i+1}^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \right] \right. \\
 &\quad \left. + \frac{1}{2} \langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle^2 \left\langle g'' \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle \right\}, \tag{2.12}
 \end{aligned}$$

$$\begin{aligned}
 (\mathbf{B})_i &= \frac{F'' \left( \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle \right)}{2} \left\{ \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_{i+1}^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle^2 \right. \\
 &\quad \left. + \langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle^2 \left\langle g' \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle^2 \right. \\
 &\quad \left. - 2 \left\langle g \circ \tau_{-\langle \text{id}, Y_{t_{i+1}^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle \left\langle g' \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_i^n \wedge t} \right\rangle \right\}, \tag{2.13}
 \end{aligned}$$

and where  $g^{(j)}$ ,  $j \in \{0, 1, 2\}$ , denotes the  $j^{\text{th}}$  derivative of  $g$ . Note that the proposed decomposition in (2.11) is intended such that each of the terms in (2.12) and (2.13) is either  $\mathcal{F}_{t_i^n \wedge t}$ -adapted or they exhibit increments of  $(Y_t)_{t \geq 0}$  between  $t_i^n \wedge t$  and  $t_{i+1}^n \wedge t$ . Several steps are described in Section 2.5 to obtain the semi-martingale decomposition of each term involving increments of  $Y_t$  of the previous sum. Note that the last terms of  $(\mathbf{A})_i$  and  $(\mathbf{B})_i$  bring out the terms in  $M_2(\mu)$ . By making the step of the subdivision tend towards 0, we obtain the expected result. By density arguments, the martingale problem (2.7) satisfied by  $F, g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$  is extended, in Section 2.5.6, to the case where  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . Once we have proved that  $\widehat{M}^F(g)$  is a martingale, using ITÔ's formula and the martingale problem (2.4) with a localisation sequence, we deduce the value of  $\langle \widehat{M}^F(g) \rangle_t$  and by FATOU's lemma that  $\widehat{M}_t^F(g) \in L^2(\widetilde{\Omega})$ .

## 2.2.3 Some extensions to the centered Fleming-Viot martingale problem

Our goal is now to give some extensions to the martingale problem (2.4) which will be useful to compute the martingale bracket between two martingales of the form (2.4) and prove uniqueness of the solution of the martingale problem of the centered FLEMING-VIOT in Section 2.3. Note that these different versions are equivalent and that it is usual to switch from one to the other.

### 2.2.3.1 Extensions to multiple variables

We firstly introduce the version of the centered FLEMING-VIOT martingale problem with  $p \in \mathbb{N}^*$  variables.

**Definition 2.2.5.** *The probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\widetilde{\Omega})$  is said to solve the centered FLEMING-VIOT martingale problem with  $p$  variables with resampling rate  $\lambda$  and with initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , if the canonical process  $(X_t)_{t \geq 0}$  on  $\widetilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$  and for each*

$F \in \mathcal{C}^2(\mathbb{R}^p, \mathbb{R})$  and  $g = (g_1, \dots, g_p) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R}^p)$ ,

$$\begin{aligned}
 \widehat{M}_t^F(g) &:= F(\langle g_1, X_t \rangle, \dots, \langle g_p, X_t \rangle) - F(\langle g_1, X_0 \rangle, \dots, \langle g_p, X_0 \rangle) \\
 &\quad - \int_0^t \sum_{k=1}^p \partial_k F(\langle g_1, X_s \rangle, \dots, \langle g_p, X_s \rangle) \times \\
 &\quad \quad \left[ \left\langle \frac{g_k''}{2}, X_s \right\rangle + \lambda (\langle g_k'', X_s \rangle M_2(X_s) - 2 \langle g_k' \times \text{id}, X_s \rangle) \right] ds \\
 &\quad - \lambda \int_0^t \sum_{i,j=1}^p \partial_{ij}^2 F(\langle g_1, X_s \rangle, \dots, \langle g_p, X_s \rangle) \times \\
 &\quad \quad \left[ \langle g_i g_j, X_s \rangle - \langle g_i, X_s \rangle \langle g_j, X_s \rangle + \langle g_i', X_s \rangle \langle g_j', X_s \rangle M_2(X_s) \right. \\
 &\quad \quad \left. - \langle g_i', X_s \rangle \langle g_j \times \text{id}, X_s \rangle - \langle g_j', X_s \rangle \langle g_i \times \text{id}, X_s \rangle \right] ds
 \end{aligned} \tag{2.14}$$

is a continuous  $\mathbb{P}_\mu$ -martingale in  $L^2(\widetilde{\Omega})$ .

We will see in Section 2.3 that this martingale problem admits a unique solution which is the same as the solution of the martingale problem (2.4) if the initial condition has all its moments finite. For the moment, we can prove:

**Theorem 2.2.6.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , the probability measure  $\mathbb{P}_\mu$  constructed in Theorem 2.2.3, satisfies the martingale problem of Definition 2.2.5.*

*Proof.* We can deduce the result from the original FLEMING-VIOT martingale problem with  $p$  variables [61] given by (2.15) below, following exactly the same method as for the proof of Theorem 2.2.3. We recall that, the probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the *original FLEMING-VIOT martingale problem* with  $p$  variables with resampling rate  $\lambda$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$ , if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$  and for each  $F \in \mathcal{C}^2(\mathbb{R}^p, \mathbb{R})$  and  $g = (g_1, \dots, g_p) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R}^p)$ ,

$$\begin{aligned}
 M_t^F(g) &:= F(\langle g_1, Y_t \rangle, \dots, \langle g_p, Y_t \rangle) - F(\langle g_1, Y_0 \rangle, \dots, \langle g_p, Y_0 \rangle) \\
 &\quad - \int_0^t \sum_{k=1}^p \partial_k F(\langle g_1, Y_s \rangle, \dots, \langle g_p, Y_s \rangle) \left\langle \frac{g_k''}{2}, Y_s \right\rangle ds \\
 &\quad - \lambda \int_0^t \sum_{i,j=1}^p \partial_{ij}^2 F(\langle g_1, Y_s \rangle, \dots, \langle g_p, Y_s \rangle) [\langle g_i g_j, Y_s \rangle - \langle g_i, Y_s \rangle \langle g_j, Y_s \rangle] ds
 \end{aligned} \tag{2.15}$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale. As mentioned in [61, Theorem 5.1], the solution  $\mathbb{P}_\nu^{\text{FV}}$  of the martingale problem (2.7) is the unique solution to the previous martingale problem (2.15). Here, (2.11) has to be replaced by the general version of Lemma A.1.1.  $\square$

This version allows us to compute the martingale bracket  $\langle \widehat{M}^G(g), \widehat{M}^H(h) \rangle_t$  in a similar form as (2.8).

**Corollary 2.2.7.** *Let  $G, H \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g, h \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . Then,*

$$\begin{aligned}
 \langle \widehat{M}^G(g), \widehat{M}^H(h) \rangle_t &= 2\lambda \int_0^t G'(\langle g, X_s \rangle) H'(\langle h, X_s \rangle) \left[ \langle gh, X_s \rangle \right. \\
 &\quad \left. - \langle g, X_s \rangle \langle h, X_s \rangle + \langle g', X_s \rangle \langle h', X_s \rangle M_2(X_s) \right. \\
 &\quad \left. - \langle g', X_s \rangle \langle h \times \text{id}, X_s \rangle - \langle h', X_s \rangle \langle g \times \text{id}, X_s \rangle \right] ds.
 \end{aligned}$$

*Proof.* Recall that for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ ,  $G_g(\nu) := G(\langle g, \nu \rangle)$  and  $H_h(\nu) := H(\langle h, \nu \rangle)$ . Using the relation (2.4) for  $\widehat{M}_t^G(g)$  and  $\widehat{M}_t^H(h)$ , we obtain that

$$\begin{aligned} \widehat{M}_t^G(g)\widehat{M}_t^H(h) &= G(\langle g, X_t \rangle) H(\langle h, X_t \rangle) - G(\langle g, X_0 \rangle) H(\langle h, X_0 \rangle) \\ &\quad - G(\langle g, X_t \rangle) \int_0^t \mathcal{L}_{\text{FVc}} H_h(X_s) ds - H(\langle h, X_t \rangle) \int_0^t \mathcal{L}_{\text{FVc}} G_g(X_s) ds \\ &\quad + \int_0^t \mathcal{L}_{\text{FVc}} G_g(X_s) ds \int_0^t \mathcal{L}_{\text{FVc}} H_h(X_s) ds \\ &\quad - G(\langle g, X_0 \rangle) \widehat{M}_t^H(h) - H(\langle h, X_0 \rangle) \widehat{M}_t^G(g) \end{aligned} \quad (2.16)$$

where  $\mathcal{L}_{\text{FVc}}$  is given by (2.5). From the martingale problem (2.14) with  $p = 2$ ,  $F(x, y) = G(x)H(y)$  and  $f = (g, h)$ , we deduce that

$$\begin{aligned} &G(\langle g, X_t \rangle) H(\langle h, X_t \rangle) - G(\langle g, X_0 \rangle) H(\langle h, X_0 \rangle) \\ &= \int_0^t G(\langle g, X_s \rangle) \mathcal{L}_{\text{FVc}} H_h(X_s) ds + \int_0^t H(\langle h, X_s \rangle) \mathcal{L}_{\text{FVc}} G_g(X_s) ds \\ &\quad + \widehat{M}_t^F(f) + 2\lambda \int_0^t G'(\langle g, X_s \rangle) H'(\langle h, X_s \rangle) \left[ \langle gh, X_s \rangle - \langle g, X_s \rangle \langle h, X_s \rangle \right. \\ &\quad \left. + \langle g', X_s \rangle \langle h', X_s \rangle M_2(X_s) - \langle g', X_s \rangle \langle h \times \text{id}, X_s \rangle - \langle h', X_s \rangle \langle g \times \text{id}, X_s \rangle \right] ds, \end{aligned}$$

where  $(\widehat{M}_t^F(f))_{t \geq 0}$  is a  $\mathbb{P}_\mu$ -martingale. Using ITÔ's formula for the third and fourth term of the right hand side (2.16) and noting that

$$\begin{aligned} \int_0^t \mathcal{L}_{\text{FVc}} G_g(X_s) ds \int_0^t \mathcal{L}_{\text{FVc}} H_h(X_s) ds &= \int_0^t \left[ \mathcal{L}_{\text{FVc}} G_g(X_s) \left( \int_0^s \mathcal{L}_{\text{FVc}} H_h(X_r) dr \right) \right] ds \\ &\quad + \int_0^t \left[ \mathcal{L}_{\text{FVc}} H_h(X_s) \left( \int_0^s \mathcal{L}_{\text{FVc}} G_g(X_r) dr \right) \right] ds, \end{aligned}$$

we deduce the DOOB-MEYER decomposition of  $\widehat{M}_t^G(g)\widehat{M}_t^H(h)$ :

$$\begin{aligned} \widehat{M}_t^G(g)\widehat{M}_t^H(h) &= \widehat{M}_t^F(f) + 2\lambda \int_0^t G'(\langle g, X_s \rangle) H'(\langle h, X_s \rangle) \left[ \langle gh, X_s \rangle - \langle g, X_s \rangle \langle h, X_s \rangle \right. \\ &\quad \left. + \langle g', X_s \rangle \langle h', X_s \rangle M_2(X_s) - \langle g', X_s \rangle \langle h \times \text{id}, X_s \rangle - \langle h', X_s \rangle \langle g \times \text{id}, X_s \rangle \right] ds \\ &\quad - \int_0^t \left( \int_0^s \mathcal{L}_{\text{FVc}} H_h(X_r) dr \right) d\widehat{M}_s^G(g) - \int_0^t \left( \int_0^s \mathcal{L}_{\text{FVc}} G_g(X_r) dr \right) d\widehat{M}_s^H(h) \\ &\quad - G(\langle g, X_0 \rangle) \widehat{M}_t^H(h) - H(\langle h, X_0 \rangle) \widehat{M}_t^G(g). \end{aligned}$$

The result follows.  $\square$

### 2.2.3.2 Extension to polynomials in $\mu$

Our goal here is to study the DOOB semi-martingale decomposition of polynomial functions of the centered FLEMING-VIOT process of the form

$$P_{f,n}(\mu) := \langle f, \mu^n \rangle := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu(dx_1) \cdots \mu(dx_n) \quad (2.17)$$

with  $n \in \mathbb{N}^*$ ,  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  and  $\mu^n$  is the  $n$ -fold product measure of  $\mu$ . To obtain the expression of  $\mathcal{L}_{\text{FVc}}$  for polynomial functions  $P_{f,n}$  with  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ , we first look for this expression when  $f$  has product form. The previous martingale problem (2.14) gives heuristics for this issue: for the choice of  $F(x_1, \dots, x_n) := \prod_{i=1}^n x_i$ , noting that  $F(\langle g_1, \mu \rangle, \dots, \langle g_n, \mu \rangle) = P_{f,n}(\mu)$  with  $f(x_1, \dots, x_n) := \prod_{i=1}^n g_i(x_i)$  and  $g_i \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,  $i \in \{1, \dots, n\}$ , we deduce that for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,

$$\begin{aligned} \mathcal{L}_{\text{FVc}} P_{f,n}(\mu) &= \sum_{i=1}^n \left( \left\langle \frac{g_i''}{2}, \mu \right\rangle + \lambda \left[ \langle g_i'', \mu \rangle M_2(\mu) - 2 \langle g_i' \times \text{id}, \mu \rangle \right] \right) \prod_{\substack{j=1 \\ j \neq i}}^n \langle g_j, \mu \rangle \\ &\quad + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \left[ \langle g_i g_j, \mu \rangle - \langle g_i, \mu \rangle \langle g_j, \mu \rangle + \langle g_i', \mu \rangle \langle g_j', \mu \rangle M_2(\mu) \right. \\ &\quad \left. - \langle g_i', \mu \rangle \langle g_j \times \text{id}, \mu \rangle - \langle g_j', \mu \rangle \langle g_i \times \text{id}, \mu \rangle \right] \prod_{\substack{k=1 \\ k \neq i,j}}^n \langle g_k, \mu \rangle. \end{aligned}$$

For all  $n \in \mathbb{N}^*$ , we denote by  $\mathbf{1} \in \mathbb{R}^n$ , the vector whose coordinates are all 1 and by  $\Delta$  the Laplacian operator on  $\mathbb{R}^n$ . The previous relation leads us to introduce, for each  $n \in \mathbb{N}^*$  and for all  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ , the operator  $B^{(n)}$  defined by

$$B^{(n)} f(x) := \frac{1}{2} \Delta f(x) - 2\lambda (\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}), \quad x \in \mathbb{R}^n. \quad (2.18)$$

Indeed, again for the choice  $f(x_1, \dots, x_n) := \prod_{i=1}^n g_i(x_i)$  with  $g_i \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,  $i \in \{1, \dots, n\}$ , we obtain

$$\begin{aligned} B^{(n)} f(x_1, \dots, x_n) &= \sum_{i=1}^n \left[ \frac{g_i''(x_i)}{2} - 2\lambda x_i g_i'(x_i) \right] \prod_{\substack{j=1 \\ j \neq i}}^n g_j(x_j) \\ &\quad - \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \left[ x_j g_j(x_j) g_i'(x_i) + x_i g_i(x_i) g_j'(x_j) \right] \prod_{\substack{k=1 \\ k \neq i,j}}^n g_k(x_k). \end{aligned}$$

Note that, for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,

$$\begin{aligned} \mathcal{L}_{\text{FVc}} P_{f,n}(\mu) &= \langle B^{(n)} f, \mu^n \rangle + \sum_{i=1}^n \lambda \langle g_i'', \mu \rangle M_2(\mu) \prod_{\substack{j=1 \\ j \neq i}}^n \langle g_j, \mu \rangle \\ &\quad + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \left[ \langle g_i g_j, \mu \rangle - \langle g_i, \mu \rangle \langle g_j, \mu \rangle + \langle g_i', \mu \rangle \langle g_j', \mu \rangle M_2(\mu) \right] \prod_{\substack{k=1 \\ k \neq i,j}}^n \langle g_k, \mu \rangle. \end{aligned} \quad (2.19)$$

This leads us to introduce another extension of the martingale problem (2.4) which will be useful in Section 2.3 to prove uniqueness.

**Definition 2.2.8.** *The probability measure  $\mathbb{P}_\mu \in \mathcal{M}_1(\tilde{\Omega})$  is said to solve the centered FLEMING-VIOT martingale problem for polynomials with resampling rate  $\lambda$  and with initial condition*

$\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , if the canonical process  $(X_t)_{t \geq 0}$  on  $\tilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$ , for all  $n \in \mathbb{N}^*$ , and for each  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\widehat{M}_t^{(n)}(f) := \langle f, X_t^n \rangle - \langle f, X_0^n \rangle - \int_0^t \mathcal{L}_{\text{FVc}} P_{f,n}(X_s) ds \quad (2.20)$$

with for all  $\varpi \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathcal{L}_{\text{FVc}} P_{f,n}(\varpi) := \langle B^{(n)} f, \varpi^n \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \varpi^{n-1} \rangle - \langle f, \varpi^n \rangle] + \lambda \sum_{i=1}^n \sum_{j=1}^n \langle K_{i,j} f, \varpi^{n+1} \rangle \quad (2.21)$$

is a continuous  $\mathbb{P}_\mu$ -martingale in  $L^2(\tilde{\Omega})$  where, for all  $i, j \in \{1, \dots, n\}$ ,

- $\Phi_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$ , with  $i \neq j$ , is the function obtained from  $f$  by inserting the variable  $x_i$  between  $x_{j-1}$  and  $x_j$  when  $i < j$  and by inserting the variable  $x_{i-1}$  between  $x_{j-1}$  and  $x_j$  when  $i > j$ :

$$\begin{aligned} \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_i, x_j, x_{j+1}, \dots, x_{n-1}) & i < j \\ \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_{i-1}, x_j, x_{j+1}, \dots, x_{n-1}) & i > j \end{aligned} \quad (2.22)$$

- $K_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^2(\mathbb{R}^{n+1}, \mathbb{R})$  is defined as

$$K_{i,j} f(x_1, \dots, x_n, x_{n+1}) := \partial_{ij}^2 f(x_1, \dots, x_n) x_{n+1}^2. \quad (2.23)$$

We will see in Section 2.3 that this martingale problem admits the same unique solution as the martingale problem (2.4) if the initial condition has all its moments finite. For the moment, we can prove:

**Theorem 2.2.9.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , the probability measure  $\mathbb{P}_\mu$  constructed in Theorem 2.2.3, satisfies the martingale problem of Definition 2.2.8.*

*Proof.* We can deduce the result from the original FLEMING-VIOT martingale problem for polynomials [91] given by (2.24) below, following exactly the same arguments as for the proof of Theorem 2.2.3. The probability measure  $\mathbb{P}_\nu^{\text{FV}} \in \mathcal{M}_1(\Omega)$  is said to solve the *original FLEMING-VIOT martingale problem for polynomials* with resampling rate  $\lambda$  and with initial condition  $\nu \in \mathcal{M}_1(\mathbb{R})$ , if the canonical process  $(Y_t)_{t \geq 0}$  on  $\Omega$  satisfies  $\mathbb{P}_\nu^{\text{FV}}(Y_0 = \nu) = 1$ , for all  $n \in \mathbb{N}^*$ , and for each  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$M_t^{(n)}(f) := \langle f, Y_t^n \rangle - \langle f, Y_0^n \rangle - \int_0^t \mathcal{L}_{\text{FV}} P_{f,n}(Y_s) ds \quad (2.24)$$

with for all  $\varpi \in \mathcal{M}_1(\mathbb{R})$

$$\mathcal{L}_{\text{FV}} P_{f,n}(\varpi) = \left\langle \frac{1}{2} \Delta f, \varpi^n \right\rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n (\langle \Phi_{i,j} f, \varpi^{n-1} \rangle - \langle f, \varpi^n \rangle)$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale. By [91, Theorem 3.2], the solution  $\mathbb{P}_\nu^{\text{FV}}$  of the martingale problem (2.7) is the unique solution to the previous martingale problem (2.24).  $\square$

## 2.2.4 Some properties of the centered Fleming-Viot process

### 2.2.4.1 Markov's property

Due to the translation invariance property of the original FLEMING-VIOT process, we can prove that the centered FLEMING-VIOT process is homogeneous MARKOV.

**Proposition 2.2.10.** *The centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$  defined by (2.9) satisfies the homogeneous MARKOV property: for all measurable bounded function  $f$ ,*

$$\forall \mu \in \mathcal{M}_1^{c,2}(\mathbb{R}), \quad \forall t, s > 0, \quad \mathbb{E}_\mu \left( f(Z_{t+s}) \middle| \mathcal{F}_t \right) = \mathbb{E}_{Z_t} (f(Z_s)) \quad \mathbb{P}_\mu\text{-a.s.}$$

*Proof.* Let  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $f$  a measurable bounded function. Let  $t, s > 0$ . Using the MARKOV property of the original FLEMING-VIOT process  $(Y_t)_{t \geq 0}$  we obtain  $\mathbb{P}_\mu$ -a.s.,

$$\mathbb{E}_\mu \left( f(Z_{t+s}) \middle| \mathcal{F}_t \right) = \mathbb{E}_\mu \left( f(\tau_{-\langle \text{id}, Y_{t+s} \rangle \# Y_{t+s}}) \middle| \mathcal{F}_t \right) = \mathbb{E}_\mu \left( g(Y_{t+s}) \middle| \mathcal{F}_t \right) = \mathbb{E}_{Y_t} (g(Y_s))$$

where the bounded measurable map  $g$  is defined on  $\mathcal{M}_1^1(\mathbb{R})$  by  $g(\nu) := f(\tau_{-\langle \text{id}, \nu \rangle \# \nu})$ . By invariance by translation of the original FLEMING-VIOT process  $(Y_t)_{t \geq 0}$  we obtain under the distribution  $\mathbb{P}_\mu$ :

$$\mathbb{E}_{Y_t} (g(Y_s)) = \mathbb{E}_{\tau_{-\langle \text{id}, Y_t \rangle \# Y_t}} (g(\tau_{\langle \text{id}, Y_t \rangle \# Y_s})) = \mathbb{E}_{\tau_{-\langle \text{id}, Y_t \rangle \# Y_t}} (g(Y_s)) = \mathbb{E}_{Z_t} (g(Y_s)) = \mathbb{E}_{Z_t} (f(Z_s)). \quad \square$$

### 2.2.4.2 Moments and some martingales

**Proposition 2.2.11.** *Let  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and let  $\mathbb{P}_\mu$  be a distribution on  $\tilde{\Omega}$  satisfying (2.4) and such that  $X_0$  is equal in law to  $\mu$ . Let  $T > 0$  and  $k \in \mathbb{N} \setminus \{0, 1\}$  be fixed.*

(1) *If  $\langle |\text{id}|^k, \mu \rangle < \infty$ , there exist two constants  $C_{k,T}, \tilde{C}_{k,T} > 0$ , such that any stochastic process  $(X_t)_{0 \leq t \leq T}$  whose law is  $\mathbb{P}_\mu$  satisfies*

$$(a) \quad \sup_{t \in [0, T]} \mathbb{E}_\mu \left( \langle |\text{id}|^k, X_t \rangle \right) \leq C_{k,T} \left( 1 + \langle |\text{id}|^k, \mu \rangle \right) \quad (2.25)$$

$$(b) \quad \forall \alpha > 0, \quad \mathbb{P}_\mu \left( \sup_{t \in [0, T]} \langle |\text{id}|^k, X_t \rangle \geq \alpha \right) \leq \frac{\tilde{C}_{k,T} \left( 1 + \langle |\text{id}|^k, \mu \rangle \right)}{\alpha}.$$

(2) *If  $\langle |\text{id}|^k, \mu \rangle < \infty$ , respectively  $\langle |\text{id}|^{k+1}, \mu \rangle < \infty$ , the process  $(\widehat{M}_t^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  defined by*

$$\begin{aligned} \widehat{M}_t^{\text{id}}(\text{id}^k) &:= \langle \text{id}^k, X_t \rangle - \langle \text{id}^k, X_0 \rangle - \int_0^t \left\langle \frac{k(k-1)}{2} \text{id}^{k-2}, X_s \right\rangle ds \\ &\quad - \lambda \int_0^t \left[ \langle k(k-1) \text{id}^{k-2}, X_s \rangle M_2(X_s) - 2k \langle \text{id}^k, X_s \rangle \right] ds \end{aligned}$$

*is a continuous  $\mathbb{P}_\mu$ -local martingale, respectively a continuous  $\mathbb{P}_\mu$ -martingale. Moreover, if  $\langle |\text{id}|^{2k}, \mu \rangle < \infty$ , then  $(\widehat{M}_t^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  is a martingale in  $L^2(\tilde{\Omega})$  whose quadratic variation is given by*

$$\begin{aligned} \langle \widehat{M}^{\text{id}}(\text{id}^k) \rangle_t &= 2\lambda \int_0^t \left[ \langle \text{id}^{2k}, X_s \rangle - \langle \text{id}^k, X_s \rangle^2 + k \langle \text{id}^{k-1}, X_s \rangle^2 M_2(X_s) \right. \\ &\quad \left. - 2k \langle \text{id}^{k-1}, X_s \rangle \langle \text{id}^{k+1}, X_s \rangle \right] ds. \end{aligned}$$



Note that the properties for  $k = 1$  fail because of  $M_2$  in the expression of  $\mathcal{L}_{\text{FVC}}$ . Note also that this result entails that  $Z_t$  is a continuous martingale.

*Proof. Step 1. Proof of (1)(a).* We prove only the case  $k \geq 3$ : the case  $k = 2$ , which is simpler because some terms disappear, is treated in the same way. Let  $t \in [0, T]$ . We consider a sequence of functions  $(g_n)_{n \in \mathbb{N}}$  of class  $\mathcal{C}^2(\mathbb{R}, \mathbb{R})$  with compact support satisfying:

- (i) for all  $n \in \mathbb{N}$ ,  $|g_n| \leq |\text{id}|$ ,
- (ii)  $\lim_{n \rightarrow +\infty} \|g_n''\|_\infty = 0$ ,
- (iii)  $g_n = \text{id}$  on  $[-n, n]$ ,
- (iv)  $g_n'$  is uniformly bounded on  $\mathbb{R}$ .

We consider the sequence of functions  $(h_n)_{n \in \mathbb{N}}$  defined by  $h_n := \sqrt{1 + g_n^2}$  and we deduce from the properties of  $g_n$  that for all  $n \in \mathbb{N}$ ,  $h_n$  is a non-negative function with compact support, that for all  $k \in \mathbb{N}$ ,

$$(h_n^k)' = k g_n g_n' h_n^{k-2} \quad \text{and} \quad (h_n^k)'' = k (g_n')^2 h_n^{k-4} (h_n^2 + (k-2)g_n^2) + k g_n g_n'' h_n^{k-2},$$

$h_n = h := \sqrt{1 + \text{id}^2}$  on the compact set  $[-n, n]$  and  $h_n \leq h$  on  $\mathbb{R}$ . We consider for all  $A \in \mathbb{N}$  and  $\ell \in \mathbb{N}$ , the stopping time  $\tau_{A, \ell} := \inf \left\{ t \geq 0 \mid \langle |\text{id}|^\ell, X_t \rangle \geq A \right\}$ . Noting that for all  $t \in [0, T]$ ,  $n \in \mathbb{N}$  and  $k \geq 3$ ,  $\langle h_n^{k-2}, X_t \rangle \leq \langle h_n^k, X_t \rangle^{\frac{k-2}{k}} \leq \langle h^k, X_t \rangle$  and  $\langle h_n^{k-2}, X_t \rangle \langle \text{id}^2, X_t \rangle \leq \langle h^k, X_t \rangle$  from HÖLDER'S inequality, we deduce from the martingale problem (2.4) that there exists constants  $C_1(k), C_2(k, A) > 0$  such that

$$\begin{aligned} \mathbb{E}_\mu \left( \langle h_n^k, X_{t \wedge \tau_{A, k}} \rangle \right) &= \mathbb{E}_\mu \left( \langle h_n^k, X_0 \rangle \right) + \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \left\langle \frac{(h_n^k)''}{2}, X_s \right\rangle ds \right) \\ &\quad + \lambda \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \left\langle (h_n^k)'', X_s \right\rangle M_2(X_s) ds \right) - 2\lambda \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \left\langle (h_n^k)' \times \text{id}, X_s \right\rangle ds \right) \\ &\leq \langle h_n^k, \mu \rangle + k(k-1) \|g_n'\|_\infty^2 \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \left[ \frac{1}{2} + \lambda \langle \text{id}^2, X_s \rangle \right] \langle h_n^{k-2}, X_s \rangle ds \right) \\ &\quad + k \|g_n''\|_\infty \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \left[ \frac{1}{2} + \lambda \langle \text{id}^2, X_s \rangle \right] \langle h_n^{k-1}, X_s \rangle ds \right) \\ &\quad + 2\lambda k \|g_n'\|_\infty \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \langle h_n^{k-1} \times |\text{id}|, X_s \rangle ds \right) \\ &\leq \langle h^k, \mu \rangle + C_1(k) \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \langle h^k, X_s \rangle ds \right) + C_2(k, A) \|g_n''\|_\infty. \end{aligned}$$

By FATOU'S lemma we obtain when  $n \rightarrow +\infty$ ,

$$\mathbb{E}_\mu \left( \langle h^k, X_{t \wedge \tau_{A, k}} \rangle \right) \leq \langle h^k, \mu \rangle + C_1(k) \mathbb{E}_\mu \left( \int_0^{t \wedge \tau_{A, k}} \langle h^k, X_s \rangle ds \right).$$

By GRONWALL'S lemma, we deduce that

$$\mathbb{E}_\mu \left( \langle |\text{id}|^k, X_{t \wedge \tau_{A, k}} \rangle \right) \leq \mathbb{E}_\mu \left( \langle h^k, X_{t \wedge \tau_{A, k}} \rangle \right) \leq \langle h^k, \mu \rangle \exp(C_1(k)t). \quad (2.26)$$

In particular, this implies that the sequence  $(\tau_{A, k})_{A \in \mathbb{N}}$  converges  $\mathbb{P}_\mu$ -a.s. to infinity. Indeed, for all  $\tilde{T} > 0$ , we have

$$\mathbb{P}_\mu \left( \sup_{A \in \mathbb{N}} \tau_{A, k} < \tilde{T} \right) \leq \frac{\sup_{t \in [0, T]} \mathbb{E}_\mu \left( \langle |\text{id}|^k, X_{t \wedge \tilde{T} \wedge \tau_{A, k}} \rangle \right)}{A}$$

which tends to 0 when  $A \rightarrow +\infty$ . We deduce by FATOU's lemma, when  $A \rightarrow +\infty$ , the first announced result.

**Step 2. Proof of (1)(b).** Let  $\alpha > 0$ . From the martingale problem (2.4), we deduce that

$$\begin{aligned} \mathbb{P}_\mu \left( \sup_{t \in [0, T \wedge \tau_{A,k}]} \langle h_n^k, X_t \rangle \geq \alpha \right) &\leq \mathbb{P}_\mu \left( \langle h_n^k, \mu \rangle \geq \frac{\alpha}{5} \right) + \mathbb{P}_\mu \left( \int_0^{T \wedge \tau_{A,k}} \left\langle \frac{(h_n^k)''}{2}, X_s \right\rangle ds \geq \frac{\alpha}{5} \right) \\ &\quad + \mathbb{P}_\mu \left( \lambda \int_0^{T \wedge \tau_{A,k}} \langle (h_n^k)'', X_s \rangle M_2(X_s) ds \geq \frac{\alpha}{5} \right) \\ &\quad + \mathbb{P}_\mu \left( 2\lambda \int_0^{T \wedge \tau_{A,k}} \langle (h_n^k)' \times \text{id}, X_s \rangle M_2(X_s) ds \geq \frac{\alpha}{5} \right) \\ &\quad + \mathbb{P}_\mu \left( \sup_{t \in [0, T \wedge \tau_{A,k}]} \left| \widehat{M}_t^{\text{id}}(h_n^k) \right| \geq \frac{\alpha}{5} \right) \end{aligned}$$

The DOOB maximal inequality allows us to write

$$\mathbb{P}_\mu \left( \sup_{t \in [0, T \wedge \tau_{A,k}]} \left| \widehat{M}_t^{\text{id}}(h_n^k) \right| \geq \frac{\alpha}{5} \right) \leq \frac{5\mathbb{E}_\mu \left( \left( \widehat{M}_{T \wedge \tau_{A,k}}^{\text{id}}(h_n^k) \right)_+ \right)}{\alpha}.$$

From the martingale problem (2.4) and the computations of Step 1, we deduce that

$$\begin{aligned} \mathbb{E}_\mu \left( \left| \widehat{M}_{T \wedge \tau_{A,k}}^{\text{id}}(h_n^k) \right| \right) &\leq 2 \langle h^k, \mu \rangle + 2C_1(k) \mathbb{E}_\mu \left( \int_0^{T \wedge \tau_{A,k}} \langle h^k, X_s \rangle ds \right) + 2C_2(k, A) \|g_n''\|_\infty \\ &\leq 2 \langle h^k, \mu \rangle [1 + \exp(C_1(k)T)] + 2C_2(k, A) \|g_n''\|_\infty, \end{aligned}$$

where we use the FUBINI-TONELLI theorem and the relation (2.26). It follows, from MARKOV's inequality, that there exists a constant  $C_k > 0$  such that

$$\mathbb{P}_\mu \left( \sup_{t \in [0, T \wedge \tau_{A,k}]} \langle h_n^k, X_t \rangle \geq \alpha \right) \leq \frac{C_k}{\alpha} [\langle h^k, \mu \rangle [1 + \exp(C_1(k)T)] + C_2(k, A) \|g_n''\|_\infty].$$

By applying the dominated convergence theorem twice, successively when  $n \rightarrow +\infty$  then when  $A \rightarrow +\infty$ , there exists a constant  $\widetilde{C}_{k,T} > 0$  such that

$$\mathbb{P}_\mu \left( \sup_{t \in [0, T]} \langle h^k, X_t \rangle \geq \alpha \right) \leq \frac{\widetilde{C}_{k,T} \langle h^k, \mu \rangle}{\alpha},$$

and thus the announced result.

**Step 3.  $\widehat{M}^{\text{id}}(\text{id}^k)$  is a continuous local martingale.** From the properties of  $(g_n)_{n \in \mathbb{N}}$ , note that there exists a constant  $\widehat{C}_k > 0$  such that for all  $n \in \mathbb{N}$ ,  $|g_n^k| \leq |\text{id}^k|$  and  $\left| (g_n^k)'' \right| \leq \widehat{C}_k (1 + |\text{id}^{k-1}|)$ . It follows from the martingale problem (2.4), the properties of  $(g_n)_{n \in \mathbb{N}}$  and the

dominated convergence theorem for conditional expectation that

$$\begin{aligned} \widehat{M}_{t \wedge \tau_{A,2}}^{\text{id}}(\text{id}^k) &:= \lim_{n \rightarrow +\infty} \widehat{M}_{t \wedge \tau_{A,2}}^{\text{id}}(g_n^k) \\ &= \langle \text{id}^k, X_{t \wedge \tau_{A,2}} \rangle - \langle \text{id}^k, X_0 \rangle - \int_0^{t \wedge \tau_{A,2}} \frac{k(k-1)}{2} \langle \text{id}^{k-2}, X_s \rangle ds \\ &\quad - \lambda \int_0^{t \wedge \tau_{A,2}} \left[ k(k-1) \langle \text{id}^{k-2}, X_s \rangle M_2(X_s) - 2k \langle \text{id}^k, X_s \rangle \right] ds \end{aligned}$$

is a continuous  $\mathbb{P}_\mu$ -martingale and thus  $(\widehat{M}^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  is a continuous  $\mathbb{P}_\mu$ -local martingale. When  $\langle |\text{id}|^{k+1}, \mu \rangle < \infty$ , using the inequality for all  $t \in [0, T]$ ,

$$\langle |\text{id}|^{k-1}, X_t \rangle \langle \text{id}^2, X_t \rangle \leq \langle |\text{id}|^{k+1}, X_t \rangle,$$

the same computation applies replacing  $t \wedge \tau_{A,2}$  by  $t$  to obtain that  $(\widehat{M}^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  is a continuous  $\mathbb{P}_\mu$ -martingale.

**Step 4.  $L^2$ -martingale and quadratic variation.** As soon as  $\langle |\text{id}|^{k+1}, \mu \rangle < \infty$ ,  $\widehat{M}^{\text{id}}(\text{id}^k) \in L^2(\widetilde{\Omega})$  as a straightforward consequence of HÖLDER's inequality. It follows from (2.6) that for all  $n \in \mathbb{N}$ , for all  $t \in [0, T]$ ,

$$\begin{aligned} \langle \widehat{M}^{\text{id}}(g_n^k) \rangle_t &= 2\lambda \int_0^t \left( \langle g_n^{2k}, X_s \rangle - \langle g_n^k, X_s \rangle^2 + \langle (g_n^k)', X_s \rangle^2 M_2(X_s) \right. \\ &\quad \left. - 2 \langle (g_n^k)', X_s \rangle \langle g_n^k \times \text{id}, X_s \rangle \right) ds. \end{aligned}$$

For all  $n \in \mathbb{N}$ , the process

$$N_{t,n} := \left[ \widehat{M}_t^{\text{id}}(g_n^k) \right]^2 - \langle \widehat{M}^{\text{id}}(g_n^k) \rangle_t$$

is a  $\mathbb{P}_\mu$ -local martingale. As  $\widehat{M}^{\text{id}}(\text{id}^k)$  is bounded on  $[0, T \wedge \tau_{A,2k}]$  for all  $A \in \mathbb{N}$ , then for all  $n \in \mathbb{N}$ ,  $(N_{t \wedge \tau_{A,2k},n})$  is a martingale. From the relation (2.25) with  $2k$ , the dominated convergence theorem for conditional expectation implies as above that  $\mathbb{P}_\mu$ -a.s.

$$\begin{aligned} \lim_{n \rightarrow +\infty} N_{t \wedge \tau_{A,2k},n} &= \left[ \widehat{M}_{t \wedge \tau_{A,2k}}^{\text{id}}(\text{id}^k) \right]^2 - 2\lambda \int_0^{t \wedge \tau_{A,2k}} \left( \langle \text{id}^{2k}, X_s \rangle - \langle \text{id}^k, X_s \rangle^2 \right. \\ &\quad \left. + k \langle \text{id}^{k-1}, X_s \rangle^2 M_2(X_s) - 2k \langle \text{id}^{k-1}, X_s \rangle \langle \text{id}^{k+1}, X_s \rangle \right) ds \end{aligned}$$

is a  $\mathbb{P}_\mu$ -martingale and we deduce the quadratic variation announced.  $\square$

### 2.2.4.3 Compact support

For all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote by  $\text{Supp } \nu$  the support of  $\nu$ . The historical reference of compact support property of the original FLEMING-VIOT process is [61, Theorem 7.1] where the authors proved that  $\text{Supp } Y_t$  is a.s. compact for each fixed  $t > 0$ . We will use a slightly stronger version based on [159].

**Proposition 2.2.12.** *For all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , for all  $\varepsilon > 0$ ,  $\bigcup_{\varepsilon \leq s \leq t} \text{Supp } Z_s$  is compact  $\mathbb{P}_\mu$ -a.s. Further, if  $\text{Supp } Z_0$  is compact, then  $\bigcup_{0 \leq s \leq t} \text{Supp } Z_s$  is compact for all  $t > 0$ ,  $\mathbb{P}_\mu$ -a.s.*

*Proof.* It is proved in [159] that the support of the  $\Lambda$ -FLEMING-VIOT process associated to a  $\Lambda$ -coalescent which comes down from infinity, is compact at all positive times. Our case corresponds to KINGMAN's coalescent. In addition, they prove that, given that the initial condition  $\nu$  has compact support,

$$\overline{\bigcup_{0 \leq s < t} \text{Supp } Y_s}$$

is compact for all  $t > 0$ ,  $\mathbb{P}_\nu^{\text{FV}}$ -a.s. MARKOV's property then entails that, if  $\nu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,  $\bigcup_{\varepsilon \leq s \leq t} \text{Supp } Y_s$  is compact for all  $0 < \varepsilon < t$ ,  $\mathbb{P}_\nu^{\text{FV}}$ -a.s. Hence, the same is true for  $Z_t = \tau_{-(\text{id}, Y_t)^\#} Y_t$ .  $\square$

## 2.3 Uniqueness for the centered Fleming-Viot process

As for the original FLEMING-VIOT martingale problem, we will prove uniqueness to the martingale problem (2.4) by relying on the *duality method* [85, 60, 89, 91]. Additional difficulties occur in our case since bounds on the dual process are much harder to obtain and the duality identity cannot be proved in its usual form. In particular, we can prove uniqueness only for initial conditions admitting finite moments.

### 2.3.1 Main result

**Theorem 2.3.1.** *The centered FLEMING-VIOT martingale problem (2.4) has a unique solution if the initial condition has all its moments finite.*

The reason why we need to assume finite initial moments will be explained at the end of Section 2.3.2. In particular, we will see in Remark 2.6.5 that we cannot hope to prove uniqueness for more general initial conditions using our duality method.

**Corollary 2.3.2.** *Existence and uniqueness hold for all the martingale problems of Definitions 2.2.1, 2.2.5 and 2.2.8 and they all admit the same solution if the initial condition has all its moments finite.*

*Proof of Corollary 2.3.2.* We proved in Section 2.2.3 that  $\mathbb{P}_\mu$  solves the martingale problems of Definitions 2.2.5 and 2.2.8. Since a solution to these martingale problems is of course also solution to Definition 2.2.1, uniqueness for Definition 2.2.1 implies uniqueness for the other martingale problems.  $\square$

### 2.3.2 Notations and outline of the uniqueness proof

Our proof of Theorem 2.3.1 is based on the *duality method* as proposed in [89, 91]. From (2.21), the operator  $\mathcal{L}_{\text{FVc}}$  applied on the function  $P_{f,n}$  defined in (2.17) with fixed  $f$  and  $n \in \mathbb{N}^*$ , satisfies the following identity:

$$\begin{aligned} \mathcal{L}_{\text{FVc}} P_{f,n}(\mu) &= \langle B^{(n)} f, \mu^n \rangle + \lambda \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle] \\ &\quad + \lambda \sum_{i=1}^n \sum_{j=1}^n [\langle K_{i,j} f, \mu^{n+1} \rangle - \langle f, \mu^n \rangle] + \lambda n^2 \langle f, \mu^n \rangle \\ &=: \tilde{\mathcal{L}}_f^* P_{f,n}(\mu) + \lambda n^2 \langle f, \mu^n \rangle \end{aligned} \tag{2.27}$$

We note that  $\tilde{\mathcal{L}}_f^*$  can be seen as an operator acting on the function  $f \mapsto P_{f,n}(\mu)$  with fixed  $\mu$ . The operator  $\tilde{\mathcal{L}}_f^*$  can be interpreted as the generator of a stochastic process on the state space  $\bigcup_{n \in \mathbb{N}^*} \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ . Following ETHIER-KURTZ's works [89, 91], this suggests to introduce a dual process  $(\xi_t)_{t \geq 0}$ , of generator  $\tilde{\mathcal{L}}_f^*$  and to prove a duality relation of the form:

$$\forall t \geq 0, \quad \mathbb{E} \left( \langle \xi_0, X_t^{M(0)} \rangle \right) = \mathbb{E} \left( \langle \xi_t, X_0^{M(t)} \rangle \exp \left( \lambda \int_0^t M^2(u) du \right) \right) \quad (2.28)$$

where  $M := (M(t))_{t \geq 0}$  is a MARKOV's birth and death process in  $\mathbb{N}$  whose transition rates  $q_{i,j}$  from  $i$  to  $j$  are given by:

$$(1) \quad q_{n,n+1} = \lambda n^2 \quad (2) \quad q_{n,n-1} = \lambda n(n-1) \quad (3) \quad q_{i,j} = 0 \text{ otherwise.}$$

It is known that the relation (2.28) implies uniqueness [94, Theorem 4.4.2]. However in our situation, it is difficult to obtain the strong version (2.28). For technical reasons, we will obtain only a weakened version. Therefore, the proof will be divided in two large steps.

**Step 1. Construction of the dual process  $(\xi_t)_{t \geq 0}$ .** The relation (2.27) suggests that the dual process  $(\xi_t)_{t \geq 0}$  jumps, for all  $i, j \in \{1, \dots, n\}$  from  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  to  $K_{i,j}f \in \mathcal{C}^2(\mathbb{R}^{n+1}, \mathbb{R})$  at rate  $\lambda$  and if  $i \neq j$ , from  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  to  $\Phi_{i,j}f \in \mathcal{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$  at rate  $\lambda$ . Moreover, note that if  $n = 1$ , the dual process can only jump from  $f \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  to  $K_{i,j}f$ . Between jumps, this dual process evolves according to the semi-group of operator  $(T^{(n)}(t))_{t \geq 0}$  associated to the generator  $B^{(n)}$  given by (2.18). We will give an explicit expression of the semi-group  $(T^{(n)}(t))_{t \geq 0}$  defined as an integral against Gaussian kernels. This representation will be derived from a probabilistic interpretation of the semi-group using a FEYNMAN-KAC's formula. We define the dual process as follows:

**Definition 2.3.3.** For all  $M(0) \in \mathbb{N}^*$ , for all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ ,

$$\xi_t := T^{(M(\tau_n))}(t - \tau_n) \Lambda_n T^{(M(\tau_{n-1}))}(\tau_n - \tau_{n-1}) \Lambda_{n-1} \cdots \Lambda_1 T^{(M(0))}(\tau_1) \xi_0, \quad \tau_n \leq t < \tau_{n+1}, \quad n \in \mathbb{N}, \quad (2.29)$$

where  $(\tau_n)_{n \in \mathbb{N}}$  is the sequence of jump times of the birth-death process  $M$  with  $\tau_0 = 0$  and where  $(\Lambda_n)_{n \in \mathbb{N}}$  is a sequence of random operators. These are conditionally independent given  $M$  and satisfy for all  $k \in \mathbb{N}$ ,  $n \geq 1$  and  $1 \leq i \neq j \leq n$ ,

$$\mathbb{P} \left( \Lambda_k = \Phi_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n-1\} \right) = \frac{1}{n(n-1)} \quad (2.30)$$

and for all  $n \geq 1$  and  $1 \leq i, j \leq n$ ,

$$\mathbb{P} \left( \Lambda_k = K_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n+1\} \right) = \frac{1}{n^2}. \quad (2.31)$$

Moreover, the random times  $(\tau_k - \tau_{k-1})_{k \geq 1}$  are independent conditionally to  $M(\tau_{k-1}) = n$  and of exponential law of parameter  $\lambda n^2 + \lambda n(n-1)$ .

Note that particle creation for the centered FLEMING-VIOT dual process is possible unlike the original FLEMING-VIOT dual process. It comes from the  $K_{i,j}$  operator which has appeared under the centering effect. Note that these terms are the ones with the  $M_2(\mu)$  factor in (2.5) and are

found by the heuristic computation in (2.19). Indeed, for the choice of  $n \in \mathbb{N}^*$ ,  $f(x_1, \dots, x_n) := \prod_{i=1}^n g_i(x_i)$  with  $g_i \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,  $i \in \{1, \dots, n\}$ , we have for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  that

$$\sum_{i=1}^n \sum_{j=1}^n \langle K_{i,j} f, \mu^{n+1} \rangle = \left( \sum_{i=1}^n \langle g_i'', \mu \rangle \prod_{\substack{j=1 \\ j \neq i}}^n \langle g_j, \mu \rangle + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \langle g_i', \mu \rangle \langle g_j', \mu \rangle \prod_{\substack{k=1 \\ k \neq i,j}}^n \langle g_k, \mu \rangle \right) M_2(\mu)$$

The terms  $\Phi_{i,j}$  are present both in the original and centered FLEMING-VIOT. Because of the operators  $K_{i,j}$ , difficulties will arise to get bounds on the dual process (see Section 2.6.2) contrary to the non-centered case. Note also that  $M$  is a non-explosive process:

$$\forall T > 0, \quad \mathbb{P} \left( \sup_{t \in [0, T]} M(t) < +\infty \right) = \mathbb{P} \left( \lim_{n \rightarrow +\infty} \tau_n = +\infty \right) = 1. \quad (2.32)$$

Indeed, we note that for the choice of  $\mu_i := \lambda i(i-1)$  and  $\lambda_i := \lambda i^2$ ,  $i \geq 1$  in [10, Theorem 2.2.], we have

$$\frac{\mu_i \dots \mu_2}{\lambda_i \dots \lambda_2 \lambda_1} = \frac{1}{2\lambda i} > 0,$$

so that

$$\sum_{i \geq 1} \left( \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i \lambda_{i-1}} + \dots + \frac{\mu_i \dots \mu_2}{\lambda_i \dots \lambda_2 \lambda_1} \right) \geq \sum_{i \geq 1} \frac{\mu_i \dots \mu_2}{\lambda_i \dots \lambda_2 \lambda_1} = +\infty.$$

Hence,  $M$  is non-explosive.

**Step 2. Weakened duality relation.** We consider fixed  $M(0) \in \mathbb{N}^*$ ,  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  and  $(X_t)_{t \geq 0}$  a stochastic process whose law  $\mathbb{P}_\mu$  is a solution of the martingale problem (2.4) with  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ . We introduce a dual process  $(\xi_t)_{t \geq 0}$  independent of  $(X_t)_{t \geq 0}$  built on the same probability space (enlarging it if necessary). We shall denote by  $\mathbb{P}_{(\mu, \xi_0)}$ , the law of  $((X_t, \xi_t))_{t \geq 0}$  on this probability space. For any  $k \in \mathbb{N}$ , we introduce the stopping time

$$\theta_k := \inf \left\{ t \geq 0 \mid M(t) \geq k \quad \text{or} \quad \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq k \right\}. \quad (2.33)$$

**Theorem 2.3.4.** *Given any  $(X_t)_{t \geq 0}$ ,  $(\xi_t)_{t \geq 0}$  as above, we have the weakened duality identity: for all  $k \in \mathbb{N}$  and any stopping time  $\theta$  such that  $\theta \leq \theta_k$ ,*

$$\forall t \geq 0, \quad \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_0, X_{t \wedge \theta}^{M(0)} \rangle \right) = \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_{t \wedge \theta}, X_0^{M(t \wedge \theta)} \rangle \exp \left( \lambda \int_0^{t \wedge \theta} M^2(u) du \right) \right). \quad (2.34)$$

Note that this result holds true for any initial measure  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ . The stopping time  $\theta_k$  ensures that each of the quantities involved in (2.34) are bounded and thus that their expectations are finite. Afterwards, we want to prove that if two solutions of the martingale problem satisfy the weakened duality identity, then their 1-dimensional marginals coincide. This is where we need stronger assumptions on  $\mu$ .

**Lemma 2.3.5.** *Assume that  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  has all its moments finite. Then, the stopping time  $\theta_k$  defined by (2.33) satisfies  $\lim_{k \rightarrow +\infty} \theta_k = +\infty$ ,  $\mathbb{P}_{(\mu, \xi_0)}$  - a.s.*

We will see in Remark 2.6.5 that the assumption on  $\mu$  is optimal in the following sense: even if  $\xi_0$  is bounded,  $\xi_t$  may have polynomial growth of any exponent  $k$  in some direction of  $\mathbb{R}^{M(t)}$  so that  $\langle |\xi_t|, \mu^{M(t)} \rangle$  is infinite if  $\mu$  has infinite  $k^{\text{th}}$  moment. This shows that we cannot expect to have  $\theta_k \rightarrow +\infty$  when  $k \rightarrow +\infty$  if  $\mu$  has not all its moments finite. This means that we cannot expect that the duality method could give uniqueness for weaker assumptions on the initial condition.

The proofs of Lemma 2.3.5 and Theorem 2.3.4 are respectively given in Sections 2.6.2 and 2.6.3. Once they are proved, the proof of uniqueness can be completed as follows.

### 2.3.3 Proof of Theorem 2.3.1 from Theorem 2.3.4 and Lemma 2.3.5

We rely on the ETHIER-KURTZ's result [94, Theorem 4.4.2]. To get the desired result, i.e. the uniqueness of the martingale problem (2.4), it is sufficient to verify that if we give ourselves two solutions to the martingale problem (2.4), they have the same 1-dimensional marginal laws.

Let  $(X_t)_{t \geq 0}$  and  $(\tilde{X}_t)_{t \geq 0}$  be two solutions of the martingale problem (2.4) with the same initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  which has all its moments finite. Let  $(\xi_t)_{t \geq 0}$  be the dual process with initial condition  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  with  $M(0) \in \mathbb{N}^*$ . We suppose that these three processes are built on the same probability space and independent of each other. We denote for all  $k \in \mathbb{N}$ ,

$$\tilde{\theta}_k := \inf \left\{ t \geq 0 \mid M(t) \geq k \quad \text{or} \quad \exists s \in [0, t], \langle \xi_s, \tilde{X}_{t-s}^{M(s)} \rangle \geq k \right\}.$$

These processes satisfy, for all  $k \in \mathbb{N}$ ,  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  the weakened duality identity (2.34):

$$\begin{aligned} \forall t \geq 0, \quad \mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_0, X_{t \wedge \theta_k \wedge \tilde{\theta}_k}^{M(0)} \right\rangle \right) \\ = \mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_{t \wedge \theta_k \wedge \tilde{\theta}_k}, \mu^{M(t \wedge \theta_k \wedge \tilde{\theta}_k)} \right\rangle \exp \left( \gamma \int_0^{t \wedge \theta_k \wedge \tilde{\theta}_k} M^2(u) du \right) \right) \\ = \mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_0, \tilde{X}_{t \wedge \theta_k \wedge \tilde{\theta}_k}^{M(0)} \right\rangle \right). \end{aligned}$$

From Lemma 2.3.5, since  $(X_t)_{t \geq 0}$  and  $(\tilde{X}_t)_{t \geq 0}$  have continuous paths for the topology of weak convergence, we have  $\mathbb{P}_{(\mu, \xi_0)}$ -a.s.,

$$\lim_{k \rightarrow +\infty} \left\langle \xi_0, X_{t \wedge \theta_k \wedge \tilde{\theta}_k}^{M(0)} \right\rangle = \left\langle \xi_0, X_t^{M(0)} \right\rangle \quad \text{and} \quad \lim_{k \rightarrow +\infty} \left\langle \xi_0, \tilde{X}_{t \wedge \theta_k \wedge \tilde{\theta}_k}^{M(0)} \right\rangle = \left\langle \xi_0, \tilde{X}_t^{M(0)} \right\rangle.$$

Therefore, we deduce from the dominated convergence theorem and (2.17), that for all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ ,

$$\forall t \geq 0, \quad \mathbb{E}_{(\mu, \xi_0)} \left( P_{\xi_0, M(0)}(X_t) \right) = \mathbb{E}_{(\mu, \xi_0)} \left( P_{\xi_0, M(0)}(\tilde{X}_t) \right).$$

As the set of test functions  $\left\{ P_{f,n} \mid f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}), n \in \mathbb{N}^* \right\}$  is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -convergence determining [60, Lemma 2.1.2], it is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -separating [94, Chapter 3, Section 4, p.112], it follows that for any  $t \geq 0$ ,  $X_t^{M(0)}$  and  $\tilde{X}_t^{M(0)}$  have the same law. In particular, for the choice  $M(0) := 1$ , [94, Theorem 4.4.2] ensures uniqueness to the martingale problem (2.4).  $\square$

## 2.4 Ergodicity for the centered Fleming-Viot process

In this section, we establish ergodicity properties with exponential convergence in total variation for the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$ . Note that in the case of the original FLEMING-VIOT process, ergodicity fails without the centering property [91, Section 9.1]. Standard duality arguments would provide weak ergodicity estimates (see (2.1) and [91]). However, using a coupling argument based on the MORAN process and its relationship with the KINGMAN coalescent, it is possible to obtain strong ergodicity bounds (see (2.2) and [91]) as done below. In addition, this will provide an explicit construction of the invariant probability measure from the DONNELLY-KURTZ modified look-down [78, 80]. For all  $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$ , we denote by

$$\|\mu - \nu\|_{TV} := \frac{1}{2} \sup_{\|f\|_\infty \leq 1} |\langle f, \mu \rangle - \langle f, \nu \rangle|,$$

the total variation distance between  $\mu$  and  $\nu$ . The main result of this section is the following.

**Theorem 2.4.1.** *There exists a unique invariant probability measure  $\pi$  for  $(Z_t)_{t \geq 0}$  and constants  $\alpha, \beta \in (0, +\infty)$  such that for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,*

$$\|\mathbb{P}_\mu(Z_T \in \cdot) - \pi\|_{TV} \leq \alpha \exp(-\beta T).$$

The main part of this section is devoted to the proof of this result (Sections 2.4.1 to 2.4.5) and in Section 2.4.6, we characterise the invariant measure of the centered FLEMING-VIOT process. In Section 2.4.1 we construct the centered MORAN process and we establish its convergence in law to the centered FLEMING-VIOT process. In Sections 2.4.2 and 2.4.3, we construct, backward in time, the MORAN process, its centered version and we exploit its relationship with the KINGMAN coalescent in order to prove in Section 2.4.4 an exponential coupling in total variation for the MORAN process. We finally deduce, in Section 2.4.5, the proof of the main result announced by letting the number of particles go to infinity.

### 2.4.1 Moran's models and Fleming-Viot's processes

In [85, 61, 99, 60], the authors construct the original FLEMING-VIOT process as a scaling limit of a particle process: the MORAN process. The aim of this section is to construct the version of the centered MORAN process and to establish its convergence in law to the centered FLEMING-VIOT process.

We consider the MORAN particle process  $Y^N$  defined by

$$Y_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)}$$

with state space  $\mathcal{M}_{1,N}(\mathbb{R})$ , the set of probability measures on  $\mathbb{R}$  consisting of  $N$  atoms of mass  $1/N$ . Moreover, if  $(X_i(0))_{i \in \mathbb{N}^*}$  is exchangeable, then for all  $t > 0$ ,  $(X_i(t))_{i \in \mathbb{N}^*}$  is exchangeable [91, Theorem 6.1]. The infinitesimal generator of the  $\mathbb{R}$ -measure-valued process  $Y^N$  is given for all  $n \in \mathbb{N}^*$ ,  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,  $\mu_N \in \mathcal{M}_{1,N}(\mathbb{R})$  by

$$\begin{aligned} L_N P_{f,n}(\mu_N) &:= P_{\Delta_f, n}(\mu_N) \\ &+ \lambda N(N-1) \int_{\mathbb{R}} \int_{\mathbb{R}} \left[ P_{f,n} \left( \mu_N - \frac{\delta_x}{N} + \frac{\delta_y}{N} \right) - P_{f,n}(\mu_N) \right] \mu_N(dx) \mu_N(dy). \end{aligned}$$



The first term of the generator describes the effect of the mutation according to the Laplacian operator. The second term describes the sampling replacement mechanism: at rate  $\lambda$  (the sampling rate) an individual of type  $x$  is immediately replaced by one of type  $y$ . Note that the population size remains constant over time.

We recall the following convergence result [60, Theorem 2.7.1]: for all initial condition  $Y_0^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i} \in \mathcal{M}_{1,N}(\mathbb{R})$  with  $(X_i)_{1 \leq i \leq N}$  exchangeable  $\mathbb{R}$ -valued random variables such that  $Y_0^N$  converges in law to  $\mu \in \mathcal{M}_1(\mathbb{R})$  as  $N \rightarrow +\infty$ , the MORAN process  $(Y_t^N)_{t \geq 0}$  converges in law on  $\mathcal{C}^0([0, +\infty), \mathcal{M}_1(\mathbb{R}))$ , as  $N \rightarrow +\infty$ , to the original FLEMING-VIOT process  $(Y_t)_{t \geq 0}$  defined as the solution to the martingale problem (2.24).

We denote  $\mathcal{M}_{1,N}^{c,2}(\mathbb{R}) := \left\{ \mu_N \in \mathcal{M}_{1,N}(\mathbb{R}) \mid \langle \text{id}^2, \mu_N \rangle < \infty, \langle \text{id}, \mu_N \rangle = 0 \right\}$ , and we define the centered MORAN process  $(Z_t^N)_{t \geq 0}$  by

$$Z_t^N := \tau_{-\langle \text{id}, Y_t^N \rangle} \# Y_t^N, \quad t \geq 0.$$

The main result of this section is the following:

**Proposition 2.4.2.** *For all initial condition  $Z_0^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_i} \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$  with  $(X_i)_{1 \leq i \leq N}$  exchangeable  $\mathbb{R}$ -valued random variables such that  $Z_0^N$  converges in law to  $Z_0 \in \mathcal{M}_1^{c,2}(\mathbb{R})$  as  $N \rightarrow +\infty$  and satisfying  $\sup_{N \in \mathbb{N}^*} \mathbb{E}(\langle \text{id}^2, Z_0^N \rangle) < \infty$ , the centered MORAN process  $(Z_t^N)_{t \geq 0}$  converges in law on  $\mathcal{C}^0([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R}))$ , as  $N \rightarrow +\infty$ , to the centered FLEMING-VIOT process  $(Z_t)_{t \geq 0}$  solution of the martingale problem (2.20) with initial condition  $Z_0$ .*

A difficulty in proving this result lies in the fact that  $\mu \mapsto \tau_{-\langle \text{id}, \mu \rangle} \# \mu$  may not be continuous on  $\mathcal{M}_1(\mathbb{R})$  because  $\text{id}$  is not bounded. Hence we need to carry out several approximations and be very careful to control the approximation error on events of large probabilities. In order to prove this proposition, we need to introduce some notations and results. For all real-valued function  $f$  on  $\mathbb{R}$ , the LIPSCHITZ seminorm is defined by  $\|f\|_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$ . We denote by

$$BL_1(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid \|f\|_{BL} \leq 1 \right\}$$

where  $\|f\|_{BL} := \|f\|_L + \|f\|_\infty$ . For all  $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$ , we denote by

$$d_{FM}(\mu, \nu) := \sup_{f \in BL_1(\mathbb{R})} |\langle f, \mu \rangle - \langle f, \nu \rangle|,$$

the FORTET-MOURIER distance. Recall that  $\mathcal{M}_1(\mathbb{R})$  endowed with the weak topology is complete for the distance of FORTET-MOURIER [81, Corollary 11.5.5]. Let  $\Lambda$  denote the class of strictly increasing, continuous mappings of  $[0, T]$  onto itself. For given metric spaces  $E$  and  $\tilde{E}$ , we denote by  $\mathbb{D}([0, T], E)$ , the space of right continuous and left limited (càd-làg) functions from  $[0, T]$  to  $E$  and by  $\mathcal{C}_b^0(E, \tilde{E})$  the space of continuous bounded functions from  $E$  to  $\tilde{E}$ . For  $x, y \in \mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R}))$ , we define the distance  $d_0(x, y)$  by:

$$d_0(x, y) := \inf_{\rho \in \Lambda} \left\{ \sup_{t \in [0, T]} d_{FM}(x \circ \rho(t), y(t)) + \sup_{s < t} \left| \log \left( \frac{\rho(t) - \rho(s)}{t - s} \right) \right| \right\}.$$

From [20, Theorem 12.2 and Remark of page 121],  $(\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R})), d_0)$  is a Polish space when  $\mathcal{M}_1(\mathbb{R})$  is endowed with the FORTET-MOURIER distance and the topology induced by the distance  $d_0$  is the SKOROHOD topology. Let us consider the following lemma whose proof is similar to that of Proposition 2.2.11 and left to the reader:

**Lemma 2.4.3.** *Let  $T > 0$  and  $k \in \mathbb{N}$  fixed. There exists a constant  $C_{k,T} > 0$  independent of  $N$  such that for all  $Y_0^N \in \mathcal{M}_{1,N}(\mathbb{R})$  satisfying  $\sup_{N \in \mathbb{N}} \mathbb{E} \left( \langle |\text{id}|^k, Y_0^N \rangle \right) < \infty$ , the MORAN process  $(Y_t^N)_{0 \leq t \leq T}$  satisfies*

$$\forall \alpha > 0, \quad \mathbb{P}_{Y_0^N} \left( \sup_{t \in [0, T]} \langle |\text{id}|^k, Y_t^N \rangle \geq \alpha \right) \leq \frac{C_{k,T} \sup_{N \in \mathbb{N}^*} \mathbb{E} \left( \langle |\text{id}|^k, Y_0^N \rangle \right)}{\alpha}.$$

*Proof of Proposition 2.4.2.* We want to establish that

$$\forall g \in \mathcal{C}_b^0(\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R})), \mathbb{R}), \quad \lim_{N \rightarrow +\infty} \mathbb{E} \left( g(Z^N) \right) = \mathbb{E} \left( g(Z) \right).$$

Let  $\varepsilon > 0$ . We consider the two following maps  $F$  and  $F_\varepsilon$  from  $\mathbb{D}([0, T], \mathcal{M}_1^1(\mathbb{R}))$  to  $\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R}))$  defined by  $F(y)(t) := \tau_{-\langle \text{id}, y(t) \rangle} \# y(t)$  and  $F_\varepsilon(y)(t) := \tau_{-\langle h_\varepsilon, y(t) \rangle} \# y(t)$  where  $h_\varepsilon$  is a map from  $\mathbb{R}$  to  $\mathbb{R}$  defined by

$$h_\varepsilon(x) := \begin{cases} x & \text{if } |x| \leq \frac{1}{\varepsilon} \\ \frac{1}{\varepsilon} & \text{if } x > \frac{1}{\varepsilon} \\ -\frac{1}{\varepsilon} & \text{if } x < -\frac{1}{\varepsilon} \end{cases}.$$

**Step 1. Continuity of  $F_\varepsilon$ .** In this step, we want to establish that

$$F_\varepsilon \in \mathcal{C}_b^0(\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R})), \mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R}))).$$

To obtain this, it is equivalent to prove that if for all  $n \in \mathbb{N}$ ,  $y_n, y \in \mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R}))$  and  $\lim_{n \rightarrow +\infty} d_0(y_n, y) = 0$ , we have  $\lim_{n \rightarrow +\infty} d_0(F_\varepsilon(y_n), F_\varepsilon(y)) = 0$ . As,  $\lim_{n \rightarrow +\infty} d_0(y_n, y) = 0$ , there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ , there exists  $\rho_n \in \Lambda$  satisfying

$$\sup_{t \in [0, T]} d_{FM}(y_n \circ \rho_n(t), y(t)) + \sup_{s < t} \left| \log \left( \frac{\rho_n(t) - \rho_n(s)}{t - s} \right) \right| \leq \frac{\varepsilon^2}{2}. \quad (2.35)$$

Note that

$$d_0(F_\varepsilon(y_n), F_\varepsilon(y)) \leq \sup_{t \in [0, T]} d_{FM}(F_\varepsilon(y_n)(\rho_n(t)), F_\varepsilon(y)(t)) + \sup_{s < t} \left| \log \left( \frac{\rho_n(t) - \rho_n(s)}{t - s} \right) \right|.$$

Now, for all  $t \in [0, T]$ ,

$$\begin{aligned} d_{FM}(F_\varepsilon(y_n)(\rho_n(t)), F_\varepsilon(y)(t)) &= \sup_{f \in BL_1(\mathbb{R})} \left| \langle f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle}, y_n \circ \rho_n(t) \rangle - \langle f \circ \tau_{-\langle h_\varepsilon, y(t) \rangle}, y(t) \rangle \right| \\ &\leq \sup_{f \in BL_1(\mathbb{R})} \left| \langle f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle}, y_n \circ \rho_n(t) - y(t) \rangle \right| \\ &\quad + \sup_{f \in BL_1(\mathbb{R})} \left| \langle f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle} - f \circ \tau_{-\langle h_\varepsilon, y(t) \rangle}, y(t) \rangle \right|. \end{aligned}$$

On the one hand, as  $f \in BL_1(\mathbb{R})$ , it follows that  $f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle} \in BL_1(\mathbb{R})$  and thus

$$\sup_{f \in BL_1(\mathbb{R})} \left| \langle f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle}, y_n \circ \rho_n(t) - y(t) \rangle \right| \leq d_{FM}(y_n \circ \rho_n(t), y(t)).$$

On the other hand, as  $f$  and  $\varepsilon h_\varepsilon$  are in  $BL_1(\mathbb{R})$ , we have

$$\begin{aligned} \left| \langle f \circ \tau_{-\langle h_\varepsilon, y_n \circ \rho_n(t) \rangle} - f \circ \tau_{-\langle h_\varepsilon, y(t) \rangle}, y(t) \rangle \right| &\leq | - \langle h_\varepsilon, y_n \circ \rho_n(t) \rangle + \langle h_\varepsilon, y(t) \rangle | \\ &= \frac{1}{\varepsilon} | \langle \varepsilon h_\varepsilon, y_n \circ \rho_n(t) - y(t) \rangle | \\ &\leq \frac{1}{\varepsilon} d_{FM}(y_n \circ \rho_n(t), y(t)). \end{aligned}$$

It follows from (2.35) that,

$$\begin{aligned} d_0(F_\varepsilon(y_n), F_\varepsilon(y)) &\leq \left(1 + \frac{1}{\varepsilon}\right) \sup_{t \in [0, T]} d_{FM}(y_n \circ \rho_n(t), y(t)) + \sup_{s < t} \left| \log \left( \frac{\rho_n(t) - \rho_n(s)}{t - s} \right) \right| \\ &\leq \varepsilon. \end{aligned}$$

**Step 2. Control in distance  $d_0$  of the difference between  $F(Y^N)$  and  $F_\varepsilon(Y^N)$ .** We consider the MORAN process  $(Y_t^N)_{0 \leq t \leq T}$  started from  $Y_0^N = Z_0^N$  and the original FLEMING-VIOT process  $(Y_t)_{0 \leq t \leq T}$  started from  $Y_0 = Z_0$ . In this step, we consider the events

$$\Omega_{\varepsilon, N} := \left\{ \sup_{t \in [0, T]} \langle \text{id}^2, Y_t^N \rangle \leq \frac{2}{\sqrt{\varepsilon}} \right\} \quad \text{and} \quad \Omega_{\varepsilon, \infty} := \left\{ \sup_{t \in [0, T]} \langle \text{id}^2, Y_t \rangle \leq \frac{2}{\sqrt{\varepsilon}} \right\}.$$

As  $\sup_{N \in \mathbb{N}} \mathbb{E}(\langle \text{id}^2, Z_0^N \rangle) < \infty$ , it follows from Lemma 2.4.3 (respectively Proposition 2.2.11), there exists a constant  $\tilde{C}_T > 0$ , independent of  $N$ , such that  $\mathbb{P}_{Y_0^N}(\Omega_{\varepsilon, N}) \geq 1 - \tilde{C}_T \sqrt{\varepsilon}$  (respectively  $\mathbb{P}_{Y_0}(\Omega_{\varepsilon, \infty}) \geq 1 - \tilde{C}_T \sqrt{\varepsilon}$ ). Moreover, on  $\Omega_{\varepsilon, N}$ , for all  $t \in [0, T]$ ,

$$\begin{aligned} d_0(F(Y^N), F_\varepsilon(Y^N)) &\leq \sup_{t \in [0, T]} d_{FM}(F(Y^N)(t), F_\varepsilon(Y^N)(t)) \\ &\leq \sup_{t \in [0, T]} \sup_{f \in BL_1(\mathbb{R})} \left\| f \circ \tau_{-\langle \text{id}, Y_t^N \rangle} - f \circ \tau_{-\langle h_\varepsilon, Y_t^N \rangle} \right\|_\infty \\ &\leq \sup_{t \in [0, T]} \left\| \langle |h_\varepsilon - \text{id}|, Y_t^N \rangle \right\|_\infty \\ &\leq \frac{\varepsilon}{2} \sup_{t \in [0, T]} \left\| \langle \text{id}^2, Y_t^N \rangle \right\|_\infty \\ &\leq \sqrt{\varepsilon}, \end{aligned}$$

where we used the inequality  $|h_\varepsilon - \text{id}| \leq \frac{\varepsilon}{2} \text{id}^2$ . Similarly on  $\Omega_{\varepsilon, \infty}$ ,

$$d_0(F(Y), F_\varepsilon(Y)) \leq \sqrt{\varepsilon}.$$

**Step 3. Conclusion.** We want to prove that for all  $g \in \mathcal{C}_b^0(\mathbb{D}([0, T], \mathcal{M}_1(\mathbb{R})), \mathbb{R})$ ,

$$\lim_{N \rightarrow +\infty} \mathbb{E}(g(Z^N)) = \mathbb{E}(g(Z)).$$

Thanks to the PORTMANTEAU theorem [20, Theorem 2.1], then it is sufficient to prove this for all  $g$  1-LIPSCHITZ. As  $(Y^N)_{N \in \mathbb{N}^*}$  converges in law to  $Y$ , we deduce that, for  $N$  large enough,

$$\left| \mathbb{E}(g \circ F_\varepsilon(Y^N)) - \mathbb{E}(g \circ F_\varepsilon(Y)) \right| \leq \sqrt{\varepsilon}.$$

Using that  $g$  is 1-LIPSCHITZ, and the inequalities of Step 2, it follows that

$$\begin{aligned} \left| \mathbb{E}(g(Z^N)) - \mathbb{E}(g(Z)) \right| &= \left| \mathbb{E}(g \circ F(Y^N)) - \mathbb{E}(g \circ F(Y)) \right| \\ &\leq \left| \mathbb{E}(g \circ F_\varepsilon(Y^N)) - \mathbb{E}(g \circ F_\varepsilon(Y)) \right| \\ &\quad + \left| \mathbb{E}(g \circ F(Y^N)) - \mathbb{E}(g \circ F_\varepsilon(Y^N)) \right| \\ &\quad + \left| \mathbb{E}(g \circ F_\varepsilon(Y)) - \mathbb{E}(g \circ F(Y)) \right| \\ &\leq \sqrt{\varepsilon} + 2 \|g\|_\infty \left[ \mathbb{P}_{Y_0^N}(\Omega_{\varepsilon, N}^c) + \mathbb{P}_{Y_0}(\Omega_{\varepsilon, \infty}^c) \right] \\ &\quad + \sqrt{\varepsilon} \left[ \mathbb{P}_{Y_0^N}(\Omega_{\varepsilon, N}) + \mathbb{P}_{Y_0}(\Omega_{\varepsilon, \infty}) \right] \\ &\leq (3 + 2 \|g\|_\infty \tilde{C}_T) \sqrt{\varepsilon}. \end{aligned}$$

The announced result follows and completes the proof.  $\square$

## 2.4.2 Backward construction of Moran's process and Kingman's coalescent

In this section, we exploit the relationship between the MORAN model and the KINGMAN coalescent to obtain in Proposition 2.4.7 a result of exponential ergodicity in total variation for the centered MORAN model uniformly in  $N$ . The genealogy of a sample from a population evolving according to the MORAN model of Section 2.4.1 is exactly determined by KINGMAN's coalescent with coalescence rate  $2\lambda$ . The state of the population at the final time is constructed from the ancestral positions by following the genealogy and adding mutations on the genealogical tree of the sample according to a standard Brownian motion. Therefore, at the final time  $T$ , the position of each individual of the sample is equal to the sum of the position of its ancestor at time 0 and the Brownian mutation steps that occurred during its ancestral branches in the coalescent. We formalise this construction below by giving some notations and illustrations of the latter ones.

Let be fixed the time  $T > 0$  and the number of particles  $N \in \mathbb{N}^*$ . We consider the probability space  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$  where  $\widehat{\Omega} := \mathcal{K}_{N,T} \times \text{Mut}_{N,T}^{2^N-1} \times \mathcal{A}_N^{\mu_N}$ ,  $\widehat{\mathcal{F}} := \mathcal{F}_{\mathcal{K}_{N,T}} \otimes \mathcal{F}_{\text{Mut}_{N,T}} \otimes \mathcal{F}_{\mathcal{A}_N^{\mu_N}}$  and  $\widehat{\mathbb{P}} := \mathbb{K}_{N,T} \otimes \mathcal{L}_{\text{Mut}}^{\otimes 2^N-1} \otimes \bigotimes_{i=1}^N \text{Ech}_i^{\mu_N}$  with the following notations.

We denote by  $\mathcal{K}_{N,T}$  the state space of the KINGMAN  $N$ -coalescent with coalescence rate  $2\lambda$  on  $[0, T]$  :

$$\mathcal{K}_{N,T} := \mathbb{D}([0, T], \Pi_N)$$

with  $\Pi_N$  the set of partitions of  $\{1, \dots, N\}$ . We denote by  $\mathbb{K}_{N,T}$  the law of the KINGMAN  $N$ -coalescent with coalescence rate  $2\lambda$  on  $[0, T]$  and  $\mathcal{F}_{\mathcal{K}_{N,T}}$  the SKOROHOD  $\sigma$ -field on  $\mathcal{K}_{N,T}$ . In the following, in order to simplify the names, the precision "at coalescence rate  $2\lambda$ " will be omitted.

We denote by  $\mathcal{A}_N^{\mu_N} := \prod_{i=1}^N \mathbb{R}^i$  the state space of the possible ancestral positions in the KINGMAN  $N$ -coalescent at time 0. We denote by  $\text{Ech}_i^{\mu_N}$  with  $\mu_N := \frac{1}{N} \sum_{j=1}^N \delta_{x_j} \in \mathcal{M}_{1,N}(\mathbb{R})$ , the law of a  $i$ -sample where we select  $i$  elements randomly and without replacement in  $\{x_1, \dots, x_N\}$  according to  $\mu_N$  and by  $\mathcal{F}_{\mathcal{A}_N^{\mu_N}}$  the BOREL  $\sigma$ -field on  $\mathcal{A}_N^{\mu_N}$ .

We denote by  $\text{Mut}_{N,T}$  the set of Brownian trajectories of  $\mathcal{C}^0([0, T], \mathbb{R})$ . We denote by  $\mathcal{L}_{\text{Mut}}$  the law of a 1-standard Brownian motion and  $\mathcal{F}_{\text{Mut}_{N,T}}$  the BOREL  $\sigma$ -field on  $\text{Mut}_{N,T}$ .

Finally, we denote an element  $\widehat{\omega}$  of  $\widehat{\Omega}$  as:

$$\widehat{\omega} := \left\{ (k_{N,t})_{0 \leq t \leq T}; \left( W^{(B)} \right)_{B \subset \{1, \dots, N\} \setminus \emptyset}; x^{(1)}, \dots, x^{(N)} \right\},$$

where we denote  $x^{(k)} := (x_1^{(k)}, x_2^{(k)}, \dots, x_k^{(k)})$ . Note that  $(k_{N,t})_{0 \leq t \leq T}$  is an increasing process in  $\mathcal{K}_{N,T}$  representing the genealogy. In addition, we denote by  $k_{N,T} := \{L_1, \dots, L_n\}$  with  $n := |k_{N,T}|$ . In other words, KINGMAN's  $N$ -coalescent contains  $n$  distinct lineages where  $L_i$  is a subset of  $\{1, \dots, N\}$  at the final time  $T$ . For each individual  $i \in \{1, \dots, N\}$ , we denote by  $a(i) \in \{1, \dots, n\}$ , the index such that  $i \in L_{a(i)}$ . In other words,  $L_{a(i)}$  is the ancestral lineage of  $i$ . For each individual  $i \in \{1, \dots, N\}$ , we denote by  $B(s, i)$  the block of the partition  $k_{N,s}$  at time  $s$  and containing the individual  $i$ . We denote by  $x^{(n)} = (x_1^{(n)}, \dots, x_n^{(n)})$  the ancestral positions at the final time  $T$  so that for all  $i \in \{1, \dots, N\}$ ,  $x_{a(i)}^{(n)} \in \mathbb{R}$  is the position of the ancestor of the individual  $i$ , at time  $T$ , in the genealogical tree. Each process  $W^{(B)} = \left( W_t^{(B)} \right)_{0 \leq t \leq T}$  governs the

dynamics of mutations occurring on the interval time where  $B \subset k_{N,T}$ .

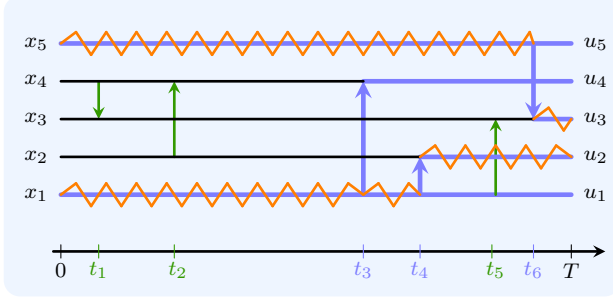


FIGURE 2.1 – Graphical representation of the MORAN model with  $N = 5$ .

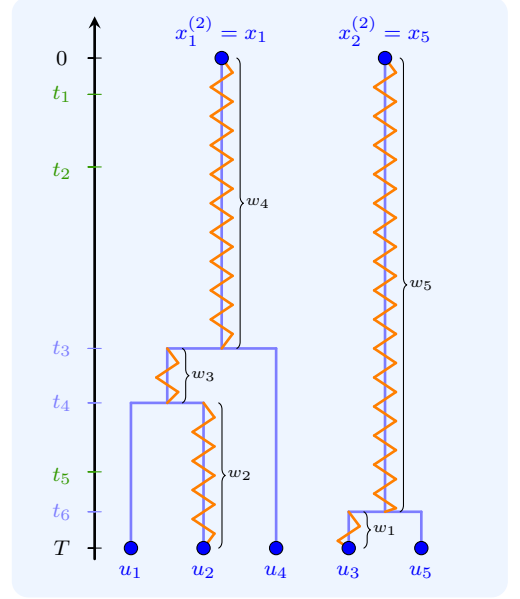


FIGURE 2.2 – KINGMAN's genealogy  $(k_{5,T-t})_{0 \leq t \leq T}$  under the MORAN model on the left, tracing back from  $T$  to  $0$ .

Let us recall the link between the MORAN model and the previous stochastic objects. At each reproduction event  $t_k$  in the MORAN model, illustrated on Figure 2.1, an ordered pair of individuals  $(i, j)$  is sampled uniformly at random from the population: one of the two individuals dies and the other reproduces with equal probabilities. In Figure 2.1, we draw an arrow between lines: the arrow  $i \rightarrow j$  indicates that  $i$  reproduced and  $j$  died. We can recover the ancestry of the sample by tracing backwards in time from the right to the left in Figure 2.1 to obtain Figure 2.2. We coalesce any pair of individuals whenever they find a common ancestor which is represented by the bold blue arrows. The other arrows do not modify the genealogical tree. The ancestral positions in Figure 2.1 are given by  $(x_i)_{1 \leq i \leq 5}$  where  $\mu_5 = \frac{1}{5} \sum_{i=1}^5 \delta_{x_i}$  and the ancestral positions in Figure 2.2 are distributed as  $x^{(2)}$ . Note that  $x_{a(1)}^{(2)} = x_{a(2)}^{(2)} = x_{a(4)}^{(2)} = x_1^{(2)} = x_1$  and  $x_{a(1)}^{(2)} = x_{a(1)}^{(2)} = x_2^{(2)} = x_5$ . Then, we add the mutations, denoted by “ $\wedge$ ”. Not all of them are shown in Figures 2.1 and 2.2 for the sake of clarity: the ones which are represented are  $w_1 := W_{T-t_6}^{\{\{3\}\}} - W_0^{\{\{3\}\}}$ ,  $w_2 := W_{T-t_4}^{\{\{2\}\}} - W_0^{\{\{2\}\}}$ ,  $w_3 := W_{T-t_3}^{\{\{1,2\}\}} - W_{T-t_4}^{\{\{1,2\}\}}$ ,  $w_4 := W_T^{\{\{1,2,4\}\}} - W_{T-t_3}^{\{\{1,2,4\}\}}$  and  $w_5 := W_T^{\{\{3,5\}\}} - W_{T-t_6}^{\{\{3,5\}\}}$ . Then the position  $u_i$  of the individual  $i \in \{1, \dots, 5\}$  at time  $T$  in Figure 2.2, are

$$\begin{aligned} u_1 &= x_1 + W_T^{\{\{1,2,4\}\}} - W_{T-t_3}^{\{\{1,2,4\}\}} + W_{T-t_3}^{\{\{1,2\}\}} - W_{T-t_4}^{\{\{1,2\}\}} + W_{T-t_4}^{\{\{1\}\}} - W_0^{\{\{1\}\}}, \\ u_2 &= x_1 + W_T^{\{\{1,2,4\}\}} - W_{T-t_3}^{\{\{1,2,4\}\}} + W_{T-t_3}^{\{\{1,2\}\}} - W_{T-t_4}^{\{\{1,2\}\}} + W_{T-t_4}^{\{\{2\}\}} - W_0^{\{\{2\}\}}, \\ u_3 &= x_2 + W_T^{\{\{3,5\}\}} - W_{T-t_6}^{\{\{3,5\}\}} + W_{T-t_6}^{\{\{3\}\}} - W_0^{\{\{3\}\}}, \\ u_4 &= x_1 + W_T^{\{\{1,2,4\}\}} - W_{T-t_3}^{\{\{1,2,4\}\}} + W_{T-t_3}^{\{\{4\}\}} - W_0^{\{\{4\}\}}, \\ u_5 &= x_2 + W_T^{\{\{3,5\}\}} - W_{T-t_6}^{\{\{3,5\}\}} + W_{T-t_6}^{\{\{5\}\}} - W_0^{\{\{5\}\}}. \end{aligned}$$

Putting everything together, we define in the general case the random variable

$$\widehat{Y}_T^{N,\mu_N} := \frac{1}{N} \sum_{i=1}^N \delta_{u_i},$$

where, for all  $i \in \{1, \dots, N\}$ ,

$$u_i := u_i^{\mu_N} := x_{a(i)}^{(n)} + \int_0^T dW_s^{(B(s^-, i))}. \quad (2.36)$$

The well-known backward construction of the MORAN model [87, Section 1.2], [86, Section 2.8] entails the following result.

**Proposition 2.4.4.** *For all initial condition  $\mu_N \in \mathcal{M}_{1,N}(\mathbb{R})$ ,  $Y_T^N \stackrel{\text{law}}{=} \widehat{Y}_T^{N,\mu_N}$  where  $Y_0^N := \mu_N$ .*

### 2.4.3 Centered variables and centering effects

We construct the centered version of the random variables  $\widehat{Y}_T^{N,\mu_N}$ . We define the random variable  $\widehat{Z}_T^{N,\mu_N}$  as follows:

$$\widehat{Z}_T^{N,\mu_N} := \tau_{-\langle \text{id}, \widehat{Y}_T^{N,\mu_N} \rangle \#} \widehat{Y}_T^{N,\mu_N}.$$

**Corollary 2.4.5.** *For all initial condition  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ ,  $Z_T^N \stackrel{\text{law}}{=} \widehat{Z}_T^{N,\mu_N} = \frac{1}{N} \sum_{i=1}^N \delta_{v_i}$  where, for all  $i \in \{1, \dots, N\}$ ,  $v_i := v_i^{\mu_N} := u_i - \frac{1}{N} \sum_{j=1}^N u_j$  and  $Z_0^N = \mu_N$ .*

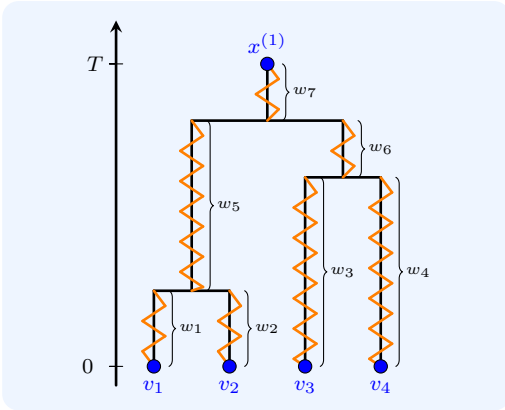


FIGURE 2.3 – Illustration of the centered MORAN process where  $|k_{4,T}| = 1$ .

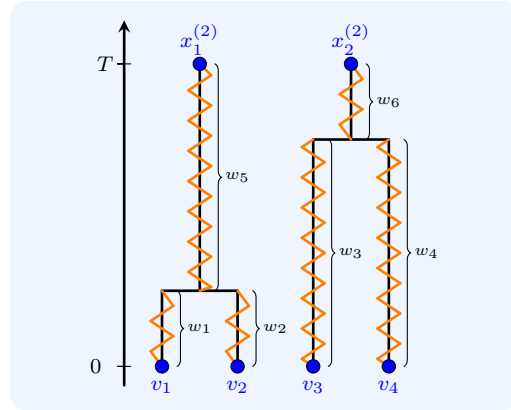


FIGURE 2.4 – Illustration of the centered MORAN process where  $|k_{4,T}| = 2$ .

In the example of Figure 2.3, we can observe that  $u_1 = x^{(1)} + w_1 + w_5 + w_7$ ,  $u_2 = x^{(1)} + w_2 + w_5 + w_7$ ,  $u_3 = x^{(1)} + w_3 + w_6 + w_7$ ,  $u_4 = x^{(1)} + w_4 + w_6 + w_7$ , so that

- $v_1 = \frac{1}{2} [w_5 - w_6] + \frac{3}{4} w_1 - \frac{1}{4} [w_2 + w_3 + w_4]$ ,
- $v_2 = \frac{1}{2} [w_5 - w_6] + \frac{3}{4} w_2 - \frac{1}{4} [w_1 + w_3 + w_4]$ ,
- $v_3 = -\frac{1}{2} [w_5 - w_6] + \frac{3}{4} w_3 - \frac{1}{4} [w_1 + w_2 + w_4]$ ,
- $v_4 = -\frac{1}{2} [w_5 - w_6] + \frac{3}{4} w_4 - \frac{1}{4} [w_1 + w_2 + w_3]$ ,

and for Figure 2.4 that  $u_1 = x_1^{(2)} + w_1 + w_5$ ,  $u_2 = x_1^{(2)} + w_2 + w_5$ ,  $u_3 = x_2^{(2)} + w_3 + w_6$ ,

$u_4 = x_2^{(2)} + w_4 + w_6$  so that

$$\begin{aligned} \bullet v_1 &= \frac{1}{2} [x_1^{(2)} - x_2^{(2)} + w_5 - w_6] + \frac{3}{4}w_1 - \frac{1}{4} [w_2 + w_3 + w_4], \\ \bullet v_2 &= \frac{1}{2} [x_1^{(2)} - x_2^{(2)} + w_5 - w_6] + \frac{3}{4}w_2 - \frac{1}{4} [w_1 + w_3 + w_4], \\ \bullet v_3 &= -\frac{1}{2} [x_1^{(2)} - x_2^{(2)} + w_5 - w_6] + \frac{3}{4}w_3 - \frac{1}{4} [w_1 + w_2 + w_4], \\ \bullet v_4 &= -\frac{1}{2} [x_1^{(2)} - x_2^{(2)} + w_5 - w_6] + \frac{3}{4}w_4 - \frac{1}{4} [w_1 + w_2 + w_3]. \end{aligned}$$

In other words, when there is just one ancestral lineage as in Figure 2.3, the random variable  $\widehat{Z}_T^{N, \mu_4}$  does not depend on the ancestral position  $x^{(1)}$ : this is the centering effect. In general, when  $n = |k_{N,T}| = 1$ ,

$$v_i = \int_0^T dW_s^{(B(s^-, i))} - \frac{1}{N} \sum_{j=1}^N \int_0^T dW_s^{(B(s^-, j))}. \quad (2.37)$$

This property is fundamental to implement coupling arguments leading to strong ergodicity (see Section 2.4.4).

For any fixed  $T > 0$ , for all  $i, j \in \{1, \dots, N\}$ , let us consider  $T_{ij}$  the coalescence time between the individuals  $i$  and  $j$  at time  $T$  in the process  $(k_{N,t})_{0 \leq t \leq T}$ . In the following proposition we establish that  $(v_i)_{1 \leq i \leq N}$  is a centered Gaussian vector whose covariance matrix is an explicit function of  $T_{ij}$ .

**Proposition 2.4.6.** (1) For any fixed  $T > 0$ , for all  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , conditionally to  $k_N := (k_{N,t})_{0 \leq t \leq T}$ , on the event  $\{|k_{N,T}| = 1\}$ , we have for all  $i, j \in \{1, \dots, N\}$ ,  $\text{Cov}(u_i, u_j | k_N) = T - T_{ij}$ .

(2) Conditionally to  $k_N$ , on the event  $\{|k_{N,T}| = 1\}$ ,  $(v_i)_{1 \leq i \leq N} \sim \mathcal{N}^{(N)}(0_{\mathbb{R}^N}, \Sigma)$  where  $\Sigma := (\Sigma_{ij})_{1 \leq i, j \leq N}$  is define by

$$\forall i, j \in \{1, \dots, N\}, \Sigma_{ij} := \text{Cov}(v_i, v_j | k_N) = \frac{1}{N} \sum_{k=1}^N (T_{ik} + T_{jk}) - \left( T_{ij} + \frac{1}{N^2} \sum_{k, \ell=1}^N T_{k\ell} \right).$$

*Proof.* (1) It is a straightforward computation from (2.36).

(2) Noting that for all  $i, j \in \{1, \dots, N\}$ ,  $T_{ij} = T_{ji}$  and  $T_{ii} = 0$  and from (2.37), we deduce the announced result by a straightforward computation.  $\square$

#### 2.4.4 Coupling arguments with two distinct initial conditions

In this section, we want to couple centered MORAN's processes from different initial conditions but with the same KINGMAN genealogy and the same mutations in order to establish the following exponential ergodicity result.

**Proposition 2.4.7.** For all  $\mu_N, \nu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ , there exist constants  $\alpha, \beta \in (0, +\infty)$ , independent of  $\mu_N, \nu_N, T$  and  $N$  such that

$$\left\| \widehat{\mathbb{P}} \left( \widehat{Z}_T^{N, \mu_N} \in \cdot \right) - \widehat{\mathbb{P}} \left( \widehat{Z}_T^{N, \nu_N} \in \cdot \right) \right\|_{TV} \leq \alpha \exp(-\beta T).$$

In particular, for all  $N \in \mathbb{N}^*$  there exists a unique invariant probability measure  $\pi_N$  for the centered MORAN process  $(Z_t^N)_{t \geq 0}$  such that for all  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,

$$\left\| \mathbb{P}_{\mu_N} \left( Z_T^N \in \cdot \right) - \pi_N \right\|_{TV} \leq \alpha \exp(-\beta T).$$

**Remark 2.4.8.** The previous result is true for all deterministic initial conditions, so also for any random initial conditions.

*Proof. Step 1. Coupling.* Let  $\widehat{\omega}_{\mu_N}$  and  $\widehat{\omega}_{\nu_N}$  be two elements of  $\widehat{\Omega}$  which have the same KINGMAN genealogy  $(k_{N,t})_{0 \leq t \leq T}$  and the same mutations but whose initial conditions  $\mu_N := \frac{1}{N} \sum_{i=1}^N \delta_{x_{a(i)}^{(n)}}$  and  $\nu_N := \frac{1}{N} \sum_{i=1}^N \delta_{y_{a(i)}^{(n)}}$  are different where  $n = |k_{N,T}|$ . We assume that the  $x_{a(i)}^{(n)}$  respectively  $y_{a(i)}^{(n)}$  are selected randomly and without replacement in  $\{x_1, \dots, x_N\}$  respectively  $\{y_1, \dots, y_N\}$ , independently. This allows us to construct, on the same probability space, two random variables

$$\widehat{Z}_T^{N,\mu_N} := \frac{1}{N} \sum_{i=1}^N \delta_{v_i^{\mu_N}} \quad \text{and} \quad \widehat{Z}_T^{N,\nu_N} := \frac{1}{N} \sum_{i=1}^N \delta_{v_i^{\nu_N}}$$

such that  $Z_T^{N,\mu_N} \stackrel{\text{law}}{=} \widehat{Z}_T^{N,\mu_N}$  and  $Z_T^{N,\nu_N} \stackrel{\text{law}}{=} \widehat{Z}_T^{N,\nu_N}$ .

**Step 2. Control in total variation.** From (2.37), on the event  $\{|k_{N,T}| = 1\}$ , we have that for all  $i \in \{1, \dots, N\}$ ,  $v_i^{\mu_N} = v_i^{\nu_N}$  a.s. and from [158] we deduce that

$$\begin{aligned} \left\| \widehat{\mathbb{P}} \left( \widehat{Z}_T^{N,\mu_N} \in \cdot \right) - \widehat{\mathbb{P}} \left( \widehat{Z}_T^{N,\nu_N} \in \cdot \right) \right\|_{TV} &\leq \widehat{\mathbb{P}} \left( \widehat{Z}_T^{N,\mu_N} \neq \widehat{Z}_T^{N,\nu_N} \right) \\ &= 1 - \mathbb{K}_{N,T}(|k_{N,T}| = 1). \end{aligned}$$

We denote by  $H_N := \sum_{k=2}^N T_k$  the height of the KINGMAN  $N$ -coalescent where  $(T_k)_{2 \leq k \leq N}$  are independent random variables such that  $T_k$  follows an exponential law of parameter  $2\lambda \binom{k}{2}$  [86, Lemma 2.20]. Now,  $\mathbb{K}_{N,T}(|k_{N,T}| = 1) \geq \mathbb{K}_{\infty,T}(|k_{\infty,T}| = 1) = \mathbb{K}_{\infty,T}(H_{\infty} \leq T)$  and by the exponential TCHEBYCHEV inequality we have

$$\mathbb{K}_{\infty,T}(H_{\infty} > T) \leq \inf_{\rho \in (0, 2\lambda)} \frac{\mathbb{E}(\exp(\rho H_{\infty}))}{\exp(\rho T)}.$$

Note that for all  $\rho \in (0, 2\lambda)$ ,

$$\mathbb{E}(\exp(\rho H_{\infty})) = \prod_{k=2}^{+\infty} \mathbb{E}(\exp(\rho T_k)) = \frac{2\lambda}{2\lambda - \rho} \prod_{k=3}^{+\infty} \frac{1}{1 - \frac{\rho}{\lambda k(k-1)}},$$

where the last product is convergent. We deduce that

$$\mathbb{K}_{\infty,T}(H_{\infty} > T) \leq C \inf_{\rho \in (0, 2\lambda)} \frac{1}{(2\lambda - \rho) \exp(\rho T)} = 2C\lambda \exp(1)T \exp(-2\lambda T),$$

where  $C := \prod_{k=3}^{+\infty} \frac{1}{1 - \frac{\rho}{\lambda k(k-1)}}$ . The result follows for  $\alpha = 2C\lambda \exp(1)T$  and  $\beta = 2\lambda$ .  $\square$



### 2.4.5 Proof of Theorem 2.4.1

Classically, it is sufficient to check that there exists constants  $\alpha, \beta \in \mathbb{R}_+$  such that for all  $\mu, \nu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,

$$\|\mathbb{P}_\mu(Z_T \in \cdot) - \mathbb{P}_\nu(Z_T \in \cdot)\|_{TV} \leq \alpha \exp(-\beta T).$$

From LUSIN's theorem [204, Corollary of Theorem 2.24], Proposition 2.4.7 and Corollary 2.4.5 there exists two constants  $\alpha, \beta \in (0, +\infty)$  such that for all  $\mu_N, \nu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , for all  $T \geq 0$ ,

$$\sup_{\substack{\|f\|_\infty \leq 1 \\ f \text{ continuous}}} \left| \mathbb{E}_{\mu_N}(f(Z_T^N)) - \mathbb{E}_{\nu_N}(f(Z_T^N)) \right| = \sup_{\|f\|_\infty \leq 1} \left| \mathbb{E}_{\mu_N}(f(Z_T^N)) - \mathbb{E}_{\nu_N}(f(Z_T^N)) \right| \leq \alpha \exp(-\beta T).$$

Now, let be fixed two deterministic initial conditions  $\mu, \nu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and consider an i.i.d. sample  $(X_i)_{1 \leq i \leq N}$  of distribution  $\mu$  and an i.i.d. sample  $(\tilde{X}_i)_{1 \leq i \leq N}$  of distribution  $\nu$ . Then, we construct two initial conditions  $\mu_N := \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$  and  $\nu_N := \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{X}_i}$  such that  $\mu_N$  and  $\nu_N$  converge in law respectively to  $\mu$  and  $\nu$ . We define  $\tilde{\mu}_N := \tau_{-\langle \text{id}, \mu_N \rangle} \# \mu_N$  and  $\tilde{\nu}_N := \tau_{-\langle \text{id}, \nu_N \rangle} \# \nu_N$  such that  $\tilde{\mu}_N, \tilde{\nu}_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ . By construction, the assumptions of exchangeability of the random variables  $(X_i)_{1 \leq i \leq N}$  and  $(\tilde{X}_i)_{1 \leq i \leq N}$  are satisfied,  $\tilde{\mu}_N$  and  $\tilde{\nu}_N$  converge in law respectively to  $\mu$  and  $\nu$  and we have

$$\mathbb{E}(\langle \text{id}^2, \tilde{\mu}_N \rangle) = \mathbb{E}(\langle \text{id}^2, \mu_N \rangle - \langle \text{id}, \mu_N \rangle^2) = \left(1 - \frac{1}{N}\right) \text{Var}(X_1) < \infty.$$

Then, we deduce from Proposition 2.4.2 that for all  $f \in \mathcal{C}_b^0(\mathbb{R}, \mathbb{R})$  satisfying  $\|f\|_\infty \leq 1$ , for all  $T \geq 0$ ,

$$|\mathbb{E}_\mu f(Z_T) - \mathbb{E}_\nu f(Z_T)| = \lim_{N \rightarrow +\infty} \left| \mathbb{E}_{\mu_N}(f(Z_T^N)) - \mathbb{E}_{\nu_N}(f(Z_T^N)) \right| \leq \alpha \exp(-\beta T)$$

which concludes the proof.  $\square$

### 2.4.6 Characterisation of the invariant probability measure

In this section, we characterise the invariant probability measure of the centered FLEMING-VIOT process  $\pi$  thanks to an adaptation of DONNELLY-KURTZ's modified look-down construction presented in Section 2.4.6.1. We give in Section 2.4.6.2 an explicit characterisation of the invariant probability measure  $\pi$ . Let us begin by giving a convergence result of the invariant probability measure  $\pi_N$  to the invariant probability measure  $\pi$ .

**Lemma 2.4.9.** *The sequence of laws  $(\pi_N)_{N \in \mathbb{N}^*}$  converges in law to  $\pi$  in  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ .*

*Proof.* Let  $T \geq 0$ ,  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$  and  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  such that  $\mu_N$  converges in law to  $\mu$ . From Proposition 2.4.7 and Theorem 2.4.1, we have for all  $f \in \mathcal{C}_b^0(\mathcal{M}_1(\mathbb{R}), \mathbb{R})$ ,

$$\begin{aligned} |\langle f, \pi_N \rangle - \langle f, \pi \rangle| &\leq \left| \langle f, \pi_N \rangle - \mathbb{E}_{\mu_N}(f(\hat{Z}_T^{N, \mu_N})) \right| + \left| \mathbb{E}_{\mu_N}(f(\hat{Z}_T^{N, \mu_N})) - \mathbb{E}_\mu(f(Z_T)) \right| \\ &\quad + |\mathbb{E}_\mu(f(Z_T)) - \langle f, \pi \rangle| \\ &\leq 2 \|f\|_\infty \alpha \exp(-\beta T) + \left| \mathbb{E}_{\mu_N}(f(\hat{Z}_T^{N, \mu_N})) - \mathbb{E}_\mu(f(Z_T)) \right|. \end{aligned}$$

The announced result follows from Proposition 2.4.2.  $\square$

### 2.4.6.1 Construction of exchangeable random variables allowing to characterise $\pi_N$ and $\pi$

We consider the probability space  $(\check{\Omega}, \check{\mathcal{F}}, \check{\mathbb{P}})$  where we define the modified look-down process on  $(-\infty, 0]$  as a population dynamics on the set  $\mathbb{N}$  of levels where one individual is assigned to each level. To each pair of levels  $(i, j) \in \mathbb{N}^2$  with  $1 \leq i < j$ , we assign an independent POISSON process  $(N_{ij}(t))_{t \geq 0}$  with intensity  $2\lambda$  and to each level  $i \in \mathbb{N}^*$ , we assign an independent standard Brownian motion  $(B_i(t))_{t \leq 0}$  on  $\mathbb{R}_-$ . Jointly with the modified look-down is constructed for all  $N \in \mathbb{N}^*$ , the so-called  $N$ -look-down process whose evolution is given as follows:

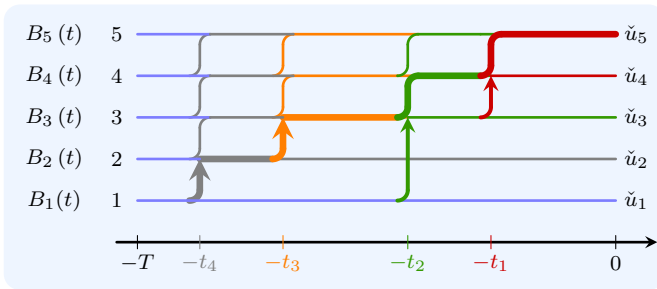


FIGURE 2.5 – Graphical representation of the modified look-down process with  $N = 5$ .

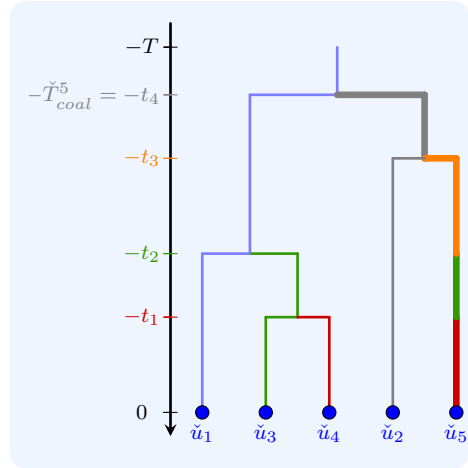


FIGURE 2.6 – KINGMAN's genealogy  $(\check{k}_{5,t})_{0 \leq t \leq T}$  under the modified look-down model on the left, tracing back from time 0 to time  $-T$ .

- (1) **Birth/Death rule.** Each jump time  $t_k$  of one of the POISSON process  $(N_{ij})_{1 \leq i < j \leq N}$  corresponds to a reproduction event at backward time  $-t_k$ . When the time  $t_k$  is the jump time of the POISSON process  $N_{ij}$ , we put an arrow from  $i$  to  $j$  as illustrated in Figure 2.5 which means that the individual at level  $i$  puts a child at level  $j$ . The offspring at level  $j$  adopts the current spatial position of its parent at level  $i$ . The parent level and position do not change. Individuals previously at level  $\ell \in \{j, \dots, N - 1\}$  are shifted one level up to  $\ell + 1$  and the individual at level  $N$  dies.
- (2) **Spatial motion.** Between reproduction events, individuals' spatial positions at each level  $i$  evolve according to the standard Brownian motion  $B_i(-t)$ . As explain below, we will fix the position of the individual at level 1 at coalescence time to 0.

Note that the  $N$ -modified look-down process is simply the first  $N$  levels of the  $(N+k)$ -modified look-down for any  $k \in \mathbb{N}^*$ . In other words, the modified look-down construction can be done with an infinite population as a projective limit of the so-called (infinite) modified look-down. From [78, 80], the genealogy  $(\check{k}_{N,t})_{t \geq 0}$  in backward time since time 0 of a sample from a population evolving according to the  $N$ -modified look-down is exactly determined by KINGMAN's  $N$ -coalescent with coalescence rate  $2\lambda$ . In Figure 2.6 we give the KINGMAN genealogy associated to the 5-modified look-down of the Figure 2.5.

We denote by  $a(i, t)$ ,  $i \in \{1, \dots, N\}$ ,  $t \in (-\infty, 0]$ , the ancestor level of the individual at level  $i$  at time  $t$ . For example, in Figure 2.5, for all  $t \in ]-t_3, -t_4]$ ,  $a(5, t) = 2$  and for all  $t \in ]-t_2, -t_4]$ ,  $a(3, t) = 1$ . Let us consider the random variables

$$\begin{aligned}\check{T}_{coal}^N &:= \inf \left\{ T \geq 0 \mid a(i, -T) = 1, \forall i \in \{1, \dots, N\} \right\}, \\ \check{T}_{coal}^\infty &:= \inf \left\{ T \geq 0 \mid a(i, -T) = 1, \forall i \in \mathbb{N}^* \right\},\end{aligned}$$

which can be interpreted respectively as the coalescence time (i.e. the first time where  $|\check{k}_{N,t}| = 1$ ) of the KINGMAN  $N$ -coalescent  $(\check{k}_{N,t})_{t \geq 0}$  and the KINGMAN coalescent  $(\check{k}_{\infty,t})_{t \geq 0}$ . Note that, for all  $N \in \mathbb{N}^*$ ,  $\check{T}_{coal}^N \leq \check{T}_{coal}^\infty$   $\mathbb{P}$ -a.s. In Proposition 2.4.10 below, we establish that  $\check{T}_{coal}^\infty$  admits moments of any order. We shall be interested in the spatial position  $\check{u}_i$  of the individual at level  $i \in \mathbb{N}^*$  at time 0 assuming that the position of its ancestor at backward time  $-\check{T}_{coal}^\infty$  is 0. For example, if we assume that, in Figure 2.6,  $\check{T}_{coal}^\infty = \check{T}_{coal}^5 = t_4$ , then the spatial position of the individual at level 5 at backward time  $-\check{T}_{coal}^\infty$ , represented by the curve in bold in Figures 2.5 and 2.6, is

$$\check{u}_5 := B_2(-t_4) - B_2(-t_3) + B_3(-t_3) - B_3(-t_2) + B_4(-t_2) - B_4(-t_1) + B_5(-t_1).$$

Similarly,  $\check{u}_1 := B_1(-t_4)$ ,  $\check{u}_2 := B_2(-t_4)$ ,  $\check{u}_3 := B_1(-t_4) - B_1(-t_2) + B_3(-t_2)$ ,  $\check{u}_4 := B_1(-t_4) - B_1(-t_2) + B_3(-t_2) - B_3(-t_1) + B_4(-t_1)$ . In general, we define for all  $i \in \mathbb{N}^*$ , the random variable

$$\check{u}_i := \int_{-\check{T}_{coal}^\infty}^0 dB_{a(i,t)}(t).$$

**Proposition 2.4.10.** *For all  $k \in \mathbb{N}$ , there exists a constant  $C_k > 0$  such that  $\mathbb{E} \left( (\check{T}_{coal}^\infty)^k \right) \leq C_k$ .*

*Proof.* Note that  $\check{T}_{coal}^\infty = \sum_{k=2}^{+\infty} T_k$  where  $(T_k)_{k \geq 2}$  are independent random variables such that  $T_k$  follows an exponential law of parameter  $2\lambda \binom{k}{2}$ . As previously in Step 2 of the proof of Proposition 2.4.7, there exists a constant  $C > 0$  such that for all  $\rho \in (0, 2\lambda)$ ,

$$\mathbb{E} \left( \exp \left( \rho \check{T}_{coal}^\infty \right) \right) = \prod_{k=2}^{+\infty} \frac{1}{1 - \frac{\rho}{\lambda k(k-1)}} \leq \exp(C\rho).$$

From the inequality: for all  $x \in \mathbb{R}_+$ ,  $x^k \leq \left( \frac{k}{\lambda \exp(1)} \right)^k \exp(\lambda x)$ , we deduce that

$$\mathbb{E} \left( (\check{T}_{coal}^\infty)^k \right) \leq \left( \frac{k}{\lambda \exp(1)} \right)^k \mathbb{E} \left( \exp \left( \lambda \check{T}_{coal}^\infty \right) \right) \leq C_k$$

where  $C_k := \left( \frac{k}{\lambda \exp(1)} \right)^k \exp(C\lambda)$  which ends the proof.  $\square$

In view of (2.37), it is natural to introduce for all  $N, i \in \mathbb{N}^*$ ,  $\check{v}_i^N := \check{u}_i - \frac{1}{N} \sum_{j=1}^N \check{u}_j$ . In the following proposition, we give analogous results to those of Proposition 2.4.6.

**Proposition 2.4.11.** (1) *Conditionally to  $\check{k}_N := (\check{k}_{N,t})_{t \geq 0}$ , on the event  $\left\{ |\check{k}_{N,T}| = 1 \right\}$ , we have for all  $i, j \in \{1, \dots, N\}$ ,  $\text{Cov} \left( \check{u}_i, \check{u}_j \mid \check{k}_N \right) = \check{T}_{coal}^\infty - \check{T}_{ij}$ , where  $\check{T}_{ij}$  is the coalescence time between individuals at level  $i$  and  $j$  at time 0.*

(2) Conditionally to  $\check{k}_N$ , on the event  $\left\{|\check{k}_{N,T}| = 1\right\}$ ,  $(\check{v}_i^N)_{1 \leq i \leq N} \sim \mathcal{N}^{(N)}(0_{\mathbb{R}^N}, \check{\Sigma})$  where  $\check{\Sigma} := (\check{\Sigma}_{ij})_{1 \leq i, j \leq N}$  is define by

$$\forall i, j \in \{1, \dots, N\}, \check{\Sigma}_{ij} := \text{Cov}\left(\check{v}_i^N, \check{v}_j^N \mid \check{k}_N\right) = \frac{1}{N} \sum_{k=1}^N (\check{T}_{ik} + \check{T}_{jk}) - \left(\check{T}_{ij} + \frac{1}{N^2} \sum_{k, \ell=1}^N \check{T}_{k\ell}\right).$$

*Proof.* The proof is similar to that of Proposition 2.4.6.  $\square$

Let us define respectively the empirical distribution of  $(\check{u}_i)_{1 \leq i \leq N}$  and its centered version by

$$\check{Y}_{coal}^N := \frac{1}{N} \sum_{i=1}^N \delta_{\check{u}_i} \quad \text{and} \quad \check{Z}_{coal}^N := \frac{1}{N} \sum_{i=1}^N \delta_{\check{v}_i^N}.$$

**Proposition 2.4.12.** *The measure-valued random variable  $\check{Z}_{coal}^N$  has the law  $\pi_N$ .*

*Proof.* The proof consists in establishing for all  $f : \mathcal{M}_1^{c,2}(\mathbb{R}) \rightarrow \mathbb{R}$  measurable real bounded function,  $\mathbb{E}\left(f\left(\check{Z}_{coal}^N\right)\right) = \int_{\mathcal{M}_1^{c,2}(\mathbb{R})} f(\mu) \pi_N(d\mu)$ . From Proposition 2.4.7, it is sufficient to establish for all  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ , that  $\lim_{T \rightarrow +\infty} \left| \mathbb{E}\left(f\left(\check{Z}_{coal}^N\right)\right) - \mathbb{E}\left(f\left(\widehat{Z}_T^{N, \mu_N}\right)\right) \right| = 0$ .

Let  $f : \mathcal{M}_1^{c,2}(\mathbb{R}) \rightarrow \mathbb{R}$  be a measurable real bounded function. Note that,

$$\begin{aligned} \left| \mathbb{E}\left(f\left(\check{Z}_{coal}^N\right)\right) - \mathbb{E}\left(f\left(\widehat{Z}_T^{N, \mu_N}\right)\right) \right| &\leq \left| \mathbb{E}\left(f\left(\check{Z}_{coal}^N\right) \mathbf{1}_{\{|\check{k}_{N,T}|=1\}}\right) - \mathbb{E}\left(f\left(\widehat{Z}_T^{N, \mu_N}\right) \mathbf{1}_{\{|\check{k}_{N,T}|=1\}}\right) \right| \\ &\quad + \|f\|_\infty \left[ \check{\mathbb{P}}\left(|\check{k}_{N,T}| > 1\right) + \widehat{\mathbb{P}}\left(|\check{k}_{N,T}| > 1\right) \right]. \end{aligned}$$

Now, from Propositions 2.4.6 and 2.4.11, it follows that for all  $\mu_N \in \mathcal{M}_{1,N}^{c,2}(\mathbb{R})$ ,

$$\widehat{Z}_T^{N, \mu_N} \mathbf{1}_{\{|\check{k}_{N,T}|=1\}} \stackrel{\text{law}}{=} \check{Z}_{coal}^N \mathbf{1}_{\{|\check{k}_{N,T}|=1\}}.$$

As established in Step 2 of the proof of Proposition 2.4.7,

$$\lim_{T \rightarrow +\infty} \check{\mathbb{P}}\left(|\check{k}_{N,T}| > 1\right) = \lim_{T \rightarrow +\infty} \widehat{\mathbb{P}}\left(|\check{k}_{N,T}| > 1\right) = 0,$$

which concludes the proof.  $\square$

In the next proposition, we establish the exchangeability property of the family  $(\check{u}_i)_{i \in \mathbb{N}^*}$  which will allow to apply the DE FINETTI representation theorem.

**Proposition 2.4.13.** (1) *The family  $(\check{u}_i)_{i \in \mathbb{N}^*}$  is exchangeable.*

(2) *There exists a random variable measure-valued  $\check{Y}_{coal}^\infty : \check{\Omega} \rightarrow \mathcal{M}_1(\mathbb{R})$  such that  $(\check{Y}_{coal}^N)_{N \in \mathbb{N}^*}$  converges  $\check{\mathbb{P}}$ -a.s. when  $N \rightarrow +\infty$  to  $\check{Y}_{coal}^\infty$  in  $\mathcal{M}_1(\mathbb{R})$  which is equipped with the weak topology. Moreover, given  $\check{Y}_{coal}^\infty$ ,  $(\check{u}_i)_{i \in \mathbb{N}^*}$  is i.i.d. of law  $\check{Y}_{coal}^\infty$ .*

*Proof.* (1) From [78, Proof of Theorem 2.2], it is enough to show for each  $N \in \mathbb{N}^*$ ,  $(\check{u}_i)_{1 \leq i \leq N}$  is exchangeable. Let  $\sigma : \mathbb{N}^* \rightarrow \mathbb{N}^*$  be a finite permutation, that is to say a bijection that leaves all but finitely many points unchanged. The well-known backward construction of the

modified look-down process [78, 80] entails that  $(\check{k}_{\infty,t})_{t \geq 0} \stackrel{\text{law}}{=} (\check{k}_{\infty,t}^{\sigma})_{t \geq 0}$  where  $\check{k}_{\infty,t}^{\sigma}$  is the partition obtained by applying the permutation  $\sigma$  to  $\check{k}_{\infty,t}$ . Therefore, for any permutation  $\sigma : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$  extended by id to  $\mathbb{N}^*$ , it is sufficient to prove that  $\left( (\check{u}_i)_{1 \leq i \leq N} \mid \check{k}_{\infty} \right) \stackrel{\text{law}}{=} \left( (\check{u}_{\sigma(i)})_{1 \leq i \leq N} \mid \check{k}_{\infty}^{\sigma} \right)$  to obtain the announced result.

We define  $\check{T}_{coal}^{\infty, \sigma} := \check{T}_{coal}^{\infty}(\check{k}_{\infty}^{\sigma})$  and for all  $i, j \in \{1, \dots, N\}$ ,  $\check{T}_{ij}^{\sigma} := \check{T}_{ij}(\check{k}_{\infty}^{\sigma})$ . Note that  $\check{T}_{coal}^{\infty, \sigma} = \check{T}_{coal}^{\infty}$   $\check{\mathbb{P}}$ -a.s. and it follows from the fact  $(\check{k}_{\infty,t})_{t \geq 0} \stackrel{\text{law}}{=} (\check{k}_{\infty,t}^{\sigma})_{t \geq 0}$  that

$$\left( (\check{T}_{ij}^{\sigma}, \check{T}_{coal}^{\infty, \sigma}) \right)_{1 \leq i, j \leq N} \stackrel{\text{law}}{=} \left( (\check{T}_{ij}, \check{T}_{coal}^{\infty}) \right)_{1 \leq i, j \leq N}.$$

From Proposition 2.4.11, for all  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  measurable real bounded function,

$$\mathbb{E} \left( f(\check{u}_1, \dots, \check{u}_N) \mid \check{k}_{\infty} \right) = F \left( (\check{T}_{ij})_{1 \leq i < j \leq N}, \check{T}_{coal}^{\infty} \right).$$

for a certain function  $F$ . So, in particular

$$\mathbb{E} \left( f(\check{u}_{\sigma(1)}, \dots, \check{u}_{\sigma(N)}) \mid \check{k}_{\infty}^{\sigma} \right) = F \left( (\check{T}_{ij}^{\sigma})_{1 \leq i < j \leq N}, \check{T}_{coal}^{\infty, \sigma} \right).$$

By taking the expectation in the previous expressions, we deduce that  $\mathbb{E} \left( f(\check{u}_{\sigma(1)}, \dots, \check{u}_{\sigma(N)}) \right) = \mathbb{E} \left( f(\check{u}_1, \dots, \check{u}_N) \right)$  which completes the proof.

(2) As the family  $(\check{u}_i)_{i \in \mathbb{N}^*}$  is exchangeable, the announced result follows from DE FINETTI's representation theorem [142, Theorem 12.26 and Remark 12.27].  $\square$

We conclude this section with a corollary which will be useful to characterise the probability measure  $\pi$ .

**Corollary 2.4.14.** (1) For all  $k \in \mathbb{N}$ , the random variable  $\check{Y}_{coal}^{\infty}$  satisfies  $\langle |\text{id}|^k, \check{Y}_{coal}^{\infty} \rangle < \infty$   $\check{\mathbb{P}}$ -a.s.

(2) The limit  $\check{u}_{\infty} := \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=1}^N \check{u}_j$  exists  $\check{\mathbb{P}}$ -a.s. and satisfies  $\check{u}_{\infty} = \langle \text{id}, \check{Y}_{coal}^{\infty} \rangle$ .

*Proof.* (1) For all  $M \in (0, +\infty)$ , let us consider  $|\text{id}|_M$  the truncation function of id at level  $M$  defined by  $|\text{id}|_M = |\text{id}|$  on  $[-M, M]$  and  $|\text{id}|_M = M$  on  $\mathbb{R} \setminus [-M, M]$ . By FATOU's lemma, we obtain that

$$\mathbb{E} \left( \langle |\text{id}|_M^k, \check{Y}_{coal}^{\infty} \rangle \right) \leq \liminf_{N \rightarrow +\infty} \mathbb{E} \left( \langle |\text{id}|_M^k, \check{Y}_{coal}^N \rangle \right) \leq \liminf_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N \left( 1 + \mathbb{E}(\check{u}_i^{2k}) \right).$$

Now, classical moment results for Gaussian random variables show that for all  $n \in \mathbb{N}$ ,  $\mathbb{E}(G^{2n}) = \frac{(2n)!}{2^n n!} \sigma^{2n}$  for  $G \sim \mathcal{N}(0, \sigma^2)$ . From Proposition 2.4.11,

$$\mathbb{E}(\check{u}_i^{2k}) = \mathbb{E} \left( \mathbb{E}(\check{u}_i^{2k} \mid \check{k}_N) \right) = \frac{(2k)!}{2^k k!} \mathbb{E} \left( (\check{T}_{coal}^{\infty})^k \right).$$

Therefore, from Proposition 2.4.10, we deduce that there exists a constant  $C_k > 0$  such that  $\mathbb{E} \left( \langle |\text{id}|_M^k, \check{Y}_{coal}^{\infty} \rangle \right) \leq 1 + \frac{(2k)!}{2^k k!} C_k$  and by the dominated convergence theorem when  $M \rightarrow +\infty$ , the announced result follows.

(2) From Proposition 2.4.13, given  $\check{Y}_{coal}^{\infty}$ ,  $(\check{u}_i)_{i \in \mathbb{N}^*}$  is i.i.d. Therefore, the announced almost surely existence limit follows from the Strong Law of Large Numbers. Moreover,

$$\langle \text{id}, \check{Y}_{coal}^{\infty} \rangle = \lim_{N \rightarrow +\infty} \langle \text{id}, \check{Y}_{coal}^N \rangle = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{j=1}^N \check{u}_j = \check{u}_{\infty}, \quad \check{\mathbb{P}} - \text{a.s.} \quad \square$$

### 2.4.6.2 Characterisation of the invariant probability measure $\pi$

Now we define the random variable  $\check{Z}_{coal}^\infty \in \mathcal{M}_1(\mathbb{R})$  as

$$\check{Z}_{coal}^\infty := \tau_{-\langle \text{id}, \check{Y}_{coal}^\infty \rangle} \# \check{Y}_{coal}^\infty.$$

The following proposition establishes the convergence of  $(\check{Z}_{coal}^N)_{N \in \mathbb{N}^*}$  to  $\check{Z}_{coal}^\infty$ . Recall from (2.17) that for all  $n \in \mathbb{N}^*$ ,  $\mu \in \mathcal{M}_1(\mathbb{R})$ ,  $P_{f,n}(\mu) := \langle f, \mu^n \rangle$ . Let us recall the following well-known fact useful for the proof below: a straightforward adaptation of the proof of [60, Lemma 2.1.2] allows us to obtain that for all  $n \in \mathbb{N}^*$ , the algebra of polynomials

$$\text{Span} \left( \left\{ P_{f,n}(\mu) \mid f : \mathbb{R}^n \rightarrow \mathbb{R} \text{ uniformly continuous, } \mu \in \mathcal{M}_1(\mathbb{R}) \right\} \right)$$

is convergence determining in  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ .

**Proposition 2.4.15.** (1) *The sequence of random variables  $(\check{Z}_{coal}^N)_{N \in \mathbb{N}^*}$  converges  $\check{\mathbb{P}}$ -a.s. when  $N \rightarrow +\infty$  to  $\check{Z}_{coal}^\infty$  in  $\mathcal{M}_1(\mathbb{R})$  for the weak convergence topology.*  
 (2) *The random variable  $\check{Z}_{coal}^\infty$  has the law  $\pi$ .*

*Proof.* (1) From the previous reminder, it is sufficient to prove that for all  $n \in \mathbb{N}^*$ , for all  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uniformly continuous,  $\lim_{N \rightarrow +\infty} P_{f,n}(\check{Z}_{coal}^N) = P_{f,n}(\check{Z}_{coal}^\infty)$ . With an argument similar to the proof of Proposition 2.4.2, we obtain that for all  $n \in \mathbb{N}^*$ , for all  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uniformly continuous,  $P_{f \circ \tau_{-\langle \text{id}, \cdot \rangle}, n}$  from  $\mathcal{M}_1^1(\mathbb{R})$  to  $\mathbb{R}$  is continuous. Let  $n \in \mathbb{N}^*$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uniformly continuous. Now,  $\check{Y}_{coal}^\infty \in \mathcal{M}_1^1(\mathbb{R})$  from Corollary 2.4.14 and

$$P_{f \circ \tau_{-\langle \text{id}, \check{Y}_{coal}^N \rangle}, n}(\check{Y}_{coal}^N) = P_{f,n}(\check{Z}_{coal}^N) \quad \text{and} \quad P_{f \circ \tau_{-\langle \text{id}, \check{Y}_{coal}^\infty \rangle}, n}(\check{Y}_{coal}^\infty) = P_{f,n}(\check{Z}_{coal}^\infty)$$

From Proposition 2.4.13,  $\lim_{N \rightarrow +\infty} P_{f \circ \tau_{-\langle \text{id}, \check{Y}_{coal}^N \rangle}, n}(\check{Y}_{coal}^N) = P_{f \circ \tau_{-\langle \text{id}, \check{Y}_{coal}^\infty \rangle}, n}(\check{Y}_{coal}^\infty)$   $\check{\mathbb{P}}$ -a.s. which concludes the proof.

(2) Let  $n \in \mathbb{N}^*$ . As for all  $N \in \mathbb{N}^*$  and  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  uniformly continuous,

$$\begin{aligned} & \left| \mathbb{E} \left( P_{f,n}(\check{Z}_{coal}^\infty) \right) - \int_{\mathcal{M}_1(\mathbb{R})} P_{f,n}(\mu) \pi(d\mu) \right| \\ & \leq \mathbb{E} \left( \left| P_{f,n}(\check{Z}_{coal}^\infty) - P_{f,n}(\check{Z}_{coal}^N) \right| \right) + \left| \mathbb{E} \left( P_{f,n}(\check{Z}_{coal}^N) \right) - \int_{\mathcal{M}_1(\mathbb{R})} P_{f,n}(\mu) \pi_N(d\mu) \right| \\ & \quad + \left| \int_{\mathcal{M}_1(\mathbb{R})} P_{f,n}(\mu) [\pi(d\mu) - \pi_N(d\mu)] \right|, \end{aligned}$$

the announced result follows from Propositions 2.4.15 (1), and 2.4.12 and Lemma 2.4.9 when  $N \rightarrow +\infty$ .  $\square$

The last characterisation of the probability measure  $\pi$  is suitable to make explicit computations. The next corollary gives an expression of the second moment under  $\pi$ .

**Corollary 2.4.16.** *We have  $\int_{\mathcal{M}_1^{c,2}(\mathbb{R})} M_2(\mu) \pi(d\mu) = 1/2\lambda$ .*

*Proof. Step 1. Uniform bound in  $N$  of  $\mathbb{E} \left( M_{2k} \left( \check{Z}_{coal}^N \right) \right)$ ,  $k \in \mathbb{N}^*$ .* In this step, we want to establish

$$\sup_{N \in \mathbb{N}^*} \mathbb{E} \left( \langle |\text{id}|^{2k}, \check{Z}_{coal}^N \rangle \right) < \infty.$$

Let  $N \in \mathbb{N}^*$ . From Proposition 2.4.13 (1),

$$\begin{aligned} \mathbb{E} \left( \langle |\text{id}|^{2k}, \check{Z}_{coal}^N \rangle \right) &= \mathbb{E} \left( \mathbb{E} \left( \left| \frac{1}{N} \sum_{i=1}^N \check{v}_i - \frac{1}{N} \sum_{j=1}^N \check{v}_j \right|^{2k} \middle| (\check{T}_{ml})_{1 \leq m, \ell \leq N}, \check{T}_{coal}^\infty \right) \right) \\ &= \mathbb{E} \left( \mathbb{E} \left( |\check{v}_1^N|^{2k} \middle| (\check{T}_{ml})_{1 \leq m, \ell \leq N}, \check{T}_{coal}^\infty \right) \right). \end{aligned}$$

From Propositions 2.4.11 and 2.4.10, we obtain that

$$\mathbb{E} \left( \mathbb{E} \left( |\check{v}_1^N|^{2k} \middle| (\check{T}_{m,\ell})_{1 \leq m, \ell \leq N}, \check{T}_{coal}^\infty \right) \right) = \frac{(2k)!}{2^k k!} \mathbb{E} \left( \check{\Sigma}_{ii}^{2k} \right) \leq \frac{(2k)!}{k!} \mathbb{E} \left( (\check{T}_{coal}^\infty)^{2k} \right) < \infty,$$

and the announced result follows.

**Step 2. Convergence result of  $\mathbb{E} \left( M_2 \left( \check{Z}_{coal}^N \right) \right)$  to  $\mathbb{E} \left( M_2 \left( \check{Z}_{coal}^\infty \right) \right)$ .** Note that for all  $N \in \mathbb{N}^*$ ,  $M \in (0, +\infty)$ ,

$$\left| \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^N \right) \right) - \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^\infty \right) \right) \right| \leq (\mathbf{A})_{N,M} + (\mathbf{B})_{N,M} + (\mathbf{C})_{N,M},$$

where

$$\begin{aligned} (\mathbf{A})_{N,M} &:= \left| \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^N \right) \right) - \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^N \rangle \right) \right|, \\ (\mathbf{B})_{N,M} &:= \left| \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^N \rangle \right) - \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^\infty \rangle \right) \right|, \\ (\mathbf{C})_{N,M} &:= \left| \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^\infty \rangle \right) - \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^\infty \right) \right) \right|. \end{aligned}$$

From the inequality  $|\text{id}^2 - \text{id}_M^2| \leq \frac{2}{3\sqrt{3}M} |\text{id}|^3$ , then the HÖLDER inequality, we obtain that

$$(\mathbf{A})_{N,M} \leq \mathbb{E} \left( \langle |\text{id}^2 - \text{id}_M^2|, \check{Z}_{coal}^N \rangle \right) \leq \frac{2}{3\sqrt{3}M} \mathbb{E} \left( \langle |\text{id}|^3, \check{Z}_{coal}^N \rangle \right) \leq \frac{2}{3\sqrt{3}M} \mathbb{E} \left( \langle \text{id}^4, \check{Z}_{coal}^N \rangle^{\frac{3}{4}} \right).$$

From Step 1, we deduce that for all  $N \in \mathbb{N}^*$ ,

$$(\mathbf{A})_{N,M} \leq \frac{2}{3\sqrt{3}M} \left( 1 + \sup_{N \in \mathbb{N}^*} \mathbb{E} \left( \langle \text{id}^4, \check{Z}_{coal}^N \rangle \right) \right) < \infty.$$

In similar way, we obtain that for all  $N \in \mathbb{N}^*$ ,  $(\mathbf{C})_{N,M} \leq \frac{2}{3\sqrt{3}M} \left( 1 + \mathbb{E} \left( \langle \text{id}^4, \check{Z}_{coal}^\infty \rangle \right) \right)$  where  $\mathbb{E} \left( \langle \text{id}^4, \check{Z}_{coal}^\infty \rangle \right) < \infty$  from Corollary 2.4.14 (1). By the monotone convergence theorem, we deduce that for all  $N \in \mathbb{N}^*$ ,  $\mathbb{E} \left( M_2 \left( \check{Z}_{coal}^N \right) \right) = \lim_{M \rightarrow +\infty} \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^N \rangle \right)$  and  $\mathbb{E} \left( M_2 \left( \check{Z}_{coal}^\infty \right) \right) = \lim_{M \rightarrow +\infty} \mathbb{E} \left( \langle |\text{id}|_M^2, \check{Z}_{coal}^\infty \rangle \right)$ . From Proposition 2.4.15, for all  $M \in (0, +\infty)$ ,  $\lim_{N \rightarrow +\infty} (\mathbf{B})_{N,M} = 0$ . From classical analysis techniques, we deduce that  $\lim_{N \rightarrow +\infty} \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^N \right) \right) = \mathbb{E} \left( M_2 \left( \check{Z}_{coal}^\infty \right) \right)$ .

**Step 3. Conclusion.** Note that for all  $N \in \mathbb{N}^*$ ,

$$M_2(\check{Z}_{coal}^N) = \langle \text{id}^2, \check{Z}_{coal}^N \rangle = \frac{1}{N} \sum_{i=1}^N \check{u}_i^2 - \frac{1}{N^2} \left[ \sum_{i=1}^N \check{u}_i^2 + 2 \sum_{1 \leq i < j \leq N} \check{u}_i \check{u}_j \right].$$

From Proposition 2.4.11, we have for all  $i, j \in \{1, \dots, N\}$ ,  $\mathbb{E}(\check{u}_i \check{u}_j) = \mathbb{E}\left(\mathbb{E}(\check{u}_i \check{u}_j \mid \check{k}_N)\right) = \mathbb{E}(\check{T}_{coal}^\infty - \check{T}_{ij})$  where  $\check{T}_{ij}$  is an exponential random variable with parameter  $2\lambda$  if  $i \neq j$ . Therefore  $\mathbb{E}(M_2(\check{Z}_{coal}^N)) = \frac{N-1}{N} \times \frac{1}{2\lambda}$ . By Step 2, we deduce that  $\mathbb{E}(M_2(\check{Z}_{coal}^\infty)) = 1/2\lambda$  which completes the proof.  $\square$

## 2.5 Proof of Theorem 2.2.3

We divide the proof of the main result into 7 steps, each of which will constitute a section (Sections 2.5.1 to 2.5.7). We recall that the aim of this proof is to prove that the law  $\mathbb{P}_{\tau_{-\langle \text{id}, \nu \rangle}^\# \nu}^{\text{FVc}}$  of the process  $(Z_t)_{0 \leq t \leq T}$  defined by

$$\forall t \geq 0, \quad Z_t := \tau_{-\langle \text{id}, Y_t \rangle}^\# Y_t,$$

under  $\mathbb{P}_\nu^{\text{FV}}$  for  $\nu \in \mathcal{M}_1^2(\mathbb{R})$  is solution of the martingale problem (2.4). We will start by considering the case with test functions  $F, g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$  and we will prove the extension to  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  in Section 2.5.6. In (2.11) there are essentially two types of terms:  $\langle \text{id}, Y_t - Y_s \rangle$  and  $\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_s \rangle}, Y_t - Y_s \rangle$ ,  $j \in \{0, 1, 2\}$ ,  $s \leq t$ . In Sections 2.5.1 and 2.5.2, we prove that the two previous quantities admit a DOOB's semi-martingale decomposition. In Sections 2.5.3 and 2.5.4, we handle all the terms in (2.11) involving respectively the first and second derivative of  $F$ . In Section 2.5.5, we deal with the different error terms involved in (2.11). Finally, in Section 2.5.7, we prove that the martingale involved in (2.4) is square integrable and we establish the relation (2.6). We conclude in Sections 2.5.8 and 2.5.9 by proving a technical lemma used in Section 2.5.2.

### 2.5.1 Doob's semi-martingale decomposition of $\langle \text{id}, Y_t - Y_s \rangle$ , $s \leq t$

In (2.7),  $M^{\text{id}}(g)$  is well-defined only for  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . The expression makes sense for more general functions  $g$ . The goal of this section is to prove that, for any  $k \in \mathbb{N}$ ,  $M^{\text{id}}(\text{id}^k)$  is the martingale part in the DOOB semi-martingale decomposition of  $\langle \text{id}^k, Y_t \rangle$ . In particular,

$$\langle \text{id}, Y_s - Y_{t_i^n \wedge t} \rangle = M_s^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}), \quad s \geq t_i^n \wedge t$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale.

**Lemma 2.5.1.** *Let  $\nu \in \mathcal{M}_1(\mathbb{R})$  and let  $\mathbb{P}_\nu$  be a distribution on  $\Omega$  satisfying (2.7) and such that  $Y_0$  is equal in law to  $\nu$ . Let  $T > 0$  and  $k \in \mathbb{N}^*$  be fixed.*

- (1) *If  $\langle |\text{id}|^k, \nu \rangle < \infty$ , then there exist two constants  $C_{k,T}, \tilde{C}_{k,T} > 0$ , such that any stochastic process  $(Y_t)_{0 \leq t \leq T}$  whose law is  $\mathbb{P}_\nu$  satisfies*

- (a) 
$$\sup_{t \in [0, T]} \mathbb{E}_\nu \left( \langle |\text{id}|^k, Y_t \rangle \right) \leq C_{k,T} \left( 1 + \langle |\text{id}|^k, \nu \rangle \right),$$
- (b) 
$$\forall \alpha > 0, \quad \mathbb{P}_\nu \left( \sup_{t \in [0, T]} \langle |\text{id}|^k, Y_t \rangle \geq \alpha \right) \leq \frac{\tilde{C}_{k,T} \left( 1 + \langle |\text{id}|^k, \nu \rangle \right)}{\alpha}.$$



(2) If  $\langle |\text{id}^k, \nu \rangle < \infty$ , then the process  $(M_t^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  defined by

$$M_t^{\text{id}}(\text{id}^k) := \langle \text{id}^k, Y_t \rangle - \langle \text{id}^k, Y_0 \rangle - \int_0^t \left\langle \frac{k(k-1)}{2} \text{id}^{k-2}, Y_s \right\rangle ds,$$

is a continuous  $\mathbb{P}_\nu$ -martingale. Moreover, if  $\langle |\text{id}^{2k}, \nu \rangle < \infty$ , then  $(M^{\text{id}}(\text{id}^k))_{0 \leq t \leq T}$  is a martingale in  $L^2(\Omega)$  whose quadratic variation is given by

$$\langle M^{\text{id}}(\text{id}^k) \rangle_t = 2\lambda \int_0^t \left[ \langle \text{id}^{2k}, Y_s \rangle - \langle \text{id}^k, Y_s \rangle^2 \right] ds.$$

*Proof.* The proof is similar to that of Proposition 2.2.11.  $\square$

### 2.5.2 Doob's semi-martingale decomposition of $\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_s \rangle}, Y_t - Y_s \rangle$ , $s \leq t$ , $j \in \{0, 1, 2\}$

Equation (2.11) involves terms of the form  $\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \rangle$  with  $j \in \{0, 1, 2\}$ . We wish to express, each of these terms using the martingale problem (2.7). However, this leads us to consider quantities of the form

$$M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \quad (2.38)$$

with  $j \in \{0, 1, 2\}$ , which are not well defined at the moment. Indeed, in (2.38) the input argument is a predictable *random* function of the process  $(Y_t)_{0 \leq t \leq T}$  while the martingale problem (2.7) defines  $M_t^{\text{id}}(g)$  only for *deterministic* functions  $g$ . Lemma 2.5.2 hereafter, allows us to give a precise meaning to (2.38) by extending the well-defined character of the martingales of (2.7) to predictable input arguments. The proof of this technical lemma, given in Section 2.5.9, is based on *regular conditional probabilities*.

**Lemma 2.5.2.** *Let  $t^* \in \mathbb{R}_+$  be a deterministic time and  $h : \Omega \rightarrow \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  be a measurable function satisfying the following property:*

$$\forall \omega, \omega' \in \Omega, \quad h(\omega) = h(\omega') \quad \text{if} \quad \omega|_{[0, t^*]} = \omega'|_{[0, t^*]}.$$

*Then, the following process defined, for all  $t \in [0, T]$ , by*

$$\mathcal{M}_t(\tilde{\omega}) := M_t^{\text{id}}(h(\tilde{\omega}))(\tilde{\omega}) - M_{t \wedge t^*}^{\text{id}}(h(\tilde{\omega}))(\tilde{\omega})$$

*is a  $\mathbb{P}_\nu^{\text{FV}}$  square integrable martingale whose quadratic variation is given by*

$$\langle \mathcal{M}(\tilde{\omega}) \rangle_t = 2\lambda \int_{t \wedge t^*}^t \left[ \langle h^2(\tilde{\omega}_s), \tilde{\omega}_s \rangle - \langle h(\tilde{\omega}_s), \tilde{\omega}_s \rangle^2 \right] ds.$$

Lemma 2.5.2 with  $t^* = t$  allows us to assert that (2.38) is a  $\mathbb{P}_\nu^{\text{FV}}(dY)$ -martingale increment. Thus, we obtain for  $j \in \{0, 1, 2\}$ :

$$\begin{aligned} & \left\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \\ &= \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \left\langle g^{(j+2)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_s \right\rangle ds \\ &+ M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \end{aligned} \quad (2.39)$$

where  $\left( M_{s \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \right)_{s \geq t_i^n}$  is a  $\mathbb{P}_\nu^{\text{FV}}$  square integrable martingale satisfying for all  $s \geq t_i^n$ ,

$$\begin{aligned} & \left\langle M_{\cdot \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \right\rangle_s \\ &= 2\lambda \int_{t_i^n \wedge t}^{s \wedge t} \left[ \left\langle \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right)^2, Y_u \right\rangle - \left\langle g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_u \right\rangle^2 \right] du. \end{aligned} \quad (2.40)$$

### 2.5.3 Expressions of the terms of (2.11) involving $F'$

In the rest of this proof, we use the following notations to simplify the writing. We denote for all  $s \geq 0$ ,  $R(s) := \langle \text{id}, Y_s \rangle$ . We assume that  $F, g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$ . Our goal is to prove the following lemma:

**Lemma 2.5.3.** *When the mesh of the subdivision  $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = T$  of  $[0, T]$  tends to 0 when  $n \rightarrow +\infty$ , we have the following convergence in probability*

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} (\mathbf{A})_i = \int_0^t F'(\langle g, Z_s \rangle) \left( \left\langle \frac{g''}{2}, Z_s \right\rangle + \lambda \langle g'', Z_s \rangle M_2(Z_s) - 2\lambda \langle g' \times \text{id}, Z_s \rangle \right) + \text{Mart}_t,$$

where  $(\text{Mart}_t)_{0 \leq t \leq T}$  is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale.

The proof of Lemma 2.5.3 is based on the following decomposition of  $(\mathbf{A})_i$  (given by the expression (2.12)) which makes use of the DOOB semi-martingale decomposition of Sections 2.5.1 and 2.5.2. We have

$$(\mathbf{A})_i = \sum_{k=1}^6 F' \left( \left\langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \right\rangle \right) (\mathbf{A})_i^k$$

where

$$\begin{aligned} (\mathbf{A})_i^1 &= \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \left\langle g'' \circ \tau_{-R(t_i^n \wedge t)}, Y_s \right\rangle ds, \\ (\mathbf{A})_i^2 &= M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right), \\ (\mathbf{A})_i^3 &= - \left[ M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right] \left\langle g' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \right\rangle, \\ (\mathbf{A})_i^4 &= - \left[ M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right] \left[ M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g' \circ \tau_{-R(t_i^n \wedge t)} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g' \circ \tau_{-R(t_i^n \wedge t)} \right) \right], \\ (\mathbf{A})_i^5 &= \frac{1}{2} \left[ M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right]^2 \left\langle g'' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \right\rangle, \\ (\mathbf{A})_i^6 &= O \left( |t_{i+1}^n \wedge t - t_i^n \wedge t| \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right| \right). \end{aligned}$$

Note that we used the following inequality

$$\int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \left\langle g^{(3)} \circ \tau_{-R(t_i^n \wedge t)}, Y_s \right\rangle ds \leq \|g^{(3)}\|_\infty (t_{i+1}^n \wedge t - t_i^n \wedge t) \quad \mathbb{P}_\nu^{\text{FV}} - \text{a.s.}$$

to bound the term  $\left[ M_{t_{i+1}^n}^{\text{id}}(\text{id}) - M_{t_i^n}^{\text{id}}(\text{id}) \right] \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \left\langle g^{(3)} \circ \tau_{-R(t_i^n \wedge t)}, Y_s \right\rangle ds$  by  $(\mathbf{A})_i^6$ . Our goal in the sequel is to write each of these six quantities as sums of finite variation terms, martingale terms and negligible terms and to study the limit of each of them.

### 2.5.3.1 Decomposition and study of $(\mathbf{A})_i^1$

Note that, for any  $i \in \{0, \dots, p_n - 1\}$ ,

$$(\mathbf{A})_i^1 = \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \langle g'' \circ \tau_{-R(s)}, Y_s \rangle ds + \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \langle g'' \circ \tau_{-R(t_i^n \wedge t)} - g'' \circ \tau_{-R(s)}, Y_s \rangle ds.$$

As a consequence of RIEMANN's sum convergences and using that  $s \mapsto \langle g'' \circ \tau_{-R(s)}, Y_s \rangle$  is continuous, we obtain  $\mathbb{P}_\nu^{\text{FV}}$ -a.s., and therefore in probability that

$$\begin{aligned} \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \langle g'' \circ \tau_{-R(s)}, Y_s \rangle ds \\ = \int_0^t F' \left( \langle g \circ \tau_{-R(s)}, Y_s \rangle \right) \frac{1}{2} \langle g'' \circ \tau_{-R(s)}, Y_s \rangle ds \\ = \int_0^t F' \left( \langle g, Z_s \rangle \right) \left\langle \frac{g''}{2}, Z_s \right\rangle ds. \end{aligned}$$

From Lemma A.2.1 (2), we deduce that, in probability,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \langle g'' \circ \tau_{-R(t_i^n \wedge t)} - g'' \circ \tau_{-R(s)}, Y_s \rangle ds = 0.$$

### 2.5.3.2 Martingale contribution of $(\mathbf{A})_i^2$ and $(\mathbf{A})_i^3$

Note that

$$\sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{A})_i^3,$$

is a stochastic integral with respect to the square integrable martingale  $(M_s^{\text{id}}(\text{id}))_{0 \leq s \leq T}$ . Since  $F'$  and  $s \mapsto \langle g' \circ \tau_{-R(s)}, Y_s \rangle$  are bounded, we deduce that

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{A})_i^3$$

is  $\mathbb{P}_\nu^{\text{FV}}$ -martingale. The term

$$\mathcal{M}_t^n := \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{A})_i^2$$

is a stochastic integral with respect to a martingale which depends on  $n$ . However, the same argument as above applies because  $(\mathcal{M}_t^n)_{0 \leq t \leq T}$  is bounded in  $L^2(\Omega)$ , hence uniformly integrable. This can be proved as follows: as  $F'$  is bounded and from Lemma 2.5.2, there exist two constants  $C_1, C_2 > 0$  such that

$$\begin{aligned} \mathbb{E} \left( [\mathcal{M}_t^n]^2 \right) &= \sum_{i=0}^{p_n-1} \mathbb{E} \left( \left[ F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{A})_i^2 \right]^2 \right) \\ &\leq C_1 \sum_{i=0}^{p_n-1} \left\langle M^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right) \right\rangle_{t_{i+1}^n \wedge t} \\ &\leq 4\lambda C_1 C_2 \sum_{i=0}^{p_n-1} (t_{i+1}^n \wedge t - t_i^n \wedge t) \\ &= 4\lambda C_1 C_2 t < \infty. \end{aligned}$$

### 2.5.3.3 Contributions of $(\mathbf{A})_i^4$ and $(\mathbf{A})_i^5$

The contribution of the next two terms corresponds to the terms due to the centering effect in the martingale problem (2.4).

**Study of the term  $(\mathbf{A})_i^4$ .** Using ITÔ's formula and the relation (2.8), we obtain that

$$(\mathbf{A})_i^4 = (\mathbf{A})_i^{41} + (\mathbf{A})_i^{42} + (\mathbf{A})_i^{43}$$

where

$$\begin{aligned} (\mathbf{A})_i^{41} &= - \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [M_s^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id})] dM_s^{\text{id}}(g' \circ \tau_{-R}(t_i^n \wedge t)), \\ (\mathbf{A})_i^{42} &= - \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [M_s^{\text{id}}(g' \circ \tau_{-R}(t_i^n \wedge t)) - M_{t_i^n \wedge t}^{\text{id}}(g' \circ \tau_{-R}(t_i^n \wedge t))] dM_s^{\text{id}}(\text{id}), \\ (\mathbf{A})_i^{43} &= -2\lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [\langle \text{id} \times g' \circ \tau_{-R}(t_i^n \wedge t), Y_s \rangle - \langle \text{id}, Y_s \rangle \langle g' \circ \tau_{-R}(t_i^n \wedge t), Y_s \rangle] ds. \end{aligned}$$

Using the same arguments as for  $(\mathbf{A})_i^2$ , we deduce, in probability, that for  $k \in \{41, 42\}$ ,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F'(\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \rangle) (\mathbf{A})_i^k$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale. Moreover, we decompose the integral of  $(\mathbf{A})_i^{43}$  in the following way:

$$(\mathbf{A})_i^{43} = (\mathbf{A})_i^{431} + (\mathbf{A})_i^{432} + (\mathbf{A})_i^{433}$$

where

$$\begin{aligned} (\mathbf{A})_i^{431} &= -2\lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \langle \text{id} \times g' \circ \tau_{-R}(t_i^n \wedge t) - \text{id} \times g' \circ \tau_{-R}(s), Y_s \rangle ds, \\ (\mathbf{A})_i^{432} &= -2\lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \langle \text{id}, Y_s \rangle \langle g' \circ \tau_{-R}(s) - g' \circ \tau_{-R}(t_i^n \wedge t), Y_s \rangle ds, \\ (\mathbf{A})_i^{433} &= -2\lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [\langle \text{id} \times g' \circ \tau_{-R}(s), Y_s \rangle - \langle \text{id}, Y_s \rangle \langle g' \circ \tau_{-R}(s), Y_s \rangle] ds. \end{aligned}$$

Using Lemma A.2.1, we deduce, in probability, that for  $k \in \{431, 432\}$ ,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F'(\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \rangle) (\mathbf{A})_i^k = 0,$$

and we deduce from the convergence of RIEMANN's sums that,  $\mathbb{P}_\nu^{\text{FV}}$ -a.s. and hence in probability,

$$\begin{aligned} &\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F'(\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \rangle) (\mathbf{A})_i^{433} \\ &= -2\lambda \int_0^t F'(\langle g \circ \tau_{-R}(s), Y_s \rangle) [\langle \text{id} \times g' \circ \tau_{-R}(s), Y_s \rangle - \langle \text{id}, Y_s \rangle \langle g' \circ \tau_{-R}(s), Y_s \rangle] ds \\ &= -2\lambda \int_0^t F'(\langle g, Z_s \rangle) \langle g' \times \text{id}, Z_s \rangle ds. \end{aligned}$$

**Study of the term  $(\mathbf{A})_i^5$ .** As  $(\mathbf{A})_i^5$  satisfies the following decomposition:

$$(\mathbf{A})_i^5 = \left( \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [M_s^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id})] dM_s^{\text{id}}(\text{id}) + \lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} [\langle \text{id}^2, Y_s \rangle - \langle \text{id}, Y_s \rangle^2] ds \right) \langle g'' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle$$

and proceeding as for  $(\mathbf{A})_i^4$  above, we obtain, in probability, that

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{A})_i^5 = \lambda \int_0^T F'(\langle g, Z_{s \wedge t} \rangle) \langle g'', Z_{s \wedge t} \rangle M_2(Z_{s \wedge t}) ds + \text{Mart}_t^{(1)}$$

where  $(\text{Mart}_t^{(1)})_{0 \leq t \leq T}$  is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale.

### 2.5.3.4 Study of the error term $(\mathbf{A})_i^6$

From the inequality: for all  $x, y \in \mathbb{R}_+$ ,  $xy \leq \frac{2}{3}(x^{\frac{3}{2}} + y^3)$  and Lemma A.3.1, we deduce, in probability, that

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} F' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) |t_{i+1}^n \wedge t - t_i^n \wedge t| \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right| \\ & \leq \lim_{n \rightarrow +\infty} \frac{2 \|F'\|_\infty}{3} \left( \sum_{i=0}^{p_n-1} |t_{i+1}^n \wedge t - t_i^n \wedge t|^{\frac{3}{2}} + \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 \right) = 0 \end{aligned}$$

and this completes the proof of Lemma 2.5.3.  $\square$

### 2.5.4 Expressions of terms of (2.11) involving $F''$

From the expression (2.13) of  $(\mathbf{B})_i$ , we have

$$(\mathbf{B})_i = \sum_{k=1}^5 \frac{1}{2} F'' \left( \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right) (\mathbf{B})_i^k,$$

where

$$\begin{aligned} (\mathbf{B})_i^1 &= \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_{i+1}^n \wedge t} \rangle^2 - \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle^2, \\ (\mathbf{B})_i^2 &= -2 \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \left[ \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_{i+1}^n \wedge t} \rangle - \langle g \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle \right], \\ (\mathbf{B})_i^3 &= [M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id})]^2 \langle g' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle^2, \\ (\mathbf{B})_i^4 &= - \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \langle g'' \circ \tau_{-R(t_i^n \wedge t)}, Y_s \rangle ds [M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id})] \langle g' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle, \\ (\mathbf{B})_i^5 &= -2 \left[ M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R(t_i^n \wedge t)} \right) \right] \\ & \quad \times [M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id})] \langle g' \circ \tau_{-R(t_i^n \wedge t)}, Y_{t_i^n \wedge t} \rangle. \end{aligned}$$

As in Section 2.5.3, we treat each of the previous terms successively to prove the following lemma:

**Lemma 2.5.4.** *When the mesh of the subdivision  $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = T$  of  $[0, T]$  tends to 0 when  $n \rightarrow +\infty$ , we obtain in probability that*

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} (\mathbf{B})_i = \lambda \int_0^t F''(\langle g, Z_s \rangle) \left[ \langle g^2, Z_s \rangle - \langle g, Z_s \rangle^2 + \langle g', Z_s \rangle^2 M_2(Z_s) - 2 \langle g', Z_s \rangle \langle g \times \text{id}, Z_s \rangle \right] ds + \widehat{\text{Mart}}_t$$

where  $(\widehat{\text{Mart}}_t)_{0 \leq t \leq T}$  is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale.

The proof is similar to Lemma 2.5.3: we use the martingale problem (2.7) to write

$$\begin{aligned} (\mathbf{B})_i^1 &= 2 \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_s \right\rangle \frac{1}{2} \left\langle g'' \circ \tau_{-R}(t_i^n \wedge t), Y_s \right\rangle ds \\ &\quad + 2\lambda \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \left[ \left\langle g^2 \circ \tau_{-R}(t_i^n \wedge t), Y_s \right\rangle - \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_s \right\rangle^2 \right] ds \\ &\quad + M_{t_{i+1}^n \wedge t}^{\text{id}^2} \left( g \circ \tau_{-R}(t_i^n \wedge t) \right) - M_{t_i^n \wedge t}^{\text{id}^2} \left( g \circ \tau_{-R}(t_i^n \wedge t) \right), \\ (\mathbf{B})_i^2 &= -2 \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \right\rangle \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \frac{1}{2} \left\langle g'' \circ \tau_{-R}(t_i^n \wedge t), Y_s \right\rangle ds \\ &\quad - 2 \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \right\rangle \left[ M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R}(t_i^n \wedge t) \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g \circ \tau_{-R}(t_i^n \wedge t) \right) \right], \end{aligned}$$

and we obtain that, in probability,

$$\begin{aligned} &\lim_{n \rightarrow +\infty} \frac{1}{2} \sum_{i=0}^{p_n-1} F'' \left( \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \right\rangle \right) (\mathbf{B})_i^1 \\ &= \int_0^t F''(\langle g, Z_s \rangle) \left\{ \langle g, Z_s \rangle \left\langle \frac{g''}{2}, Z_s \right\rangle + \lambda \left[ \langle g^2, Z_s \rangle - \langle g, Z_s \rangle^2 \right] \right\} ds + \widehat{\text{Mart}}_t^{(1)}, \\ &\lim_{n \rightarrow +\infty} \frac{1}{2} \sum_{i=0}^{p_n-1} F'' \left( \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \right\rangle \right) (\mathbf{B})_i^2 \\ &= - \int_0^t F''(\langle g, Z_s \rangle) \langle g, Z_s \rangle \left\langle \frac{g''}{2}, Z_s \right\rangle ds + \widehat{\text{Mart}}_t^{(2)}, \\ &\lim_{n \rightarrow +\infty} \frac{1}{2} \sum_{i=0}^{p_n-1} F'' \left( \left\langle g \circ \tau_{-R}(t_i^n \wedge t), Y_{t_i^n \wedge t} \right\rangle \right) \left( (\mathbf{B})_i^3 + (\mathbf{B})_i^4 + (\mathbf{B})_i^5 \right) \\ &= \lambda \int_0^t F''(\langle g, Z_s \rangle) \left( \langle g', Z_s \rangle^2 M_2(Z_s) - 2 \langle g', Z_s \rangle \langle g \times \text{id}, Z_s \rangle \right) ds + \widehat{\text{Mart}}_t^{(3)}, \end{aligned}$$

where  $(\widehat{\text{Mart}}_t^{(j)})_{0 \leq t \leq T}$ ,  $j \in \{1, 2, 3\}$  are  $\mathbb{P}_\nu^{\text{FV}}$ -martingales.

### 2.5.5 Error terms

In this section, we examine the different error terms involved in the approximation (2.11). From Lemma A.3.1, we deduce that, in probability,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \left| \left\langle \text{id}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \right|^3 = \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 = 0.$$

Using the relations (2.39) and (2.40), we deduce for any  $k \in \{0, 1, 2\}$ ,

$$\begin{aligned} & \sum_{i=0}^{p_n-1} \left| \left\langle g^{(k)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \right|^3 \\ & \leq \max \left\{ \frac{1}{2} \|g^{(k)}\|_\infty^3, 4 \right\} \sum_{i=0}^{p_n-1} \left( |t_{i+1}^n \wedge t - t_i^n \wedge t|^3 \right. \\ & \quad \left. + \left| M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(k)} \circ \tau_{-R(t_i^n \wedge t)} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(k)} \circ \tau_{-R(t_i^n \wedge t)} \right) \right|^3 \right). \end{aligned}$$

Hence,

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \left| \left\langle g^{(k)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle}, Y_{t_{i+1}^n \wedge t} - Y_{t_i^n \wedge t} \right\rangle \right|^3 = 0$$

in probability. Combining all the previous results, we deduce that (2.10) is a martingale for all  $F, g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$ .

### 2.5.6 Extension to test functions $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ and $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$

For all  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  and for all  $t \geq 0$ , we have  $\langle g, Y_t \rangle \in [-\|g\|_\infty, \|g\|_\infty]$ , so we can assume without loss of generality that  $F \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  in the martingale problem (2.4) with  $X_t$  replaced by  $Z_t$ . Let  $F, g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . Then, by density arguments, there exists  $(F_n)_{n \in \mathbb{N}}, (g_n)_{n \in \mathbb{N}} \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})^{\mathbb{N}}$  such that for all  $t \geq 0$ , for all  $i \in \{0, 1, 2\}$ , we have  $\mathbb{P}_\mu$ -a.s.

$$g_n^{(i)} \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_\infty} g^{(i)} \quad \text{and} \quad F_n^{(i)} \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_\infty} F^{(i)}. \quad (2.41)$$

For all  $n \in \mathbb{N}$ , for all  $t \geq 0$ ,  $\widehat{M}_t^{F_n}(g_n)$  given by the martingale problem (2.4) is a  $\mathbb{P}_\mu$ -martingale. Let  $T > 0$ . Note that, there exists a constant  $C$  for all  $n \in \mathbb{N}$ , for all  $t \in [0, T]$  such that

$$\left| \widehat{M}_t^{F_n}(g_n) - \widehat{M}_t^F(g) \right| \leq C \left( 1 + \int_0^T M_2(Z_s) ds \right).$$

Thanks to Proposition 2.2.11, (2.41) and the dominated convergence theorem, we have for all  $t \in [0, T]$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left( \left| \widehat{M}_t^{F_n}(g_n) - \widehat{M}_t^F(g) \right| \right) = 0,$$

and  $\widehat{M}_t^F(g) \in L^1(\widetilde{\Omega})$ . Then, using the dominated convergence theorem for conditional expectation, we obtain that  $(\widehat{M}_t^F(g))_{0 \leq t \leq T}$  is a  $\mathbb{P}_\mu$ -martingale.

### 2.5.7 $L^2$ -martingale and quadratic variation

As  $(\widehat{M}_t^F(g))_{0 \leq t \leq T}$  is a  $\mathbb{P}_\mu$ -martingale for  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  we deduce that  $(\widehat{M}_t^F(g)^2 - \langle \widehat{M}^F(g) \rangle_t)_{0 \leq t \leq T}$  is a  $\mathbb{P}_\mu$ -local martingale. In consequence, there exists a increasing sequence of stopping times  $(\tau_n)_{n \in \mathbb{N}}$  satisfying  $\lim_{n \rightarrow +\infty} \tau_n = +\infty$  such that for all  $n \in \mathbb{N}$ ,

$$\mathbb{E} \left( \widehat{M}_{t \wedge \tau_n}^F(g)^2 \right) = \mathbb{E} \left( \langle \widehat{M}^F(g) \rangle_{t \wedge \tau_n} \right).$$

We apply ITÔ's formula to compute  $F^2(\langle g, X_{t \wedge \tau_n} \rangle)$  from DOOB's semi-martingale decomposition of  $F(\langle g, X_{t \wedge \tau_n} \rangle)$  and apply the martingale problem (2.4) to test functions  $F^2$  and  $g$  to deduce (2.6) at time  $t \wedge \tau_n$ . By FATOU's lemma, letting  $n \rightarrow +\infty$ , we deduce that

$$\mathbb{E} \left( \widehat{M}_t^F(g)^2 \right) \leq \mathbb{E} \left( \left\langle \widehat{M}^F(g) \right\rangle_t \right) < \infty.$$

This ends the proof of Theorem 2.2.3.

### 2.5.8 Technical result for Lemma 2.5.2

As the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$  is Polish (see Section 2.2.1), we deduce from [134, Theorem 3.18 of Section 5.3 (p 307)] there exists, for all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , a unique family  $(\mathbb{Q}_\omega)_{\omega \in \Omega}$  of regular conditional probability of  $\mathbb{P}_\nu^{\text{FV}}$  given  $\mathcal{F}_{t^*}$  and a  $\mathbb{P}_\nu^{\text{FV}}$ -null event  $N \in \mathcal{F}_{t^*}$  such that for all  $\omega \in \Omega \setminus N$ ,

$$\mathbb{Q}_\omega \left( \left\{ \tilde{\omega} \in \Omega \mid \tilde{\omega}_{t^*} = \omega_{t^*} \right\} \right) = 1. \quad (2.42)$$

The following Theorem 2.5.5 ensures that time shifts of regular conditional probabilities of  $\mathbb{P}_\nu^{\text{FV}}$  remain solutions to the FLEMING-VIOT martingale problem (2.7). The proof of this result is given hereafter and is based on the proof of [134, Lemma 4.19 of Section 5.4 (p 321)]. We introduce, for  $\omega \in \Omega$ , the time-shift operator  $\theta$  defined by

$$[\theta_s \omega]_t := \omega_{s+t}, \quad 0 \leq t < +\infty, \quad s \geq 0.$$

**Theorem 2.5.5.** *Let  $t^* \in \mathbb{R}_+$  be a deterministic time. Then there exists a  $\mathbb{P}_\nu^{\text{FV}}$ -null event  $N \in \mathcal{F}_{t^*}$  such that, for every  $\omega \in \Omega \setminus N$ , the probability measure*

$$\mathbb{P}_\omega(d\tilde{\omega}) := \theta_{t^*} \# \mathbb{Q}_\omega(d\tilde{\omega}) \quad (2.43)$$

*solves the martingale problem (2.7) with  $\nu := \omega_{t^*}$ .*

*Proof. Step 0. Preliminary results.* We denote by  $\mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$  the space of real functions of class  $\mathcal{C}^2(\mathbb{R}, \mathbb{R})$  with compact support. It is well-known that the formulation of the martingale problem (2.7) for  $F, g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  is equivalent to the one for  $F, g \in \mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$  [61]. The space  $\mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$  equipped with the norm  $\|f\|_{W_0^{2,\infty}} := \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty$  is separable. So, we can choose a dense countable family  $\mathcal{B} \subset \mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$ , for the topology associated to the norm, that is to say

$$\forall F, g \in \mathcal{C}_K^2(\mathbb{R}, \mathbb{R}), \quad \exists (F_n)_{n \in \mathbb{N}}, (g_n)_{n \in \mathbb{N}} \in \mathcal{B}^{\mathbb{N}}, \quad F_n \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_{W_0^{2,\infty}}} F, \quad g_n \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_{W_0^{2,\infty}}} g.$$

Hence, if we denote by  $\mathcal{L}_{\text{FV}}$  the generator of the original FLEMING-VIOT process, we deduce that  $\mathcal{L}_{\text{FV}}(F_n)_{g_n} \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_\infty} \mathcal{L}_{\text{FV}} F_g$ .

**Step 1. Reformulation of the goal.** Let  $\nu \in \mathcal{M}_1(\mathbb{R})$ . From (2.42), it follows that  $\mathbb{P}_\omega(\tilde{\omega}_0 = \nu) = 1$  is satisfied with  $\nu := \omega_{t^*}$ . The rest of the proof is devoted to construct a  $\mathbb{P}_\nu^{\text{FV}}$ -null event  $N_4$  such that

$$\mathbb{E}_{\mathbb{P}_{\tilde{\omega}}^{\text{FV}}} \left[ F_g(\omega_t) - F_g(\omega_s) - \int_s^t \mathcal{L}_{\text{FV}} F_g(\omega_r) dr \mid \mathcal{F}_s \right] = 0,$$



is satisfied for all  $\omega \in \Omega \setminus N_4$ . This means that for all  $0 \leq s < t < \infty$ ,  $A \in \mathcal{F}_s$ ,  $F, g \in \mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$ ,

$$\forall \omega \in \Omega \setminus N_4, \quad \int_{\Omega} \left[ M_t^{Fg}(\tilde{\omega}) - M_s^{Fg}(\tilde{\omega}) \right] \mathbb{1}_A(\tilde{\omega}) \mathbb{P}_{\omega}(d\tilde{\omega}) = 0, \quad (2.44)$$

where

$$M_t^{Fg}(\tilde{\omega}) := F_g(\tilde{\omega}_t) - F_g(\tilde{\omega}_0) - \int_0^t \mathcal{L}_{\text{FV}} F_g(\tilde{\omega}_r) dr.$$

Let  $\omega \in \Omega$ ,  $0 \leq s < t < \infty$ ,  $A \in \mathcal{F}_s$ ,  $F, g \in \mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$  be fixed.

**Step 2. Property (2.44) satisfied except on a  $\mathbb{P}_{\nu}^{\text{FV}}$ -null event  $N_1(s, t, A, F, g) \in \mathcal{F}_{t^*}$ .** As  $\mathcal{L}_{\text{FV}} F_g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ , the random variable  $M_t^{Fg} - M_s^{Fg}$  is bounded. Note that,

$$\begin{aligned} & \int_{\Omega} \left[ M_t^{Fg}(\tilde{\omega}) - M_s^{Fg}(\tilde{\omega}) \right] \mathbb{1}_A(\tilde{\omega}) \mathbb{P}_{\omega}(d\tilde{\omega}) \\ &= \mathbb{E}_{\mathbb{Q}_{\omega}(d\tilde{\omega})} \left( \left[ M_t^{Fg} - M_s^{Fg} \right] \circ \theta_{t^*}(\tilde{\omega}) \mathbb{1}_{\theta_{t^*}^{-1}A}(\tilde{\omega}) \right) \\ &= \mathbb{E}_{\mathbb{P}_{\nu}^{\text{FV}}} \left( \left[ M_t^{Fg} - M_s^{Fg} \right] \circ \theta_{t^*} \mathbb{1}_{\theta_{t^*}^{-1}A} \middle| \mathcal{F}_{t^*} \right) (\omega) \\ &= \mathbb{E}_{\mathbb{P}_{\nu}^{\text{FV}}} \left[ \mathbb{E}_{\mathbb{P}_{\nu}^{\text{FV}}} \left( \left[ M_t^{Fg} - M_s^{Fg} \right] \circ \theta_{t^*} \mathbb{1}_{\theta_{t^*}^{-1}A} \middle| \mathcal{F}_{t^*+s} \right) \middle| \mathcal{F}_{t^*} \right] (\omega) \\ &= \mathbb{E}_{\mathbb{P}_{\nu}^{\text{FV}}} \left[ \mathbb{1}_{\theta_{t^*}^{-1}A} \mathbb{E}_{\mathbb{P}_{\nu}^{\text{FV}}} \left( \left[ M_t^{Fg} - M_s^{Fg} \right] \circ \theta_{t^*} \middle| \mathcal{F}_{t^*+s} \right) \middle| \mathcal{F}_{t^*} \right] (\omega) \\ &= 0, \end{aligned}$$

where the last equality follows from martingale property (2.7). This chain of equalities shows that the random variable  $\omega \mapsto \int_A \left[ M_t^{Fg}(\tilde{\omega}) - M_s^{Fg}(\tilde{\omega}) \right] \mathbb{P}_{\omega}(d\tilde{\omega})$  is null except on a  $\mathbb{P}_{\nu}^{\text{FV}}$ -null event  $N_1(s, t, A, F, g) \in \mathcal{F}_{t^*}$  which depends on  $s, t, A, F$  and  $g$ .

**Step 3. Property (2.44) satisfied except on a  $\mathbb{P}_{\nu}^{\text{FV}}$ -null event  $N_2(s, t, F, g) \in \mathcal{F}_{t^*}$ .** We consider a countable subcollection  $\mathcal{E}$  of  $\mathcal{F}_s$  which generates  $\mathcal{F}_s$  [134, Definition 3.17 of Section 5.3 (p 306)] and a  $\mathbb{P}_{\nu}$ -null event  $N_2(s, t, F, g) \in \mathcal{F}_{t^*}$  such that for  $\omega \in \Omega \setminus N_2(s, t, F, g)$ ,

$$\forall A \in \mathcal{E}, \quad \int_A \left[ M_t^{Fg}(\tilde{\omega}) - M_s^{Fg}(\tilde{\omega}) \right] \mathbb{P}_{\omega}(d\tilde{\omega}) = 0.$$

Therefore, the two measures

$$v_{\omega}^+(A) := \int_A \left[ M_t^{Fg} - M_s^{Fg} \right]^+(\tilde{\omega}) \mathbb{P}_{\omega}(d\tilde{\omega}) \quad \text{and} \quad v_{\omega}^-(A) := \int_A \left[ M_t^{Fg} - M_s^{Fg} \right]^-(\tilde{\omega}) \mathbb{P}_{\omega}(d\tilde{\omega}),$$

coincide on  $\mathcal{E}$ , hence on  $\mathcal{F}_s$ . Therefore, for  $\omega \in \Omega \setminus N_2(s, t, F, g)$ , we have proved that for all  $A \in \mathcal{F}_s$ ,  $\mathbb{E}_{\mathbb{P}_{\omega}} \left( \mathbb{1}_A \left[ M_t^{Fg} - M_s^{Fg} \right] \right) = 0$ .

**Step 4. Property (2.44) satisfied except on a  $\mathbb{P}_{\nu}^{\text{FV}}$ -null event  $N_3(F, g) \in \mathcal{F}_{t^*}$ .** We may set now, the  $\mathbb{P}_{\nu}^{\text{FV}}$ -null event

$$N_3(F, g) := \bigcup_{\substack{s, t \in \mathbb{Q} \\ 0 \leq s < t < \infty}} N_2(s, t, F, g).$$

Due to the boundedness and continuity of  $t \mapsto M_t^{Fg}$ , it follows from the dominated convergence theorem that for  $\omega \in \Omega \setminus N_3(F, g)$

$$\forall s < t, \forall A \in \mathcal{F}_s, \quad \mathbb{E}_{\mathbb{P}_\omega} \left( \mathbf{1}_A \left[ M_t^{Fg} - M_s^{Fg} \right] \right) = 0,$$

in other words, for all  $\omega \in \Omega \setminus N_3(F, g)$ ,  $(M_t^{Fg}(\tilde{\omega}))_{t \geq 0}$  is a  $(\mathcal{F}_s, \mathbb{P}_\omega(d\tilde{\omega}))$ -martingale.

**Step 5. Conclusion.** Now we define the  $\mathbb{P}_\nu^{\text{FV}}$ -null event

$$N_4 := \bigcup_{F, g \in \mathcal{B}} N_3(F, g)$$

From the Step 4, we have for all  $s \leq t$ ,

$$\forall \omega \in \Omega \setminus N_4, \quad \forall A \in \mathcal{F}_s, \quad \forall F, g \in \mathcal{B}, \quad \mathbb{E}_{\mathbb{P}_\omega} \left[ \mathbf{1}_A \left( M_t^{Fg} - M_s^{Fg} \right) \right] = 0.$$

From Step 0, for all  $F, g \in \mathcal{C}_K^2(\mathbb{R}, \mathbb{R})$ , there exist two sequences  $(F_n)_{n \in \mathbb{N}}, (g_n)_{n \in \mathbb{N}} \in \mathcal{B}^{\mathbb{N}}$  such that

$$F_n \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_{W_0^{2,\infty}}} F, \quad g_n \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_{W_0^{2,\infty}}} g, \quad \text{and} \quad \mathcal{L}_{\text{FV}}(F_n)_{g_n} \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_\infty} \mathcal{A}Fg.$$

By the dominated convergence theorem, we deduce that for all  $\omega \in \Omega$ ,  $s \leq t$ , and  $A \in \mathcal{F}_s$ ,

$$\mathbb{E}_{\mathbb{P}_\omega} \left[ \mathbf{1}_A \left( M_t^{Fg} - M_s^{Fg} \right) \right] = \lim_{n \rightarrow +\infty} \mathbb{E}_{\mathbb{P}_\omega} \left[ \mathbf{1}_A \left( M_t^{(F_n)_{g_n}} - M_s^{(F_n)_{g_n}} \right) \right] = 0.$$

which concludes the proof.  $\square$

### 2.5.9 Proof of Lemma 2.5.2

By abuse of notation, we note  $h(\omega|_{[0, t^*]}) = h(\omega)$ . We want to prove that for all  $0 \leq s \leq t$ , for all  $\mathcal{F}_s$ -measurable bounded random variable  $Z$ ,

$$\mathbb{E}_{\mathbb{P}_\nu^{\text{FV}}(d\tilde{\omega})} ([\mathcal{M}_t(\tilde{\omega}) - \mathcal{M}_s(\tilde{\omega})] Z(\tilde{\omega})) = 0.$$

Using [127, Definition 3.2 (iii)' of Section 1] we deduce that

$$\mathbb{E}_{\mathbb{P}_\nu^{\text{FV}}(d\tilde{\omega})} ([\mathcal{M}_t(\tilde{\omega}) - \mathcal{M}_s(\tilde{\omega})] Z(\tilde{\omega})) = \mathbb{E}_{\mathbb{P}_\nu(d\omega)} \left[ \mathbb{E}_{\mathbb{Q}_\omega(d\tilde{\omega})} ([\mathcal{M}_t(\tilde{\omega}) - \mathcal{M}_s(\tilde{\omega})] Z(\tilde{\omega})) \right].$$

Thus, it is sufficient to prove that for  $\mathbb{P}_\nu^{\text{FV}}$ -almost every  $\omega \in \Omega$ ,  $(\mathcal{M}_t(\tilde{\omega}))_{0 \leq t \leq T}$  is a  $\mathbb{Q}_\omega(d\tilde{\omega})$ -martingale and this is what we propose to establish in the rest of this proof.

For fixed  $\omega$ , the function  $h(\omega) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  can be considered as deterministic. We deduce from Theorem 2.5.5 that there exists a  $\mathbb{P}_\nu^{\text{FV}}$ -null event  $N \in \mathcal{F}_{t^*}$  such that for all  $\omega \in \Omega \setminus N$ ,  $(M_t^{\text{id}}(h(\omega))(\tilde{\omega}))_{0 \leq t \leq T}$  is a  $\mathbb{P}_\omega(d\tilde{\omega})$ -martingale. We deduce from [134, Theorem 3.18 of Section 5.3 (p 307)] that  $\mathbb{P}_\nu^{\text{FV}}$ -almost every  $\omega \in \Omega$ ,

$$\tilde{\omega}|_{[0, t^*]} = \omega|_{[0, t^*]}, \quad \mathbb{Q}_\omega(d\tilde{\omega})\text{-a.s.} \quad (2.45)$$

This implies that,  $\mathbb{Q}_\omega(d\tilde{\omega})$ -almost surely,

$$\begin{aligned} \mathcal{M}_t(\tilde{\omega}) &= M_t^{\text{id}}(h(\omega))(\tilde{\omega}) - M_{t \wedge t^*}^{\text{id}}(h(\omega))(\tilde{\omega}) = \begin{cases} M_{t-t^*}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) & \text{if } t > t^* \\ 0 & \text{if } t \leq t^* \end{cases} \\ &= M_{(t-t^*)^+}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) \end{aligned}$$

where  $(a)^+$  designates the non-negative part of  $a \in \mathbb{R}$ . Let  $n \in \mathbb{N}^*$  and  $0 \leq s \leq T$ . To prove the martingale property for all  $\mathcal{F}_s$ -measurable bounded random variable  $Z$ , it is sufficient to prove it on elementary events. Then, we consider a random variable  $Z$  of the form

$$Z(\omega) := \mathbb{1}_{\{\omega_{t_1} \in \Gamma_1, \dots, \omega_{t_n} \in \Gamma_n\}}$$

where for all  $i \in \{1, \dots, n\}$ ,  $t_i \leq s$  and  $\Gamma_i \subset \mathcal{M}_1(\mathbb{R})$  measurable. We define

$$\tilde{Z}(\omega, \tilde{\omega}) := \mathbb{1}_{\{\omega_{t_i} \in \Gamma_i, \forall i \in \{1, \dots, n\} \text{ such that } t_i \leq t^*\}} \mathbb{1}_{\{\tilde{\omega}_{t_j} \in \Gamma_j, \forall j \in \{1, \dots, n\} \text{ such that } t_j > t^*\}}.$$

By (2.45),  $\tilde{Z}(\omega, \tilde{\omega}) = Z(\omega)$ ,  $\mathbb{Q}_\omega(d\tilde{\omega})$  - a.s. Therefore, for  $\mathbb{P}_\nu^{\text{FV}}$ -almost every  $\omega \in \Omega$ ,

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}_\omega}([\mathcal{M}_t - \mathcal{M}_s]Z) \\ &= \mathbb{E}_{\mathbb{Q}_\omega(d\tilde{\omega})} \left( \left[ M_{(t-t^*)+}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) - M_{(s-t^*)+}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) \right] \tilde{Z}(\omega, \tilde{\omega}) \right) \\ &= \mathbb{1}_{\{\omega_{t_i} \in \Gamma_i, \forall i \in \{1, \dots, n\} \text{ such that } t_i \leq t^*\}} \times \\ & \quad \mathbb{E}_{\mathbb{Q}_\omega(d\tilde{\omega})} \left( \left[ M_{t-t^*}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) - M_{s-t^*}^{\text{id}}(h(\omega))(\theta_{t^*}(\tilde{\omega})) \right] \right. \\ & \quad \left. \times \mathbb{1}_{\{[\theta_{t^*}(\tilde{\omega})]_{t_j-t^*} \in \Gamma_j, \forall j \in \{1, \dots, n\} \text{ such that } t_j > t^*\}} \right) \\ & \stackrel{(2.43)}{=} \mathbb{1}_{\{\omega_{t_i} \in \Gamma_i, \forall i \in \{1, \dots, n\} \text{ such that } t_i \leq t^*\}} \mathbb{E}_{\mathbb{P}_\omega(d\hat{\omega})} \left( \left[ M_{t-t^*}^{\text{id}}(h(\omega))(\hat{\omega}) - M_{s-t^*}^{\text{id}}(h(\omega))(\hat{\omega}) \right] \right. \\ & \quad \left. \times \mathbb{1}_{\{\hat{\omega}_{t_j-t^*} \in \Gamma_j, \forall j \in \{1, \dots, n\} \text{ such that } t_j > t^*\}} \right) \\ &= 0, \end{aligned}$$

using that  $(M_{t-t^*}^{\text{id}}(h(\omega))(\hat{\omega}))_{t^* \leq t \leq T+t^*}$  is a  $\mathbb{P}_\omega(d\hat{\omega})$ -martingale if the internal indicator is non zero. Thus, for  $\mathbb{P}_\nu^{\text{FV}}$ -almost every  $\omega \in \Omega$ ,  $(\mathcal{M}_t(\tilde{\omega}))_{0 \leq t \leq T}$  is a  $\mathbb{Q}_\omega(d\tilde{\omega})$ -martingale which completes the first part of this proof. In similar way, we can prove

$$M_t^{\text{id}^2}(h(\tilde{\omega}))(\tilde{\omega}) - M_{t \wedge t^*}^{\text{id}^2}(h(\tilde{\omega}))(\tilde{\omega})$$

is a  $\mathbb{P}_\nu^{\text{FV}}$ -martingale. Applying ITO's formula to compute  $\langle h(\tilde{\omega}_t), \tilde{\omega}_t \rangle^2$  and comparing it to the previous result, we obtain the announced result.  $\square$

## 2.6 Proof of the results of Section 2.3

### 2.6.1 Study of a semi-group

In this section, we devote a specific study to the semi-group  $(T^{(n)}(t))_{t \geq 0}$  generated by the operator  $B^{(n)}$ . In Section 2.6.1.1, we provide an explicit expression of  $(t, x) \mapsto T^{(n)}(t)f(x)$  and prove that it is a strong solution to the semi-group PDE associated with  $B^{(n)}$ , by means of FEYNMAN-KAC's formula. With the aim of subsequently obtaining fairly fine bounds on this operator (see Corollary 2.6.2), we give all the necessary details. In Section 2.6.1.2, we give a MILD formulation of the martingale problem (2.20) using the semigroup  $(T^{(n)}(t))_{t \geq 0}$  in Proposition 2.6.3.

### 2.6.1.1 Construction of the semi-group

Recall that we denote by  $\mathbf{1} \in \mathbb{R}^n$ , the vector whose coordinates are all 1. For any real vector-valued function  $f$  and  $g$  of  $L^1(\mathbb{R}^n)$ , we denote by  $(f * g)(x) := \int_{\mathbb{R}^n} f(t)g(x-t)dt$  the convolution product of  $f$  and  $g$ . For any function  $f$  whose second partial derivatives exist, we denote by  $\text{Hess}(f) := \left( \partial_{ij}^2 f \right)_{1 \leq i, j \leq n}$  the Hessian matrix of  $f$ . We denote by  $\mathcal{C}_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n, \mathbb{R})$  the space of real functions on  $\mathbb{R}_+ \times \mathbb{R}^n$  of class  $\mathcal{C}^1(\mathbb{R}_+, \mathbb{R})$  with respect to the first variable and of class  $\mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  to the second variable.

**Theorem 2.6.1.** *The family of operators  $(T^{(n)}(t))_{t \geq 0}$  defined as:*

$$\begin{aligned} \forall t > 0, \forall x \in \mathbb{R}^n, \quad T^{(n)}(t)f(x) &:= \int_{\mathbb{R}^n} f(u)g_{t,x}^X(u)du, \\ \forall x \in \mathbb{R}^n, \quad T^{(n)}(0)f(x) &:= f(x), \end{aligned} \quad (2.46)$$

where  $g_{t,x}^X$  is a density of the Gaussian distribution  $\mathcal{N}^{(n)}(m_{t,x}, \Sigma_t)$  where  $\Sigma_t := P\sigma_t P^{-1}$  and  $m_{t,x} := P\mu_{t,P^{-1}x} = x - \frac{(1-\exp(-2\lambda nt))}{n}(x \cdot \mathbf{1})\mathbf{1}$  with

$$\mu_{t,y} := \begin{pmatrix} y_1 \exp(-2\lambda nt) \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \text{and} \quad \sigma_t := \begin{pmatrix} e_4(t) & 0 & \dots & \dots & 0 \\ 0 & t & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & t \end{pmatrix},$$

where  $e_4(t) := \frac{1-\exp(-4\lambda nt)}{4\lambda n}$  and  $P$  is an explicit change of orthonormal basis matrix defined in the proof below, is a semi-group of bounded operators on  $L^\infty(\mathbb{R}^n)$ . In addition, for all  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

- (1) *The application  $(t, x) \mapsto T^{(n)}(t)f(x)$  is of class  $\mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n, \mathbb{R})$  and is a strong solution of the PDE*

$$\forall t \geq 0, \forall x \in \mathbb{R}^n, \quad \partial_t u(t, x) = \frac{1}{2}\Delta u(t, x) - 2\lambda(\nabla u(t, x) \cdot \mathbf{1})(x \cdot \mathbf{1}) \quad (2.47)$$

$$\forall x \in \mathbb{R}^n, \quad u(0, x) = f(x), \quad (2.48)$$

and

- (2)  $\nabla T^{(n)}(t)f(x) = (\partial_{x_i} m_{t,x} \cdot (\nabla f * g_{t,0}^X)(m_{t,x}))_{1 \leq i \leq n}^t$   
 (3)  $\forall i, j \in \{1, \dots, n\}, \quad \partial_{x_i x_j}^2 T^{(n)}(t)f(x) = (\partial_{x_j} m_{t,x})^t [(f * \text{Hess}(g_{t,0}^X))(m_{t,x})\partial_{x_i} m_{t,x}]$

where  $\partial_{x_i} m_{t,x} = \epsilon_i - \frac{1-\exp(-2\lambda nt)}{n}\mathbf{1}$  with  $(\epsilon_1, \dots, \epsilon_n)$  the canonical basis of  $\mathbb{R}^n$ .

As we will see in the proof, everything follows quite directly from the FEYNMAN-KAC formula, except the fact that  $(t, x) \mapsto T^{(n)}(t)f(x)$  is a strong solution of the PDE up to time  $t = 0$ . This technical point will be useful for the MILD formulation and this is why we make a detailed proof.

*Proof.* In view of the operator  $B^{(n)}$  given by (2.18), it is natural to define the semi-group  $T^{(n)}(t)$  using the FEYNMAN-KAC formula: for any  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$T^{(n)}(t)f(x) := \mathbb{E}_x f(X_t)$$

where  $(X_t)_{t \geq 0}$  is solution to the following SDE:

$$X_0 = x, \quad dX_t = dB_t - 2\lambda(X_t \cdot \mathbf{1}) \mathbf{1} dt, \quad X_t \in \mathbb{R}^n, \quad t > 0 \quad (2.49)$$

where  $(B_t)_{t \geq 0}$  is a  $n$ -standard Brownian motion and  $x \in \mathbb{R}^n$ .

In Step 1, we check that this definition of  $T^{(n)}(t)$  coincides with the one given in the statement of Theorem 2.6.1. In Step 2, we verify that  $(x, t) \mapsto \mathbb{E}_x f(X_t)$  is indeed a solution of the PDE (2.47) for all  $t > 0$ . In Step 3, we treat the case  $t = 0$ . In Step 4, we prove the announced expressions of the derivatives of  $T^{(n)}(t)f(x)$ .

**Step 1. Change of basis in the SDE (2.49).** We consider the orthonormal basis  $(v_1, \dots, v_n)$  of  $\mathbb{R}^n$  defined by  $v_1 := \frac{1}{\sqrt{n}}(1 \ \dots \ 1)^t$  and for  $2 \leq i \leq n$ ,

$$v_i := \sqrt{\frac{i-1}{i}} \left( \underbrace{\frac{1}{i-1}, \dots, \frac{1}{i-1}}_{i-1 \text{ terms}}, -1, 0, \dots, 0 \right)^t.$$

We denote by  $P$  the change of basis matrix from the canonical basis to the orthonormal basis  $(v_1, \dots, v_n)$ . We define for all  $t \geq 0$ ,  $Z_t = P^{-1}X_t$ , i.e.  $Z_t := (Z_t^{(1)}, \dots, Z_t^{(n)})$  where for all  $i \in \{1, \dots, n\}$ ,  $Z_t^{(i)} := (X_t \cdot v_i)$ . It is standard to check that  $W_t := (W_t^{(1)}, \dots, W_t^{(n)})$  where for all  $i \in \{1, \dots, n\}$ ,  $W_t^{(i)} := (B_t \cdot v_i)$  is a  $n$ -standard Brownian motion and that  $(Z_t)_{t \geq 0}$  is solution to the SDE

$$Z_0 = y = P^{-1}x, \quad \begin{cases} dZ_t^{(1)} = dW_t^{(1)} - 2\lambda n Z_t^{(1)} dt, \\ dZ_t^{(j)} = dW_t^{(j)}, \end{cases} \quad j \in \{2, \dots, n\}. \quad (2.50)$$

All coordinates in (2.50) are independent and solve a one-dimensional SDE whose solution is explicit (ORNSTEIN-UHLENBECK for  $Z^{(1)}$ , standard Brownian motion for  $Z^{(i)}, i \geq 2$ ). It follows that  $Z_t$  is a Gaussian vector of law  $\mathcal{N}^{(n)}(\mu_{t,y}, \sigma_t)$ . Therefore, for any  $t > 0$  and all  $y \in \mathbb{R}^n$ ,  $\mathbb{P}_{Z_t}$  has a density with respect to the LEBESGUE measure on  $\mathbb{R}^n$  given by:

$$g_{t,y}^Z(z_1, \dots, z_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\sigma_t)}} \exp \left( -\frac{[z_1 - y_1 \exp(-2\lambda nt)]^2}{2e_4(t)} - \frac{1}{2t} \sum_{j=2}^n (z_j - y_j)^2 \right). \quad (2.51)$$

Since,  $X_t = PZ_t$ , we deduce that for all  $x \in \mathbb{R}^n$  and for all  $t > 0$ ,  $X_t$  follows the normal distribution  $\mathcal{N}^{(n)}(m_{t,x}, \Sigma_t)$ , with density

$$g_{t,x}^X(r) := g_{t,P^{-1}x}^Z(P^{-1}r) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma_t)}} \exp \left( -\frac{(r - m_{t,x})^t \Sigma_t^{-1} (r - m_{t,x})}{2} \right). \quad (2.52)$$

Hence,  $\mathbb{E}_x f(X_t)$  coincides with (2.46).

**Step 2.  $T^{(n)}(t)f$  is solution to (2.47) on  $(0, +\infty) \times \mathbb{R}^n$ .** Without difficulty we verify that for any  $y \in \mathbb{R}^n$ ,  $g_{t,y}^Z$  satisfies the following FOKKER-PLANCK PDE:

$$\forall t > 0, \forall z \in \mathbb{R}^n, \quad \partial_t g_{t,y}^Z(z) = \frac{1}{2} \Delta_y g_{t,y}^Z(z) - 2\lambda n \partial_{y_1} g_{t,y}^Z(z). \quad (2.53)$$

We deduce from (2.52) that

$$\forall y \in \mathbb{R}^n, \forall r \in \mathbb{R}^n, \quad \partial_t g_{t,y}^Z(P^{-1}r) = \partial_t g_{t,Py}^X(r),$$

and, for all  $y, r \in \mathbb{R}^n$ ,  $\partial_{y_i} g_{t,y}^Z(P^{-1}r) = \sum_{k=1}^n P_{ki} \partial_{x_k} g_{t,Py}^X(r)$ . In particular,

$$\partial_{y_1} g_{t,y}^Z(P^{-1}r) = \frac{1}{\sqrt{n}} \left( \nabla_x g_{t,Py}^X(r) \cdot \mathbf{1} \right).$$

In an analogous way, we deduce that

$$\Delta_y g_{t,y}^Z(P^{-1}r) = \sum_{i=1}^n \sum_{k,\ell=1}^n P_{ki} P_{\ell i} \partial_{x_\ell, x_k}^2 g_{t,Py}^X(r) = \Delta_x g_{t,Py}^X(r),$$

because  $P$  is an orthonormal matrix. From (2.53) and since  $(P^{-1}x)_1 = \frac{(x \cdot \mathbf{1})}{\sqrt{n}}$ , we deduce that the density  $g_{t,x}^X$  satisfies:

$$\forall t > 0, \forall x \in \mathbb{R}^n, \forall r \in \mathbb{R}^n, \quad \partial_t g_{t,x}^X(r) - \frac{1}{2} \Delta_x g_{t,x}^X(r) + 2\lambda(x \cdot \mathbf{1}) \nabla_x g_{t,x}^X(r) = 0.$$

Now, the fact that for all  $f \in L^\infty(\mathbb{R}^n, \mathbb{R})$ ,

$$T^{(n)}(t)f(x) = \int_{\mathbb{R}^n} f(r) g_{t,x}^X(r) dr$$

is  $\mathcal{C}_b^{1,2}((0, +\infty) \times \mathbb{R}^n, \mathbb{R})$  and is a solution of (2.47) on  $(0, +\infty) \times \mathbb{R}^n$  follows from the theorem of differentiation under the integral sign. Note that, if  $f$  is continuous,

$$T^{(n)}(t)f(x) = \mathbb{E}_x f(X_t) \xrightarrow[t \rightarrow 0]{} f(x)$$

by the dominated convergence theorem which leads to (2.48).

**Step 3. Verification of (2.47) up to  $t = 0$ .** Assume that  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ . This is equivalent to prove that for all  $x \in \mathbb{R}^n$ ,

$$\lim_{t \rightarrow 0} \frac{\int_{\mathbb{R}^n} f(u) g_{t,x}^X(u) du - f(x)}{t} = \frac{1}{2} \Delta f(x) - 2\lambda(\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}).$$

Let be  $x \in \mathbb{R}^n$  fixed. Using TAYLOR's formula we obtain that

$$\int_{\mathbb{R}^n} f(u) g_{t,x}^X(u) du - f(x) = (\mathbf{A})_t + (\mathbf{B})_t + (\mathbf{C})_t,$$

where

$$(\mathbf{A})_t := \int_{\mathbb{R}^n} ([u - x] \cdot \nabla f(x)) g_{t,x}^X(u) du, \quad (\mathbf{B})_t := \frac{1}{2} \int_{\mathbb{R}^n} (u - x)^t \text{Hess}(f)(x) (u - x) g_{t,x}^X(u) du,$$

$$(\mathbf{C})_t := \int_{\mathbb{R}^n} R_x(u) g_{t,x}^X(u) du,$$

where  $R_x(u) := o(\|x - u\|_2^2)$ . As  $g_{t,x}^X$  is a Gaussian density of  $\mathcal{N}^{(n)}(m_{t,x}, \Sigma_t)$ , where

$$\begin{aligned} x - m_{t,x} &= (1 - \exp(-2\lambda nt)) \frac{(x \cdot \mathbf{1})}{\sqrt{n}} \times \frac{\mathbf{1}}{\sqrt{n}} \underset{t \rightarrow 0}{\sim} -2\lambda(x \cdot \mathbf{1})t, \\ \forall i \in \{1, \dots, n\}, \quad (\Sigma_t)_{ii} &\underset{t \rightarrow 0}{\sim} t, \end{aligned} \quad (2.54)$$

it follows that

$$\begin{aligned} (\mathbf{A})_t &= (\nabla f(x) \cdot [m_{t,x} - x]) = -\frac{(1 - \exp(-2\lambda nt))}{n} (\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}). \\ (\mathbf{B})_t &= \frac{1}{2} \int_{\mathbb{R}^n} (u - m_{t,x})^t \text{Hess}(f)(x) (u - m_{t,x}) g_{t,x}^X(u) du \\ &\quad + \frac{1}{2} (x - m_{t,x})^t \text{Hess}(f)(x) (x - m_{t,x}) \\ &= \frac{1}{2} \sum_{i,j=1}^n \partial_{x_i x_j}^2 f(x) (\Sigma_t)_{ij} + (1 - \exp(-2\lambda nt))^2 \frac{(x \cdot \mathbf{1})^2}{2n^2} \mathbf{1}^t \text{Hess}(f)(x) \mathbf{1}. \end{aligned}$$

Therefore,

$$\lim_{t \rightarrow 0} \frac{(\mathbf{A})_t + (\mathbf{B})_t}{t} = -2\lambda (\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}) + \frac{1}{2} \Delta f(x).$$

Now, it remains to manage the  $(\mathbf{C})_t$  error term. Note that,

$$\begin{aligned} \forall \varepsilon > 0, \exists \alpha > 0, \forall u \in B(x, \alpha), \quad |R_x(u)| &\leq \varepsilon \|u - x\|_2^2. \\ \forall u \in \mathbb{R}^n \setminus B(x, \alpha), \quad |R_x(u)| &\leq 2\|f\|_\infty + \|\nabla f\|_\infty \|u - x\|_2 \\ &\quad + \frac{1}{2} \|\text{Hess}f(x)\|_\infty \|u - x\|_2. \end{aligned}$$

Let  $\varepsilon > 0$ ,  $\alpha > 0$  and  $t_0 \geq 0$  such that for all  $t \in [0, t_0]$ ,  $\|x - m_{t,x}\|_2 \leq \frac{\alpha}{2}$ . Let  $t \in [0, t_0]$ . Separating the domain of integration of the integral of  $(\mathbf{C})_t$  into  $B(x, \alpha)$  and  $\mathbb{R}^n \setminus B(x, \alpha)$ , it follows from the YOUNG and previous inequalities that there exists a constant  $C > 0$  such that

$$\begin{aligned} (\mathbf{C})_t &\leq \varepsilon \int_{\mathbb{R}^n} \|u - x\|_2^2 g_{t,x}^X(u) du + C \int_{\mathbb{R}^n \setminus B(x, \alpha)} (1 + \|u - x\|_2^2) g_{t,x}^X(u) du \\ &\leq 2\varepsilon \sum_{i=1}^n (\Sigma_t)_{ii} + 2(\varepsilon + C) \|x - m_{t,x}\|_2^2 + 2C \int_{\mathbb{R}^n \setminus B(x, \alpha)} (1 + \|u - m_{t,x}\|_2^2) g_{t,x}^X(u) du \end{aligned}$$

Now, for the choice of  $\alpha$  and then the MARKOV inequality, we obtain that

$$\begin{aligned} \int_{\mathbb{R}^n \setminus B(x, \alpha)} (1 + \|u - m_{t,x}\|_2^2) g_{t,x}^X(u) du &\leq \left(1 + \frac{4}{\alpha^2}\right) \int_{\mathbb{R}^n \setminus B(m_{t,x}, \frac{\alpha}{2})} \|u - m_{t,x}\|_2^2 g_{t,x}^X(u) du \\ &\leq \left(1 + \frac{4}{\alpha^2}\right) \frac{\int_{\mathbb{R}^n} \|u - m_{t,x}\|_2^4 g_{t,x}^X(u) du}{\left(\frac{\alpha}{2}\right)^4} \\ &\leq \frac{16n}{\alpha^4} \left(1 + \frac{4}{\alpha^2}\right) \sum_{i=1}^n \int_{\mathbb{R}^n} (u_i - (m_{t,x})_i)^4 g_{t,x}^X(u) du. \end{aligned}$$

As for all  $i \in \{1, \dots, n\}$ , the fourth moment of a random variable of law  $\mathcal{N}(0, (\Sigma_t)_{ii})$  is smaller than  $3(\Sigma_t)_{ii}^2$ , it follows from (2.54) that there exists a constant  $\tilde{C} > 0$  such that

$$\frac{(\mathbf{C})_t}{t} \leq 2\varepsilon n + \varepsilon \tilde{C}$$

for  $t$  small enough and then the conclusion.

**Step 4. Expression of the derivatives of  $T^{(n)}(t)f$ .** Noting that for all  $u \in \mathbb{R}^n$ ,  $g_{t,x}^X(u) = g_{t,0}^X(u - m_{t,x})$  and using the symmetry property of this density, we obtain that

$$T^{(n)}(t)f(x) = (f * g_{t,0}^X)(m_{t,x}).$$

By the chain rule formula, we deduce the properties **(2)** and for all  $i, j \in \{1, \dots, n\}$ ,

$$\begin{aligned} \partial_{x_i x_j}^2 T^{(n)}(t)f(x) &= (\partial_{x_j} m_{t,x})^t \left[ (f * \text{Hess}(g_{t,0}^X))(m_{t,x}) \partial_{x_i} m_{t,x} \right] \\ &\quad + (\partial_{x_i x_j}^2 m_{t,x} \cdot (f * \nabla g_{t,0}^X)(m_{t,x})). \end{aligned}$$

Now,

$$\partial_{x_i} m_{t,x} = \partial_{x_i} P\mu_{t,P^{-1}x} = \epsilon_i - (1 - \exp(-2\lambda nt)) (P^{-1})_{1i} P\epsilon_1.$$

The property **(3)** follows. □

The following corollary is useful for bounding the dual process in Section 2.6.2.

**Corollary 2.6.2.** *Let  $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ . We assume that there exists a constant  $C_1 > 0$  such that for all  $x \in \mathbb{R}^n$ ,*

$$|f(x)| \leq C_1 (1 + \|x\|_2^{2n}).$$

*Then, for all  $t > 0$  and  $x \in \mathbb{R}^n$ , there exists two constants  $C_2(t, n) > 0$  locally bounded on  $\mathbb{R}_+ \times \mathbb{N}$  such that  $t \mapsto C_2(t, n)$  is non-decreasing and  $C_3(t, n) > 0$  locally bounded on  $(0, +\infty) \times \mathbb{N}$  satisfying*

$$(1) \quad \left| T^{(n)}(t)f(x) \right| \leq C_2(t, n) (1 + \|x\|_2^{2n})$$

$$(2) \quad \left\| (\text{Hess}(g_{t,0}^X) * f)(m_{t,x}) \right\| \leq C_3(t, n) (1 + \|x\|_2^{2n})$$

*Proof. Step 1. Proof of (1).* From (2.51) and (2.52), note that for all  $x, r \in \mathbb{R}^n$ ,  $t \geq 0$ ,  $g_{t,x}^X(r) = \prod_{j=1}^n g_{t,[P^{-1}x]_j}^{Z_j}([P^{-1}r]_j)$ , where

$$\begin{aligned} g_{t,y_1}^{Z(1)}(z_1) &:= \frac{1}{\sqrt{2\pi e_4(t)}} \exp\left(-\frac{[z_1 - y_1 \exp(-2\lambda nt)]^2}{2e_4(t)}\right), \\ g_{t,y_j}^{Z(j)}(z_j) &:= \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{[z_j - y_j]^2}{2t}\right), \quad j \in \{2, \dots, n\}. \end{aligned} \tag{2.55}$$

Since  $\|Pz\|_2 = \|z\|_2$  for all  $z \in \mathbb{R}^n$ , we also have

$$\int_{\mathbb{R}^n} \|u\|_2^{2n} g_{t,x}^X(u) du = \int_{\mathbb{R}^n} \|z\|_2^{2n} g_{t,P^{-1}x}^Z(z) dz.$$

Hence,

$$\int_{\mathbb{R}^n} \|u\|_2^{2n} g_{t,x}^X(u) du \leq n^{n-1} \sum_{i=1}^n \int_{\mathbb{R}^n} z_i^{2n} \prod_{j=1}^n g_{t,[P^{-1}x]_j}^{Z(j)}(z_j) dz = n^{n-1} \sum_{i=1}^n \mathbb{E} \left( [Z^{(i)}]^{2n} \right).$$



Classical moment bounds for Gaussian random variables show that  $\mathbb{E}(G^{2n}) \leq C(n)t^{2n}$  for  $G \sim \mathcal{N}(0, t)$  and  $C(n) > 0$ . Since  $e_4(t) \leq t$  and using (2.55), we deduce that there exists two constants  $\tilde{C}_1(n)$  and  $\tilde{C}_2(t, n)$  such that for all  $i \in \{1, \dots, n\}$

$$\mathbb{E} \left( [Z^{(i)}]^{2n} \right) \leq \tilde{C}_1(n) \left( (P^{-1}x)_i^{2n} + t^{2n} \right) \leq \tilde{C}_2(t, n) \left( 1 + \|P^{-1}x\|_2^{2n} \right).$$

The result (1) follows.

**Step 2. Proof of (2).** Now, we want to control, for all  $i, j \in \{1, \dots, n\}$ ,

$$\left| \left( \text{Hess} \left( g_{t,0}^X \right) * f \right) (m_{t,x})_{ij} \right| \leq C_1(n) \int_{\mathbb{R}^n} \left| \partial_{r_i r_j}^2 g_{t,0}^X(r) \right| \left( 1 + \|m_{t,x} - r\|_2^{2n} \right) dr.$$

For all  $k \in \{1, \dots, n\}$ , we consider:

$$V_k(t) := \begin{cases} e_4(t) & \text{if } k = 1 \\ t & \text{if } k \neq 1 \end{cases}.$$

From (2.55), we deduce that for all  $i, j, k \in \{1, \dots, n\}$ , for all  $t > 0$ ,

$$\begin{aligned} \partial_{r_i} g_{t,0}^{Z^{(k)}} \left( (P^{-1}r)_k \right) &= - \frac{(P^{-1})_{ki} (P^{-1}r)_k}{V_k(t)} g_{t,0}^{Z^{(k)}} \left( (P^{-1}r)_k \right), \\ \partial_{r_j r_i}^2 g_{t,0}^{Z^{(k)}} \left( (P^{-1}r)_k \right) &= \frac{(P^{-1})_{kj} (P^{-1})_{ki}}{V_k(t)} \left( \frac{(P^{-1}r)_k^2}{V_k(t)} - 1 \right) g_{t,0}^{Z^{(k)}} \left( (P^{-1}r)_k \right). \end{aligned}$$

Hence, for all  $i, j \in \{1, \dots, n\}$ , for all  $t > 0$ ,

$$\begin{aligned} \partial_{r_j r_i}^2 g_{t,0}^X(r) &= \sum_{k=1}^n \frac{(P^{-1})_{kj} (P^{-1})_{ki}}{V_k(t)} \left( \frac{(P^{-1}r)_k^2}{V_k(t)} - 1 \right) g_{t,0}^X(r) \\ &\quad + \sum_{k=1}^n \sum_{\substack{\ell=1 \\ \ell \neq k}}^n \frac{(P^{-1})_{kj} (P^{-1})_{\ell j}}{V_k(t) V_\ell(t)} (P^{-1}r)_k (P^{-1}r)_\ell g_{t,0}^X(r) \end{aligned}$$

Noting that for all  $i, j, k \in \{1, \dots, n\}$ ,  $|(P^{-1})_{ij}| \leq 1$  and  $|(P^{-1}r)_k| \leq n \|r\|_2$ ,

$$\|m_{t,x} - r\|_2^{2n} \leq 2^{2n-1} \left( \|r\|_2^{2n} + 2^{2n-1} n^n (2 - \exp(-2\lambda nt)) \|x\|_2^{2n} \right),$$

we deduce that for all  $i, j \in \{1, \dots, n\}$ , there exists a constant  $\tilde{C}_3(t, n) > 0$  locally bounded on  $(0, +\infty) \times \mathbb{N}$  such that

$$\begin{aligned} &\int_{\mathbb{R}^n} \left| \partial_{r_i r_j}^2 g_{t,0}^X(r) \right| \|m_{t,x} - r\|_2^{2n} dr \\ &\leq \int_{\mathbb{R}^n} \left[ \frac{1}{e_4(t)} \left( \frac{n^2 \|r\|_2^2}{e_4(t)} + 1 \right) + \frac{n-1}{t} \left( \frac{n^2 \|r\|_2^2}{t} + 1 \right) + n^3 \left( \frac{1}{e_4(t)} + \frac{n-1}{t} \right) \|r\|_2^2 \right] \\ &\quad \times g_{t,0}^X(r) \|m_{t,x} - r\|_2^{2n} dr \\ &\leq \tilde{C}_3(t, n) \left( 1 + \|x\|_2^{2n} \right). \end{aligned}$$

The announced result (2) follows. □

### 2.6.1.2 MILD formulation

In this section, we establish the MILD formulation associated with the martingale problem (2.20).

**Proposition 2.6.3.** *Let  $(X_t)_{t \geq 0}$  be a stochastic process whose law  $\mathbb{P}_\mu$  is solution to the martingale problem (2.4) with initial value  $\mu$ . Then, for all  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,*

$$\begin{aligned} \langle T^{(n)}(t_0 - t)f, X_t^n \rangle - \lambda \int_0^t \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n & \left[ \langle \Phi_{i,j} T^{(n)}(t_0 - s)f, X_s^{n-1} \rangle - \langle T^{(n)}(t_0 - s)f, X_s^n \rangle \right] ds \\ - \lambda \int_0^t \sum_{i=1}^n \sum_{j=1}^n & \langle K_{i,j} T^{(n)}(t_0 - s)f, X_s^{n+1} \rangle ds \end{aligned}$$

is a  $\mathbb{P}_\mu$ -martingale for  $0 \leq t \leq t_0$ .

*Proof.* Let  $t_0 \geq 0$ . Using (2.17), let  $u, v, w : [0, t_0] \times \mathcal{M}_1^{c,2}(\mathbb{R}) \times \tilde{\Omega} \rightarrow \mathbb{R}$  be  $\mathcal{B}([0, t_0]) \otimes \mathcal{B}(\mathcal{M}_1^{c,2}(\mathbb{R})) \otimes \tilde{\mathcal{F}}$ -measurable defined by

$$\begin{aligned} \bullet u(r, \mu) &:= \langle T^{(n)}(t_0 - r)f, \mu^n \rangle, & \bullet v(r, \mu) &:= - \langle \partial_t T^{(n)}(t_0 - r)f, \mu^n \rangle, \\ \bullet w(r, \mu) &:= \mathcal{L}_{\text{FVc}} P_{T^{(n)}(t_0 - r)f, n}(\mu). \end{aligned}$$

The expected result is a direct consequence of a version of [94, Lemma 4.3.4] where we replace the assumption of boundedness on  $w$  by an assumption of domination. Hence, we need to check the following assumptions of this lemma:

(i) The process  $(u(t, X_t))_{t \geq 0}$  is  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$ -adapted and the processes  $(v(t, X_t))_{t \geq 0}$  and  $(w(t, X_t))_{t \geq 0}$  are  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$ -progressive. These properties are standard in our case.

(ii) The functions  $u, v$  are bounded on  $[0, t_0] \times \mathcal{M}_1^{c,2}(\mathbb{R})$  and there exists  $C > 0$  such that for all  $t \in [0, t_0]$ , for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,  $w(t, \mu) \leq C(1 + M_2(\mu))$  where we recall that  $M_2(\mu) = \langle \text{id}^2, \mu \rangle$ .

(iii) The function  $\mu \mapsto v(t, \mu, \tilde{\omega})$  is continuous for fixed  $t$  and  $\tilde{\omega}$ .

(iv) For all  $t_0 \geq t_2 > t_1 \geq 0$ ,

$$\mathbb{E} \left( u(t_2, X_{t_2}) - u(t_1, X_{t_2}) \middle| \tilde{\mathcal{F}}_{t_1} \right) = \mathbb{E} \left( \int_{t_1}^{t_2} v(s, X_{t_2}) ds \middle| \tilde{\mathcal{F}}_{t_1} \right), \quad (2.56)$$

and

$$\mathbb{E} \left( u(t_1, X_{t_2}) - u(t_1, X_{t_1}) \middle| \tilde{\mathcal{F}}_{t_1} \right) = \mathbb{E} \left( \int_{t_1}^{t_2} w(t_1, X_s) ds \middle| \tilde{\mathcal{F}}_{t_1} \right). \quad (2.57)$$

(v) The process  $(X_t)_{t \geq 0}$  is right continuous (here, it is continuous) and

$$\lim_{\delta \rightarrow 0_+} \mathbb{E} (|w(t - \delta, X_t) - w(t, X_t)|) = 0, \quad t_0 > t > 0. \quad (2.58)$$

**Step 1. Verification of Assumptions (ii) and (iii).** From Theorem 2.6.1, for all  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,  $(t, x) \mapsto T^{(n)}(t)f(x)$  is bounded on  $[0, t_0] \times \mathbb{R}^n$ . The boundedness of  $u$  follows. Moreover, as  $(t, x) \mapsto T^{(n)}(t)f(x)$  is solution of the PDE (2.47), we obtain for all  $r \in [0, t_0]$  and

$\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  that

$$v(r, \mu) = -\frac{1}{2} \langle \Delta T^{(n)}(t_0 - r)f, \mu^n \rangle + 2\lambda \langle (\nabla T^{(n)}(t_0 - r)f \cdot \mathbf{1}) (\text{id} \cdot \mathbf{1}), \mu^n \rangle.$$

Since  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , the second term of the right hand side is well-defined. Using the properties **(2)** and **(3)** of Theorem 2.6.1, we deduce that  $v$  is bounded on  $[0, t_0] \times \mathcal{M}_1^{c,2}(\mathbb{R})$ . In addition,  $\Delta T^{(n)}(t_0 - r)f$  and  $(\nabla T^{(n)}(t_0 - r)f \cdot \mathbf{1})$  are continuous bounded, hence  $\mu \mapsto v(r, \mu)$  is continuous on  $\mathcal{M}_1(\mathbb{R})$  for the topology of weak convergence. Now, using (2.18), (2.21), (2.22), (2.23) and Theorem 2.6.1 **(2)** and **(3)**, note that for all  $r \in [0, t_0]$  and  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , there exists a constant  $C_f > 0$  such that,

$$\begin{aligned} w(r, \mu) &= \langle \partial_t T^{(n)}(t_0 - r)f, \mu^n \rangle + \lambda \sum_{i,j=1}^n \langle K_{ij} T^{(n)}(t_0 - r)f, \mu^{n+1} \rangle \\ &\quad + \lambda \sum_{\substack{i,j=1 \\ j \neq i}}^n \left[ \langle \Phi_{ij} T^{(n)}(t_0 - r)f, \mu^{n-1} \rangle - \langle T^{(n)}(t_0 - r)f, \mu^n \rangle \right] \\ &\leq C_f (1 + M_2(\mu)), \end{aligned} \tag{2.59}$$

where the bound  $M_2(\mu)$  comes from  $K_{ij}$ .

**Step 2. Verification of Assumptions (iv) and (v).** Let  $t_2 > t_1 \geq 0$ . On the one hand,

$$u(t_2, X_{t_2}) - u(t_1, X_{t_2}) = \left\langle -\int_{t_1}^{t_2} \partial_t T^{(n)}(t_0 - s)f ds, X_{t_2}^n \right\rangle = \int_{t_1}^{t_2} v(s, X_{t_2}) ds$$

and the relation (2.56) follows. On the other hand, as the martingale problem (2.4) classically involves the martingale problem (2.20) [91], we obtain that

$$\begin{aligned} u(t_1, X_{t_2}) - u(t_1, X_{t_1}) &= \langle T^{(n)}(t_0 - t_1)f, X_{t_2}^n - X_{t_1}^n \rangle \\ &= \int_{t_1}^{t_2} \mathcal{L}_{\text{FVc}} P_{T^{(n)}(t_0 - t_1)f, n}(X_s) ds \\ &\quad + \widehat{M}_{t_2}^{(n)}(T^{(n)}(t_0 - t_1)f) - \widehat{M}_{t_1}^{(n)}(T^{(n)}(t_0 - t_1)f). \end{aligned}$$

The relation (2.57) follows. Finally, from (2.59), Proposition 2.2.11 and the LEBESGUE dominated convergence theorem, we deduce the relation (2.58).  $\square$

### 2.6.2 Proof of Lemma 2.3.5

Recall that, our goal is to prove that the stopping time  $\theta_k$ , defined by

$$\forall k \in \mathbb{N}, \quad \theta_k := \inf \left\{ t \geq 0 \mid M(t) \geq k \quad \text{or} \quad \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq k \right\},$$

satisfies  $\lim_{k \rightarrow +\infty} \theta_k = +\infty$ ,  $\mathbb{P}_{(\mu, \xi_0)}$ -a.s. with  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ . Before to prove Lemma 2.3.5, we introduce the following lemma, whose proof will be given at Section 2.6.2.2. We denote by  $S_t$  the number of jumps of the process  $M$  on the time interval  $[0, t]$ .

**Lemma 2.6.4.** *If  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  then there exists a function  $C_0$  on  $\bigcup_{k \in \mathbb{N}} (0, +\infty)^k \times \{k\}$  to  $\mathbb{R}_+$ , locally bounded, such that for all  $(t_j)_{j \in \mathbb{N}} \in (0, +\infty)^{\mathbb{N}}$ ,  $k \mapsto C_0((t_i)_{0 \leq i \leq k}, k)$  is non-decreasing and satisfying*

$$\forall t \in [0, T], \forall x \in \mathbb{R}^{M(t)}, \quad |\xi_t(x)| \leq C_0(\tau_1, \tau_2 - \tau_1, \dots, \tau_{S_T+1} - \tau_{S_T}, S_T) (1 + \|x\|_2^{2S_T}).$$

The bound obtained above will only allow us to show that  $\theta_k \rightarrow +\infty$   $\mathbb{P}_{(\mu, \xi_0)}$ -a.s. under the assumption that the initial condition  $X_0$  has all its finite moments. The following remark shows that we cannot expect that  $\theta_k \rightarrow +\infty$  under weaker assumptions on the initial condition.

**Remark 2.6.5.** *Let  $\xi_0 : x \mapsto \sin(x) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ . Let us assume that  $\langle |\text{id}|^4, \mu \rangle = +\infty$  and  $\xi_t$  successively jumps at times  $\tau_1, \tau_2$  and  $\tau_3$  with respective jump operator  $K_{11}, K_{11}$  and  $\Phi_{13}$ . If we denote by  $\tau_{1,2} := \tau_2 - \tau_1$ , straightforward but tedious computations give*

$$\xi_{\tau_1}(x, y) = K_{11}T^{(1)}(\tau_1)\xi_0(x, y) = -\exp\left(-\frac{e_4(\tau_1)}{2} - 4\lambda\tau_1\right) \sin(x \exp(-2\lambda\tau_1)) y^2$$

and

$$\begin{aligned} \xi_{\tau_2}(x, y, z) &= K_{11}T^{(2)}(\tau_{1,2})\xi_{\tau_1}(x, y, z) \\ &= \frac{1}{4} \exp\left(-\frac{e_4(\tau_1)}{2} - 4\lambda\tau_1 - \exp(-4\lambda\tau_1) - \frac{\exp(-4\lambda\tau_1)(e_4(\tau_{1,2}) - \tau_{1,2})^2}{4(e_4(\tau_{1,2}) + \tau_{1,2})}\right) \\ &\times \left[ \left\{ \frac{\exp(-4\lambda\tau_1)}{4} (1 + \exp(-4\lambda\tau_{1,2}))^2 \left[ 2(e_4(\tau_{1,2}) + \tau_{1,2}) - \exp(-4\lambda\tau_1)(e_4(\tau_{1,2}) - \tau_{1,2})^2 \right. \right. \right. \\ &\quad \left. \left. \left. + (x[1 - \exp(-4\lambda\tau_{1,2})] - y[1 + \exp(-4\lambda\tau_{1,2})])^2 \right] - 2(1 + \exp(-8\lambda\tau_1) - \exp(-4\lambda\tau_1)) \right. \right. \\ &\quad \left. \left. - 2\exp(-4\lambda\tau_1)(1 - \exp(-8\lambda\tau_{1,2}))(e_4(\tau_{1,2}) - \tau_{1,2}) \right\} \right. \\ &\quad \left. \times \sin\left(\frac{\exp(-2\lambda\tau_1)}{2} [x - y + (x + y)\exp(-4\lambda\tau_{1,2})]\right) \right. \\ &\quad \left. - 2\exp(-2\lambda\tau_1)(1 + \exp(-4\lambda\tau_{1,2})) \left\{ x - y + (x + y)\exp(-8\lambda\tau_{1,2}) - x\exp(-4\lambda\tau_{1,2}) \right. \right. \\ &\quad \left. \left. - \frac{\exp(-4\lambda\tau_{1,2})}{2} (e_4(\tau_{1,2}) - \tau_{1,2}) [-(x - y) + (x + y)\exp(-4\lambda\tau_{1,2})] \right\} \right. \\ &\quad \left. \times \cos\left(\frac{\exp(-2\lambda\tau_1)}{2} [x - y + (x + y)\exp(-4\lambda\tau_{1,2})]\right) \right] z^2. \end{aligned}$$

Note that the leading order term in  $\xi_{\tau_2}(x, y, z)$  is of the form  $(ax - by)^2 z^2 \sin(cx + dy)$ . Now,

$$\xi_{\tau_3}(x, y) = \Phi_{13}T^{(3)}(\tau_3 - \tau_2)\xi_{\tau_2}(x, y).$$

If  $\tau_3 = \tau_2$ , we obtain as leading order term in  $\xi_{\tau_3}(x, y)$  the term  $(ax - by)^2 x^2 \sin(cx + dy)$ , which is not integrable with respect to  $\mu^2(dx, dy)$ . If  $\tau_3 > \tau_2$ , one can check that the leading order term in  $T^{(3)}(\tau_3 - \tau_2)\xi_{\tau_2}(x, y, z)$  is of the form  $P_4(x, y, z) \sin(\tilde{c}x + \tilde{d}y + \tilde{e}z)$  where  $P_4(X, Y, Z)$  is a homogeneous polynomial of degree 4 such that  $P_4(X, Y, Z) \rightarrow (aX - bY)^2 Z^2$ ,  $\tilde{c} \rightarrow c$ ,  $\tilde{d} \rightarrow d$  and  $\tilde{e} \rightarrow 0$  when  $\tau_3 \rightarrow \tau_2$ . Therefore, for  $\tau_3$  close enough to  $\tau_2$ ,  $\xi_{\tau_3}(x, y)$  has a non-zero term proportional to  $x^4 \sin(\tilde{c}x + \tilde{e}y)$  which is not compensated by another term. Hence,  $\langle |\xi_{\tau_3}|, \mu^2 \rangle = +\infty$  if  $\tau_3 - \tau_2$  is small enough, for any values of  $\tau_1$  and  $\tau_2$ . Given  $T$  large enough, we have proved that  $\theta_k \leq \tau_3 \leq T$  with positive probability.

**2.6.2.1 Proof that Lemma 2.6.4 implies Lemma 2.3.5**

Note that  $\theta_k = \widehat{\theta}_k \wedge \widetilde{\theta}_k$  where

$$\widehat{\theta}_k := \inf \left\{ t \geq 0 \mid M(t) \geq k \right\} \quad \text{and} \quad \widetilde{\theta}_k := \inf \left\{ t \geq 0 \mid \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq k \right\}.$$

Thanks to (2.32) it follows that  $\widehat{\theta}_k \rightarrow +\infty$  when  $k \rightarrow +\infty$ . In order to prove that  $\widetilde{\theta}_k \rightarrow +\infty$  when  $k \rightarrow +\infty$ , we rely on the control of the dual process obtained in Lemma 2.6.4. So we need to control  $\langle \|\cdot\|_2^{2S_T}, X_{t-s}^{M(s)} \rangle$ . Let  $T > 0$  and  $\varepsilon > 0$  be arbitrary. From (2.32), we choose  $A(T, \varepsilon) > 0$  such that  $\mathbb{P}_{(\mu, \xi_0)}(S_T \leq A) \geq 1 - \varepsilon/3$ . Then, using Proposition 2.2.11, we choose  $B(T, \varepsilon, A) > 0$  such that  $\mathbb{P}_{(\mu, \xi_0)}(\forall k \leq 2A, \forall t \leq T, \langle \text{id}^k, X_t \rangle \leq B) \geq 1 - \varepsilon/3$ . Finally, from Lemma 2.6.4 we choose  $\overline{C}_0(T, \varepsilon, A) > 0$  such that  $\mathbb{P}_{(\mu, \xi_0)}(C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq A}, A) \leq \overline{C}_0) \geq 1 - \varepsilon/3$ . We recall that for any  $m \in \mathbb{N}^*$ , for all  $x \in \mathbb{R}^m$ ,  $(\sum_{i=1}^m x_i)^n \leq m^{n-1} \sum_{i=1}^m x_i^n$ . Thus, the following inequality

$$\begin{aligned} \langle \|\cdot\|_2^{2S_T}, X_{t-s}^{M(s)} \rangle &\leq M(s)^{S_T-1} \sum_{i=1}^{M(s)} \int_{\mathbb{R}^{M(s)}} x_i^{2S_T} X_{t-s}^{M(s)}(dx) = M(s)^{S_T} \langle \text{id}^{2S_T}, X_{t-s} \rangle \\ &\leq (M(0) + A)^A B, \end{aligned}$$

takes place with probability  $1 - 2\varepsilon/3$ . Therefore, we deduce from Lemma 2.6.4 that for all  $s \leq t \leq T$ ,

$$\langle \xi_s, X_{t-s}^{M(s)} \rangle \leq \overline{C}_0 (M(0) + A)^A B$$

In particular, for  $k \geq \overline{C}_0 (1 + (M(0) + A)^A B)$ , it follows that

$$\begin{aligned} \mathbb{P}_{(\mu, \xi_0)}(\widetilde{\theta}_k \geq T) &\geq \mathbb{P}_{(\mu, \xi_0)}(\{S_T \leq A\} \cap \{\forall k \leq 2A, \forall t \leq T, \langle \text{id}^k, X_t \rangle \leq B\} \\ &\quad \cap \{C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq A}, A)\}) \\ &\geq 1 - \varepsilon. \end{aligned}$$

The conclusion follows.  $\square$

**2.6.2.2 Proof of Lemma 2.6.4**

By mathematical induction on  $k \in \mathbb{N}$ , we prove the property

$$(\mathcal{P}_k) : \quad \forall t \in [\tau_k, \tau_{k+1}[ , \forall x \in \mathbb{R}^{M(t)}, \quad |\xi_t(x)| \leq C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k}, k) \left(1 + \|x\|_2^{2k}\right),$$

where  $C_0$  is locally bounded on  $\bigcup_{k \in \mathbb{N}} (0, +\infty)^k \times \{k\}$ .

**Initial case.** For  $k = 0$ ,  $S_0 = 0$  and  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ . Hence, the property  $(\mathcal{P}_0)$  is satisfied.

**Inductive step.** We assume that, for  $k \in \mathbb{N}^*$ , the property  $(\mathcal{P}_{k-1})$  is satisfied and prove that  $(\mathcal{P}_k)$  is also. Let  $t \in [\tau_k, \tau_{k+1}[$  and note that  $M(t) = M(\tau_k)$ . We make a partition of cases according to whether the dual process loses or gains a variable. Let  $i, j \in \{1, \dots, M(\tau_{k-1})\}$  be fixed.

**Step 1. Case  $\Lambda_k = \Phi_{i,j}$  at the  $k^{\text{th}}$  jump.** In this case,  $M(\tau_k) = M(\tau_{k-1}) - 1$  and we deduce from the explicit expression (2.29) of the dual process that for all  $x \in \mathbb{R}^{M(\tau_{k-1})-1}$ ,

$$\xi_t(x) = T^{(M(\tau_{k-1})-1)}(t - \tau_k) \Phi_{i,j} T^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x).$$

By using expression (2.22) of  $\Phi_{i,j}$  and the property  $(\mathcal{P}_{k-1})$ , we deduce from Corollary 2.6.2 (1) that for all  $x \in \mathbb{R}^{M(\tau_{k-1})-1}$ ,

$$\begin{aligned} & \left| \Phi_{i,j} T^{(M(\tau_{k-1}))} (\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ & \leq C_2 (\tau_k - \tau_{k-1}, M(\tau_{k-1})) C_0 ((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k-1) (1 + \|x\|_2^{2(k-1)}), \end{aligned}$$

where  $C_2 C_0$  is locally bounded. Using again Corollary 2.6.2 (1) and the fact that  $t \mapsto C_2(t, M(\tau_{k-1}))$  is non-decreasing, we deduce the property  $(\mathcal{P}_k)$ .

**Step 2. Case  $\Lambda_k = K_{i,j}$  at the  $k^{\text{th}}$  jump.** In this case,  $M(\tau_k) = M(\tau_{k-1}) + 1$  and the explicit expression (2.29) of dual process that for all  $x \in \mathbb{R}^{M(\tau_{k-1})+1}$ ,

$$\xi_t(x) = T^{(M(\tau_{k-1})+1)}(t - \tau_k) K_{i,j} T^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x).$$

From the expression (2.23) of  $K_{i,j}$  and Theorem 2.6.1 (3), we have for all  $x \in \mathbb{R}^{M(\tau_{k-1})+1}$ ,

$$\begin{aligned} & \left| K_{i,j} T^{(M(\tau_{k-1})+1)} (\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ & = \left| (\partial_{x_j} m_{\tau_k - \tau_{k-1}, \tilde{x}})^t \left[ (\xi_{\tau_{k-1}} * \text{Hess}(g_{\tau_k - \tau_{k-1}, 0}^X)) (m_{\tau_k - \tau_{k-1}, \tilde{x}}) \partial_{x_i} m_{\tau_k - \tau_{k-1}, \tilde{x}} \right] \right| x_{M(\tau_{k-1})+1}^2, \end{aligned}$$

where  $\tilde{x} = (x_1, \dots, x_{M(\tau_{k-1})})^t \in \mathbb{R}^{M(\tau_{k-1})}$ . From the property  $(\mathcal{P}_{k-1})$  and Corollary 2.6.2 (2), we deduce that

$$\begin{aligned} & \left| K_{i,j} T^{(M(\tau_{k-1})+1)} (\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ & \leq C_3 (\tau_k - \tau_{k-1}, M(\tau_{k-1})) C_0 ((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k-1) (1 + \|x\|_2^{2k}), \end{aligned}$$

where  $C_3 C_0$  is locally bounded. Using Corollary 2.6.2 (1), we deduce the property  $(\mathcal{P}_k)$ . We conclude by the principle of induction.  $\square$

### 2.6.3 Proof of Theorem 2.3.4

Recall that  $(X_t)_{t \geq 0}$  is a stochastic process whose law  $\mathbb{P}_\mu$  is a solution of the martingale problem (2.4) with  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $(\xi_t)_{t \geq 0}$  a dual process independent of  $(X_t)_{t \geq 0}$  built on the same probability space. To simplify, we will note  $\mathbb{P} = \mathbb{P}_{(\mu, \xi_0)}$  the distribution of  $((X_t, \xi_t))_{t \geq 0}$ . As  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  and for the choice of the stopping time  $\theta_k$  given by (2.33), the set of quantities, involved in the expectations of the weakened duality identity (2.34), are bounded.

**Step 1. Approximation reasoning.** To establish the relation (2.34) we introduce an increasing sequence  $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = t$  of subdivisions of  $[0, t]$  such that  $t_{i+1}^n = t_i^n + h$  with  $h$  tending to 0. Note that

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{t \wedge \theta_k}, X_0^{M(t \wedge \theta_k)} \right\rangle \exp \left( \lambda \int_0^{t \wedge \theta_k} M^2(u) du \right) \right) - \mathbb{E} \left( \left\langle \xi_0, X_{t \wedge \theta_k}^{M(0)} \right\rangle \right) \\ & = \sum_{i=0}^{p_n-1} \left[ \mathbb{E} \left( \left\langle \xi_{t_{i+1}^n \wedge \theta_k}, X_{t \wedge \theta_k - t_{i+1}^n \wedge \theta_k}^{M(t_{i+1}^n \wedge \theta_k)} \right\rangle \exp \left( \lambda \int_0^{t_{i+1}^n \wedge \theta_k} M^2(u) du \right) \right) \right. \\ & \quad \left. - \mathbb{E} \left( \left\langle \xi_{t_i^n \wedge \theta_k}, X_{t \wedge \theta_k - t_i^n \wedge \theta_k}^{M(t_i^n \wedge \theta_k)} \right\rangle \exp \left( \lambda \int_0^{t_i^n \wedge \theta_k} M^2(u) du \right) \right) \right]. \end{aligned}$$

We are therefore interested in terms of the form

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{(s+h)\wedge\theta_k}, X_{t\wedge\theta_k - (s+h)\wedge\theta_k}^{M((s+h)\wedge\theta_k)} \right\rangle \exp \left( \lambda \int_0^{(s+h)\wedge\theta_k} M^2(u) du \right) \right) \\ & - \mathbb{E} \left( \left\langle \xi_{s\wedge\theta_k}, X_{t\wedge\theta_k - s\wedge\theta_k}^{M(s\wedge\theta_k)} \right\rangle \exp \left( \lambda \int_0^{s\wedge\theta_k} M^2(u) du \right) \right), \quad \text{for } s \in [0, t-h]. \end{aligned} \quad (2.60)$$

It is sufficient to prove that these quantities are  $O(h^2)$ . The procedure to be adopted is as follows. First of all, we consider separately the two terms which constitute (2.60) in Steps 2 and 3. Then, we prove that the sum of these terms is  $O(h^2)$  in Steps 4 and 5. Throughout this section, in order to simplify the writing, the following notations are introduced:

$$t_k := t \wedge \theta_k, \quad s_k := s \wedge \theta_k, \quad \text{and} \quad s_k^h := (s+h) \wedge \theta_k.$$

We note respectively  $\tau_1, \tau_2$  the first and second jump times after  $s_k$  for the process  $M$ . We denote by  $\tau_{1,k} := \tau_1 \wedge \theta_k$  and  $\tau_{2,k} := \tau_2 \wedge \theta_k$ .

**Step 2. First term of (2.60).** We exploit the explicit expression (2.29) of the dual process and make the following partition:

- (a) If there has been *no* jump of  $M$  on the interval  $[s_k, s_k^h]$ .
- (b) If there was *only one* jump of  $M$  on the interval  $[s_k, s_k^h]$  and distinguish according to the events  $\{\Lambda = \Phi_{i,j}\}$  and  $\{\Lambda = K_{i,j}\}$  where  $\Lambda$  is the first  $\Lambda_k$  defined by (2.30) and (2.31) after  $s_k$ .
- (c) If there are *two or more than two* jumps of  $M$  on the interval  $[s_k, s_k^h]$ .

Then

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ & = \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} > s_k^h\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ & \quad + \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} \leq s_k^h, \tau_{2,k} > s_k^h, \Lambda = \Phi_{i,j}\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ & \quad + \sum_{i,j=1}^{M(s_k)} \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} \leq s_k^h, \tau_{2,k} > s_k^h, \Lambda = K_{i,j}\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ & \quad + \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{2,k} \leq s_k^h\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right). \end{aligned}$$

We consider successively each term in the right-hand side.

**Firts term: no jump.** As there is no jump on  $[s_k, s_k^h]$ , we have  $M(s_k^h) = M(s_k)$  and  $\xi_{s_k^h} = T^{(M(s_k))}(s_k^h - s_k) \xi_{s_k}$ . Thus,

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} > s_k^h\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ & = \left\langle T^{(M(s_k))}(s_k^h - s_k) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du - \left[ \lambda (s_k^h - s_k) M(s_k) (M(s_k) - 1) \right] \right) \end{aligned}$$

where we used the fact that  $\tau_{1,k} - s_k$  given  $\widetilde{\mathcal{F}}_{s_k}$  follows an exponential law of parameter  $\lambda M^2(s_k) + \lambda M(s_k)(M(s_k) - 1)$ .

**Second and third terms: only one jump.** These terms are treated in an analogous way, so we only give the details for the first one.

If there is a jump on  $[s_k, s_k^h]$  and for  $i \neq j \in \{1, 2, \dots, M(s_k)\}$  fixed,  $\Lambda = \Phi_{i,j}$ , then  $M(s_k^h) = M(s_k) - 1$  and  $\xi_{s_k^h} = T^{(M(s_k)-1)}(s_k^h - \tau_{1,k}) \Phi_{i,j} T^{(M(s_k))}(\tau_{1,k} - s_k) \xi_{s_k}$ . Thus,

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} \leq s_k^h, \tau_{2,k} > s_k^h, \Lambda = \Phi_{i,j}\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ &= \left\langle T^{(M(s_k)-1)}(s_k^h - \tau_{1,k}) \Phi_{i,j} T^{(M(s_k))}(\tau_{1,k} - s_k) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)-1} \right\rangle \\ & \quad \times \exp \left( \lambda \int_0^{s_k} M^2(u) du + \lambda [\tau_{1,k} - s_k] M^2(s_k) + \lambda [s_k^h - \tau_{1,k}] (M(s_k) - 1)^2 \right) \\ & \quad \times \mathbb{1}_{\{\tau_{1,k} - s_k \leq s_k^h - s_k\}} \mathbb{E} \left[ \mathbb{1}_{\{\Lambda = \Phi_{i,j}\}} \mathbb{1}_{\{\tau_{2,k} - \tau_{1,k} > s_k^h - \tau_{1,k}\}} \middle| \sigma(\tau_{1,k}) \vee \widetilde{\mathcal{F}}_{s_k} \right]. \end{aligned}$$

Now, using that, given  $\Lambda = \Phi_{i,j}$  and  $\sigma(\tau_{1,k}) \vee \widetilde{\mathcal{F}}_{s_k}$ ,  $\tau_{2,k} - \tau_{1,k}$  follows an exponential law of parameter  $\lambda (M(s_k) - 1)^2 + \lambda (M(s_k) - 1)(M(s_k) - 2)$ , we deduce that

$$\begin{aligned} & \mathbb{P} \left( \{\Lambda = \Phi_{i,j}\} \cap \{\tau_{2,k} - \tau_{1,k} > s_k^h - \tau_{1,k}\} \middle| \sigma(\tau_{1,k}) \vee \widetilde{\mathcal{F}}_{s_k} \right) \\ &= \frac{\exp \left( -\lambda \left[ (M(s_k) - 1)^2 + (M(s_k) - 1)(M(s_k) - 2) \right] [s_k^h - \tau_{1,k}] \right)}{M^2(s_k) + M(s_k)(M(s_k) - 1)}. \end{aligned}$$

Then using that, given  $\sigma(\tau_{1,k}) \vee \widetilde{\mathcal{F}}_{s_k}$ ,  $\tau_{1,k} - s_k$  follows an exponential law of parameter  $\lambda M^2(s_k) + \lambda M(s_k)(M(s_k) - 1)$ , we deduce that

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} \leq s_k^h, \tau_{2,k} > s_k^h, \Lambda = \Phi_{i,j}\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ &= \lambda \int_{s_k}^{s_k^h} \left\{ \left\langle T^{(M(s_k)-1)}(s_k^h - r) \Phi_{i,j} T^{(M(s_k))}(r - s_k) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)-1} \right\rangle \exp \left( \lambda \int_0^{s_k} M^2(u) du \right) \right. \\ & \quad \left. \times \exp \left( -2\lambda (r - s_k)(M(s_k) - 1) - \lambda (s_k^h - s_k)(M(s_k) - 1)(M(s_k) - 2) \right) \right\} dr. \end{aligned}$$

Similarly,

$$\begin{aligned} & \mathbb{E} \left( \left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp \left( \lambda \int_0^{s_k^h} M^2(u) du \right) \mathbb{1}_{\{\tau_{1,k} \leq s_k^h, \tau_{2,k} > s_k^h, \Lambda = K_{i,j}\}} \middle| \widetilde{\mathcal{F}}_{s_k} \right) \\ &= \lambda \int_{s_k}^{s_k^h} \left\{ \left\langle T^{(M(s_k)+1)}(s_k^h - r) K_{i,j} T^{(M(s_k))}(r - s_k) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)+1} \right\rangle \right. \\ & \quad \left. \times \exp \left( \lambda \int_0^{s_k} M^2(u) du - 2\lambda (r - s_k)M(s_k) - \lambda (s_k^h - s_k)M(s_k)(M(s_k) + 1) \right) \right\} dr. \end{aligned}$$



**Fourth term: at least two jumps.** Note that,  $\mathbb{P}\left(\cdot \mid \widetilde{\mathcal{F}}_{s_k}\right)$  -a.s.,  $\tau_{2,k} - \tau_{1,k} \geq E_{s_k}$  where  $E_{s_k}$  is an exponential random variable with parameter  $\lambda_2(s_k) := \lambda(M(s_k) + 1)^2 + \lambda(M(s_k) + 1)M(s_k)$ . We denote by  $\lambda_1(s_k) := \lambda(M(s_k) + 1)^2 + \lambda(M(s_k) + 1)M(s_k)$ . Using the strong MARKOV property at time  $\tau_{1,k}$ , we obtain that

$$\mathbb{P}\left(\tau_{2,k} \leq s_k^h \mid \widetilde{\mathcal{F}}_{s_k}\right) \leq \mathbb{P}\left(\{\tau_{1,k} \leq s_k^h\} \cap \{\tau_{2,k} - \tau_{1,k} \leq s_k^h\} \mid \widetilde{\mathcal{F}}_{s_k}\right) \leq \lambda_1(s_k)\lambda_2(s_k) [s_k^h - s_k]^2.$$

It follows that there exists a constant  $C(k)$  such that,

$$\mathbb{E}\left(\left\langle \xi_{s_k^h}, X_{t_k - s_k^h}^{M(s_k^h)} \right\rangle \exp\left(\lambda \int_0^{s_k^h} M^2(u) du\right) \mathbb{1}_{\{\tau_{2,k} \leq s_k^h\}} \mid \widetilde{\mathcal{F}}_{s_k}\right) \leq C(k)h^2.$$

**Step 3. Second term of (2.60).** It follows from the MILD formulation of Proposition 2.6.3 that

$$\begin{aligned} & \mathbb{E}\left(\left\langle \xi_{s_k}, X_{t_k - s_k}^{M(s_k)} \right\rangle \exp\left(\lambda \int_0^{s_k} M^2(u) du\right) \mid \widetilde{\mathcal{F}}_{s_k}\right) \\ &= \mathbb{E}\left(\left\langle T^{(M(s_k))}(s_k^h - s_k)\xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)} \right\rangle \exp\left(\lambda \int_0^{s_k} M^2(u) du\right) \mid \widetilde{\mathcal{F}}_{s_k}\right) \\ &+ \lambda \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \mathbb{E}\left(\int_0^{s_k^h - s_k} \left(\left\langle \Phi_{i,j} T^{(M(s_k))}(r)\xi_{s_k}, X_{t_k - s_k - r}^{M(s_k) - 1} \right\rangle \right. \right. \\ &\quad \left. \left. - \left\langle T^{(M(s_k))}(r)\xi_{s_k}, X_{t_k - s_k - r}^{M(s_k)} \right\rangle\right) dr \exp\left(\lambda \int_0^{s_k} M^2(u) du\right) \mid \widetilde{\mathcal{F}}_{s_k}\right) \\ &+ \lambda \sum_{i,j=1}^{M(s_k)} \mathbb{E}\left(\int_0^{s_k^h - s_k} \left\langle K_{i,j} T^{(M(s_k))}(r)\xi_{s_k}, X_{t_k - s_k - r}^{M(s_k) + 1} \right\rangle dr \exp\left(\lambda \int_0^{s_k} M^2(u) du\right) \mid \widetilde{\mathcal{F}}_{s_k}\right). \end{aligned}$$

**Step 4. Conclusion.** Putting together all the previous equations, we deduce that

$$\begin{aligned} & \mathbb{E}\left(\left\langle \xi_{(s+h)\wedge\theta_k}, X_{t\wedge\theta_k - (s+h)\wedge\theta_k}^{M([s+h]\wedge\theta_k)} \right\rangle \exp\left(\lambda \int_0^{[s+h]\wedge\theta_k} M^2(u) du\right)\right) \\ & \quad - \mathbb{E}\left(\left\langle \xi_{s\wedge\theta_k}, X_{t\wedge\theta_k - s\wedge\theta_k}^{M(s\wedge\theta_k)} \right\rangle \exp\left(\lambda \int_0^{s\wedge\theta_k} M^2(u) du\right)\right) \\ &= \mathbf{(A)} + \mathbf{(B)} + \mathbf{(C)} + \mathbf{(D)} + \mathbf{(E)} + O(h^2), \end{aligned}$$

where

$$\begin{aligned} \mathbf{(A)} &= \mathbb{E}\left(\left\langle T^{(M(s_k))}(s_k^h - s_k)\xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)} \right\rangle \right. \\ &\quad \left. \times \exp\left(\lambda \int_0^{s_k} M^2(u) du - \lambda [s_k^h - s_k] M(s_k) (M(s_k) - 1)\right)\right) \\ &+ \lambda \mathbb{E}\left(\int_0^{s_k^h - s_k} \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \left\langle T^{(M(s_k))}(r)\xi_{s_k}, X_{t_k - s_k - r}^{M(s_k)} \right\rangle dr \exp\left(\lambda \int_0^{s_k} M^2(u) du\right)\right), \end{aligned}$$

$$\begin{aligned}
 \mathbf{(B)} &= \lambda \mathbb{E} \left( \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \int_0^{s_k^h - s_k} \left\{ \left\langle T^{(M(s_k)-1)}(s_k^h - s_k - r) \Phi_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)-1} \right\rangle \right. \right. \\
 &\quad \left. \left. - \left\langle T^{(M(s_k)-1)}(0) \Phi_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k - r}^{M(s_k)-1} \right\rangle \right\} \exp \left( \lambda \int_0^{s_k} M^2(u) du \right) \right), \\
 \mathbf{(C)} &= \lambda \mathbb{E} \left( \sum_{i,j=1}^{M(s_k)} \int_0^{s_k^h - s_k} \left\{ \left\langle T^{(M(s_k)+1)}(s_k^h - s_k - r) K_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)+1} \right\rangle \right. \right. \\
 &\quad \left. \left. - \left\langle T^{(M(s_k)+1)}(0) K_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k - r}^{M(s_k)+1} \right\rangle \right\} \exp \left( \lambda \int_0^{s_k} M^2(u) du \right) \right), \\
 \mathbf{(D)} &= \lambda \mathbb{E} \left( \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \int_0^{s_k^h - s_k} \left\{ \left\langle T^{(M(s_k)-1)}(s_k^h - s_k - r) \Phi_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)-1} \right\rangle \right. \right. \\
 &\quad \left. \left. \times \exp \left( -2\lambda r [M(s_k) - 1] - \lambda (s_k^h - s_k) [M(s_k) - 1] [M(s_k) - 2] \right) \right\} dr \right), \\
 \mathbf{(E)} &= \lambda \mathbb{E} \left( \sum_{i,j=1}^{M(s_k)} \int_0^{s_k^h - s_k} \left\{ \left\langle T^{(M(s_k)+1)}(s_k^h - s_k - r) K_{i,j} T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)+1} \right\rangle \right. \right. \\
 &\quad \left. \left. \times \exp \left( -2\lambda r M(s_k) - \lambda (s_k^h - s_k) M(s_k) [M(s_k) - 1] \right) \right\} dr \right).
 \end{aligned}$$

All these terms are  $O(h^2)$  uniformly with respect to the other parameters which is sufficient to conclude the proof. This can be proved similarly for each term, so we only give the details for  $\mathbf{(A)}$ . We first notice that when  $h \rightarrow 0$ ,

$$\begin{aligned}
 \mathbf{(A)} &= \mathbb{E} \left( M(s_k) (M(s_k) - 1) \int_0^{s_k^h - s_k} \left\{ \left\langle T^{(M(s_k))}(s_k^h - s_k) \xi_{s_k}, X_{t_k - s_k^h}^{M(s_k)} \right\rangle \right. \right. \\
 &\quad \left. \left. - \left\langle T^{(M(s_k))}(r) \xi_{s_k}, X_{t_k - s_k - r}^{M(s_k)} \right\rangle \right\} dr \exp \left( \lambda \int_0^{s_k} M^2(u) du \right) \right) + O(h^2).
 \end{aligned}$$

By conditioning with respect to  $\tilde{\mathcal{F}}_{s_k}$  in the previous expression and using the MILD formulation of the Proposition 2.6.3 again we obtain that

$$\begin{aligned}
 \mathbf{(A)} &= \lambda \mathbb{E} \left( \exp \left( \lambda \int_0^{s_k} M^2(u) du \right) M(s_k) (M(s_k) - 1) \int_0^{s_k^h - s_k} \left[ \int_{t_k - s_k - r}^{t_k - s_k^h} \left\{ \right. \right. \right. \\
 &\quad \sum_{\substack{i,j=1 \\ i \neq j}}^{M(s_k)} \left[ \left\langle \Phi_{i,j} T^{(M(s_k))}(t_k - s_k - v) \xi_{s_k}, X_v^{M(s_k)-1} \right\rangle - \left\langle T^{(M(s_k))}(t_k - s_k - v) \xi_{s_k}, X_v^{M(s_k)} \right\rangle \right] \\
 &\quad \left. \left. \left. + \sum_{i,j=1}^{M(s_k)} \left\langle K_{i,j} T^{(M(s_k))}(t_k - s_k - v) \xi_{s_k}, X_v^{M(s_k)+1} \right\rangle \right\} dv \right] dr \right) + O(h^2).
 \end{aligned}$$

Both integrals are on intervals of length at most  $h$ . Since all the quantities in the integrals are bounded by definition of the stopping time  $\theta_k$ , we deduce that  $\mathbf{(A)} \underset{h \rightarrow 0}{=} O(h^2)$ .  $\square$

## Chapter 3

# Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics

This chapter corresponds to the preprint [40] “Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics”, written in collaboration with NICOLAS CHAMPAGNAT and submitted to *Annals of Applied Probability*.

### Contents

---

<b>3.1</b>	<b>Introduction</b>	<b>150</b>
<b>3.2</b>	<b>A trait structured IBM</b>	<b>154</b>
3.2.1	Parameters and assumptions of the model	155
3.2.2	Main result	156
3.2.3	Sketch and outline of the proof	157
3.2.3.1	Slow-fast asymptotic analysis	157
3.2.3.2	Difficulties encountered	159
3.2.3.3	Outline of the proof	161
3.2.3.4	Prospects	161
<b>3.3</b>	<b>Infinitesimal generator approximation</b>	<b>162</b>
3.3.1	Generators and martingales	163
3.3.2	Asymptotic expansions	164
3.3.2.1	Slow component	164
3.3.2.2	Fast component	165
3.3.3	Generators and martingales in the case of polynomials in $\mu$	167
3.3.3.1	For bounded test functions	167
3.3.3.2	For some unbounded test functions	168
3.3.4	Proof of Lemma 3.3.4	171
<b>3.4</b>	<b>Moments estimates</b>	<b>173</b>
3.4.1	Estimates of the moment of order 6	174
3.4.2	Some inequalities on the moment of order 2	177
3.4.3	Convergence to $+\infty$ of the stopping time $\tau^K$ and support concentration property	178

3.4.4	Proof of Lemma 3.4.9 . . . . .	179
3.4.4.1	One-step transitions . . . . .	180
3.4.4.2	Construction of the coupling . . . . .	184
3.4.4.3	Estimates of exit from an attractive domain for random walks . . . . .	185
3.4.4.4	Conclusion . . . . .	188
<b>3.5</b>	<b>Tightness on the torus . . . . .</b>	<b>188</b>
<b>3.6</b>	<b>Characterisation of the occupation measure limit on the torus . . . . .</b>	<b>194</b>
3.6.1	Proof of Proposition 3.6.3 . . . . .	195
3.6.2	Proof of Lemma 3.6.4 . . . . .	197
3.6.2.1	Extension of the duality result for the centered FLEMING-VIOT process . . . . .	197
3.6.2.2	Proof of Lemma 3.6.4 . . . . .	202
<b>3.7</b>	<b>Characterisation of the limiting values of the slow component . . . . .</b>	<b>205</b>
3.7.1	Convergence of the slow component on the torus . . . . .	205
3.7.2	End of the proof of Theorem 3.2.1 . . . . .	205

---

**Abstract.** In this chapter, a stochastic individual-based model describing Darwinian evolution of asexual, phenotypic trait-structured population, is studied. We consider a large population with constant population size characterised by a resampling rate modelling competition pressure driving selection and a mutation rate where mutations occur during life. In this model, the population state at fixed time is given as a measure on the space of phenotypes and the evolution of the population is described by a continuous time, measure-valued MARKOV process. We investigate the asymptotic behaviour of the system, where mutations are frequent, in the double simultaneous limit of large population ( $K \rightarrow +\infty$ ) and small mutational effects ( $\sigma \rightarrow 0$ ) proving convergence to an ODE known as the canonical equation of adaptive dynamics. This result holds only for a certain range of  $\sigma$  parameters (as a function of  $K$ ) which must be small enough but not too small either. The canonical equation describes the evolution in time of the dominant trait in the population driven by a fitness gradient. This result is based on a slow-fast asymptotic analysis. We use an averaging method, inspired by KURTZ [144], which exploits a martingale approach and compactness-uniqueness arguments. The contribution of the fast component, which converges to the centered FLEMING-VIOT process, is obtained by averaging according to its invariant measure, recently characterised in Chapter 2.

### 3.1 Introduction

In this chapter we study, at the individual level and in the interplay between ecology and Darwinian evolution, a population model, structured by a 1-dimensional quantitative phenotypic trait. The Darwinian evolution is based on three basic mechanisms. Firstly, *heredity* which allows the transmission of the individual phenotypic characteristics from one generation to another. Secondly, a source of variation in the individual phenotypic characteristics: in our case it is only *mutations*. Finally, a *selection* mechanism which can result from interaction between individuals in the population such as competition. Our model is an *individual-based model* (in short, IBM) which involves a finite and asexual population with constant population size in which each individual's birth, death and mutation events are described. IBMs were first introduced in ecology as a tool to describe local interactions or complex phenomena at the individual level [186, 26, 27, 77, 118, 63]. Ecological studies using IBMs are mainly numerical and the models are mostly posed in discrete space as systems of interacting particles [186] and more

rarely in continuous space [26, 27, 77]. Many IBMs (with non-constant population size) have been proposed and studied in the context of Darwinian evolutionary by the biology community [171, 73, 96] and the mathematical community [37, 39, 10]. Others are dispersal models in spatially structured populations where the trait is viewed as a spatial location and mutations as dispersal [26, 27, 149, 102]. Other models, structured in age, were developed in [50] and studied mathematically in [187, 209], or structured in age and traits in [97, 166, 167].

We consider an IBM with fixed population size, so that births and deaths occur simultaneously in so-called *resampling* (or swap phenomenon) events. The mutation and resampling rate of an individual depends on its phenotype. When a mutation occurs, the new mutant trait is close to its parent's one yielding a slow variation of the trait. In population genetics, the WRIGHT-FISHER model (and its extensions with selection, mutation or immigration), CANNINGS model or the MORAN model are interested in the evolution of allele frequencies according to various mechanisms [86]. In [99, 61, 60, 85], the authors construct the FLEMING-VIOT process as a scaling limit of large population from the MORAN process. Several extensions exist, including frequency-dependent selection, recombination, other reproduction mechanisms [91, 117]. In particular, the last article provides a bridge between population genetics and eco-evolutionary models. Others, based on KERMACK-MCKENDRICK's model, are interested in epidemiological questions [180, 182].

The aim of this chapter is to describe the evolutionary dynamics of the dominant trait, at the population level, on a long time scale, as a solution to an ordinary differential equation (in short, ODE) called *Canonical Equation of the Adaptive Dynamics* (in short, CEAD). Canonical equations are well-known tools in evolutionary biology, used to predict the evolutionary fate of ecological communities. More precisely, such equations describe the evolution of dominant traits in a biological population as driven by mutations and a fitness gradient which describes the strength of selection that pushes the population to locally increase its fitness [214, 148, 203, 72].

Fitness measures the selective value of a given individual in a given environment including the population under consideration itself. This individual can be any (fictitious) mutant individual that can be born in the population at the time under consideration. The way to construct a fitness landscape depends on the ecological context [172]. For continuous time homogeneous MARKOV models as studied below, it is the instantaneous growth rate (birth rate minus death rate) of the individual considered in the environment considered. If we further assume that the population constituting the environment is in a stationary state, then we can consider that the fitness of a given mutant individual in this population governs the possibility of invasion of the descendants of this mutant in the population.

The canonical equation has been studied and derived in different contexts such as game theory [125] and quantitative genetics [148]. In the branch of evolutionary biology called *adaptive dynamics* [126, 185, 161], the CEAD has been introduced heuristically in [72]. The theory of adaptive dynamics studies the links between ecology and Darwinian evolution, more precisely, it investigates the effects of the ecological aspects of population dynamics on the evolutionary process, and so describes the population dynamics on the phenotypic level instead of the genotypic level. The theory of adaptive dynamics is based on biological assumptions of *rare* and *small mutations* and of *large population* under which the CEAD was proposed.

Two mathematical approaches were developed to give a proper mathematical justification of this theory: a *deterministic* one, and a *stochastic* one. All these approaches are based on the use of IBMs and different combinations of the three previous biological assumptions and consider a parameter scaling under which the population distribution over the trait space concentrates to DIRAC masses, i.e. to subpopulations in which all the individuals have the same trait. The cano-

nical equation corresponds to the motion of DIRAC masses. Let us introduce three parameters corresponding to the different biological assumptions:  $p$  for the mutation probability,  $\sigma$  for the mutation size and  $K$  for the population size.

In the deterministic approach, [102] first establishes that, under the asymptotic of large population ( $K \rightarrow +\infty$ ), the IBM converges in law to a deterministic process which is a weak solution of a partial differential equation. Then by adding the small mutations assumption ( $\sigma \rightarrow 0$ ) and an appropriate time scaling  $1/\sigma$ , [75, 194, 160] establish the convergence to a version of the canonical equation different from the one of [72] and described by a HAMILTON-JACOBI equation with constraint.

The stochastic approach was developed in [37, 47]. In [37] it is proved that the IBM converges, for finite dimensional distributions, under the double simultaneous asymptotic of large population ( $K \rightarrow +\infty$ ) and rare mutations ( $p \rightarrow 0$ ) to a stochastic process: the *TSS* for *Trait Substitution Sequence* introduced in [171, Section 6.4]. This is a pure jump Markov process in the trait space where the population is at all times monomorphic and where the jumps describe the invasion and then the fixation of a mutant  $y$  in a monomorphic resident population of trait  $x$ . By adding the small mutations assumption ( $\sigma \rightarrow 0$ ) to the TSS, its convergence to the canonical equation proposed in [72] is established in [47]. The time scale involved in observing the canonical equation phenomenon is  $1/K\sigma^2p$ . Extensions of these results were obtained in [43] for chemostat models, in [84, 166] for age-structured populations models, in [154] for spatial-structured populations models and [6] which studies the simultaneous application of the three limits ( $K \rightarrow +\infty, p \rightarrow 0, \sigma \rightarrow 0$ ) in order to determine precisely the range of parameters leading to the canonical equation.

Despite their success, the proposed approaches are criticised by biologists [212, 195]. Among the biological assumptions of adaptive dynamics, the assumption of rare mutations is the most criticised as unrealistic [212]. The rate of molecular mutation is relatively well known and generally involves several nucleotide substitutions per generation. The adaptive dynamics response is based on the fact that only non-synonymous mutations (changing phenotypes) producing viable individuals should be taken into account [3] [171, Section 6.4]. Since only a small fraction of DNA codes for proteins and many mutations produce non-functional proteins, and thus non-viable individuals, it is not unreasonable to assume that mutations are rare, but probably not as rare as assumed in the stochastic approach [37, 47, 6]. Another criticism of stochastic approaches is that the phenomenon of the canonical equation takes place on a too long evolutionary time scale. In order to solve these problems, we propose to apply a double simultaneous asymptotic of small mutations ( $\sigma \rightarrow 0$ ) and large population ( $K \rightarrow +\infty$ ), but *frequent mutations* ( $p \equiv 1$ ). After conveniently scaling the population state, this leads to a *slow-fast dynamics* where the phenomenon of the canonical equation is visible on a time scale  $1/K\sigma^2$ . So, there is some consistency between the previous stochastic approaches developed in [47, 6] and ours: it is exactly the same canonical equation. However in our situation, the CEAD is visible on a shorter evolutionary scale than theirs, and therefore is biologically more reasonable.

As the mutation size parameter is small, the population distribution tends to be close to a DIRAC mass. Our aim is to describe the evolution of the support of this DIRAC mass on a large time scale. The mean trait appears as the natural slow component. It will be proved to act on the time scale  $1/K\sigma^2$ . In our case, the fast component acts on the time scale  $K$  and is given by the dynamics of the centered and dilated distribution of traits corresponding to a discrete version of the centered FLEMING-VIOT process (see Chapter 2). Since the centered FLEMING-VIOT process is ergodic (see Chapter 2), we expect the centered and dilated distribution of traits to stabilise on the slow time scale, hence the distribution of traits should stabilise to a DIRAC mass.

Therefore, the dynamics of the dominant trait corresponds to that of the mean trait.

The reason for not considering the same IBM as in [37, 47, 6] is because it involves three time scales, a slow one corresponding to the dynamics of the mean trait acting on the evolutionary time scale  $1/K\sigma^2$ , a fast one corresponding to the dynamics of the centered, dilated distribution of traits acting on the evolutionary time scale  $K$  and a very fast one corresponding to the population size dynamics acting on the ecological time scale 1. For simplicity, we focus here on a model with constant population size to reduce the number of time scales to two. We expect our results to extend to general IBMs as in [37, 47, 6] and we leave this for future works.

To prove convergence in the framework of slow-fast dynamics (also called *stochastic singular perturbations*), different techniques can be used.

Firstly, the method of the *perturbed test function* initially proposed by PAPANICOLAOU, STROOCK and VARADHAN in [192] identifies the generator of the limit process with a martingale approach whose idea is the following. If we consider a family of stochastic processes  $((X^\varepsilon, Y^\varepsilon))_{\varepsilon>0}$  of generator  $L^\varepsilon$  where  $X^\varepsilon$  is the *slow component* and  $Y^\varepsilon$  is the *fast component* in the form  $Y^\varepsilon(t) = Y(t/\varepsilon)$  where  $Y$  is a MARKOV process, the slow-fast problem consists in identifying the limit process of  $X^\varepsilon$  using ergodicity properties of the fast dynamics. Assuming that the family  $(X^\varepsilon)_{\varepsilon>0}$  is tight, we consider  $X$  a limiting value. We can expect to characterise  $X$  with a martingale problem derived from the martingale problem of  $(X^\varepsilon, Y^\varepsilon)$  provided that the solution to this problem is unique. We would then obtain the convergence of  $X^\varepsilon$  to  $X$  in law. However, for multiscale singular problems, the convergence of  $L^\varepsilon\varphi((X_t^\varepsilon, Y_t^\varepsilon))$  when  $\varepsilon \rightarrow 0$ , for a test function  $\varphi$  depending only on the  $X$  component, does not (in general) take place because it contains diverging terms in  $\varepsilon$ . To overcome this difficulty, firstly the idea is to decompose the generator  $L^\varepsilon$  in the following form  $L^\varepsilon = \frac{1}{\varepsilon}L_1 + L_0$  where  $L_1$  is the infinitesimal generator of the MARKOV process  $Y$  in the variable  $y$  and  $L_0$  the operator of the slow component depending on slow and fast variables. Secondly, the idea is to perturb the initial test function  $\varphi(x)$  into a test function  $\varphi^\varepsilon(x, y) := \varphi(x) + \varepsilon\varphi_1(x, y)$  such that

$$L^\varepsilon\varphi^\varepsilon = \left[ \frac{1}{\varepsilon}L_1 + L_0 \right] (\varphi + \varepsilon\varphi_1) = (L_1\varphi_1 + L_0\varphi - \bar{L}_0\varphi) + \bar{L}_0\varphi + \varepsilon L_0\varphi_1 \quad (3.1)$$

because of  $L_1\varphi = 0$  and where  $\bar{L}_0$  is the operator of the limit slow component *averaged* by the unique invariant probability measure of the limit fast component  $Y$ . Provided that  $\varphi_1$  solves the POISSON equation (with respect to  $L_1$  and the variable  $y$ )

$$L_1\varphi_1 + L_0\varphi - \bar{L}_0\varphi = 0, \quad (3.2)$$

we deduce that  $L^\varepsilon\varphi^\varepsilon = \bar{L}_0\varphi + O(\varepsilon)$ . Then,  $\varphi^\varepsilon(x, y) \rightarrow \varphi(x)$  and  $L^\varepsilon\varphi^\varepsilon(x, y) \rightarrow \bar{L}_0\varphi(x)$  when  $\varepsilon \rightarrow 0$  as expected.

The perturbed test function method has been extended in [22, 145, 146] and in the books [100, Section 6.3], [147]. Many other references apply the perturbed test function strategy in various settings: in finance [101] in transport problems as in [192] or in [56] where the tightness of the fast component is established using its occupation measure. Similar methods are used in homogenisation [192, 17, 193] (see also [16] which exploits spectral and semi-group properties in addition), and in stochastic stability and control [147].

Finally, an important method is the stochastic averaging, in a generic framework, that was proposed by KURTZ in [144]. Many approaches [8, 167, 119, 54, 108, 28, 9, 19] including ours, are based on it. The main idea consists in exploiting the *occupation measure*  $\Gamma^\varepsilon$  of the fast

component  $Y^\varepsilon$  which is formally defined for all  $t \geq 0$ , for any Borelian  $B$  by

$$\Gamma^\varepsilon([0, t] \times B) = \int_0^t \mathbb{1}_B(Y_s^\varepsilon) ds,$$

and to establish the convergence of the couple  $(X^\varepsilon, \Gamma^\varepsilon)$  when  $\varepsilon \rightarrow 0$  using *compactness-uniqueness* arguments. Proceeding in this way allows us to escape the difficulties created by the fluctuations of the fast component and avoids to obtain tightness result of the “slow-fast” couple  $(X^\varepsilon, Y^\varepsilon)$ . The proof can be divided into four steps.

The first one consists in establishing uniform tightness of the sequence of laws of the couple  $(X^\varepsilon, \Gamma^\varepsilon)$ . The second one consists in establishing a martingale problem for any limiting value of the family of laws of the couple. The third one is based on establishing the uniqueness of the limiting value of  $\Gamma^\varepsilon$  expressed in terms of the limit  $X$  of  $X^\varepsilon$  which is given. The characterisation of the limiting value  $\Gamma$  of  $\Gamma^\varepsilon$  is usually based on ergodicity arguments. Finally, the last step is to establish the uniqueness of the martingale problem for the slow component limit  $X$  when  $\Gamma$  is given as above.

The KURTZ approach seems to be better adapted to our situation than the perturbed test function method. Indeed, the generator  $L_1$  being that of a measure-valued diffusion, it is thus delicate to find the good class of test function  $\varphi_1$  satisfying (3.2) and the computations are difficult because of the moment terms (see Section 3.2.3). In addition, the remainder terms generated by (3.1) are difficult to control and the fast component  $Y^\varepsilon$  does not take the form  $Y(t/\varepsilon)$ . However, KURTZ’s approach can neither be implemented directly because of difficulties inherent to our model described in Section 3.2. Therefore, our result required a complete reworking of the classical arguments. In particular, in our case, we do not have uniform moment estimates but only a uniform control, in probability, of the second moment of the fast variable up to a stopping time  $\check{\tau}^K$  where the diameter of the support of the fast component becomes large.

This chapter is organised as follows. In Section 3.2, we define our trait structured IBM, state the central theorem about the CEAD characterising the limit model which consists of an ODE ruling the dynamics of the limit slow component. We establish in this section the sketch and outline of the main proof and the difficulties encountered. We give also some prospects that let us believe that we can improve our main result by relaxing some assumptions. The rest of this chapter is devoted to prove the central theorem of Section 3.2. In Section 3.3 some approximations of the infinitesimal generator of the slow-fast process are proved. In Section 3.4, we establish estimates of moments, in particular of the sixth and second order moment up to time  $\check{\tau}^K$ . We also prove the convergence of  $\check{\tau}^K$  to  $+\infty$  when  $K \rightarrow +\infty$  exploiting coupling arguments between the moment of order 2 and a certain biased random walk for which large deviations estimates are established. In Section 3.5, we prove the tightness of the couple “slow-occupation measure fast” in the torus case. In Section 3.6, we establish, in the torus case, the uniqueness of the limit occupation measure which is described by the unique invariant probability measure of the centered FLEMING-VIOT process. This result is used in Section 3.7 to characterise the limit slow component as the unique solution of the CEAD with values in the torus. Thanks to the non-explosion of this ODE and choosing the torus large enough we are able to conclude the proof of our main result given in Section 3.2.

## 3.2 A trait structured IBM

Let us describe the microscopic model which models population evolution, in a Darwinian sense, at the individual level.



### 3.2.1 Parameters and assumptions of the model

We consider a discrete population of constant size in continuous time in which the survival and reproductive capacity of each individual is characterised by a continuous quantitative phenotypic trait  $x \in \mathbb{R}$ , i.e. an overall characteristic subject to selection such as body size at maturity. The individuals reproduce asexually during their lives, i.e. the reproduction scheme is assumed clonal simultaneously with death of another individual, with frequency-dependent rates. Each birth of an individual occurs simultaneously with the death of another individual in the population. A mutation occurs during life at the individual level at rate 1: in this sense mutations are considered *frequent*. Each mutation has an amplitude of the order of magnitude  $\sigma \in (0, +\infty)$ . Small  $\sigma$  means mutations have a small phenotypic effect, i.e. evolution acts slowly on the individual phenotypic characteristics.

We are interested in approximating the long term dynamics of a large population. We assume that the number of individuals alive at each time  $t \geq 0$  is always equal to  $K$ . Let us denote by  $x_1(t), \dots, x_K(t)$ , the phenotypic trait values of these individuals at time  $t$ . The state of the population at time  $t$ , can be described by the finite point measure on  $\mathbb{R}$  rescaled by  $K$

$$\nu_t^{K,\sigma} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i(t)}$$

where  $\delta_x$  is the DIRAC measure at  $x$ .

For all  $x, y \in \mathbb{R}$ , let us introduce the following biological parameters:

- $b(x, y) \in \mathbb{R}_+$  is the *resampling* rate, i.e. the rate of simultaneous birth of an individual holding trait  $y$  and death of an individual holding trait  $x$  and it can be interpreted as modelling a competitive pressure driving selection.
- $\theta(x) \in \mathbb{R}_+$  is the rate of mutation of an individual holding trait  $x$ .
- $m(x, dh)$  is the mutation law of the scaled mutation step  $h$ , born from an individual with trait  $x$ , where the mutant trait is given by  $y := x + \sigma h \in \mathbb{R}$ . It is a probability measure on  $\mathbb{R}$ .

Let us also introduce the following notations, used throughout this chapter:

- $\beta(x) := \frac{\theta(x)}{b(x,x)}$ , the ratio between the mutation rate and the resampling rate in a monomorphic population with trait  $x$ , which can be interpreted (in the scaling limit considered below) as the mean number of mutations between two resampling events.
- $\text{Fit}(y, x) := b(x, y) - b(y, x)$  is the adaptive value or *fitness* of a mutant individual with trait  $y$  in the population of  $(K - 1)$  individuals of trait  $x$ . The fitness can be interpreted as the initial growth rate of a mutant individual  $y$  in a resident monomorphic population with trait  $x$ .

Indeed, the total birth rate of a mutant individual  $y$  in this population is  $b(x, y)(K - 1)$  and the total death rate of a mutant individual  $y$  in this population is  $b(y, x)(K - 1)$ . Hence, the (initial) growth rate of the mutant population is  $\text{Fit}(y, x)(K - 1)$ .

Let  $\mathcal{M}_1(\mathbb{R})$  be the set of probability measures on  $\mathbb{R}$ , endowed with the topology of weak convergence making it a Polish space [20]. For a measurable real bounded function  $f$ , and a measure  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote  $\langle f, \nu \rangle := \int_{\mathbb{R}} f(x)\nu(dx)$ . We denote by  $\text{id}$  the identity function. We denote also  $\mathbb{N} := \{0, 1, 2, \dots\}$  and  $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ . If  $\mathcal{I}$  is an interval of  $\mathbb{R}$ , then for all  $\ell \in \mathbb{N}$ ,

$p \in \mathbb{N}^*$ , we denote by  $\mathcal{C}^\ell(\mathcal{I}^p, \mathbb{R})$  the space of functions of class  $\mathcal{C}^\ell$  from  $\mathcal{I}^p$  to  $\mathbb{R}$  and by  $\mathcal{C}_b^\ell(\mathcal{I}^p, \mathbb{R})$  the space of functions of class  $\mathcal{C}^\ell(\mathcal{I}^p, \mathbb{R})$  with bounded derivatives. Finally, we define for all  $K \in \mathbb{N}^*$ ,

$$\mathcal{M}_{1,K}(\mathbb{R}) := \left\{ \frac{1}{K} \sum_{i=1}^K \delta_{x_i} \mid (x_i)_{1 \leq i \leq K} \in \mathbb{R}^K \right\}$$

the set of probability measures on  $\mathbb{R}$  of  $K$  atoms of mass  $1/K$ .

**Assumptions.** Let us denote by **(A)** the following two assumptions:

**(A1)** The maps  $b : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ , and  $\theta : \mathbb{R} \rightarrow \mathbb{R}_+$  are respectively in  $\mathcal{C}_b^2(\mathbb{R}^2, \mathbb{R})$  and  $\mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$  and there exists  $0 < \underline{b}, \bar{b}, \underline{\theta}, \bar{\theta} < \infty$  such that:

$$\underline{b} \leq b(\cdot, \cdot) \leq \bar{b}, \quad \text{and} \quad \underline{\theta} \leq \theta(\cdot) \leq \bar{\theta}.$$

**(A2)** **(a)** There exists  $A_m \in (0, +\infty)$  such that the law  $m(x, dh)$  is absolutely continuous with respect to the LEBESGUE measure on  $\mathbb{R}$  with density  $m(x, h)$  centered and satisfies

$$\forall x \in \mathbb{R}, \forall |h| \geq A_m, \quad m(x, h) = 0.$$

**(b)** For all  $\alpha \in \mathbb{R}$ , for all  $\ell \in \mathbb{N}$ , we denote by

$$m_\ell(\alpha) := \int_{\mathbb{R}} |h|^\ell m(\alpha, h) dh$$

the  $\ell^{\text{th}}$  moment of the mutation law. We assume for  $\ell \in \{2, 4, 6\}$  that  $m_\ell$  is LIPSCHITZ and there exists  $\underline{m}_\ell, \bar{m}_\ell \in (0, +\infty)$  such that for all  $\alpha \in \mathbb{R}$ ,  $\underline{m}_\ell \leq m_\ell(\alpha) \leq \bar{m}_\ell$ .

Let us now give the infinitesimal generator  $L^{K,\sigma}$  of the  $\mathcal{M}_{1,K}(\mathbb{R})$ -valued MARKOV process  $\nu^{K,\sigma} := (\nu_t^{K,\sigma})_{t \geq 0}$  describing the ecological dynamics of the population with resampling. The generator  $L^{K,\sigma}$  is defined for any bounded measurable map  $\phi$  from  $\mathcal{M}_{1,K}(\mathbb{R})$  to  $\mathbb{R}$ , by

$$\begin{aligned} L^{K,\sigma} \phi(\nu) &= K \int_{\mathbb{R}} \nu(dx) \int_{\mathbb{R}} \nu(dy) b(x, y) \left[ \phi \left( \nu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right) - \phi(\nu) \right] \\ &\quad + K \int_{\mathbb{R}} \theta(x) \nu(dx) \int_{\mathbb{R}} m(x, h) \left[ \phi \left( \nu - \frac{\delta_x}{K} + \frac{\delta_{x+\sigma h}}{K} \right) - \phi(\nu) \right] dh. \end{aligned} \tag{3.3}$$

The first term describes the resampling event of one individual by another and the second term describes the effect of mutations over the lifetime of individuals.

### 3.2.2 Main result

The main result of this chapter is the following theorem. For all  $\nu \in \mathcal{M}_1(\mathbb{R})$ , let us denote by  $\text{Diam}(\text{Supp } \nu)$  the diameter of the support of  $\nu$ .

**Theorem 3.2.1.** *Assume **(A)** and for all  $K \in \mathbb{N}^*$ ,  $\sigma \in (0, +\infty)$ , for some  $x_0 \in \mathbb{R}$ , for all  $\ell \in \{1, 2, 3\}$ ,*

$$\langle \text{id}, \nu_0^{K,\sigma} \rangle := x_0, \quad \text{and} \quad \langle (\text{id} - x_0)^{2\ell}, \nu_0^{K,\sigma} \rangle \leq C_{2\ell}^* K^\ell \sigma^{2\ell} \tag{3.4}$$

for some  $C_{2\ell}^* > 0$ . Assume that  $K \rightarrow +\infty$ ,  $\sigma \rightarrow 0$  and  $\sigma = \sigma(K)$  such that

$$\exists \varepsilon > 0, \quad \forall C > 0, \quad K^{-C \log(K)} \ll \sigma \ll K^{-(2+\varepsilon)}. \quad (3.5)$$

Then, for the SKOROHOD topology, the sequence of the mean trait processes  $\left( \left( \langle \text{id}, \nu_{t/K\sigma^2}^{K,\sigma} \rangle \right)_{t \geq 0} \right)_{K \in \mathbb{N}^*}$  converges in law in  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$  when  $K \rightarrow +\infty$ ,  $\sigma \rightarrow 0$ , to the unique deterministic process  $(\zeta_t)_{t \geq 0} \in \mathcal{C}^0(\mathbb{R}_+, \mathbb{R})$  with initial condition  $\zeta_0 := x_0$ , which is solution to the Canonical Equation of Adaptive Dynamics (CEAD):

$$\dot{y} = \partial_1 \text{Fit}(y, y) \times \frac{\beta(y)m_2(y)}{2}.$$

In addition, we have the following support concentration property: for all  $T > 0$ ,

$$\sup_{t \in [0, T]} \text{Diam} \left( \text{Supp } \nu_{t/K\sigma^2}^{K,\sigma} \right) \leq \frac{1}{K} \quad (3.6)$$

holds with probability which tends to 1 when  $K \rightarrow +\infty$ .

The canonical equation is composed of two terms. The first term: the fitness gradient, describes the strength of selection that pushes the population to increase its adaptive value locally. The second term describes the effect of mutations. Note that (3.6) describes, in some sense, when  $K \rightarrow +\infty$ , the convergence of the population distribution  $(\nu_{t/K\sigma^2}^{K,\sigma})$  to a DIRAC mass.

**Remark 3.2.2.** (1) We conjecture that the assumption  $\sigma \ll K^{-(2+\varepsilon)}$  given by (3.5) is too restrictive and that it can be weakened by  $\sigma \ll K^{-(\frac{3}{2}+\varepsilon)}$ . This conjecture is supported by the discussion in Section 3.2.3.4 below.

(2) We actually prove Theorem 3.2.1 (see Section 3.4.4.3) under a slightly weaker assumption than the left bound of (3.5): for a universal constant  $C_0$  depending only on  $b, \theta, m$  and not on  $K$  and  $\sigma$ ,

$$\exists \varepsilon > 0, \quad K^{-\varepsilon^2 C_0 \log(K)} \ll \sigma \ll K^{-(2+\varepsilon)}.$$

(3) Assumption **(A2)(a)** is only used in Section 3.4.4 and could be weakened with some tail bounds (typically exponential) on  $m(x, \cdot)$ .

The aim of the rest of the chapter is to prove Theorem 3.2.1. Since  $\sigma = \sigma(K)$ , we will use the notation  $\nu^K$  to designate  $\nu^{K,\sigma}$ .

### 3.2.3 Sketch and outline of the proof

Note that, the convergence result of Theorem 3.2.1 takes place on the time scale  $1/K\sigma^2$ . The idea of the proof is based on slow-fast asymptotic analysis techniques developed by KURTZ [144]. Let us begin by introducing the slow and fast dynamics involved in our model, then the difficulties encountered and finally the outline of the proof.

#### 3.2.3.1 Slow-fast asymptotic analysis

Let  $a, \alpha \in \mathbb{R}$  and  $\mathcal{B}(\mathbb{R})$  the BOREL  $\sigma$ -field on  $\mathbb{R}$ . Let us define respectively by  $\tau_\alpha$  and  $h_a$ , the translation operator of vector  $\alpha$  and the homothety of ratio  $a$ . For all  $x \in \mathbb{R}$ , for all  $A \in \mathcal{B}(\mathbb{R})$ ,

$$(h_a \circ \tau_\alpha) \# \delta_x(A) = \delta_{a(x+\alpha)}(A) \quad \text{and} \quad (\tau_\alpha \circ h_a) \# \delta_x(A) = \delta_{ax+\alpha}(A).$$

Note that for all  $K \in \mathbb{N}^*$ , any  $\nu \in \mathcal{M}_{1,K}(\mathbb{R})$  has all its moments finite. We define for all  $K \in \mathbb{N}^*$ ,  $\mathcal{M}_{1,K}^c(\mathbb{R})$  the set of centered probability measures of  $\mathcal{M}_{1,K}(\mathbb{R})$ . From the population process  $(\nu_{t/K\sigma^2}^K)_{t \geq 0}$ , we introduce two evolutionary dynamics:

- The slow dynamics, with value in  $\mathbb{R}$ , corresponds to that of the mean trait defined by

$$z_t^K := \langle \text{id}, \nu_{t/K\sigma^2}^K \rangle.$$

- The fast dynamics, with value in  $\mathcal{M}_{1,K}^c(\mathbb{R})$ , corresponds to that of the centered and dilated distribution of traits defined by

$$\mu_t^K := \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-z_t^K} \right) \# \nu_{t/K\sigma^2}^K = \frac{1}{K} \sum_{i=1}^K \delta_{\frac{1}{\sigma\sqrt{K}}(x_i(t) - z_t^K)}. \quad (3.7)$$

Note that  $\nu_t^K = \left( \tau_{z_{t/K\sigma^2}^K} \circ h_{\sigma\sqrt{K}} \right) \# \mu_{t/K\sigma^2}^K$ .

For all  $\ell \in \mathbb{N}$  and  $\nu \in \mathcal{M}_1(\mathbb{R})$ , we denote by  $M_\ell(\nu) := \langle |\text{id}|^\ell, \nu \rangle$  the  $\ell^{\text{th}}$  moment of the measure  $\nu$ . Let us consider for all  $\ell \in \mathbb{N}^*$ ,

$$\mathcal{M}_1^{c,\ell}(\mathbb{R}) := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \mid M_\ell(\mu) < \infty, \langle \text{id}, \mu \rangle = 0 \right\}$$

and the convention  $M_\ell(\mu) = +\infty$  if  $\mu \notin \mathcal{M}_1^{c,\ell}(\mathbb{R})$ . Let us denote by  $\mathcal{B}(E, \mathbb{R})$  the set of BOREL functions on the metric space  $E$ .

In the following, for all  $K \in \mathbb{N}^*$ , we study the couple of processes  $((z_t^K, \mu_t^K))_{t \geq 0}$ . Denoting  $\mathcal{L}^K$  the infinitesimal generator of the process  $((z_t^K, \mu_t^K))_{t \geq 0}$ , computed in Proposition 3.3.1, we will check the following assertions, proved respectively in Propositions 3.3.2 and 3.3.3.

- (1) If we consider a function  $f : (z, \mu) \mapsto f(z) \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ , the generator  $\mathcal{L}^K$  satisfies the following decomposition

$$\mathcal{L}^K f(z, \mu) = \mathcal{L}_{\text{SLOW}} f(z, \mu) + O\left(\frac{1}{K^2} + \frac{M_2(\mu)}{K} + \sigma\sqrt{K}M_3(\mu)\right) \quad (3.8)$$

where the operator  $\mathcal{L}_{\text{SLOW}}$ , of the limit slow component, is defined from  $\mathcal{C}_b^1(\mathbb{R}, \mathbb{R})$  to  $\mathcal{B}(\mathbb{R} \times \mathcal{M}_1^{c,2}(\mathbb{R}), \mathbb{R})$  by

$$\mathcal{L}_{\text{SLOW}} f(z, \mu) = f'(z)M_2(\mu) \partial_1 \text{Fit}(z, z). \quad (3.9)$$

- (2) If we consider a function  $F_\varphi : (z, \mu) \mapsto F_\varphi(\mu) := F(\langle \varphi, \mu \rangle)$  with  $F \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$  and  $\varphi \in \mathcal{C}_b^3(\mathbb{R}, \mathbb{R})$ , the generator  $\mathcal{L}^K$  satisfies the following decomposition

$$\mathcal{L}^K F_\varphi(z, \mu) = \frac{\theta(z)m_2(z)}{K^2\sigma^2} \left[ \mathcal{L}_{\text{FVc}}^{\lambda(z)} F_\varphi(z, \mu) + O\left(\frac{1}{\sqrt{K}} + \sigma K^{\frac{3}{2}}M_2(\mu) + \frac{M_3(\mu)}{K}\right) \right] \quad (3.10)$$

where  $\mathcal{L}_{\text{FVc}}^{\lambda(z)}$  is the generator of the *centered FLEMING-VIOT process* with resampling rate  $\lambda(z) := \frac{b(z,z)}{\theta(z)m_2(z)}$ , as studied in Chapter 2 and whose definition is recalled below. For all  $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$  and  $g \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ ,

$$\begin{aligned} \mathcal{L}_{\text{FVc}}^\lambda F_g(\mu) &:= F'(\langle g, \mu \rangle) \left( \left\langle \frac{g''}{2}, \mu \right\rangle + \lambda \left[ \langle g'', \mu \rangle M_2(\mu) - 2 \langle g' \times \text{id}, \mu \rangle \right] \right) \\ &\quad + \lambda F''(\langle g, \mu \rangle) \left[ \langle g^2, \mu \rangle - \langle g, \mu \rangle^2 + \langle g', \mu \rangle^2 M_2(\mu) - 2 \langle g', \mu \rangle \langle g \times \text{id}, \mu \rangle \right]. \end{aligned} \quad (3.11)$$

From (3.10) and (3.8), note that the fast component  $(\mu_{K\sigma^2 t}^K)_{t \geq 0} = \left( \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-\langle \text{id}, \nu_t^K \rangle} \right) \# \nu_t^K \right)_{t \geq 0}$  moves on the evolutionary time scale  $K$  whereas the slow component  $(z_{K\sigma^2 t}^K)_{t \geq 0} = (\langle \text{id}, \nu_t^K \rangle)_{t \geq 0}$  moves over a much longer evolutionary time scale  $1/K\sigma^2$ . The different time scales involved in this model can be represented as in Figure 3.1. It follows that the fast component will be the first to stabilise in its equilibrium state. Note that the operator of the slow component (3.9) depends on the fast variable through the second-order moment  $M_2(\mu)$ . This is a standard difficulty in slow-fast analysis, usually solved by assuming that once the fast component is stabilised in its equilibrium state,  $M_2(\mu)$  can be characterised in terms of the slow component leading to an autonomous slow dynamics in the limit  $K \rightarrow +\infty$ .

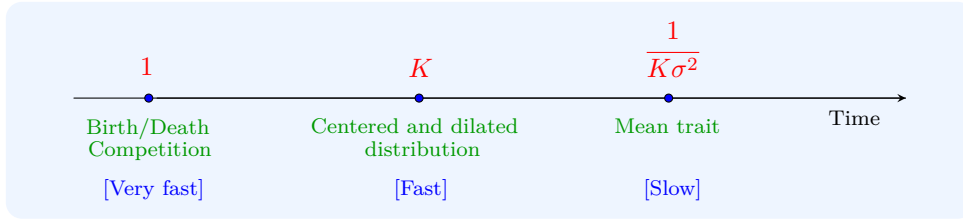


FIGURE 3.1 – Different time scales involved in our individual-based model expected in the original time scale of the process  $(\nu_t^K)_{t \geq 0}$ .

As the limit fast component is driven by a centered FLEMING-VIOT process, we expect that it will inherit ergodicity properties as stated in Section 2.4. This reference establishes the existence of a unique invariant probability measure  $\pi^\lambda$  for the centered FLEMING-VIOT process and characterises it. In particular, we expect that the averaging principle applied to the drift coefficient in (3.9) will lead to the averaged drift

$$\partial_1 \text{Fit}(z_t, z_t) \int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^{\lambda(z_t)}(d\mu) = \frac{\partial_1 \text{Fit}(z_t, z_t)}{2\lambda(z_t)} = \partial_1 \text{Fit}(z_t, z_t) \frac{\beta(z_t) m_2(z_t)}{2}$$

where  $(z_t)_{t \geq 0}$  is the limit slow component (see Corollary 2.4.16 of Chapter 2 for the computation of the integral). Formally, by replacing  $M_2(\mu)$  by its mean value  $\int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^{\lambda(z_t)}(d\mu)$  in (3.9), we obtain that (see details in Section 3.7)

$$\mathcal{L}_{\text{CEAD}} f(z, \mu) = f'(z) \partial_1 \text{Fit}(z, z) \frac{\beta(z) m_2(z)}{2}$$

and we recognise the generator of the announced CEAD.

### 3.2.3.2 Difficulties encountered

The slow-fast analysis relies on a stochastic averaging result that exploits tightness arguments. The classical approach to prove tightness of sequences of laws [10, 102] requires to have uniform pathwise estimates on the moments of the process. In our situation, we note the presence of moments in (3.8) and (3.10). However, we could not establish control in expectation of  $\sup_{0 \leq t \leq T} M_2(\mu_t^K)$  for all  $T \geq 0$  because of the long time scale which does not allow to exploit the martingale problem associated to the decomposition of  $M_2(\mu_t^K)$ . This is an important difference, for example with [167] which studies similar forms of processes (individual-based models).

Instead, we will first use fine pathwise estimates and expectation bounds on  $M_2(\mu_t^K)$  up to the stopping time  $\tau^K$  defined for all  $K \in \mathbb{N}^*$  by

$$\tau^K := \check{\tau}^K \wedge \hat{\tau}^K. \quad (3.12)$$

where

$$\hat{\tau}^K := \inf \left\{ t \geq 0 \mid M_2(\mu_t^K) \geq K^\varepsilon \right\} \quad (3.13)$$

with  $\varepsilon > 0$  given in Theorem 3.2.1 and where

$$\check{\tau}^K := \inf \left\{ t \geq 0 \mid \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}} \right\} \quad (3.14)$$

which allows to control the diameter of the support of the centered and dilated distribution of the traits  $\mu^K$ . An important difficulty arises in Lemma 3.3.6 from the presence of third order moment in the DOOB semi-martingale decomposition of  $M_2(\mu_t^K)$ :

$$\begin{aligned} M_2(\mu_t^K) &= M_2(\mu_0^K) - \frac{1}{K^2 \sigma^2} \int_0^t \left\{ 2b(z_s^K, z_s^K) M_2(\mu_s^K) - \theta(z_s^K) m_2(z_s^K) \right\} ds \\ &\quad + \frac{1}{K^2 \sigma^2} \int_0^t O\left(\frac{1}{K} + \sigma K M_2(\mu_s^K) + \sigma K^{\frac{3}{2}} M_3(\mu_s^K)\right) ds + M_t^{K, P_{\text{id}^2, 1}} \end{aligned}$$

where  $(M_t^{K, P_{\text{id}^2, 1}})_{t \geq 0}$  is a local-martingale. This leads to introduce the stopping time  $\check{\tau}^K$  which guarantees that the error term in  $M_3(\mu_t^K)$  remains under control. This leads to a new difficulty: establishing that  $\check{\tau}^K$  tends to  $+\infty$  in probability when  $K \rightarrow +\infty$ . To prove this, we use pathwise estimates on  $M_2(\mu_t^K)$  up to time  $\check{\tau}^K$  to estimate the different transitions of  $M_2(\mu_t^K)$  between thresholds of the form  $3^{\ell-1} K^{\frac{\varepsilon}{2}}$  and  $3^{\ell+1} K^{\frac{\varepsilon}{2}}$  and to construct a coupling between these transitions and biased random walks. This allows us to use large deviations results on random walks and estimates on the exit from an attracting domain (see e.g. [104, 64]) to prove that the stopping time  $\hat{\tau}^K$ , defined by (3.13), converges to  $+\infty$  in probability when  $K \rightarrow +\infty$ . This in turn implies that  $\check{\tau}^K \rightarrow +\infty$  in probability when  $K \rightarrow +\infty$  (see Section 3.2.3.4 for a discussion of the method used for this step).

The implementation of the slow-fast method of KURTZ [144] is done in two steps. To establish the tightness of the sequence of laws of the slow component, we exploit criteria developed by ETHIER-KURTZ [94, Theorems 3.9.1 and 3.9.4] by restricting ourselves to the torus case. This strategy allows to overcome the difficulty related to the verification of the compact containment condition, on the real line, of the slow component stopped at the stopping time  $\tau^K$ . The second step consists in characterising each accumulation point of the sequence of laws of the slow dynamics and the occupation measure of the fast dynamics. For this, we need to check that any measure  $\gamma$  in  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$  satisfying

$$\int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^\lambda \phi(\mu) \gamma(d\mu) = 0 \quad (3.15)$$

for a certain class  $\mathcal{F}$  of functions  $\phi$  must be  $\pi^\lambda$ . Equation (3.10) suggests to take  $\mathcal{F}$  as the set of functions of the form  $F_\varphi(\mu) := F(\langle \varphi, \mu \rangle)$ . However, the last property seems hard to prove for this choice of  $\mathcal{F}$ . Instead, we adapt a result of DAWSON [60, Theorem 2.7.1] that applies to the set  $\mathcal{F}$  of so-called *polynomial functions* of the form

$$P_{f,n}(\mu) := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu(dx_1) \cdots \mu(dx_n).$$

However the duality property used by DAWSON in [60] does not hold in our case. In Chapter 2, it is proved only a weak duality relation involving stopping times. This difficulty will be solved by proving that the measure  $\gamma$  in (3.15) gives mass only to measures having its first five moments finite. This will allow us to characterise the limit fast component and hence the limit slow component on the torus, as solution to an ODE. Since this ODE is non-explosive, we choose the torus large enough to conclude the proof on the real line.

### 3.2.3.3 Outline of the proof

In Section 3.3, we begin by giving an approximation of the infinitesimal generator of the slow-fast process  $((z_t^K, \mu_t^K))_{t \geq 0}$  for a class of test functions on  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$ , large enough to be convergence determining and we characterise martingales associated to our process. We give also martingale problems of the slow-fast process for polynomials in  $\mu$ . In Section 3.4, we prove some moment estimates. In Section 3.4.1, we give estimates of the moment of order 6. In Section 3.4.2, we establish some inequalities on  $M_2(\mu_t^K)$  and we control its bracket that we will use in Section 3.4.3 to prove  $\mathbb{P}$ -a.s. that  $M_2(\mu_t^K)$  takes superlinear (of the order of  $K^{C \log(K)}$ ) long time before hitting  $K^\varepsilon$ . The fact that  $\tau^K \rightarrow +\infty$  in probability when  $K \rightarrow +\infty$  is proved in Section 3.4.3. The rest of the proof deals with the compactness-uniqueness argument associated to our slow-fast problem. Firstly, in Section 3.5, we establish uniform tightness of the sequence of laws of  $((z^K, \Gamma^K))_{K \in \mathbb{N}^*}$  stopped at time  $\check{\tau}^K$  in the torus case where  $\Gamma^K$  designates the occupation measure of the process  $\mu^K$ . In Section 3.6, we identify and characterise in a unique way the limiting distribution of the fast component on the torus. In Section 3.7.1, we proceed similarly for the slow component. Thanks to the uniqueness on the torus of the limit slow component, we deduce in Section 3.7.2 the announced result of Theorem 3.2.1.

### 3.2.3.4 Prospects

In the first part of this section, we explain the origin of exponent 2 in the inequality  $\sigma \ll K^{-(2+\varepsilon)}$  of Assumption (3.5). In the second part, it is explained how Assumption (3.5) may be improved into the assumption  $\sigma \ll K^{-(\frac{3}{2}+\varepsilon)}$  by using estimates for moments of order  $2\ell$ . Finally, in the last part of this section, we explain the difficulties in obtaining these moment estimates. Note that the assumption  $\sigma \ll K^{-\frac{3}{2}}$  is the best we can expect because of the error term  $\sigma K^{\frac{3}{2}}$  in (3.10).

**(a) Origin of the exponent 2 in Assumption (3.5).** On the one hand, as for all  $t \geq 0$  and  $K \in \mathbb{N}^*$ ,  $\text{Diam}(\text{Supp } \mu_t^K)^2 \leq \left(2 \max\{|x| \mid x \in \text{Supp } \mu_t^K\}\right)^2 \leq 4K M_2(\mu_t^K)$ , we deduce that

$$\begin{aligned} \mathbb{P}(\check{\tau}^K < \hat{\tau}^K \wedge T) &\leq \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}}\right) \\ &\leq \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, M_2(\mu_t^K) > \frac{1}{4\sigma^2 K^{4+\varepsilon}}\right) \end{aligned}$$

On the other hand, note that thanks to the definition (3.13) of the stopping time  $\hat{\tau}^K$ ,

$$\mathbb{P}(\exists t < T \wedge \hat{\tau}^K, M_2(\mu_t^K) > K^\varepsilon) = 0.$$

Hence, Assumption (3.5) implies that  $\mathbb{P}(\check{\tau}^K < \hat{\tau}^K \wedge T) = 0$  for  $K$  large enough. Using this relation, we prove in Section 3.4.3 that  $\mathbb{P}(\tau^K < T) \leq \mathbb{P}(\hat{\tau}^K \leq T \wedge \check{\tau}^K)$ . Hence, to establish

that  $\tau^K \rightarrow +\infty$  when  $K \rightarrow +\infty$  in probability, we need to prove that

$$\lim_{K \rightarrow +\infty} \mathbb{P} \left( \widehat{\tau}^K \leq T \wedge \check{\tau}^K \right) = \lim_{K \rightarrow +\infty} \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \check{\tau}^K} M_2(\mu_t^K) \geq K^\varepsilon \right) = 0$$

as established in Section 3.4.4.

**(b) Improvement of Assumption (3.5) with moments of order  $2\ell$ .** Assume now that  $\sigma \ll K^{-(\frac{3}{2}+\varepsilon)}$ . Then, there exists  $\ell \in \mathbb{N}^*$ ,  $\tilde{\varepsilon} > 0$  such that

$$\forall C > 0, \quad K^{-C \log(K)} \ll \sigma \ll K^{-(\frac{3}{2} + \frac{1}{2\ell} + \frac{\ell+1}{2\ell} \tilde{\varepsilon})} \quad (3.16)$$

and we consider the stopping times

$$\widehat{\tau}_{\text{bis}}^K := \inf \left\{ t \geq 0 \mid M_2(\mu_t^K) \geq K^{\tilde{\varepsilon}} \right\} \quad \text{and} \quad \check{\tau}_{\text{bis}}^K := \inf \left\{ t \geq 0 \mid \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\tilde{\varepsilon}}{2}}} \right\}.$$

Using now that for all  $t \geq 0$ , for all  $K \in \mathbb{N}^*$ , for all  $\ell \in \mathbb{N}^*$ ,  $\text{Diam}(\text{Supp } \mu_t^K)^{2\ell} \leq 2^{2\ell} K M_{2\ell}(\mu_t^K)$ , we deduce that

$$\begin{aligned} \mathbb{P} \left( \exists t < T \wedge \widehat{\tau}_{\text{bis}}^K, \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\tilde{\varepsilon}}{2}}} \right) \\ \leq \mathbb{P} \left( \exists t < T \wedge \widehat{\tau}_{\text{bis}}^K, M_{2\ell}(\mu_t^K) > \frac{1}{2^{2\ell} \sigma^{2\ell} K^{3\ell+1+\ell\tilde{\varepsilon}}} \right), \end{aligned}$$

which is zero for  $K$  large enough by (3.16). As previously, if we can prove for all  $T \geq 0$  that

$$\lim_{K \rightarrow +\infty} \mathbb{P} \left( \widehat{\tau}_{\text{bis}}^K \leq T \wedge \check{\tau}_{\text{bis}}^K \right) = \lim_{K \rightarrow +\infty} \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \check{\tau}_{\text{bis}}^K} M_2(\mu_t^K) \geq K^{\tilde{\varepsilon}} \right) = 0, \quad (3.17)$$

then, we can conclude as before.

**(c) Difficulties encountered.** The main difficulty consists in proving (3.17). For this we could seek for an extension of Lemma 3.4.2 to the framework of moments of order  $2\ell$ . Lemma 3.4.2 relies on the inequality for all  $t > s \geq 0$ ,  $K$  large enough,

$$\begin{aligned} & \frac{3}{4} M_6(\mu_{t \wedge \check{\tau}^K}) + 3M_4(\mu_{t \wedge \check{\tau}^K}^K) M_2(\mu_{t \wedge \check{\tau}^K}^K) + \frac{3}{2} M_2^3(\mu_{t \wedge \check{\tau}^K}^K) \\ & \leq \frac{3}{4} M_6(\mu_0^K) + 7M_4(\mu_0^K) M_2(\mu_0^K) + \frac{3}{2} M_2^3(\mu_0) \\ & \quad - \frac{C_1}{K^2 \sigma^2} \int_0^{t \wedge \check{\tau}^K} \left( \frac{3}{4} M_6(\mu_r^K) + 3M_4(\mu_r^K) M_2(\mu_r^K) + \frac{3}{2} M_2^3(\mu_r^K) - C_2 \right) dr + \text{Mart}_{t \wedge \check{\tau}^K} \end{aligned}$$

for some constants  $C_1, C_2 > 0$  and where  $(\text{Mart}_{t \wedge \check{\tau}^K})_{t \geq 0}$  is a martingale. However a similar calculation does not seem to be successful for moments of order 8 or more.

### 3.3 Infinitesimal generator approximation

Let us begin this section by giving the generator of the couple of process  $(z^K, \mu^K)$ . Then, we give an approximation of the previous formula for a class of test functions which is convergence determining. Finally, we give an extension to this result in the case of polynomials in  $\mu$ .



### 3.3.1 Generators and martingales

For all  $K \in \mathbb{N}^*$ , let us introduce the filtration  $(\mathcal{F}_t^K)_{t \geq 0}$  defined by  $\mathcal{F}_t^K := \sigma \left( z_s^K, \mu_s^K \mid s \leq t \right)$ .

**Proposition 3.3.1.** *The infinitesimal generator of the  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$ -valued MARKOV process  $(z^K, \mu^K)$ , defined for all bounded measurable function  $\Phi$  from  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$  to  $\mathbb{R}$ , is given by*

$$\begin{aligned} \mathcal{L}^K \Phi(z, \mu) &= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) \\ &\quad \times \left[ \Phi \left( z + \frac{\sigma\sqrt{K}}{K}(y-x), \tau_{-\frac{y-x}{K}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right) - \Phi(z, \mu) \right] \\ &+ \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \theta(\sigma\sqrt{K}x + z) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z, h) \\ &\quad \times \left[ \Phi \left( z + \frac{\sigma\sqrt{K}h}{K^{\frac{3}{2}}}, \tau_{-\frac{h}{K^{3/2}}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right] \right) - \Phi(z, \mu) \right] dh. \end{aligned}$$

Moreover, for all  $K \in \mathbb{N}^*$ , for all bounded measurable function  $\Phi$  from  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$  to  $\mathbb{R}$ , the following process  $(M_t^{K,\Phi})_{t \geq 0}$  defined by

$$M_t^{K,\Phi} := \Phi(z_t^K, \mu_t^K) - \Phi(z_0^K, \mu_0^K) - \int_0^t \mathcal{L}^K \Phi(z_s^K, \mu_s^K) ds \quad (3.18)$$

is a  $(\mathcal{F}_t^K)_{t \geq 0}$ -martingale, square integrable, with quadratic variation:

$$\begin{aligned} \langle M_t^{K,\Phi} \rangle_t &= \frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu_s^K(dx) \int_{\mathbb{R}} \mu_s^K(dy) b(\sigma\sqrt{K}x + z_s^K, \sigma\sqrt{K}y + z_s^K) \\ &\quad \times \left[ \Phi \left( z_s^K + \frac{\sigma\sqrt{K}}{K}(y-x), \tau_{-\frac{y-x}{K}} \# \left[ \mu_s^K - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right) - \Phi(z_s^K, \mu_s^K) \right]^2 \\ &+ \frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu_s^K(dx) \theta(\sigma\sqrt{K}x + z_s^K) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z_s^K, h) \\ &\quad \times \left[ \Phi \left( z_s^K + \frac{\sigma\sqrt{K}h}{K^{\frac{3}{2}}}, \tau_{-\frac{h}{K^{3/2}}} \# \left[ \mu_s^K - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right] \right) - \Phi(z_s^K, \mu_s^K) \right]^2 dh. \end{aligned}$$

Note that the factor  $1/K\sigma^2$  corresponds to the time scaling of  $\nu_t^K$  used to define  $z_t^K$  and  $\mu_t^K$ .

*Proof. Step 1. About the generators.* On the one hand, from (3.3), note that for all  $\Phi : \mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}$  and  $\phi : \mathcal{M}_{1,K}(\mathbb{R}) \rightarrow \mathbb{R}$  such that  $\phi(\nu) = \Phi \left( \langle \text{id}, \nu \rangle, \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-\langle \text{id}, \nu \rangle} \right) \# \nu \right)$ ,

$$\mathcal{L}^K \Phi(z, \mu) = \frac{1}{K\sigma^2} \times L^{K,\sigma} \phi \left( (\tau_z \circ h_{\sigma\sqrt{K}}) \# \mu \right). \quad (3.19)$$

On the other hand, note that for all  $\nu \in \mathcal{M}_{1,K}(\mathbb{R})$ ,  $u, v \in \mathbb{R}$ ,  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  and  $\Psi : \mathcal{M}_{1,K}(\mathbb{R}) \rightarrow \mathbb{R}$  measurable bounded functions, we have

$$\begin{aligned} &\int_{\mathbb{R}} \psi(x) \Psi \left( \nu + \frac{\delta_{x+\sigma\sqrt{K}u} - \delta_v}{K} \right) \nu(dx) \\ &= \int_{\mathbb{R}} \psi(\sigma\sqrt{K}y + \langle \text{id}, \nu \rangle) \Psi \left( (\tau_{\langle \text{id}, \nu \rangle} \circ h_{\sigma\sqrt{K}}) \# \left[ \mu + \frac{\delta_{y+u}}{K} - \frac{\delta_v}{K} \right] \right) \mu(dy). \end{aligned} \quad (3.20)$$

Finally, by noting that for all  $u \in \mathbb{R}$  the mean trait of  $(\tau_{\langle \text{id}, \nu \rangle} \circ h_{\sigma\sqrt{K}}) \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_u}{K} \right]$  is  $\langle \text{id}, \nu \rangle + \frac{\sigma\sqrt{K}}{K}(u-x)$  and the centered and dilated distribution of traits of  $(\tau_{\langle \text{id}, \nu \rangle} \circ h_{\sigma\sqrt{K}}) \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_u}{K} \right]$  is  $\tau_{-\frac{u-x}{K}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_u}{K} \right]$ , the announced result follows from (3.19) and (3.20).

**Step 2. About the martingales.** The martingale property follows from classical arguments since  $\Phi$  and  $\mathcal{L}^K \Phi$  are bounded [94, Chapter 4]. To compute the bracket, we proceed according to the following classical method (see e.g. [102]). We apply (3.18) replacing  $\Phi$  by  $\Phi^2$  to obtain the martingale  $M^{K, \Phi^2}$ . Then, we apply the ITÔ formula to compute  $\Phi^2(z_t^K, \mu_t^K)$  from (3.18). We deduce that

$$\text{Mart}_t := \Phi^2(z_t^K, \mu_t^K) - \Phi^2(z_0^K, \mu_0^K) - 2 \int_0^t \Phi(z_s^K, \mu_s^K) \mathcal{L}^K \Phi(z_s^K, \mu_s^K) ds - \langle M^{K, \Phi} \rangle_t$$

is a martingale. Hence, the martingale  $M_t^{K, \Phi^2} - \text{Mart}_t$  has finite variation, so it is null [150, Theorem 4.1], leading to

$$\langle M^{K, \Phi} \rangle_t = \int_0^t \left[ \mathcal{L}^K \Phi^2(z_s^K, \mu_s^K) - 2\Phi(z_s^K, \mu_s^K) \mathcal{L}^K \Phi(z_s^K, \mu_s^K) \right] ds$$

then the announced conclusion.  $\square$

### 3.3.2 Asymptotic expansions

Recall that, a set  $S \subset \mathcal{C}_b^0(\mathbb{R}, \mathbb{R})$  is called  $\mathcal{M}_{1,K}^c(\mathbb{R})$ -convergence determining if whenever  $(P_n)_{n \in \mathbb{N}} \in \mathcal{M}_{1,K}^c(\mathbb{R})^{\mathbb{N}}$ ,  $P \in \mathcal{M}_{1,K}^c(\mathbb{R})$  and  $\lim_{n \rightarrow +\infty} \langle f, P_n \rangle = \langle f, P \rangle$  for all  $f \in S$ , we have that  $(P_n)_{n \in \mathbb{N}}$  converges weakly to  $P$  [94, Chapter 3, Section 4, p.112]. As developed in [60, Theorem 3.2.6], the class of functions on  $\mathcal{M}_{1,K}^c(\mathbb{R})$ ,

$$\mathbb{F}_b^2 := \left\{ F_\varphi \mid F_\varphi(\mu) := F(\langle \varphi, \mu \rangle), F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R}), \varphi \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R}) \right\}$$

is  $\mathcal{M}_{1,K}^c(\mathbb{R})$ -convergence determining.

In the following sections, we prove the assertions (1) and (2) of Section 3.2.3.1.

#### 3.3.2.1 Slow component

**Proposition 3.3.2.** *For all  $f \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ , understood as a function of  $(z, \mu)$  which depends only on  $z$ , the infinitesimal generator  $\mathcal{L}^K$  defined in Proposition 3.3.1, satisfies the following decomposition*

$$\mathcal{L}^K f(z, \mu) = \mathcal{L}_{\text{SLOW}} f(z, \mu) + O\left(\frac{1}{K^2} + \frac{M_2(\mu)}{K} + \sigma\sqrt{K} M_3(\mu)\right)$$

where the operator  $\mathcal{L}_{\text{SLOW}}$  is given by (3.9).

*Proof.* From Proposition 3.3.1 for the choice of test functions  $\Phi(z, \mu) := f(z)$  we obtain that  $\mathcal{L}^K f(z, \mu) = (\mathbf{A})^K + (\mathbf{B})^K$  where

$$\begin{aligned} (\mathbf{A})^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) \left[ f\left(z + \frac{\sigma\sqrt{K}}{K}(y-x)\right) - f(z) \right], \\ (\mathbf{B})^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \theta(\sigma\sqrt{K}x + z) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z, h) \left[ f\left(z + \frac{\sigma\sqrt{K}}{K^{\frac{3}{2}}}h\right) - f(z) \right] dh. \end{aligned}$$

Now, we want to decompose and study  $(\mathbf{A})^K$ . Denoting  $(\mathbf{C})_{x,y}^K(z) := f\left(z + \frac{\sigma\sqrt{K}}{K}(y-x)\right) - f(z)$ , from TAYLOR's formula, note that  $(\mathbf{A})^K = (\mathbf{A})_1^K + (\mathbf{A})_2^K + (\mathbf{A})_3^K$  where

$$\begin{aligned} (\mathbf{A})_1^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) (\mathbf{C})_{x,y}^K(z), \\ (\mathbf{A})_2^K &:= \frac{K}{K\sigma^2} \times \sigma\sqrt{K} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) (x\partial_1 b(z, z) + y\partial_2 b(z, z)) (\mathbf{C})_{x,y}^K(z), \\ (\mathbf{A})_3^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) \left[ b\left(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z\right) - b(z, z), \right. \\ &\quad \left. - \sigma\sqrt{K} (x\partial_1 b(z, z) + y\partial_2 b(z, z)) \right] (\mathbf{C})_{x,y}^K(z). \end{aligned}$$

Using that  $(\mathbf{C})_{x,y}^K(z) = f'(z) \frac{\sigma\sqrt{K}}{K}(x-y) + O\left(\frac{\sigma^2}{K} |y-x|^2\right)$ , we deduce that

$$\begin{aligned} (\mathbf{A})_1^K &= O\left(\frac{M_2(\mu)}{K}\right), \\ (\mathbf{A})_2^K &= f'(z) (\partial_2 b(z, z) - \partial_1 b(z, z)) M_2(\mu) + O\left(\frac{\sigma\sqrt{K} M_3(\mu)}{K}\right), \end{aligned}$$

and noting that  $|\mathbf{C}_{x,y}^K| \leq C_1 \frac{\sigma\sqrt{K}}{K} |y-x|$  for some constant  $C_1 > 0$  and

$$\left| b\left(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z\right) - b(z, z) - \sigma\sqrt{K} (x\partial_1 b(z, z) + y\partial_2 b(z, z)) \right| \leq C_2 \sigma^2 K (x^2 + y^2)$$

for some constant  $C_2 > 0$ , we deduce that  $(\mathbf{A})_3^K = O\left(\sigma\sqrt{K} M_3(\mu)\right)$ . Therefore,

$$(\mathbf{A})^K = f'(z) (\partial_2 b(z, z) - \partial_1 b(z, z)) M_2(\mu) + O\left(\frac{M_2(\mu)}{K} + \sigma\sqrt{K} M_3(\mu)\right).$$

As previously, we obtain that

$$(\mathbf{B})^K = O\left(\frac{\bar{\theta} m_2}{K^2}\right).$$

and the announced result follows.  $\square$

### 3.3.2.2 Fast component

**Proposition 3.3.3.** *For all  $F \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$  and  $\varphi \in \mathcal{C}_b^3(\mathbb{R}, \mathbb{R})$ , the infinitesimal generator  $\mathcal{L}^K$  defined in Proposition 3.3.1, satisfies the following decomposition*

$$\mathcal{L}^K F_\varphi(z, \mu) = \frac{\theta(z)m_2(z)}{K^2\sigma^2} \left[ \mathcal{L}_{\text{FVc}}^{\lambda(z)} F_\varphi(z, \mu) + O\left(\frac{1}{\sqrt{K}} + \sigma K^{\frac{3}{2}} M_2(\mu) + \frac{M_3(\mu)}{K}\right) \right]$$

where the operator  $\mathcal{L}_{\text{FVc}}^{\lambda(z)}$  is given by (3.11) and  $F_\varphi$  is understood as a function of  $(z, \mu)$  which depends only on  $\mu$ .

*Proof.* Noting that

$$\begin{aligned} (\mathbf{D})_{x,y}^K &:= \left\langle \varphi, \tau_{-\frac{y-x}{K}} \sharp \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right\rangle - \langle \varphi, \mu \rangle \\ &= \left\langle \varphi \circ \tau_{-\frac{y-x}{K}}, \mu \right\rangle + \frac{1}{K} \left\{ \varphi\left(y - \frac{y-x}{K}\right) - \varphi\left(x - \frac{y-x}{K}\right) \right\} - \langle \varphi, \mu \rangle, \end{aligned}$$

setting  $(\mathbf{C})_{x,y}^K := F(\langle \varphi, \mu \rangle + (\mathbf{D})_{x,y}^K) - F(\langle \varphi, \mu \rangle)$  and from Proposition 3.3.1 we obtain that  $\mathcal{L}^K F_\varphi(z, \mu) = (\mathbf{A})^K + (\mathbf{B})^K$  where

$$\begin{aligned} (\mathbf{A})^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) (\mathbf{C})_{x,y}^K, \\ (\mathbf{B})^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \theta(\sigma\sqrt{K}x + z) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z, h) (\mathbf{C})_{x,x+h/\sqrt{K}}^K dh. \end{aligned}$$

**Step 1. Decomposition and study of  $(\mathbf{A})^K$ .** From TAYLOR's formula, we obtain that  $(\mathbf{A})^K = (\mathbf{A})_1^K + (\mathbf{A})_2^K$  where

$$\begin{aligned} (\mathbf{A})_1^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) (\mathbf{C})_{x,y}^K, \\ (\mathbf{A})_2^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) [b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) - b(z, z)] (\mathbf{C})_{x,y}^K. \end{aligned}$$

Denoting  $(\mathbf{E})_{x,y}^K := \frac{1}{K}(\varphi(y) - \varphi(x) - (y-x)\langle \varphi', \mu \rangle)$  and from TAYLOR's formula again, we obtain that  $(\mathbf{A})_1^K = (\mathbf{A})_{11}^K + (\mathbf{A})_{12}^K + (\mathbf{A})_{13}^K + (\mathbf{A})_{14}^K$  where

$$\begin{aligned} (\mathbf{A})_{11}^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) F'(\langle \varphi, \mu \rangle) (\mathbf{D})_{x,y}^K, \\ (\mathbf{A})_{12}^K &:= \frac{K}{2K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) F''(\langle \varphi, \mu \rangle) [(\mathbf{E})_{x,y}^K]^2, \\ (\mathbf{A})_{13}^K &:= \frac{K}{2K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) F''(\langle \varphi, \mu \rangle) [[(\mathbf{D})_{x,y}^K]^2 - [(\mathbf{E})_{x,y}^K]^2], \\ (\mathbf{A})_{14}^K &:= \frac{K}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(z, z) \left( (\mathbf{C})_{x,y}^K - F'(\langle \varphi, \mu \rangle) (\mathbf{D})_{x,y}^K - \frac{F''(\langle \varphi, \mu \rangle)}{2} [(\mathbf{D})_{x,y}^K]^2 \right). \end{aligned}$$

From TAYLOR's formula again and the centered condition of  $\mu$ , we obtain that

$$\begin{aligned} (\mathbf{A})_{11}^K &= \frac{Kb(z, z)}{K\sigma^2} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) F'(\langle \varphi, \mu \rangle) \left[ (\mathbf{E})_{x,y}^K - \frac{1}{K^2}(y-x)(\varphi'(y) - \varphi'(x)) \right. \\ &\quad \left. + \frac{(y-x)^2 \langle \varphi'', \mu \rangle}{2K^2} + O\left(\frac{|y-x|^2 + |y-x|^3}{K^3}\right) \right] \\ &= \frac{b(z, z)}{K^2\sigma^2} F'(\langle \varphi, \mu \rangle) [-2\langle \varphi' \times \text{id}, \mu \rangle + \langle \varphi'', \mu \rangle M_2(\mu)] + \frac{1}{K^2\sigma^2} \times O\left(\frac{M_2(\mu) + M_3(\mu)}{K}\right). \end{aligned}$$

In a straightforward way, we obtain that

$$(\mathbf{A})_{12}^K = \frac{b(z, z)}{K^2\sigma^2} F''(\langle \varphi, \mu \rangle) [\langle \varphi^2, \mu \rangle - \langle \varphi, \mu \rangle^2 + M_2(\mu) \langle \varphi', \mu \rangle^2 - 2\langle \varphi \times \text{id}, \mu \rangle \langle \varphi', \mu \rangle].$$

As  $\varphi$  is bounded and so  $F$  too, we deduce that there exists two constants  $C_1, C_2 > 0$  such that

$$\begin{aligned} |[(\mathbf{D})_{x,y}^K]^2 - [(\mathbf{E})_{x,y}^K]^2| &= |(\mathbf{D})_{x,y}^K - (\mathbf{E})_{x,y}^K| |(\mathbf{D})_{x,y}^K + (\mathbf{E})_{x,y}^K| \leq C_1 \frac{|y-x| + (y-x)^2}{K^2} \left( \frac{1}{K} + \frac{|y-x|}{K} \right), \\ \left| (\mathbf{C})_{x,y}^K - F'(\langle \varphi, \mu \rangle) (\mathbf{D})_{x,y}^K - \frac{F''(\langle \varphi, \mu \rangle)}{2} [(\mathbf{D})_{x,y}^K]^2 \right| &\leq C_2 \frac{1 + |y-x|^3}{K^3}. \end{aligned}$$

Hence,

$$(\mathbf{A})_{13}^K = \frac{1}{K^2\sigma^2} \times O\left(\frac{1 + M_3(\mu)}{K}\right) \quad \text{and} \quad (\mathbf{A})_{14}^K = \frac{1}{K^2\sigma^2} \times O\left(\frac{1 + M_3(\mu)}{K}\right).$$

Finally, as  $\left| b\left(\sigma\sqrt{K}x+z, \sigma\sqrt{K}y+z\right) - b(z, z) \right| \leq C_3\sigma\sqrt{K}(|x|+|y|)$  and  $\left| (\mathbf{C})_{x,y}^K \right| \leq C_4\frac{1+|y-x|}{K}$  for some constants  $C_3, C_4 > 0$ , we deduce that

$$(\mathbf{A})_2^K = \frac{1}{K^2\sigma^2} \times O\left(\sigma K^{\frac{3}{2}} [M_1(\mu) + M_2(\mu)]\right).$$

Therefore, using HÖLDER's inequalities to bound  $M_1(\mu) \leq \sqrt{M_2(\mu)} \leq 1 + M_2(\mu)$  and  $M_2(\mu) \leq M_3^{2/3}(\mu) \leq 1 + M_3(\mu)$  we have that

$$\begin{aligned} (\mathbf{A})^K &= \frac{b(z, z)}{K^2\sigma^2} F'(\langle \varphi, \mu \rangle) \left[ -2\langle \varphi' \times \text{id}, \mu \rangle + \langle \varphi'', \mu \rangle M_2(\mu) \right] \\ &\quad + \frac{b(z, z)}{K^2\sigma^2} F''(\langle \varphi, \mu \rangle) \left[ \langle \varphi^2, \mu \rangle - \langle \varphi, \mu \rangle^2 + M_2(\mu) \langle \varphi', \mu \rangle^2 - 2\langle \varphi \times \text{id}, \mu \rangle \langle \varphi', \mu \rangle \right] \\ &\quad + \frac{1}{K^2\sigma^2} \times O\left(\frac{1}{K} + \sigma K^{\frac{3}{2}} [1 + M_2(\mu)] + \frac{M_3(\mu)}{K}\right). \end{aligned}$$

**Step 2. Conclusion.** In similar way to Step 1, we obtain that

$$(\mathbf{B})^K = \frac{\theta(z)m_2(z)}{K^2\sigma^2} F'(\langle \varphi, \mu \rangle) \left\langle \frac{\varphi''}{2}, \mu \right\rangle + \frac{1}{K^2\sigma^2} \times O\left(\frac{1}{\sqrt{K}} + \sigma K M_1(\mu)\right).$$

Since, by Assumption (3.5),  $\sigma K^{\frac{3}{2}} \ll \frac{1}{\sqrt{K}}$ , the announced result follows.  $\square$

### 3.3.3 Generators and martingales in the case of polynomials in $\mu$

#### 3.3.3.1 For bounded test functions

In this section, we extend in Lemma 3.3.4 the result of Proposition 3.3.1 to test functions of the form

$$P_{f,n}(\mu) := \langle f, \mu^n \rangle := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu(dx_1) \cdots \mu(dx_n) \quad (3.21)$$

with  $n \in \mathbb{N}^*$ ,  $\mu \in \mathcal{M}_{1,K}^c(\mathbb{R})$ ,  $f \in \mathcal{C}_b^3(\mathbb{R}^n, \mathbb{R})$  and where  $\mu^n$  is the  $n$ -fold product measure of  $\mu$ . In Lemmas 3.3.6 and 3.3.7 we extend this result to the case of specific unbounded test functions.

For all  $n \in \mathbb{N}^*$ , we denote by  $\mathbf{1} \in \mathbb{R}^n$ , the vector whose coordinates are all 1 and by  $\Delta$  the Laplacian operator on  $\mathbb{R}^n$ . Let us introduce, for all  $n \in \mathbb{N}^*$ ,  $\lambda > 0$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ , the operator  $B_\lambda^{(n)}$  defined by

$$B_\lambda^{(n)} f(x) := \frac{1}{2} \Delta f(x) - 2\lambda (\nabla f(x) \cdot \mathbf{1})(x \cdot \mathbf{1}), \quad x \in \mathbb{R}^n. \quad (3.22)$$

Let us consider for all  $n \in \mathbb{N}^*$  for all  $i, j \in \{1, \dots, n\}$ ,

- $\Phi_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^2(\mathbb{R}^{n-1}, \mathbb{R})$ , with  $i \neq j$ , is the function obtained from  $f$  by inserting the variable  $x_i$  between  $x_{j-1}$  and  $x_j$  when  $i < j$  and by inserting the variable  $x_{i-1}$  between  $x_{j-1}$  and  $x_j$  when  $i > j$ :

$$\begin{aligned} \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_i, x_j, x_{j+1}, \dots, x_{n-1}) & i < j \\ \Phi_{i,j} f(x_1, \dots, x_{n-1}) &= f(x_1, \dots, x_{j-1}, x_{i-1}, x_j, x_{j+1}, \dots, x_{n-1}) & i > j \end{aligned} \quad (3.23)$$

- $K_{i,j} : \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_b^2(\mathbb{R}^{n+1}, \mathbb{R})$  is defined as

$$K_{i,j} f(x_1, \dots, x_n, x_{n+1}) := \partial_{ij}^2 f(x_1, \dots, x_n) x_{n+1}^2. \quad (3.24)$$

We recall from Definition 2.2.8 of Chapter 2 that the extended generator (in the sense of DYNKIN, see (3.26) below)  $\mathcal{L}_{\text{FVc}}^\lambda$  of the centered FLEMING-VIOT process with resampling rate  $\lambda$  is defined for any  $n \in \mathbb{N}^*$ , for any test functions  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  by

$$\mathcal{L}_{\text{FVc}}^\lambda P_{f,n}(\mu) = \langle B_\lambda^{(n)} f, \mu^n \rangle + \lambda \sum_{\substack{i,j=1 \\ i \neq j}}^n [\langle \Phi_{i,j} f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle] + \lambda \sum_{i,j=1}^n \langle K_{i,j} f, \mu^{n+1} \rangle. \quad (3.25)$$

From Definition 2.2.8 and denoting

$$\tilde{\Omega} := \left\{ X \in \mathcal{C}^0([0, +\infty), \mathcal{M}_1^{c,2}(\mathbb{R})) \mid \forall T > 0, \sup_{0 \leq t \leq T} M_2(X_t) < \infty \right\},$$

we recall that a probability measure  $\mathbb{P}_\mu$  on  $\tilde{\Omega}$  is said to solve the *centered FLEMING-VIOT martingale problem for polynomials* with initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , if the canonical process  $(X_t)_{t \geq 0}$  on  $\tilde{\Omega}$  satisfies  $\mathbb{P}_\mu(X_0 = \mu) = 1$  and, for all  $n \in \mathbb{N}^*$  and  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\widehat{M}_t^{P_{f,n}} := P_{f,n}(X_t) - P_{f,n}(X_0) - \int_0^t \mathcal{L}_{\text{FVc}}^\lambda P_{f,n}(X_s) ds \quad (3.26)$$

is a  $\mathbb{P}_\mu$ -martingale.

**Lemma 3.3.4.** *The infinitesimal generator  $\mathcal{L}^K$  of the  $\mathbb{R} \times \mathcal{M}_{1,K}^c(\mathbb{R})$ -valued MARKOV process  $(z^K, \mu^K)$  given by Proposition 3.3.1, satisfies for all  $n \in \mathbb{N}^*$  and  $f \in \mathcal{C}_b^3(\mathbb{R}^n, \mathbb{R})$ , the following relations:*

$$\mathcal{L}^K P_{f,n}(z, \mu) = \frac{\theta(z)m_2(z)}{K^2\sigma^2} \mathcal{L}_{\text{FVc}}^{\lambda(z)} P_{f,n}(\mu) + \frac{1}{K^2\sigma^2} \times O\left(\frac{1}{\sqrt{K}} + \sigma K^{\frac{3}{2}} M_2(\mu) + \frac{M_3(\mu)}{K}\right)$$

where  $\lambda(z) := \frac{b(z,z)}{\theta(z)m_2(z)}$ . Moreover, for all  $K, n \in \mathbb{N}^*$  and test functions  $f \in \mathcal{C}_b^3(\mathbb{R}^n, \mathbb{R})$ , the process  $(M_t^{K,P_{f,n}})_{t \geq 0}$  defined by

$$M_t^{K,P_{f,n}} := P_{f,n}(\mu_t^K) - P_{f,n}(\mu_0^K) - \int_0^t \mathcal{L}^K P_{f,n}(z_s^K, \mu_s^K) ds$$

is a square integrable martingale started at 0.

The proof of this result is given in Section 3.3.4.

### 3.3.3.2 For some unbounded test functions

The following results are particular extensions of Lemma 3.3.4 when the test functions are no longer bounded but the processes are stopped at time  $\check{\tau}^K$ . Then, we give in Lemmas 3.3.6 and 3.3.7 DOOB's semi-martingale decomposition of moments of order 2 and those whose degree is 6.

**Proposition 3.3.5.** *For all  $n \in \mathbb{N}^*$ , for all  $f \in \mathcal{C}^3(\mathbb{R}^n, \mathbb{R})$ , the process  $(M_{t \wedge \check{\tau}^K}^{K,P_{f,n}})_{t \geq 0}$  defined*

by

$$\begin{aligned}
 M_{t \wedge \check{\tau}^K}^{K, P_{f,n}} &:= P_{f,n}(\mu_{t \wedge \check{\tau}^K}^K) - P_{f,n}(\mu_0^K) - \frac{K}{K\sigma^2} \int_0^{t \wedge \check{\tau}^K} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z_s^K, \sigma\sqrt{K}x + z_s^K) \\
 &\quad \times \left[ P_{f \circ \tau_{-\frac{y-x}{K}}, n} \left( \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right) - P_{f,n}(\mu) \right] ds \\
 &\quad + \frac{K}{K\sigma^2} \int_0^{t \wedge \check{\tau}^K} \int_{\mathbb{R}} \mu(dx) \theta(\sigma\sqrt{K}x + z_s^K) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z_s^K, h) dh \\
 &\quad \times \left[ P_{f \circ \tau_{-\frac{h}{K^{\frac{3}{2}}}, n} \left( \mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right) - P_{f,n}(\mu) \right] ds
 \end{aligned}$$

is a bounded martingale.

*Proof.* Thanks to the stopping time  $\check{\tau}^K$ ,  $(M_{t \wedge \check{\tau}^K}^{K, P_{f,n}})_{t \geq 0}$  is bounded. Hence the martingality of this process follows from Lemma 3.3.4.  $\square$

(a) **Moment of order 2.** The following result will be useful in Section 3.4.

**Lemma 3.3.6.** *For all  $K \in \mathbb{N}^*$ , the process  $(M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}})_{t \geq 0}$  defined in Proposition 3.3.5 satisfies for all  $t \geq 0$*

$$\begin{aligned}
 M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} &= M_2(\mu_{t \wedge \check{\tau}^K}^K) - M_2(\mu_0^K) + \frac{1}{K^2\sigma^2} \int_0^{t \wedge \check{\tau}^K} \{2b(z_s^K, z_s^K) M_2(\mu_s^K) - \theta(z_s^K) m_2(z_s^K)\} ds \\
 &\quad + \frac{1}{K^2\sigma^2} \int_0^{t \wedge \check{\tau}^K} O\left(\frac{1}{K} + \sigma\sqrt{K}M_1(\mu_s^K) + \sigma K^{\frac{3}{2}}M_3(\mu_s^K)\right) ds
 \end{aligned}$$

is a square integrable martingale with quadratic variation:

$$\begin{aligned}
 \langle M^{K, P_{\text{id}^2, 1}} \rangle_{t \wedge \check{\tau}^K} &= \frac{2}{K^2\sigma^2} \int_0^{t \wedge \check{\tau}^K} b(z_s^K, z_s^K) \{M_4(\mu_s^K) - M_2^2(\mu_s^K)\} ds \\
 &\quad + \frac{1}{K^2\sigma^2} \int_0^{t \wedge \check{\tau}^K} O\left(\frac{\sigma\sqrt{K}}{K^2} + \frac{M_2(\mu_s^K)}{K} + \frac{\sigma\sqrt{K}M_3(\mu_s^K)}{K} + \sigma\sqrt{K}M_5(\mu_s^K)\right) ds.
 \end{aligned} \tag{3.27}$$

*Proof. Step 1. Approximation of the Doob decomposition.* Note that for all  $x, y \in \mathbb{R}$ ,

$$P_{\text{id}^2 \circ \tau_{-\frac{y-x}{K}, 1} \left( \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right) - P_{\text{id}^2, 1}(\mu) = \frac{1}{K} (y^2 - x^2) - \frac{1}{K^2} (y - x)^2.$$

From TAYLOR's formula, it follows that

$$\begin{aligned}
 &\frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z_s^K, \sigma\sqrt{K}y + z_s^K) \\
 &\quad \times \left[ P_{\text{id}^2 \circ \tau_{-\frac{y-x}{K}, 1} \left( \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right) - P_{\text{id}^2, 1}(\mu) \right] \\
 &= \frac{1}{K^2\sigma^2} \int_0^t [-2b(z_s^K, z_s^K) M_2(\mu_s^K) + O(\sigma K^{\frac{3}{2}} M_3(\mu_s^K))] ds
 \end{aligned}$$

and

$$\begin{aligned} & \frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z_s^K, h) \theta(\sigma\sqrt{K}x + z_s^K) \\ & \quad \times \left[ P_{\text{id}^2 \circ \tau_{-\frac{h}{K^{3/2}}}, 1} \left( \mu - \frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right) - P_{\text{id}^2, 1}(\mu) \right] dh \\ & = \frac{1}{K^2\sigma^2} \int_0^t \left[ \theta(z_s^K) m_2(z_s^K) + O\left(\frac{1}{K} + \sigma\sqrt{K}M_1(\mu_s^K)\right) \right] ds. \end{aligned}$$

We deduce the first announced result.

**Step 2. Quadratic variation.** In similar way to the proof of Proposition 3.3.1, we have

$$\begin{aligned} \langle M^{K, P_{\text{id}^2, 1}} \rangle_t & = \frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu_s^K(dx) \int_{\mathbb{R}} \mu_s^K(dy) b(\sigma\sqrt{K}x + z_s^K, \sigma\sqrt{K}y + z_s^K) \\ & \quad \times \left[ \frac{1}{K} (y^2 - x^2) - \frac{1}{K^2} (y - x)^2 \right]^2 \\ & \quad + \frac{K}{K\sigma^2} \int_0^t ds \int_{\mathbb{R}} \mu_s^K(dx) \int_{\mathbb{R}} m(\sigma\sqrt{K}x + z_s^K, h) \theta(\sigma\sqrt{K}x + z_s^K) \left[ \frac{h^2}{K^2} - \frac{h^2}{K^3} + \frac{2xh}{K^{\frac{3}{2}}} \right]^2 dh. \end{aligned}$$

The announced result follows from TAYLOR's formula and straightforward computations.  $\square$

**(b) Moments of degree 6.** Let us denote for all  $\ell \in \mathbb{N}$ ,  $\widetilde{M}_\ell(\mu) := \langle \text{id}^\ell, \mu \rangle$ .

**Lemma 3.3.7.** *Let us consider the functions  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $h : \mathbb{R}^3 \rightarrow \mathbb{R}$  respectively defined by  $f(u, v) := u^4 v^2$ ,  $g(u, v) := u^3 v^3$  and  $h(u, v, w) := u^2 v^2 w^2$ . For all  $K \in \mathbb{N}^*$ , the processes  $(M_{t \wedge \tau^K}^{K, P_{\text{id}^6, 1}})_{t \geq 0}$ ,  $(M_{t \wedge \tau^K}^{K, P_{f, 2}})_{t \geq 0}$ ,  $(M_{t \wedge \tau^K}^{K, P_{g, 2}})_{t \geq 0}$  and  $(M_{t \wedge \tau^K}^{K, P_{h, 3}})_{t \geq 0}$  defined in Proposition 3.3.5 satisfy for all  $t \geq 0$*

- $M_{t \wedge \tau^K}^{K, P_{\text{id}^6, 1}} = M_6(\mu_{t \wedge \tau^K}^K) - M_6(\mu_0^K) + \frac{3}{K^2\sigma^2} \int_0^{t \wedge \tau^K} \left\{ b(z_s^K, z_s^K) [4M_6(\mu_s^K) - 10M_4(\mu_s^K)M_2(\mu_s^K)] - 5\theta(z_s^K) m_2(z_s^K) M_4(\mu_s^K) \right\} ds$   
 $+ \frac{1}{K^2\sigma^2} \int_0^{t \wedge \tau^K} O\left(\frac{M_3(\mu_s^K)}{\sqrt{K}} + \frac{1 + M_4(\mu_s^K)}{K} + \sigma\sqrt{K}M_5(\mu_s^K) + \frac{M_6(\mu_s^K)}{K^2} + \sigma K^{\frac{3}{2}}M_7(\mu_s^K)\right) ds,$
- $M_{t \wedge \tau^K}^{K, P_{f, 2}} = M_4(\mu_{t \wedge \tau^K}^K) M_2(\mu_{t \wedge \tau^K}^K) - M_4(\mu_0^K) M_2(\mu_0^K)$   
 $+ \frac{1}{K^2\sigma^2} \int_0^{t \wedge \tau^K} \left\{ 2b(z_s^K, z_s^K) [6M_4(\mu_s^K) M_2(\mu_s^K) - M_6(\mu_s^K) + 4\widetilde{M}_3^2(\mu_s^K) - 6M_2^3(\mu_s^K)] \right.$   
 $\left. - \theta(z_s^K) m_2(z_s^K) [6M_2^2(\mu_s^K) + M_4(\mu_s^K)] \right\} ds$   
 $+ \frac{1}{K^2\sigma^2} \int_0^{t \wedge \tau^K} O\left(\frac{M_6(\mu_s^K)}{K^3} + \sigma K^{\frac{3}{2}}M_7(\mu_s^K)\right) ds,$
- $M_{t \wedge \tau^K}^{K, P_{g, 2}} = \widetilde{M}_3^2(\mu_{t \wedge \tau^K}^K) - \widetilde{M}_3^2(\mu_0^K) + \frac{1}{K^2\sigma^2} \int_0^{t \wedge \tau^K} \left\{ b(z_s^K, z_s^K) [-2M_6(\mu_s^K) + 14\widetilde{M}_3^2(\mu_s^K) \right.$   
 $\left. + 12M_4(\mu_s^K) M_2(\mu_s^K) - 18M_2^3(\mu_s^K)] \right\} ds$   
 $+ \frac{1}{K^2\sigma^2} \int_0^{t \wedge \tau^K} O\left(\frac{M_6(\mu_s^K)}{K^3} + \sigma K^{\frac{3}{2}}M_7(\mu_s^K)\right) ds,$



$$\begin{aligned}
 \bullet M_{t \wedge \tau^K}^{K, P_{h,3}} &= M_2^3(\mu_{t \wedge \tau^K}^K) - M_2^3(\mu_0^K) + \frac{1}{K^2 \sigma^2} \int_0^{t \wedge \tau^K} \left\{ b(z_s^K, z_s^K) \left[ -6M_4(\mu_s^K) M_2(\mu_s^K) \right. \right. \\
 &\quad \left. \left. + 12M_2^3(\mu_s^K) \right] - 3\theta(z_s^K) m_2(z_s^K) M_4(\mu_s^K) \right\} ds \\
 &\quad + \frac{1}{K^2 \sigma^2} \int_0^{t \wedge \tau^K} O\left(\frac{M_6(\mu_s^K)}{K^3} + \sigma K^{\frac{3}{2}} M_7(\mu_s^K)\right) ds.
 \end{aligned}$$

*Proof.* The approximation proposed is obtained in a similar way to the proof of Lemma 3.3.6  $\square$

### 3.3.4 Proof of Lemma 3.3.4

From Proposition 3.3.1, for the choice of test functions  $\Phi(z, \mu) := P_{f,n}(\mu)$  given by (3.21) and noting that for all  $x, y \in \mathbb{R}$ ,

$$\left\langle f, \left( \tau_{-\frac{y-x}{K}} \# \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right)^n \right\rangle = \left\langle f \circ \tau_{-\frac{y-x}{K}}, \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^n \right\rangle,$$

we obtain that  $\mathcal{L}^K P_{f,n}(z, \mu) = \frac{K}{K\sigma^2} ((\mathbf{A})^K + (\mathbf{B})^K)$  where

$$\begin{aligned}
 (\mathbf{A})^K &:= \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) (\mathbf{C})_{x,y}^K, \\
 (\mathbf{B})^K &:= \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta(\sigma\sqrt{K}x + z) m(z + \sigma\sqrt{K}x, h) (\mathbf{C})_{x, x+h/\sqrt{K}}^K dh
 \end{aligned}$$

where  $(\mathbf{C})_{x,y}^K := \left\langle f \circ \tau_{-\frac{y-x}{K}}, \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^n \right\rangle - \langle f, \mu^n \rangle$ . Note that, from TAYLOR's formula, we obtain that

$$\begin{aligned}
 (\mathbf{C})_{x,y}^K &= \sum_{k=1}^n \binom{n}{k} \left\langle f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k \right\rangle - \frac{y-x}{K} \sum_{i=1}^n \left\langle \partial_i f, \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^n \right\rangle \\
 &\quad + \frac{(y-x)^2}{2K^2} \sum_{i=1}^n \sum_{j=1}^n \left\langle \partial_{ij}^2 f, \left[ \mu - \frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^n \right\rangle + O\left(\frac{|y-x|^3}{K^3}\right).
 \end{aligned}$$

By abuse of notation, we do not indicate the orders of the products of measure  $\mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k$ .

**Step 1. Decomposition and study of  $(\mathbf{A})^K$ .** Note that  $(\mathbf{A})^K = \sum_{i=1}^5 (\mathbf{A})_i^K$  where

$$\begin{aligned}
 (\mathbf{A})_1^K &:= b(z, z) \sum_{k=1}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) \left\langle f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k \right\rangle, \\
 (\mathbf{A})_2^K &:= -\frac{b(z, z)}{K} \sum_{i=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) (y-x) \left\langle \partial_i f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k \right\rangle, \\
 (\mathbf{A})_3^K &:= \frac{b(z, z)}{2K^2} \sum_{i,j=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) (y-x)^2 \left\langle \partial_{ij}^2 f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k \right\rangle, \\
 (\mathbf{A})_4^K &:= O\left(\frac{M_3(\mu)}{K^3}\right), \\
 (\mathbf{A})_5^K &:= \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) \left[ b(\sigma\sqrt{K}x + z, \sigma\sqrt{K}y + z) - b(z, z) \right] (\mathbf{C})_{x,y}^K.
 \end{aligned}$$

We first study precisely the  $(\mathbf{A})_2^K$  term by explaining the case  $k = 0$ ,  $k = 1$  and  $k \geq 2$ . We will just give the result for the other  $(\mathbf{A})_i^K$ ,  $i \in \{1, 3, 5\}$ : the approach remains the same. By the centered assumption of  $\mu$ , note that for all  $i \in \{1, \dots, n\}$ ,  $\int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy)(y-x) \langle \partial_i f, \mu^n \rangle = 0$ . Then, using again the centered assumption of  $\mu$ , we obtain that

$$\begin{aligned}
 & \binom{n}{1} \sum_{i=1}^n \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy)(y-x) \left\langle \partial_i f, \mu^{n-1} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right] \right\rangle \\
 &= \frac{1}{K} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy) \\
 & \quad \times \left[ x \partial_i f(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n) - x \partial_i f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n) \right. \\
 & \quad \left. - y \partial_i f(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n) + y \partial_i f(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n) \right] \prod_{\substack{\ell=1 \\ \ell \neq j}}^n \mu(dx_\ell) \\
 &= \frac{2}{K} \sum_{i=1}^n \sum_{j=1}^n \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \mu(dx) x \partial_i f(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n) \prod_{\substack{\ell=1 \\ \ell \neq j}}^n \mu(dx_\ell) \\
 &= \frac{2}{K} \langle (\nabla f \cdot \mathbf{1})(\bullet \cdot \mathbf{1}), \mu^n \rangle.
 \end{aligned}$$

As  $\nabla f$  is bounded, we deduce that for all  $k \geq 2$ , for all  $i \in \{1, \dots, n\}$ ,

$$\int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \mu(dy)(y-x) \left\langle \partial_i f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_y}{K} \right]^k \right\rangle = O\left(\frac{M_1(\mu)}{K^2}\right).$$

Therefore, we deduce that

$$(\mathbf{A})_2^K = -\frac{2b(z, z)}{K^2} \langle (\nabla f \cdot \mathbf{1})(\bullet \cdot \mathbf{1}), \mu^n \rangle + O\left(\frac{M_1(\mu)}{K^3}\right).$$

In similar way and using Assumptions  $(\mathbf{A})$  we have that

$$\begin{aligned}
 (\mathbf{A})_1^K &= \frac{b(z, z)}{K^2} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n [\langle \Phi_{i,j} f, \mu^{n-1} \rangle - \langle f, \mu^n \rangle] + O\left(\frac{1}{K^3}\right), \\
 (\mathbf{A})_3^K &= \frac{b(z, z)}{K^2} \sum_{i=1}^n \sum_{j=1}^n \langle K_{i,j} f, \mu^{n+1} \rangle + O\left(\frac{M_2(\mu)}{K^3}\right), \\
 (\mathbf{A})_5^K &= O\left(\frac{\sigma\sqrt{K}}{K} [M_1(\mu) + M_2(\mu)]\right).
 \end{aligned}$$

**Step 2. Decomposition and study of  $(\mathbf{B})^K$ .** For all  $\alpha, h \in \mathbb{R}$ , we denote  $\theta_m(\alpha, h) := \theta(\alpha) m(\alpha, h)$ . Note that  $(\mathbf{B})^K = \sum_{i=1}^5 (\mathbf{B})_i^K$  where

$$\begin{aligned}
 (\mathbf{B})_1^K &:= \sum_{k=1}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta_m(z, h) \left\langle f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle dh, \\
 (\mathbf{B})_2^K &:= -\frac{1}{K^{\frac{3}{2}}} \sum_{i=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta_m(z, h) h \left\langle \partial_i f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle dh,
 \end{aligned}$$

$$\begin{aligned}
 (\mathbf{B})_3^K &:= \frac{1}{2K^3} \sum_{i,j=1}^n \sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta_m(z, h) h^2 \left\langle \partial_{ij}^2 f, \mu^{n-k} \left[ -\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right]^k \right\rangle dh, \\
 (\mathbf{B})_4^K &:= O\left( \frac{1}{K^{\frac{9}{2}}} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta_m(z, h) |h|^3 dh \right), \\
 (\mathbf{B})_5^K &:= \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} [\theta_m(\sigma\sqrt{K}x + z, h) - \theta_m(z, h)] (\mathbf{C})_{x, x+h/\sqrt{K}}^K dh.
 \end{aligned}$$

Note that, from TAYLOR's formula and Assumptions **(A)**, we have that

$$\begin{aligned}
 &\binom{n}{1} \int_{\mathbb{R}} \mu(dx) \int_{\mathbb{R}} \theta_m(z, h) \left\langle f, \mu^{n-1} \left[ -\frac{\delta_x}{K} + \frac{\delta_{x+\frac{h}{\sqrt{K}}}}{K} \right] \right\rangle dh \\
 &= \frac{1}{K} \sum_{i=1}^n \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \mu(dx_i) \int_{\mathbb{R}} \theta_m(z, h) \left[ \frac{h}{\sqrt{K}} \partial_i f(x_1, \dots, x_i, \dots, x_n) \right. \\
 &\quad \left. + \frac{h^2}{2K} \partial_{ii}^2 f(x_1, \dots, x_i, \dots, x_n) + O\left( \frac{|h|^3}{K^{\frac{3}{2}}} \right) \right] \prod_{\substack{\ell=1 \\ \ell \neq i}}^n \mu(dx_\ell) \\
 &= \frac{\theta(z)m_2(z)}{K^2} \left\langle \frac{\Delta f}{2}, \mu^n \right\rangle + O\left( \frac{1}{K^{\frac{5}{2}}} \right).
 \end{aligned}$$

In similar way to Step 1, we obtain that

$$(\mathbf{B})_1^K = \frac{\theta(z)m_2(z)}{K^2} \left\langle \frac{\Delta f}{2}, \mu^n \right\rangle + O\left( \frac{1}{K^{\frac{5}{2}}} \right),$$

$$(\mathbf{B})_2^K = (\mathbf{B})_3^K = O\left( \frac{1}{K^3} \right), \quad (\mathbf{B})_4^K = O\left( \frac{1}{K^{\frac{9}{2}}} \right) \quad \text{and} \quad (\mathbf{B})_5^K = O\left( \frac{\sigma\sqrt{K}}{K^{\frac{3}{2}}} M_1(\mu) \right).$$

and using Assumption (3.5) the announced result follows.

**Step 3. Martingality.** By classical arguments, since  $M^{K, P_{f,n}}$  is bounded, we obtain that  $M^{K, P_{f,n}}$  is a square integrable martingale started at 0.

### 3.4 Moments estimates

Lemmas 3.3.6 and 3.3.7 lead us to look for moment estimates. In Lemma 3.3.6 we have given the DOOB semi-martingale decomposition of the second-order moment and note that the latter involves the third-order moment in the error term. The presence of the higher order moment generates a difficulty to obtain a fine control of the second order moment. This difficulty is overcome by introducing the stopping time  $\check{\tau}^K$ , given by (3.14). In Section 3.4.1, we give estimates of the moment of order 6 and some corollaries useful for Sections 3.5 and 3.6. In Section 3.4.2 we establish estimates in expectation and probability of the second order moment up to the stopping time  $\check{\tau}^K$  and we also control the martingale bracket of the DOOB decomposition given in Lemma 3.3.6. In Sections 3.4.3 and 3.4.4, we prove that the stopping time  $\tau^K := \check{\tau}^K \wedge \hat{\tau}^K$ , given by (3.12), converges in probability to  $+\infty$  when  $K \rightarrow +\infty$  using coupling arguments between  $M_2(\mu^K)$  and a biased random walk on  $\mathbb{N}^*$ , reflected in 1, and large deviations estimates for this random walk.

### 3.4.1 Estimates of the moment of order 6

In this section, we establish a technical lemma useful for Lemma 3.4.2 below. The idea is to obtain an explicit upper bound of the moments of order 6.

**Lemma 3.4.1.** *Given any continuous function  $y(t)$  satisfying*

$$y(t) \leq y(s) + \int_s^t F(y(r)) dr, \quad \forall s \leq t \quad (3.28)$$

for some LIPSCHITZ function  $F : \mathbb{R} \rightarrow \mathbb{R}$ , then  $y(s) \leq z(s)$  for all  $s \geq 0$ , where  $z(t)$  is the unique solution of

$$z(t) = z(s) + \int_s^t F(z(r)) dr, \quad \forall s \leq t \quad (3.29)$$

such that  $z(0) = y(0)$ .

*Proof.* First, it is sufficient to prove that, for all  $\varepsilon > 0$ ,  $y(s) \leq z_\varepsilon(s)$  for all  $s \geq 0$ , where  $z_\varepsilon$  is the unique solution of

$$z_\varepsilon(t) = z_\varepsilon(s) + \int_s^t [F(z_\varepsilon(r)) + \varepsilon] dr, \quad \forall s \leq t \quad (3.30)$$

such that  $z_\varepsilon(0) = y(0) + \varepsilon$ . Indeed, it follows from (3.29) and (3.30) that, for all  $t \geq 0$  and  $\varepsilon > 0$

$$|z(t) - z_\varepsilon(t)| \leq \varepsilon(1+t) + C_{\text{Lip}} \int_0^t |z(s) - z_\varepsilon(s)| ds$$

where  $C_{\text{Lip}}$  is the LIPSCHITZ constant of  $F$ . GRONWALL's lemma then entails

$$|z(t) - z_\varepsilon(t)| \leq \varepsilon(1+t) \exp(C_{\text{Lip}}t) \xrightarrow{\varepsilon \rightarrow 0} 0$$

so that  $y(t) \leq z_\varepsilon(t)$ , for all  $\varepsilon > 0$  indeed implies  $y(t) \leq z(t)$ .

So, let  $\varepsilon > 0$  be fixed and let us prove that  $y(t) \leq z_\varepsilon(t)$  for all  $t \geq 0$ . Define

$$\vartheta_\varepsilon := \inf \left\{ t \geq 0 \mid y(t) \geq z_\varepsilon(t) \right\}.$$

Our goal is to prove that  $\vartheta_\varepsilon = +\infty$ .

By continuity of  $y$ ,  $\vartheta_\varepsilon > 0$  and, if  $\vartheta_\varepsilon < \infty$ ,  $y(\vartheta_\varepsilon) = z_\varepsilon(\vartheta_\varepsilon)$ . Assume by contradiction that  $\vartheta_\varepsilon < \infty$ . Then, using (3.28) and (3.30), for all  $\delta \in (0, \vartheta_\varepsilon)$ ,

$$\frac{z_\varepsilon(\vartheta_\varepsilon) - y(\vartheta_\varepsilon) - (z_\varepsilon(\vartheta_\varepsilon - \delta) - y(\vartheta_\varepsilon - \delta))}{h} \geq \varepsilon + \frac{1}{\delta} \int_{s-\delta}^s [F(z_\varepsilon(r)) - F(y(r))] dr.$$

By continuity of  $z_\varepsilon$  and  $y$ , the last term of the right-hand side converges to 0 when  $\delta \rightarrow 0$ . Hence the left-hand side is positive for all  $\delta$  small enough. Since  $z_\varepsilon(\vartheta_\varepsilon) = y(\vartheta_\varepsilon)$ , this implies that

$$z_\varepsilon(\vartheta_\varepsilon - \delta) - y(\vartheta_\varepsilon - \delta) < 0$$

for such  $\delta$ . This contradicts the definition of  $\vartheta_\varepsilon$  as the first time  $t$  such that  $z_\varepsilon(t) \leq y(t)$ .  $\square$

**Lemma 3.4.2.** *There exists  $K_0 \in \mathbb{N}^*$  large enough and two constants  $C_1, C_2 > 0$  such that for all  $K \geq K_0$  and  $t \geq 0$ ,*

$$\mathbb{E} \left( M_6 \left( \mu_t^K \right) \mathbf{1}_{t \leq \tau^K} \right) \leq C_1 \exp \left( -\frac{9b}{28K^2\sigma^2} t \right) + C_2$$

*Proof. Step 1. Pathwise inequality.* Note that for all  $t \geq 0$ ,

$$M_7(\mu_{t \wedge \check{\tau}^K}^K) \leq \text{Diam}(\text{Supp } \mu_{t \wedge \check{\tau}^K}^K) M_6(\mu_{t \wedge \check{\tau}^K}^K) \leq \frac{M_6(\mu_{t \wedge \check{\tau}^K}^K)}{\sigma K^{\frac{3+\varepsilon}{2}}}. \quad (3.31)$$

Introducing for all  $t \geq 0$ ,  $A_t := \frac{3}{4}M_6(\mu_t^K) + 7M_4(\mu_t^K)M_2(\mu_t^K) - 4\widetilde{M}_3^2(\mu_t^K) + \frac{3}{2}M_2^3(\mu_t^K)$ , we obtain from Lemma 3.3.7, (3.31) and Assumptions **(A)** that there exists a constant  $C > 0$  such that for all  $t > s \geq 0$ ,

$$\begin{aligned} A_{t \wedge \check{\tau}^K} &\leq A_{s \wedge \check{\tau}^K} - \frac{1}{K^2 \sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left\{ \left( \underline{b} - \frac{C}{K^{\frac{\varepsilon}{2}}} \right) \left[ 3M_6(\mu_r^K) + \frac{9}{2}M_4(\mu_r^K)M_2(\mu_r^K) + 6M_2^3(\mu_r^K) \right] \right. \\ &\quad \left. - \bar{\theta} \bar{m}_2 \left[ \frac{91}{4}M_4(\mu_r^K) + 42M_2^2(\mu_r^K) \right] dr \right\} + \text{Mart}_{t \wedge \check{\tau}^K} - \text{Mart}_{s \wedge \check{\tau}^K} \end{aligned}$$

where  $(\text{Mart}_t)_{t \geq 0}$  is a martingale. Note that, choosing  $K_0$  large enough, for all  $K \geq K_0$ ,  $\underline{b} - CK^{-\varepsilon/2} > \underline{b}/2$ . From the inequalities  $M_2^2(\mu) \leq M_4(\mu)$  and  $M_4(\mu) \leq M_6^{2/3}(\mu) \leq \alpha M_6(\mu) + 1/\alpha^2$  with  $\alpha := 2\underline{b}/259\bar{\theta}\bar{m}_2$ , we deduce that

$$\frac{3\underline{b}}{2}M_6(\mu) - \bar{\theta}\bar{m}_2 \left[ \frac{91}{4}M_4(\mu) + 42M_2^2(\mu) \right] \geq \frac{3\underline{b}}{2}M_6(\mu) - \frac{259\bar{\theta}\bar{m}_2}{4}M_4(\mu) \geq \underline{b}M_6(\mu) - \frac{259^3\bar{\theta}^3\bar{m}_2^3}{16\underline{b}^2}.$$

Thus, for all  $t > s \geq 0$ ,

$$\begin{aligned} A_{t \wedge \check{\tau}^K} &\leq A_{s \wedge \check{\tau}^K} + \text{Mart}_{t \wedge \check{\tau}^K} - \text{Mart}_{s \wedge \check{\tau}^K} \\ &\quad - \frac{\underline{b}}{2K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left[ 2M_6(\mu_r^K) + \frac{9}{2}M_4(\mu_r^K)M_2(\mu_r^K) + 6M_2^3(\mu_r^K) - \left( \frac{259\bar{\theta}\bar{m}_2}{2\underline{b}} \right)^3 \right] dr \\ &\leq A_{s \wedge \check{\tau}^K} - \frac{9\underline{b}}{28K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left( A_r - \frac{14}{9} \left( \frac{259\bar{\theta}\bar{m}_2}{2\underline{b}} \right)^3 \right) dr + \text{Mart}_{t \wedge \check{\tau}^K} - \text{Mart}_{s \wedge \check{\tau}^K} \end{aligned}$$

**Step 2. Conclusion.** Let us consider the stochastic process  $(N(\mu_t^K))_{t \geq 0}$  defined by

$$N(\mu_t^K) := \begin{cases} A_t & \text{if } s < t \leq \check{\tau}^K \\ (A_{\check{\tau}^K} - C_0) \exp\left(-\frac{3\underline{b}}{4K^2\sigma^2}(t - \check{\tau}^K)\right) + C_0 & \text{if } s < \check{\tau}^K < t \\ 0 & \text{if } \check{\tau}^K < s < t \end{cases}$$

where  $C_0 := \frac{14}{9} \left( \frac{259\bar{\theta}\bar{m}_2}{2\underline{b}} \right)^3$ . From Step 1, for all  $t > s \geq 0$ ,

$$N(\mu_t^K) \leq N(\mu_s^K) - \frac{9\underline{b}}{28K^2\sigma^2} \int_s^t [N(\mu_r^K) - C_0] dr + \text{Mart}_t - \text{Mart}_s$$

which entails

$$y(t) \leq y(s) - \frac{9\underline{b}}{28K^2\sigma^2} \int_s^t (y(r) - C_0) dr$$

where  $y(t) := \mathbb{E}(N(\mu_t^K))$ . From Lemma 3.4.1, we deduce that for all  $t \geq 0$ ,  $y(t) \leq z(t)$  where  $z(t)$  is the unique solution of

$$z(t) = z(s) - \frac{9\underline{b}}{28K^2\sigma^2} \int_s^t (z(r) - C_0) dr$$

such that  $z(0) = A_0$ , i.e.  $z(t) = (z(0) - C_0) \exp\left(-\frac{9b}{28K^2\sigma^2}t\right) + C_0$ . From (3.7) then (3.4), note that

$$M_6(\mu_0^K) = \langle \text{id}^6, \mu_0^K \rangle = \frac{1}{K^3\sigma^6} \langle (\text{id} - x_0)^6, \nu_0^K \rangle \leq C_6^*.$$

From CAUCHY-SCHWARZ's inequality  $\widetilde{M}_3^2(\mu) \leq M_4(\mu)M_2(\mu)$ , note that  $z(0) \leq 21C_6^*/4$  and so, for all  $t \geq 0$ ,

$$\begin{aligned} \mathbb{E}\left(M_6(\mu_t^K) \mathbf{1}_{t \leq \check{\tau}K}\right) &\leq \frac{4}{3} \mathbb{E}\left(\left[\frac{3}{4}M_6(\mu_t^K) + 3M_4(\mu_t^K)M_2(\mu_t^K) + \frac{3}{2}M_2^3(\mu_t^K)\right] \mathbf{1}_{t \leq \check{\tau}K}\right) \\ &\leq \frac{4}{3}z(t) \leq \frac{4}{3}\left(\frac{21}{4}(C_6^* - C_0) \exp\left(-\frac{9b}{28K^2\sigma^2}t\right) + C_0\right) \end{aligned}$$

and the announced result follows.  $\square$

The following lemma will be useful in the proof of Theorem 3.5.1.

**Lemma 3.4.3.** *For all  $t \geq 0$ ,*

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}\left(\int_0^{t \wedge \check{\tau}K} M_6(\mu_s^K) ds\right) < \infty.$$

*Proof.* Let  $t \geq 0$  be fixed. By FUBINI's theorem and Lemma 3.4.2, there exists two constants  $C_1, C_2 > 0$  such that

$$\mathbb{E}\left(\int_0^t M_6(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds\right) \leq C_1 \int_0^t \exp\left(-\frac{9b}{28K^2\sigma^2}s\right) ds + C_2 t \leq (C_1 + C_2)t$$

which ends the proof.  $\square$

The last lemma has the following consequence, used in Sections 3.5 and 3.6.

**Corollary 3.4.4.** *For all  $T > 0$ , the family*

$$\left(\int_0^{t \wedge \check{\tau}K} M_5(\mu_s^K) ds\right)_{t \in [0, T], K \in \mathbb{N}^*}$$

*is uniformly integrable.*

*Proof.* Note that for all  $K \in \mathbb{N}^*$ , for all  $t \in [0, T]$ , for all  $A \in (0, +\infty)$ ,

$$\mathbf{1}_{\int_0^t M_5(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds \geq A} \int_0^t M_5(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds \leq \frac{1}{\sqrt[5]{A}} \left(\int_0^t M_5(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds\right)^{6/5},$$

we deduce from HÖLDER'inequality that, for all  $t \in [0, T]$ , for all  $A \in (0, +\infty)$ ,

$$\begin{aligned} \sup_{K \in \mathbb{N}^*} \mathbb{E}\left(\mathbf{1}_{\int_0^t M_5(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds \geq A} \int_0^t M_5(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds\right) &\leq \frac{T}{\sqrt[5]{A}} \sup_{K \in \mathbb{N}^*} \mathbb{E}\left(\int_0^t M_5^{6/5}(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds\right) \\ &\leq \frac{T}{\sqrt[5]{A}} \sup_{K \in \mathbb{N}^*} \mathbb{E}\left(\int_0^t M_6(\mu_s^K) \mathbf{1}_{s \leq \check{\tau}K} ds\right). \end{aligned}$$

From Lemma 3.4.3, we deduce that the right hand side of the previous inequality goes to 0 when  $A \rightarrow +\infty$  and so the announced result.  $\square$

### 3.4.2 Some inequalities on the moment of order 2

**Lemma 3.4.5.** *There exists  $K_0 \in \mathbb{N}^*$  large enough such that for all  $K \geq K_0$ , the process  $M_2(\mu^K)$  satisfies for all  $t \geq 0$  the following inequalities:*

$$\begin{aligned} M_2(\mu_{s \wedge \check{\tau}^K}^K) - \frac{5\bar{b}}{2K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left[ M_2(\mu_r^K) - \frac{2\theta m_2}{5\bar{b}} + \frac{2\bar{\theta}\bar{m}_2}{5\bar{b}K} \right] dr + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} \\ \leq M_2(\mu_{t \wedge \check{\tau}^K}^K) \leq M_2(\mu_{s \wedge \check{\tau}^K}^K) - \frac{\underline{b}}{K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left[ M_2(\mu_r^K) - \frac{\bar{\theta}\bar{m}_2}{\underline{b}} \right] dr + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} \end{aligned} \quad (3.32)$$

where  $M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}}$  is defined in Proposition 3.3.5. Moreover, for  $t \geq 0$ ,

$$\mathbb{E} \left( M_2(\mu_t^K) \mathbf{1}_{t \leq \check{\tau}^K} \right) \leq \mathbb{E} \left( M_2(\mu_0^K) - \frac{\bar{\theta}\bar{m}_2}{\underline{b}} \right) \exp \left( -\frac{\underline{b}}{K^2\sigma^2} t \right) + \frac{\bar{\theta}\bar{m}_2}{\underline{b}}.$$

*Proof.* Note that for all  $t \geq 0$ ,

$$M_3(\mu_{t \wedge \check{\tau}^K}^K) \leq \text{Diam}(\text{Supp } \mu_{t \wedge \check{\tau}^K}^K) M_2(\mu_{t \wedge \check{\tau}^K}^K) \leq \frac{M_2(\mu_{t \wedge \check{\tau}^K}^K)}{\sigma K^{\frac{3+\varepsilon}{2}}}. \quad (3.33)$$

From Lemma 3.3.6, (3.33) and Assumptions **(A)**, there exists two constants  $C, \tilde{C} > 0$  such that for all  $t > s \geq 0$

$$\begin{aligned} M_2(\mu_{s \wedge \check{\tau}^K}^K) - \frac{1}{K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left( \left[ 2\bar{b} + \frac{C}{K^{\frac{\varepsilon}{2}}} \right] M_2(\mu_r^K) - \theta m_2 + \frac{\bar{\theta}\bar{m}_2}{K} \right) dr + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} \\ \leq M_2(\mu_{t \wedge \check{\tau}^K}^K) \\ \leq M_2(\mu_{s \wedge \check{\tau}^K}^K) - \frac{1}{K^2\sigma^2} \int_{s \wedge \check{\tau}^K}^{t \wedge \check{\tau}^K} \left( \left[ 2\underline{b} - \frac{\tilde{C}}{K^{\frac{\varepsilon}{2}}} \right] M_2(\mu_r^K) - \bar{\theta}\bar{m}_2 + \frac{\theta m_2}{K} \right) dr + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} - M_{s \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}}. \end{aligned}$$

Noting that for all  $K \geq K_0$ ,  $2\underline{b} - \tilde{C}K^{-\varepsilon/2} > \underline{b}$  and  $2\bar{b} + CK^{-\varepsilon/2} < 5\bar{b}/2$ , the announced first result follows. In similar way to Step 2 of the proof of Lemma 3.4.2, we obtain the second part of the announced result.  $\square$

**Corollary 3.4.6.** *For all  $t \geq 0$ ,  $\mathbb{E}(M_2(\mu_t^K) \mathbf{1}_{t \leq \check{\tau}^K}) \leq \max\{C_2^*, \bar{\theta}\bar{m}_2/\underline{b}\}$ .*

*Proof.* From (3.7) and (3.4), note that

$$M_2(\mu_0^K) = \langle \text{id}^2, \mu_0^K \rangle = \frac{1}{K\sigma^2} \langle (\text{id} - x_0)^2, \nu_0^K \rangle \leq \frac{C_2^*}{K\sigma^2}.$$

The announced result follows from Lemma 3.4.5.  $\square$

The goal of the following result is to bound the martingale bracket  $\langle M^{K, P_{\text{id}^2, 1}} \rangle_{t \wedge \check{\tau}^K}$ . To do this, we exploit the martingale approximation of the moment of order 6 in order to control the dominant term  $M_4(\mu_s^K) - M_2^2(\mu_s^K)$  in (3.27).

**Lemma 3.4.7.** *There exists a constant  $C > 0$  such that for all  $K \in \mathbb{N}^*$ , for all  $t \geq 0$  the martingale bracket  $\langle M^{K, P_{\text{id}^2, 1}} \rangle_t$  satisfies*

$$\mathbb{E} \left( \langle M^{K, P_{\text{id}^2, 1}} \rangle_{t \wedge \check{\tau}^K} \right) \leq \frac{C}{K^2\sigma^2} \mathbb{E}(t \wedge \check{\tau}^K).$$

*Proof.* From Lemma 3.3.6 and HÖLDER's inequalities  $M_2^2(\mu) \leq M_4(\mu) \leq M_6^{2/3} \leq 1 + M_6(\mu)$  and  $M_5(\mu) \leq M_6(\mu)^{5/6} \leq 1 + M_6(\mu)$ , we deduce that there exists a constant  $C > 0$  such that for all  $t \geq 0$

$$\langle M^{K, P_{\text{id}^2, 1}} \rangle_{t \wedge \hat{\tau}^K} \leq \frac{C}{K^2 \sigma^2} \int_0^{t \wedge \hat{\tau}^K} (1 + M_6(\mu_s^K)) \left(1 + \frac{1}{K}\right) ds.$$

The announced result follows from FUBINI's theorem and Lemma 3.4.2.  $\square$

### 3.4.3 Convergence to $+\infty$ of the stopping time $\tau^K$ and support concentration property

The main result of this section is the following convergence result for the stopping time  $\tau^K$ , defined by (3.12), when  $K \rightarrow +\infty$ .

**Proposition 3.4.8.** *Under Assumption (3.5),  $\tau^K$  converges in probability to  $+\infty$  when  $K \rightarrow +\infty$ .*

The proof is based on the next lemma proved in Section 3.4.4 below. Thanks to Proposition 3.4.8, we deduce the support concentration property of  $(\nu_{t/K\sigma^2}^K)$  given by (3.6). Indeed, with probability which tends to 1 when  $K \rightarrow +\infty$ ,  $\text{Diam}(\mu_t^K) \leq \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}}$  for all  $t \in [0, T]$ . Hence,

$$\text{Diam}(\nu_{t/K\sigma^2}^K) = \sigma \sqrt{K} \text{Diam}(\mu_t^K) \leq \frac{1}{K^{1+\frac{\varepsilon}{2}}}$$

and (3.6) follows.

**Lemma 3.4.9.** *Under Assumption (3.5), for all  $T \geq 0$ ,*

$$\lim_{K \rightarrow +\infty} \mathbb{P}(\hat{\tau}^K \leq \check{\tau}^K \wedge T) = \lim_{K \rightarrow +\infty} \mathbb{P}\left(\sup_{0 \leq t \leq T \wedge \hat{\tau}^K} M_2(\mu_t^K) \geq K^\varepsilon\right) = 0.$$

*Proof of Proposition 3.4.8.* Note that for all  $T \geq 0$ ,

$$\mathbb{P}(\check{\tau}^K < T \wedge \hat{\tau}^K) \leq \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}}\right).$$

As for all  $t \geq 0$ , for all  $K \in \mathbb{N}^*$ ,  $\text{Diam}(\text{Supp } \mu_t^K)^2 \leq \left(2 \max\{|x| \mid x \in \text{Supp } \mu_t^K\}\right)^2 \leq 4K M_2(\mu_t^K)$ , we deduce that

$$\begin{aligned} \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, \text{Diam}(\text{Supp } \mu_t^K) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}}\right) \\ \leq \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, M_2(\mu_t^K) > \frac{1}{4\sigma^2 K^{4+\varepsilon}}\right). \end{aligned}$$

According to (3.5), we deduce that the following inequality  $4\sigma^2 K^{4+\varepsilon} \leq 1/K^\varepsilon$  holds true when  $K \rightarrow +\infty$  and  $\sigma \rightarrow 0$ . Hence, we deduce that

$$\mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, M_2(\mu_t^K) > \frac{1}{4\sigma^2 K^{4+\varepsilon}}\right) \leq \mathbb{P}\left(\exists t < T \wedge \hat{\tau}^K, M_2(\mu_t^K) > K^\varepsilon\right) = 0.$$



Now, for all  $T > 0$ ,  $\{\check{\tau}^K \geq T \wedge \hat{\tau}^K\} = \{\check{\tau}^K \geq \hat{\tau}^K\} \cup \{\check{\tau}^K \geq T\}$ , and so

$$\begin{aligned} \mathbb{P}\left(\{\check{\tau}^K \geq T \wedge \hat{\tau}^K\} \cap \{\tau^K < T\}\right) &\leq \mathbb{P}\left(\left[\{\hat{\tau}^K < T\} \cap \{\hat{\tau}^K \leq \check{\tau}^K\}\right] \cup \{\hat{\tau}^K < T \leq \check{\tau}^K\}\right) \\ &\leq \mathbb{P}\left(\hat{\tau}^K \leq T \wedge \check{\tau}^K\right). \end{aligned}$$

Note that for all  $T > 0$ ,

$$\mathbb{P}\left(\tau^K < T\right) \leq \mathbb{P}\left(\check{\tau}^K < T \wedge \hat{\tau}^K\right) + \mathbb{P}\left(\hat{\tau}^K \leq T \wedge \check{\tau}^K\right).$$

Hence, from Lemma 3.4.9, we have for all  $T \geq 0$ ,  $\lim_{K \rightarrow +\infty} \mathbb{P}\left(\tau^K < T\right) = 0$  and the announced result follows.  $\square$

### 3.4.4 Proof of Lemma 3.4.9

Note that if  $M_2(\mu_s^K)$  becomes larger than  $\bar{\theta}\bar{m}_2/b$ , the drift term in the right-hand side of (3.32) is then negative preventing  $M_2(\mu_s^K)$  to become excessively large unless the martingale  $M^{K, P_{\text{id}^2, 1}}$  has an exceptional path with large increments on a small time interval. We then expect to have large deviations estimates and bounds on the time of exit from attractive domains for  $M_2(\mu^K)$ . However, we cannot consider establishing directly a large deviation principle on  $M_2(\mu^K)$ , in particular because of the slow-fast limit. The approach considered below is based on coupling arguments between  $M_2(\mu^K)$  and a simpler process for which large deviations estimates are known.

Let us introduce  $u_0 := 0$  and for all  $\ell \in \mathbb{N}^*$  the real number  $u_\ell := 3^\ell K^{\frac{\varepsilon}{2}}$  and the interval  $\mathcal{I}_\ell := [u_{\ell-1}, u_{\ell+1})$ . We will look at the process  $M_2(\mu^K)$  at successive exit times of  $(\mathcal{I}_\ell)_{\ell \in \mathbb{N}^*}$ . We set  $L_0^K := 1$  when  $M_2(\mu_0^K) \in [0, 2u_1 + 1]$  and  $L_0^K := \ell_0 \in \mathbb{N}^*$  when  $M_2(\mu_0^K) \in (2u_{\ell_0-1} + 1, 2u_{\ell_0} + 1]$ . We also set  $T_0^K := 0$  and by induction on  $k \in \mathbb{N}^*$ , if  $(L_i^K)_{1 \leq i \leq k}$  and  $(T_i^K)_{1 \leq i \leq k}$  are constructed,

$$T_{k+1}^K := \inf \left\{ t \geq T_k^K \mid M_2(\mu_t^K) \notin \mathcal{I}_{L_k^K} \right\} \quad \text{and} \quad L_{k+1}^K := \min \left\{ \ell \in \mathbb{N} \mid M_2(\mu_{T_{k+1}^K}^K) \leq 2u_\ell + 1 \right\}.$$

Note that from (3.4), for  $K$  large enough,  $L_0^K = 1$  because  $M_2(\mu_0^K) \leq C_2^*$ . Since  $(\mu_t^K)_{t \geq 0}$  is a pure jump process, note that  $(M_2(\mu_t^K))_{t \geq 0}$  is also a pure jump process. The definition of  $u_\ell$  is motivated by the fact that, for all  $k \geq 0$ ,  $L_{k+1}^K \leq L_k^K + 1$ , which follows from the next lemma.

**Lemma 3.4.10.** *Let  $\vartheta_1$  be the first jump time of  $\mu^K$ . Then, for  $K$  large enough*

$$M_2(\mu_{\vartheta_1}^K) \leq 2M_2(\mu_0^K) + 1.$$

*In particular, for all  $t > 0$ ,  $\Delta M_2(\mu_t^K) := M_2(\mu_t^K) - M_2(\mu_{t-}^K) \leq M_2(\mu_{t-}^K) + 1$ .*

*Proof.* Let us consider  $\mu_0^K := \frac{1}{K} \sum_{k=1}^K \delta_{x_k}$  and  $y = (y_k)_{1 \leq k \leq K}$  defined by

$$y = \begin{cases} (x_1, \dots, x_{i-1}, x_i + \frac{H}{\sqrt{K}}, x_{i+1}, \dots, x_K) & \text{on } E_1, \\ (x_1, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_K) & \text{on } E_2, \end{cases}$$

where  $H$  follows the mutation law  $m(x_i, dh)$ ,  $E_1$  is the event ‘‘a mutation occurs at time  $\vartheta_1$ ’’ and  $E_2$  is the event ‘‘a resampling between  $i$  and  $j$  occurs at time  $\vartheta_1$ ’’. Note that  $\mu_{\vartheta_1}^K = \tau_{-\langle \text{id}, \frac{1}{K} \sum_{k=1}^K \delta_{y_k} \rangle} \# \left( \frac{1}{K} \sum_{k=1}^K \delta_{y_k} \right)$ . Hence, noting that for all  $k \in \{1, \dots, K\}$ ,

$$|x_k| \leq \sqrt{K} \left( 1 + M_2(\mu_0^K) \right)$$

and using Assumption **(A2)**, for  $K$  large enough we have that

$$\begin{aligned}
 M_2(\mu_{\vartheta_1}^K) &\leq M_2\left(\frac{1}{K}\sum_{i=1}^K\delta_{y_i}\right) = \frac{1}{K}\sum_{i=1}^K y_i^2 \leq \begin{cases} \frac{1}{K}\sum_{k=1}^K x_k^2 + \frac{2A_m x_k}{K^{\frac{3}{2}}} + \frac{A_m^2}{K^2} & \text{on } E_1, \\ \frac{1}{K}\sum_{k=1}^K x_k^2 + \frac{x_j^2 - x_i^2}{K} & \text{on } E_2, \end{cases} \\
 &\leq \begin{cases} M_2(\mu_0^K) + \frac{2A_m(1 + M_2(\mu_0^K))}{K} + \frac{A_m^2}{K^2} & \text{on } E_1, \\ 2M_2(\mu_0^K) & \text{on } E_2, \end{cases} \\
 &\leq 2M_2(\mu_0^K) + 1. \quad \square
 \end{aligned}$$

Each of the previous steps will be called *transitions of  $M_2(\mu_t^K)$* . The idea of the proof of Lemma 3.4.9 is to estimate the transition probabilities of the sequence  $(L_k^K)_{k \in \mathbb{N}}$  in order to construct a coupling between  $(L_k^K)_{k \in \mathbb{N}}$  and a biased random walk on  $\mathbb{N}^*$  reflected in 1.

At time  $\hat{\tau}^K$ ,  $M_2(\mu_{\hat{\tau}^K}^K) \in \mathcal{I}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}$  where  $\lfloor x \rfloor$  is the lower integer part of  $x$ . So, the number of transitions of  $M_2(\mu_t^K)$  before  $\hat{\tau}^K$  is greater than or equal to  $T_{k_0}^K$  where  $k_0$  is the first integer such that  $L_{k_0}^K = \lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$ .

By using estimates on the number of steps that a biased random walk takes to reach  $\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$  and estimates on the durations between two transitions  $T_{k+1}^K - T_k^K$ , we will deduce a lower bound, exponential in  $K$ , on  $\hat{\tau}^K$ . An additional difficulty comes from the fact that the previous coupling argument is only valid up to time  $\check{\tau}^K$ . Hence we will construct a coupling that takes into account the possibility that  $\check{\tau}^K$  happens during each transition step. The proof is divided into four steps: in Section 3.4.4.1, we characterise the behaviour of the first transition step; in Section 3.4.4.2, the proposed coupling is constructed; in Section 3.4.4.3, we give estimates on the first exit time from an attractive domain for random walks; finally, we conclude in Section 3.4.4.4.

### 3.4.4.1 One-step transitions

In this section we look at only one transition: we suppose that  $M_2(\mu_0^K) \in [2u_{\ell-1} + 1, 2u_{\ell} + 1)$  for  $\ell \in \mathbb{N}^*$  fixed and we look for bounds on the first transition probabilities of  $M_2(\mu_t^K)$  and on time  $T_1^K$ . The main result is the following.

**Lemma 3.4.11.** *For all  $\ell \in \mathbb{N}$ , for all  $\mu_0^K \in \mathcal{M}_{1,K}^{c,2}(\mathbb{R})$  such that  $M_2(\mu_0^K) \in [2u_{\ell-1} + 1, 2u_{\ell} + 1)$ , there exist a constant  $C > 0$ ,  $K_0 \in \mathbb{N}^*$  such that for all  $K \geq K_0$ , we have*

$$\mathbb{P}_{\mu_0^K} \left( \left( \{L_1^K = \ell - 1\} \cap \{\check{\tau}^K > T_1^K \geq K\sigma^2\} \right) \cup \{T_1^K \geq \check{\tau}^K\} \right) \geq \frac{1}{2} + \eta_K, \quad (3.34)$$

where  $\eta_K := \frac{1}{2} - \frac{C}{K^{\varepsilon/2}}$ .

The proof of the previous result is based on Corollary 3.4.12 which is obtained as a straightforward consequence of Lemma 3.4.5.

**Corollary 3.4.12.** *Let  $\ell \in \mathbb{N}^*$  be fixed. There exists  $K_0 \in \mathbb{N}^*$  large enough such that for all  $K \geq K_0$ , for all  $\mu_0^K \in \mathcal{M}_{1,K}^{c,2}(\mathbb{R})$  such that  $M_2(\mu_0^K) \in [2u_{\ell-1} + 1, 2u_{\ell} + 1)$ , we have for all  $t \leq T_1^K$  that*

$$2u_{\ell-1} + 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K^2\sigma^2} t \wedge \check{\tau}^K + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}} \leq M_2(\mu_{t \wedge \check{\tau}^K}^K) \leq 2u_{\ell} + 1 - \frac{bu_{\ell-1}}{K^2\sigma^2} t \wedge \check{\tau}^K + M_{t \wedge \check{\tau}^K}^{K, P_{\text{id}^2, 1}}.$$

*Proof of Lemma 3.4.11.* By passing on the complementary of (3.34), it is equivalent to prove that

$$\mathbb{P}_{\mu_0^K} \left( \left[ \{L_1^K = \ell + 1\} \cup \{T_1^K < K\sigma^2\} \right] \cap \{T_1^K < \check{\tau}^K\} \right) < \frac{1}{2} - \eta_K.$$

As for all events  $A, B$ ,  $\mathbb{P}(A \cup B) = \mathbb{P}(A \cap B^c) + \mathbb{P}(B)$ , we have

$$\begin{aligned} & \mathbb{P}_{\mu_0^K} \left( \left[ \{L_1^K = \ell + 1\} \cup \{T_1^K < K\sigma^2\} \right] \cap \{T_1^K < \check{\tau}^K\} \right) \\ &= \mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \{L_1^K = \ell + 1\} \cap \{T_1^K \geq K\sigma^2\} \right) \\ & \quad + \mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \{T_1^K < K\sigma^2\} \right) \\ & \leq \mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \{L_1^K = \ell + 1\} \right) + \mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \{T_1^K < K\sigma^2\} \right). \end{aligned} \quad (3.35)$$

**Step 1. Control of the first right-hand term of (3.35).** Let us consider the martingale  $(\text{Mart}_t^{K,+})_{t \geq 0}$  defined by

$$\text{Mart}_t^{K,+} := \begin{cases} M_t^{K,P_{\text{id}^2,1}} & \text{if } t \leq \check{\tau}^K \\ M_{\check{\tau}^K}^{K,P_{\text{id}^2,1}} + \frac{1}{K\sigma} B_{t-\check{\tau}^K} & \text{if } t > \check{\tau}^K \end{cases}$$

where  $(B_t)_{t \geq 0}$  is a standard Brownian motion independent of  $(\mu_t^K)_{t \geq 0}$ . Note that from Lemma 3.4.7, there exists a constant  $C > 0$  such that for all  $t \geq 0$  we have

$$\mathbb{E} \left( \langle \text{Mart}^{K,+} \rangle_t \right) \leq \frac{C}{K^2\sigma^2} t. \quad (3.36)$$

Let us consider the process  $(M_2^{K,+}(t))_{t \geq 0}$  defined by

$$M_2^{K,+}(t) := \begin{cases} M_2(\mu_t^K) & \text{if } t \leq \check{\tau}^K \\ M_2(\mu_{\check{\tau}^K}^K) - \frac{bu_{\ell-1}}{K^2\sigma^2} (t - \check{\tau}^K) + (\text{Mart}_t^{K,+} - \text{Mart}_{\check{\tau}^K}^{K,+}) & \text{if } t > \check{\tau}^K \end{cases}.$$

Note that from Corollary 3.4.12, for all  $t \geq 0$ ,

$$2u_{\ell-1} + 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K^2\sigma^2} t + \text{Mart}_t^{K,+} \leq M_2^{K,+}(t) \leq 2u_{\ell} + 1 - \frac{bu_{\ell-1}}{K^2\sigma^2} t + \text{Mart}_t^{K,+}. \quad (3.37)$$

By GIRSANOV's theorem, the process  $(E_t^{K,+})_{t \geq 0}$  defined for all  $t \geq 0$  by

$$E_t^{K,+} := \exp \left( \text{Mart}_t^{K,+} - \frac{1}{2} \langle \text{Mart}^{K,+} \rangle_t \right)$$

is a local martingale. Let  $t_0^K := \frac{6K^2\sigma^2}{b}$  be fixed and let us consider

$$T_1^{K,+} := \inf \left\{ t \geq 0 \mid M_2^{K,+}(t) \notin \mathcal{I}_{\ell} \right\} \quad \text{and} \quad L_1^{K,+} := \min \left\{ \ell \in \mathbb{N} \mid M_2^{K,+}(T_1^{K,+}) \leq 2u_{\ell} - 1 \right\}.$$

From (3.37) and denoting  $\Gamma := \left\{ \langle \text{Mart}^{K,+} \rangle_{T_1^{K,+} \wedge t_0^K} \leq \frac{2bu_{\ell-1}}{K^2\sigma^2} T_1^{K,+} \wedge t_0^K \right\}$ , we have

$$\begin{aligned} 1 &= \mathbb{E}_{\mu_0^K} \left( E_{T_1^{K,+} \wedge t_0^K}^{K,+} \right) \geq \mathbb{E}_{\mu_0^K} \left( E_{T_1^{K,+} \wedge t_0^K}^{K,+} \mathbf{1}_{\Gamma} \right) \\ &\geq \exp(-2u_{\ell} - 1) \mathbb{E}_{\mu_0^K} \left( \exp \left( M_2^{K,+}(T_1^{K,+} \wedge t_0^K) + \frac{bu_{\ell-1}}{K^2\sigma^2} T_1^{K,+} \wedge t_0^K - \frac{1}{2} \langle \text{Mart}^{K,+} \rangle_{T_1^{K,+} \wedge t_0^K} \right) \mathbf{1}_{\Gamma} \right) \\ &\geq \exp(-2u_{\ell} - 1) \mathbb{E}_{\mu_0^K} \left( \exp \left( M_2^{K,+}(T_1^{K,+} \wedge t_0^K) \right) \mathbf{1}_{\Gamma} \right) \\ &\geq \exp(-2u_{\ell} - 1) \left[ \exp(u_{\ell+1}) \mathbb{P}_{\mu_0^K} \left( \{L_1^{K,+} = \ell + 1\} \cap \Gamma \cap \{T_1^{K,+} \leq t_0^K\} \right) \right. \\ & \quad \left. + \mathbb{P}_{\mu_0^K} \left( \{L_1^{K,+} \leq \ell - 1\} \cap \Gamma \cap \{T_1^{K,+} \leq t_0^K\} \right) \right]. \end{aligned}$$

As for all events  $A, B, C$ ,  $\mathbb{P}(A \cap B \cap C) \geq \mathbb{P}(A) - \mathbb{P}(A \cap B^c) - \mathbb{P}(A \cap C^c)$ , and denoting  $p := \mathbb{P}_{\mu_0^K}(L_1^{K,+} = \ell + 1)$  we have that

$$\begin{aligned} 1 &\geq \exp(u_\ell - 1) \left[ p - \mathbb{P}_{\mu_0^K}(\{L_1^{K,+} = \ell + 1\} \cap \Gamma^c) - \mathbb{P}_{\mu_0^K}(\{L_1^{K,+} = \ell + 1\} \cap \{T_1^{K,+} > t_0^K\}) \right] \\ &\quad + \exp(-2u_\ell - 1) \left[ 1 - p - \mathbb{P}_{\mu_0^K}(\{L_1^{K,+} \leq \ell - 1\} \cap \Gamma^c) - \mathbb{P}_{\mu_0^K}(\{L_1^{K,+} \leq \ell - 1\} \cap \{T_1^{K,+} > t_0^K\}) \right] \\ &\geq (\exp(u_\ell - 1) - \exp(-2u_\ell - 1))p + \exp(-2u_\ell - 1) \\ &\quad - (\exp(u_\ell - 1) + \exp(-2u_\ell - 1)) \left[ \mathbb{P}_{\mu_0^K}(\Gamma^c) + \mathbb{P}_{\mu_0^K}(T_1^{K,+} > t_0^K) \right]. \end{aligned}$$

Then, considering  $s_0^K := \frac{u_{\ell-1}K^2\sigma^2}{5bu_{\ell+1} + \theta m_2} < t_0^K$  be fixed, we deduce that

$$\begin{aligned} p &\leq \frac{1 - \exp(-2u_\ell - 1)}{\exp(u_\ell - 1) - \exp(-2u_\ell - 1)} + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)} \left[ \mathbb{P}_{\mu_0^K}(\Gamma^c) + \mathbb{P}_{\mu_0^K}(T_1^{K,+} > t_0^K) \right] \\ &\leq \exp(-u_\ell + 1) \times \frac{1 - \exp(-2u_\ell - 1)}{1 - \exp(-u_{\ell+1})} + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)} \\ &\quad \times \left[ \mathbb{P}_{\mu_0^K}(\Gamma^c \cap \{T_1^{K,+} \geq s_0^K\}) + \mathbb{P}_{\mu_0^K}(T_1^{K,+} < s_0^K) + \mathbb{P}_{\mu_0^K}(T_1^{K,+} > t_0^K) \right]. \end{aligned}$$

Denoting  $[a]$  the upper integer part of  $a \in \mathbb{R}$ , we obtain on the one hand that there exists a constant  $C_1 > 0$  such that

$$\begin{aligned} &\mathbb{P}_{\mu_0^K}(\Gamma^c \cap \{T_1^{K,+} \geq s_0^K\}) \\ &= \mathbb{P}_{\mu_0^K} \left( \left\{ \langle \text{Mart}^{K,+} \rangle_{T_1^{K,+} \wedge t_0^K} > \frac{2bu_{\ell-1}}{K^2\sigma^2} T_1^{K,+} \wedge t_0^K \right\} \cap \{T_1^{K,+} \geq s_0^K\} \right) \\ &= \sum_{k=1}^{\lceil t_0^K/s_0^K \rceil - 1} \mathbb{P}_{\mu_0^K} \left( \left\{ \langle \text{Mart}^{K,+} \rangle_{T_1^{K,+} \wedge t_0^K} > \frac{2bu_{\ell-1}}{K^2\sigma^2} T_1^{K,+} \wedge t_0^K \right\} \cap \{ks_0^K \leq T_1^{K,+} \leq (k+1)s_0^K\} \right) \\ &\quad + \mathbb{P}_{\mu_0^K} \left( \left\{ \langle \text{Mart}^{K,+} \rangle_{T_1^{K,+} \wedge t_0^K} > \frac{2bu_{\ell-1}}{K^2\sigma^2} T_1^{K,+} \wedge t_0^K \right\} \cap \{T_1^{K,+} \geq \lceil t_0^K/s_0^K \rceil s_0^K\} \right) \\ &\leq \sum_{k=1}^{\lceil t_0^K/s_0^K \rceil - 1} \mathbb{P}_{\mu_0^K} \left( \langle \text{Mart}^{K,+} \rangle_{(k+1)s_0^K} > \frac{2bu_{\ell-1}}{K^2\sigma^2} ks_0^K \right) + \mathbb{P}_{\mu_0^K} \left( \langle \text{Mart}^{K,+} \rangle_{t_0^K} > \frac{2bu_{\ell-1}}{K^2\sigma^2} t_0^K \right) \\ &\leq K^2\sigma^2 \left( \frac{\sum_{k=1}^{\lceil t_0^K/s_0^K \rceil - 1} \mathbb{E} \left( \langle \text{Mart}^{K,+} \rangle_{(k+1)s_0^K} \right)}{2bu_{\ell-1}ks_0^K} + \frac{\mathbb{E} \left( \langle \text{Mart}^{K,+} \rangle_{t_0^K} \right)}{2bu_{\ell-1}t_0^K} \right) \\ &\leq \frac{C_1}{K^{\frac{\varepsilon}{2}}} \end{aligned}$$

where we use MARKOV's inequality in the second inequality and (3.37) in the last inequality.

On the other hand, we control  $\mathbb{P}_{\mu_0^K}(T_1^{K,+} < s_0^K)$  as follows:

$$\begin{aligned} \mathbb{P}_{\mu_0^K}(T_1^{K,+} < s_0^K) &\leq \mathbb{P}_{\mu_0^K} \left( \{T_1^{K,+} < s_0^K\} \cap \left\{ \sup_{0 \leq t \leq s_0^K} M_2^{K,+}(t) \geq u_{\ell+1} \right\} \right) \\ &\quad + \mathbb{P}_{\mu_0^K} \left( \{T_1^{K,+} < s_0^K\} \cap \left\{ \inf_{0 \leq t \leq s_0^K} M_2^{K,+}(t) < u_{\ell-1} \right\} \right). \end{aligned}$$

Let us consider  $K_1 \in \mathbb{N}^*$  independent of  $\ell$  satisfying  $K_1^{\varepsilon/2} \geq 2$ . There exists a constant  $C_1 > 0$  such that for all  $K \geq K_1$ , we have that

$$\begin{aligned}
 & \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \sup_{0 \leq t \leq s_0^K} M_2^{K,+}(t) \geq u_{\ell+1} \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \sup_{0 \leq t \leq s_0^K} \left\{ 2u_{\ell} + 1 - \frac{bu_{\ell-1}}{K^2\sigma^2}t + \text{Mart}_t^{K,+} \right\} \geq u_{\ell+1} \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \sup_{0 \leq t \leq s_0^K} \text{Mart}_t^{K,+} \geq u_{\ell} - 1 \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \sup_{0 \leq t \leq s_0^K} \left| \text{Mart}_t^{K,+} \right| \geq \frac{3^\ell K^{\frac{\varepsilon}{2}}}{2} \right) \\
 & \leq 4\mathbb{E}_{\mu_0^K} \left( \left\langle \text{Mart}^{K,+} \right\rangle_{s_0^K} \right) \times \frac{4}{3^{2\ell} K^\varepsilon} \\
 & \leq \frac{C_1}{K^\varepsilon}
 \end{aligned}$$

where we use (3.37) in the first inequality, DOOB's maximal inequality in the fourth inequality and (3.36) in the last inequality. Now, let us consider  $K_2 \in \mathbb{N}^*$  independent of  $\ell$  satisfying  $K_2^{\varepsilon/2} \geq 4$ . Using again (3.37), DOOB's maximal inequality, (3.36), and the definition of  $s_0^K$ , we deduce that there exists a constant  $C_2 > 0$  such that for all  $K \geq K_2$ , we have that

$$\begin{aligned}
 & \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \inf_{0 \leq t \leq s_0^K} M_2^{K,+}(t) < u_{\ell-1} \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \inf_{0 \leq t \leq s_0^K} \left\{ 2u_{\ell-1} - 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K^2\sigma^2}t + \text{Mart}_t^{K,+} \right\} < u_{\ell-1} \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \left\{ T_1^{K,+} < s_0^K \right\} \cap \left\{ \inf_{0 \leq t \leq s_0^K} \text{Mart}_t^{K,+} < -\frac{u_{\ell-1}}{2} + 1 \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \sup_{0 \leq t \leq s_0^K} \left| \text{Mart}_t^{K,+} \right| > \frac{u_{\ell-1}}{2} - 1 \right) \\
 & \leq \frac{C_2}{K^\varepsilon}.
 \end{aligned}$$

Hence, for all  $K \geq \max\{K_1, K_2\}$ ,  $\mathbb{P}_{\mu_0^K} (T_1^{K,+} < s_0^K) \leq \frac{C_1+C_2}{K^\varepsilon}$ . Finally, from (3.37) and MARKOV's inequality and the definition of  $t_0^K$ , note that there exists a constant  $C_3 > 0$  such that

$$\begin{aligned}
 \mathbb{P}_{\mu_0^K} (T_1^{K,+} > t_0^K) & \leq \mathbb{P}_{\mu_0^K} \left( \left\{ \sup_{0 \leq t \leq t_0^K} M_2^{K,+}(t) < u_{\ell+1} \right\} \cap \left\{ \inf_{0 \leq t \leq t_0^K} M_2^{K,+}(t) \geq u_{\ell-1} \right\} \right) \\
 & \leq \mathbb{P}_{\mu_0^K} \left( \text{Mart}_{t_0^K}^{K,+} \geq u_{\ell-1} - 2u_{\ell} - 1 + \frac{bu_{\ell-1}}{K^2\sigma^2}t_0^K \right) \\
 & \leq \frac{\mathbb{E}_{\mu_0^K} \left( \left\langle \text{Mart}^{K,+} \right\rangle_{t_0^K} \right)}{(u_{\ell-1} - 1)^2} \\
 & \leq \frac{C_3}{K^\varepsilon}.
 \end{aligned}$$

Hence, for all  $K \geq \max\{K_1, K_2\}$ , we deduce that  $p \leq \varepsilon_K$  where

$$\varepsilon_K := \exp(-u_\ell + 1) \times \frac{1 - \exp(-2u_\ell - 1)}{1 - \exp(-u_{\ell+1})} + \frac{\exp(u_\ell) + \exp(-2u_\ell)}{\exp(u_\ell) - \exp(-2u_\ell)} \times \frac{C_1}{K^{\frac{\varepsilon}{2}}}$$

and thus, for all  $K \geq \max\{K_1, K_2\}$ ,

$$\mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \{L_1^K = \ell + 1\} \right) = \mathbb{P}_{\mu_0^K} \left( \{T_1^{K,+} < \check{\tau}^K\} \cap \{L_1^K = \ell + 1\} \right) \leq p \leq \varepsilon_K.$$

**Step 2. Control of the second right-hand term of (3.35).** Similarly, let us consider  $K_3 \in \mathbb{N}^*$  independent of  $\ell$  satisfying  $K_3^{\varepsilon/2} \geq 2$ . Using Corollary 3.4.12, DOOB's maximal inequality and Lemma 3.4.7 in similar way to Step 1, there exists a constant  $C_4 > 0$  such that for all  $K \geq K_3$  we have that

$$\mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \left\{ \sup_{0 \leq t \leq K\sigma^2 \wedge \check{\tau}^K} M_2(\mu_t^K) \geq u_{\ell+1} \right\} \right) \leq \frac{16\mathbb{E}_{\mu_0^K} \left( \langle M^{K, P_{\text{id}^2, 1}} \rangle_{K\sigma^2 \wedge \check{\tau}^K} \right)}{3^{2\ell} K^\varepsilon} \leq \frac{C_4}{K^{1+\varepsilon}}.$$

Let us consider  $K_4 \in \mathbb{N}^*$  independent of  $\ell$  satisfying the relation  $u_{\ell-1} > 1 + \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K_4}$ . As previously, we establish that there exists a constant  $C_5 > 0$  such that for all  $K \geq K_4$  we have that

$$\mathbb{P}_{\mu_0^K} \left( \{T_1^K < \check{\tau}^K\} \cap \left\{ \inf_{0 \leq t \leq K\sigma^2 \wedge \check{\tau}^K} M_2(\mu_t^K) < u_{\ell-1} \right\} \right) \leq \frac{4\mathbb{E}_{\mu_0^K} \left( \langle M^{K, P_{\text{id}^2, 1}} \rangle_{K\sigma^2 \wedge \check{\tau}^K} \right)}{\left( u_{\ell-1} - 1 - \frac{5\bar{b}u_{\ell+1} + \theta m_2}{2K} \right)^2} \leq \frac{C_5}{K^{1+\varepsilon}}.$$

To conclude this proof, there exists a constant  $C > 0$  such that for all  $K \geq K_0$  where  $K_0 := \max_{i \in \{1, \dots, 4\}} K_i$ ,

$$\mathbb{P}_{\mu_0^K} \left( \left( \{L_1^K = \ell + 1\} \cup \{T_1^K < K\sigma^2\} \right) \cap \{T_1^K < \check{\tau}^K\} \right) < \frac{1}{2} - \eta_K$$

for any  $\eta_K \geq \frac{1}{2} - \varepsilon_K - \frac{C_4 + C_5}{K^{1+\varepsilon}}$ . A convenient choice is given by  $\eta_K := \frac{1}{2} - \frac{C}{K^{\varepsilon/2}}$  for  $K$  large enough independent of  $\ell$  which completes the proof.  $\square$

### 3.4.4.2 Construction of the coupling

The goal of this section is to construct a coupling between  $(L_k^K)_{k \in \mathbb{N}}$  and a biased random walk on  $\mathbb{N}^*$  reflected in 1. To do this, we will construct a sequence  $(\xi_k^K)_{k \in \mathbb{N}}$  of i.i.d. random variables with values in  $\{-1, 1\}$  as follows.

**Step 1.** First of all, thanks to Lemma 3.4.11 we can construct the random variable  $\xi_0^K \in \{-1, 1\}$  such that

$$\{\xi_0^K = -1\} \subset \left[ \{L_1^K = L_0^K - 1\} \cap \{K\sigma^2 \leq T_1^K < \check{\tau}^K\} \right] \cup \{T_1^K \geq \hat{\tau}^K\}$$

and  $\mathbb{P}(\xi_0^K = -1) = \frac{1}{2} + \eta_K$ .

**Step 2.** Then, thanks to Lemma 3.4.11 again and after having applied the MARKOV property at time  $T_1^K$ , we can construct the random variable  $\xi_1^K \in \{-1, 1\}$  such that, conditionally to  $\mathcal{F}_{T_1^K} \vee \sigma(\xi_0^K)$ ,

$$\{\xi_1^K = -1\} \subset \left[ \{L_2^K = L_1^K - 1\} \cap \{K\sigma^2 + T_1^K \leq T_2^K < \check{\tau}^K\} \right] \cup \{T_2^K \geq \hat{\tau}^K\}$$

and  $\mathbb{P}\left(\xi_1^K = -1 \mid \mathcal{F}_{T_1^K} \vee \sigma(\xi_0^K)\right) = \frac{1}{2} + \eta_K$ . Note that this implies in particular that  $\xi_1^K$  is independent of  $\xi_0^K$ .

**Step 3.** By induction on  $k \in \mathbb{N}$ , if  $(\xi_i^K)_{0 \leq i \leq k}$  are constructed, then thanks to Lemma 3.4.11 again and after having applied the MARKOV property at time  $T_{k+1}^K$ , we can construct the random variable  $\xi_{k+1}^K \in \{-1, 1\}$  such that, conditionally to  $\mathcal{F}_{T_{k+1}^K} \vee \sigma(\xi_0^K, \dots, \xi_k^K)$ ,

$$\{\xi_{k+1}^K = -1\} \subset \left[ \{L_{k+2}^K = L_{k+1}^K - 1\} \cap \{K\sigma^2 + T_{k+1}^K \leq T_{k+2}^K < \check{\tau}^K\} \right] \cup \{T_{k+2}^K \geq \hat{\tau}^K\}$$

and  $\mathbb{P}\left(\xi_1^K = -1 \mid \mathcal{F}_{T_{k+1}^K} \vee \sigma(\xi_0^K, \dots, \xi_k^K)\right) = \frac{1}{2} + \eta_K$ .

To conclude, we set

$$\forall K \in \mathbb{N}^*, \quad \forall k \in \mathbb{N}, \quad Z_{k+1}^K - Z_k^K := \xi_k^K \quad \text{and} \quad Z_0^K = L_0^K \quad (3.38)$$

and the sequence  $(Z_k^K)_{k \in \mathbb{N}}$  satisfies by construction the following lemma:

**Lemma 3.4.13.** *The sequence  $(Z_k^K)_{k \in \mathbb{N}}$ , given by (3.38) is a biased simple random walk on  $\mathbb{N}^*$ , reflected in 1 and for all  $k \in \mathbb{N}$  as long as  $T_{k+1}^K < \check{\tau}^K$ , that*

- (1)  $Z_{k+1}^K - Z_k^K \geq L_{k+1}^K - L_k^K,$
- (2)  $T_{k+1}^K - T_k^K \geq K\sigma^2 \quad \text{when} \quad Z_{k+1}^K - Z_k^K = -1.$

On the one hand, note that for all  $k \in \mathbb{N}$ ,  $L_k^K \leq Z_k^K$ . On the other hand, note that Lemma 3.4.10 implies for all  $k \in \mathbb{N}$ ,  $L_{k+1}^K - L_k^K \in \{\dots, -3, -2, -1, 1\}$  which justifies (1) of Lemma 3.4.13. In Figure 3.2, we illustrate the coupling between  $(L_k^K)_{0 \leq k \leq 7}$  and  $(Z_k^K)_{0 \leq k \leq 7}$ . In this figure, we observe that  $\xi_k^K = L_{k+1}^K - L_k^K$  for all  $k \in \{0, 1, 2, 4, 5, 6\}$  and  $\xi_3^K = -1 \geq -2 = L_4^K - L_3^K$ . This illustrates that  $Z_{k+1}^K - Z_k^K$  is always 1 when  $L_{k+1}^K - L_k^K$  is equal to 1. Note that  $Z_{k+1}^K - Z_k^K$  can be 1 when  $L_{k+1}^K - L_k^K$  is equal to  $-1$ .

### 3.4.4.3 Estimates of exit from an attractive domain for random walks

Let us consider  $N^K$  the number of transitions before reaching  $\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$  for the random walk  $(Z_k^K)_{k \in \mathbb{N}}$ .

**Remark 3.4.14.** *By Lemma 3.4.13,  $N^K \leq k_0$  where  $k_0$  is the first integer such that  $L_{k_0}^K = \lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$ , so  $\hat{\tau}^K \geq T_{k_0}^K \geq T_{N^K}^K$ .*

The following lemma gives an estimate on the problem of exit from a domain for  $Z^K$ . For all  $k \in \mathbb{N}^*$ , we denote by  $\mathbf{P}_k$  the law of the MARKOV chain  $Z^K$  given  $Z_0^K = L_0^K = k$ .

**Lemma 3.4.15.** *We have,*

$$\lim_{K \rightarrow +\infty} \mathbf{P}_1 \left( N^K \geq \exp \left( \frac{\varepsilon^2}{16 \log(3)} \log^2(K) \right) \right) = 1.$$

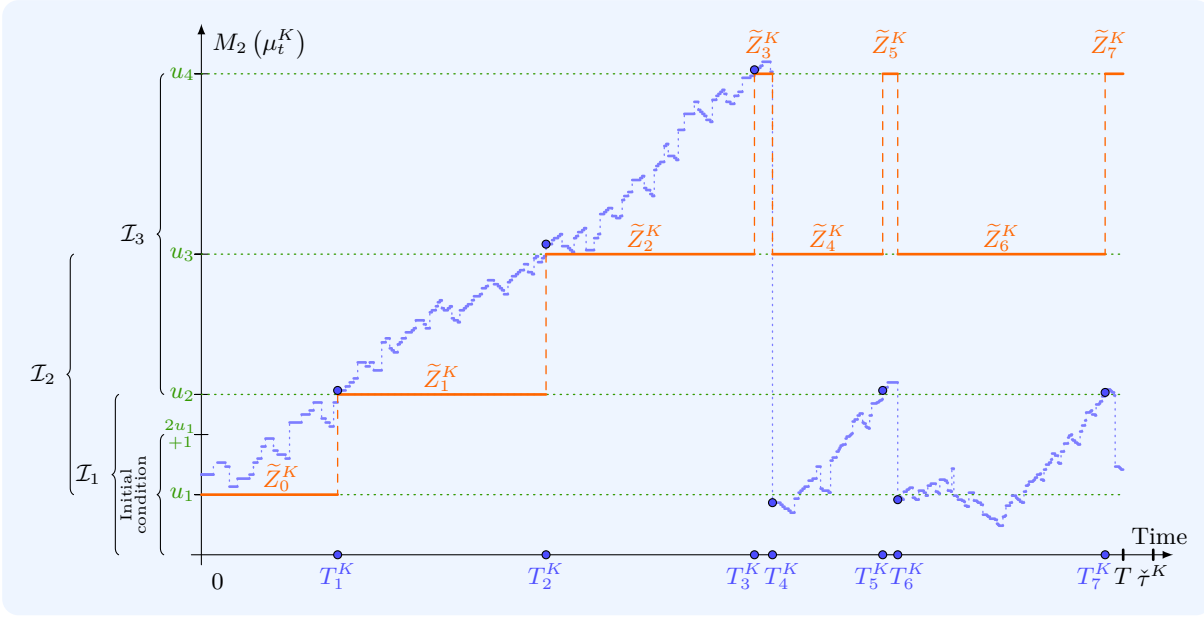


FIGURE 3.2 – For  $K$  large enough, coupling between  $(L_k^K)_{0 \leq k \leq 7}$  and  $(Z_k^K)_{0 \leq k \leq 7}$  up to time  $T$  and before  $\tilde{\tau}^K$  and where  $M_2(\mu_0^K) \in [0, 2u_1 + 1)$  so that  $L_0^K = 1$  and  $\tilde{Z}_k^K := u_{Z_k^K} = 3^{Z_k^K} K^{\varepsilon/2}$  for all  $k \in \mathbb{N}$ .

*Proof.* For all  $k \in \mathbb{N}^*$ , let us consider the stopping time  $\mathcal{T}_k^K := \inf \left\{ n \in \mathbb{N} \mid Z_n^K = k \right\}$  and we set  $v_k^K := \mathbf{P}_k \left( \mathcal{T}_1^K < \mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K \right)$ . Note that  $v_1^K = 1$ . By a first step analysis and MARKOV's property, we obtain that

$$\begin{aligned} v_k^K &= \mathbf{P}_k \left( \left\{ \mathcal{T}_1^K < \mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K \right\} \cap \{Z_1^K = k+1\} \right) + \mathbf{P}_k \left( \left\{ \mathcal{T}_1^K < \mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K \right\} \cap \{Z_1^K = k-1\} \right) \\ &= \mathbf{P}_k(Z_1^K = k+1) \mathbf{P}_{k+1} \left( \mathcal{T}_1^K < \mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K \right) + \mathbf{P}_k(Z_1^K = k-1) \mathbf{P}_{k-1} \left( \mathcal{T}_1^K < \mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K \right) \\ &= \left( \frac{1}{2} - \eta_K \right) v_{k+1}^K + \left( \frac{1}{2} + \eta_K \right) v_{k-1}^K \end{aligned}$$

The characteristic equation associated to the equation on  $v_k^K$  is

$$\left( \frac{1}{2} - \eta_K \right) X^2 - X + \left( \frac{1}{2} + \eta_K \right) = 0.$$

As  $\eta_K \in (0, \varepsilon/2)$ , the previous equation has exactly two distinct roots :

$$r_K := \frac{1 + 2\eta_K}{1 - 2\eta_K} > 1 \quad \text{and} \quad \tilde{r}_K := 1.$$

For all  $k \in \mathbb{N}^*$ , we look for  $v_k^K$  in the form  $v_k^K = ar_K^k + b\tilde{r}_K^k$  with  $a, b \in \mathbb{R}$  to be determined. As,  $v_1^K = 1$  and  $v_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K = 0$ , we deduce that

$$a = \frac{1}{r_K \left( 1 - r_K^{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1} \right)} \quad \text{and} \quad b = \frac{r_K^{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1}}{1 - r_K^{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1}}.$$



Hence, for all  $k \in \mathbb{N}$ ,

$$\begin{aligned} v_k^K &= \frac{1 - r_K^{-\left(\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - k\right)}}{1 - r_K^{-\left(\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1\right)}} = 1 - r_K^{-\left(\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1\right)} \frac{r_K^{k-1} - 1}{1 - r_K^{-\left(\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor - 1\right)}} \\ &\leq 1 - \left(r_K^{k-1} - 1\right) \exp\left(-\left(\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1\right) \log(r_K)\right). \end{aligned}$$

Note that  $N^K$  is greater than the number of transitions from 2 to 1 that occur before  $\mathcal{T}_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}^K$ .

Hence, under  $\mathbf{P}_1$ ,  $N^K \geq X$  where  $X$  is random variable with geometric law of parameter  $p_K := (r_K - 1) \exp\left(-\left(\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1\right) \log(r_K)\right)$ . Since  $r_K \sim_{K \rightarrow +\infty} \frac{K^{\varepsilon/2}}{C}$ , with  $C > 0$  given by Lemma 3.4.11,  $p_K \rightarrow 0$  when  $K \rightarrow +\infty$  and using that for all  $x > 0$  small enough,  $\log(1 - x) \geq -2x$  we deduce that for all  $K$  large enough,

$$\begin{aligned} \mathbf{P}_1 &\left(X \geq \exp\left(\frac{\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1}{2} \log(r_K)\right)\right) \\ &= \exp\left(\exp\left(\frac{\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1}{2} \log(r_K)\right) \log(1 - p_K)\right) \\ &\geq \exp\left(-2p_K \exp\left(-\frac{\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1}{2} \log(r_K)\right)\right). \end{aligned}$$

Using again that  $r_K \sim_{K \rightarrow +\infty} \frac{K^{\varepsilon/2}}{C}$ , we deduce that for  $K$  large enough,

$$\frac{\left\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \right\rfloor - 1}{2} \log(r_K) \geq \frac{\varepsilon^2}{16 \log(3)} \log^2(K)$$

and the announced result follows.  $\square$

It is in the next Corollary 3.4.16 that we see the importance of Assumption (3.5) where  $\sigma$  must not be too small.

**Corollary 3.4.16.** *For all  $T > 0$ ,*

$$\lim_{K \rightarrow +\infty} \mathbb{P}\left(\frac{K\sigma^2}{2} \left(N^K - \frac{\varepsilon}{2 \log(3)} \log(K)\right) < T\right) = 0.$$

*Proof.* From Lemma 3.4.15 and Assumption (3.5) we deduce in a straightforward manner that

$$\lim_{K \rightarrow +\infty} \mathbf{P}_1\left(\frac{K\sigma^2}{2} \left(N^K - \frac{\varepsilon}{2 \log(3)} \log(K)\right) < T\right) = 0.$$

As  $L_0^K$  depends on  $\mu_0^K$  note that  $L_0^K = Z_0^K$ , is function of the initial condition  $\mu_0^K$ . However, by Assumption (3.4) and MARKOV's inequality, we have that

$$\mathbb{P}\left(L_0^K > 1\right) \leq \mathbb{P}\left(M_2\left(\mu_0^K\right) \geq K^{\frac{\varepsilon}{2}}\right) \leq \frac{C_2^*}{K^\varepsilon}$$

which tends to 0 when  $K \rightarrow +\infty$  and the announced result follows.  $\square$

### 3.4.4.4 Conclusion

We denote respectively by  $N_+^K$  and  $N_-^K$  the number of upward transitions and downward transitions for  $(Z_k^K)_{k \in \mathbb{N}}$  before reaching  $\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$ . Let  $T \geq 0$  be fixed. Note that

$$\begin{aligned} \mathbb{P}(\hat{\tau}^K \leq \check{\tau}^K \wedge T) &\leq \mathbb{P}(\{\hat{\tau}^K \leq \check{\tau}^K \wedge T\} \cap \{\check{\tau}^K > T_{N^K}^K\}) \\ &\quad + \mathbb{P}(\{\hat{\tau}^K \leq \check{\tau}^K \wedge T\} \cap \{\check{\tau}^K \leq T_{N^K}^K\}). \end{aligned} \quad (3.39)$$

**Step 1. Control of the first right-hand term of (3.39).** As  $Z_0^K = 1$  with probability converging to 1 and  $Z_{N^K}^K = \lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$ , we have  $N_+^K - N_-^K \leq \lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor$  and  $N_+^K + N_-^K = N^K$ . Thus,

$$N_-^K \geq \frac{N^K - \lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor}{2}.$$

Hence, if  $\check{\tau}^K > T_{N^K}^K$ , then by Lemma 3.4.13 (2),

$$\hat{\tau}^K \geq T_{N^K}^K \geq \sum_{i=1}^{N^K} (T_i^K - T_{i-1}^K) \geq N_-^K K \sigma^2 \geq \frac{K \sigma^2}{2} \left( N^K - \frac{\varepsilon}{2 \log(3)} \log(K) \right).$$

Therefore,

$$\mathbb{P}(\{\hat{\tau}^K \leq \check{\tau}^K \wedge T\} \cap \{\check{\tau}^K > T_{N^K}^K\}) \leq \mathbb{P}\left(\frac{K \sigma^2}{2} \left( N^K - \frac{\varepsilon}{2 \log(3)} \log(K) \right) \leq T\right)$$

which tends to 0 when  $K \rightarrow +\infty$  according to Corollary 3.4.16.

**Step 2. Control of the second right-hand term of (3.39).** If  $\check{\tau}^K \leq T_{N^K}^K$ , let us consider  $\tilde{N}^K$  the last index  $k$  such that  $T_k^K < \check{\tau}^K$ . We have  $\tilde{N}^K < N^K$  and so  $\tilde{N}^K + 1 \leq \tilde{N}^K$  which implies, thanks to Lemma 3.4.10, that for all  $t \leq T_{\tilde{N}^K+1}^K$  and  $K$  large enough,

$$M_2(\mu_t^K) \leq 2u_{\lfloor \frac{\varepsilon}{2 \log(3)} \log(K) \rfloor} + 1 \leq 2K^{\frac{\varepsilon}{2}} + 1 < K^\varepsilon.$$

As the jump times of  $\nu^K$  are isolated, we deduce that  $\hat{\tau}^K > T_{\tilde{N}^K+1}^K$ . Now, by definition of  $\tilde{N}^K$ , we have  $T_{\tilde{N}^K+1}^K \geq \check{\tau}^K$  and thus  $\hat{\tau}^K > \check{\tau}^K$ . Hence, we obtain that

$$\mathbb{P}(\{\hat{\tau}^K \leq \check{\tau}^K \wedge T\} \cap \{\check{\tau}^K \leq T_{N^K}^K\}) = \mathbb{P}(\emptyset) = 0.$$

Therefore,  $\lim_{K \rightarrow +\infty} \mathbb{P}(\hat{\tau}^K \leq \check{\tau}^K \wedge T) = 0$  which concludes the proof of Lemma 3.4.9.

## 3.5 Tightness on the torus

The main result of this section is given by Theorem 3.5.1. This is a stochastic averaging result inspired by KURTZ (see Theorem 1.2.19 of Chapter 1) establishing a tightness result, in the torus case, of the sequence of laws of  $((z^K, \Gamma^K))_{K \in \mathbb{N}^*}$  where  $\Gamma^K$  is the occupation measure of the fast component  $\mu^K$ . We will use criteria proposed by ETHIER-KURTZ in [94, Theorems 3.9.1 and 3.9.4]. In Sections 3.6 and 3.7 we identify the limit and we prove its uniqueness.

Let introduce the torus  $\mathbb{T}_R := [x_0 - 2R, x_0 + 2R]$  of length  $4R$  with  $R > 0$  fixed and  $x_0$  is the value of the mean trait of  $\nu_0^K$  as in Theorem 3.2.1. We define  $\mathcal{M}_1(\mathbb{T}_R)$  the set of probability measure on  $\mathbb{T}_R$ . We denote by  $\mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$  the set of measures  $\Gamma$  on  $\mathbb{R}_+ \times \mathcal{M}_1(\mathbb{R})$  such that for all  $t \geq 0$ ,  $\Gamma([0, t] \times \mathcal{M}_1(\mathbb{R})) = t$ . For any  $t \geq 0$ , we denote by  $\mathcal{M}_m^t(\mathcal{M}_1(\mathbb{R}))$  the set of measures  $\Gamma \in \mathcal{M}_m(\mathcal{M}_1(\mathbb{R}))$  restricted to  $[0, t] \times \mathcal{M}_1(\mathbb{R})$ . For all  $T > 0$ , we denote by  $\mathbb{D}([0, T], \mathbb{T}_R)$  the space of càd-làg functions on  $[0, T]$  with values in  $\mathbb{T}_R$ .

Let  $b_R \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{R})$  and  $\theta_R \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{R})$  be two functions satisfying  $b_R = b$  on  $[x_0 - R, x_0 + R]^2$  and  $\theta_R = \theta$  on  $[x_0 - R, x_0 + R]$ . Let  $T > 0$  be fixed. We define on  $\mathbb{T}_R$ , in similar way as in Section 3.2, the processes  $\nu^{K,R} := (\nu_t^{K,R})_{t \in [0, T]}$ ,  $z^{K,R} := (z_t^{K,R})_{t \in [0, T]}$  and  $\mu^{K,R} = (\mu_t^{K,R})_{t \in [0, T]}$  as follows

$$\nu_t^{K,R} := \frac{1}{K} \sum_{i=1}^K \delta_{x_i^R(t)}, \quad z_t^{K,R} := \langle \text{id}, \nu_{t/K\sigma^2}^{K,R} \rangle \quad \text{and} \quad \mu_t^{K,R} := \left( h_{\frac{1}{\sigma\sqrt{K}}} \circ \tau_{-z_t^K} \right) \# \nu_{t/K\sigma^2}^{K,R}$$

where for all  $i \in \{1, \dots, K\}$ ,  $\nu_t^{K,R}$  is defined on  $\mathcal{M}_1(\mathbb{T}_R)$  from  $b_R$  and  $\theta_R$  as  $\nu_t^K$  was defined from  $b$  and  $\theta$  and  $z_t^{K,R} \in \mathbb{T}_R$ . Note that  $\mu_t^{K,R}$  takes values in  $\mathcal{M}_1(\mathbb{T}_R^{K,\sigma})$  where  $\mathbb{T}_R^{K,\sigma}$  is the torus corresponding to the interval  $\left[ \frac{1}{\sigma\sqrt{K}}(x_0 - 2R - z_t^K), \frac{1}{\sigma\sqrt{K}}(x_0 + 2R - z_t^K) \right]$ . However, we will identify in the sequel  $\mathcal{M}_1(\mathbb{T}_R^{K,\sigma})$  as a subset of  $\mathcal{M}_1(\mathbb{R})$  using the natural embedding of  $\mathbb{T}_R^{K,\sigma}$  in  $\mathbb{R}$ . So,  $\mu_t^{K,R} \in \mathcal{M}_1(\mathbb{R})$ . We define also the stopping times  $\tau^{K,R} := \check{\tau}^{K,R} \wedge \hat{\tau}^{K,R}$  where

$$\hat{\tau}^{K,R} := \inf \left\{ t \geq 0 \mid M_2(\mu_t^{K,R}) \geq K^\varepsilon \right\} \quad \text{and} \quad \check{\tau}^{K,R} := \inf \left\{ t \geq 0 \mid \text{Diam}(\text{Supp } \nu_t^{K,R}) > \frac{1}{\sigma K^{\frac{3+\varepsilon}{2}}} \right\}.$$

We define on  $\mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  the sequence of random measures  $(\Gamma^{K,R})_{K \in \mathbb{N}^*}$  as follows

$$\Gamma^{K,R}(dt, d\mu) = \delta_{\mu_t^{K,R}}(d\mu) dt.$$

In the sequel, we study the limit as  $K$  tends to  $+\infty$  of the sequence of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$ .

**Theorem 3.5.1.** *Let  $T, R > 0$ . The sequence of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  is tight in the set of probability measures on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  and for any limiting value  $\mathbb{Q}$  of this sequence, the canonical process  $(\zeta^R, \Gamma^R)$  on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  satisfies that for all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ ,*

$$N_t^{f,R} := f(\zeta_t^R) - f(\zeta_0^R) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\zeta_s^R, \mu) \Gamma^R(ds, d\mu) = 0 \quad (3.40)$$

$\mathbb{Q}$ -a.s. Moreover, for all  $t \in [0, T]$ ,

$$\mathbb{E}_{\mathbb{Q}} \left( \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} M_5(\mu) \Gamma^R(ds, d\mu) \right) < \infty, \quad (3.41)$$

so that the definition (3.40) makes sense.

This result is based on Propositions 3.3.1 and 3.4.8, Lemma 3.4.3. The proof is divided in eight steps. In Step 1, we establish the tightness of the family of laws of the stopped slow component  $(z_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$ . In Step 2, we establish a compact containment condition for the

stopped *fast* component  $(\mu_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$ . In Step 3, we prove the tightness of the family of laws of the occupation measure. In Step 4, we deduce the tightness of the family of laws of the couple (slow, occupation measure fast). In Step 5, we prove uniform integrability results for a family  $(N_{t \wedge \tau^{K,R}}^{f,K,R})_{t \in [0,T], K \in \mathbb{N}^*}$  constructed from  $z^{K,R}$  and  $\Gamma^{K,R}$  similarly as  $N_t^{f,R}$  in (3.40). Step 6 is devoted to establish the convergence in distribution of  $(N_{\bullet \wedge \tau^{K,R}}^{f,K,R})_{K \in \mathbb{N}^*}$  to  $N^{f,R}$ . In Steps 7 and 8, we prove that the limit  $N^{f,R}$  is null  $\mathbb{Q}$ -a.s.

The main modification of KURTZ's setting of Theorem 1.2.19 is that we have to work with stopped times and need to be careful with moment estimates and uniform integrability properties. This led us to rewrite the proof.

**Proof of Theorem 3.5.1. Step 1. Tightness of the family of laws of  $(z_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$  on  $\mathbb{D}([0, T], \mathbb{T}_R)$ .** Let  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ . For all  $K \in \mathbb{N}^*$  and for all  $t \in [0, T]$ , let us consider the two processes  $Y_t^{K,R} := f(z_{t \wedge \tau^{K,R}}^{K,R})$  and  $Z_t^{K,R}$  defined by the relation

$$Y_t^{K,R} = Y_0^{K,R} + \int_0^t Z_s^{K,R} ds + M_{t \wedge \tau^{K,R}}^{f,K,R}$$

where  $M_{t \wedge \tau^{K,R}}^{f,K,R}$  is the martingale given by (3.18) in the torus case. Note that, for all  $s > \tau^{K,R}$ ,  $Z_s^{K,R} = 0$ . From Proposition 3.3.2, (3.33) and Lemma 3.4.3, note that there exists a constant  $C > 0$  such that

$$\begin{aligned} & \sup_{K \in \mathbb{N}^*} \mathbb{E} \left[ \left( \int_0^T |Z_t^{K,R}|^2 dt \right)^{\frac{1}{2}} \right] \\ & \leq \sup_{K \in \mathbb{N}^*} \mathbb{E} \left[ 1 + \int_0^T |Z_t^{K,R}|^2 dt \right] \\ & \leq C \left\{ 1 + \sup_{K \in \mathbb{N}^*} \mathbb{E} \left[ \int_0^{T \wedge \tau^{K,R}} \left( M_2^2(\mu_t^{K,R}) + \frac{1}{K^2} [1 + M_2^2(\mu_t^{K,R})] \right) dt \right] \right\} \\ & < \infty. \end{aligned}$$

Hence, from [94, Theorem 3.9.4], the family of laws of  $(Y^{K,R})_{K \in \mathbb{N}^*}$ , on  $\mathbb{D}([0, T], \mathbb{T}_R)$ , is tight. Let us observe that the compact containment condition is satisfied by the stopped *slow* component  $(z_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$  since  $\mathbb{T}_R$  is compact and

$$\forall T > 0, \quad \inf_{K \in \mathbb{N}^*} \mathbb{P} \left( \sup_{0 \leq t \leq T} z_{t \wedge \tau^{K,R}}^{K,R} \in \mathbb{T}_R \right) = 1. \quad (3.42)$$

As  $\mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$  is a dense subset of  $\mathcal{C}_b^0(\mathbb{T}_R, \mathbb{T}_R)$  in the topology of the uniform norm, we deduce from (3.42) and [94, Theorem 3.9.1] that the family of laws of  $(z_{\bullet \wedge \tau^{K,R}}^{K,R})_{K \in \mathbb{N}^*}$  is tight on  $\mathbb{D}([0, T], \mathbb{T}_R)$ .

### Step 2. Compact containment condition.

**Lemma 3.5.2.** *Let  $R > 0$ . For all  $T \geq 0$ , the family of laws of the marginal random variables of the stopped fast process  $(\mu_{t \wedge \tau^{K,R}}^{K,R})_{t \in [0,T], K \in \mathbb{N}^*}$  is tight on  $\mathcal{M}_1(\mathbb{R})$  i.e.*

$$\forall \eta > 0, \exists D_{\varepsilon, \eta}^{R,T} \subset \mathcal{M}_1(\mathbb{R}) \text{ compact, } \forall t \in [0, T], \forall K \in \mathbb{N}^*, \quad \mathbb{P} \left( \mu_{t \wedge \tau^{K,R}}^{K,R} \in D_{\varepsilon, \eta}^{R,T} \right) \geq 1 - \eta.$$

*Proof.* Let  $\eta > 0$  be fixed. From Proposition 3.4.8, there exists  $K_0 \in \mathbb{N}^*$  large enough such that for all  $K \geq K_0$ , for all  $t \in [0, T]$ ,  $\mathbb{P}(t > \tau^{K,R}) \leq \frac{\eta}{2}$ . We consider the  $\mathbb{R}$ -valued sequence  $(a_{p,\eta})_{p \in \mathbb{N}}$  satisfying for all  $p \in \mathbb{N}$ ,  $a_{p,\eta} > 1$ , and  $\sum_{p \in \mathbb{N}^*} \frac{p}{a_{p,\eta}^2} < \frac{\eta}{2M_2}$  where  $\overline{M}_2 := \max\{C_2^*, \overline{\theta} \overline{m}_2 / \underline{b}\}$  is a uniform upper bound of  $\mathbb{E}(M_2(\mu_{t \wedge \tau^{K,R}}^{K,R}))$  given by Corollary 3.4.6. Let  $(\kappa_{p,\eta})_{p \in \mathbb{N}}$  be a sequence of compact intervals on  $\mathbb{R}$ , increasing for inclusion, of the form  $[-a_{p,\eta}, a_{p,\eta}]$ . Let  $t \in [0, T]$ . For all  $p \in \mathbb{N}^*$ , and  $K \geq K_0$ , note that

$$M_2(\mu_t^{K,R}) \mathbb{1}_{t \leq \tau^{K,R}} \geq \mathbb{1}_{t \leq \tau^{K,R}} \int_{\kappa_{p,\eta}^c} x^2 \mu_t^{K,R}(dx) \geq a_{p,\eta}^2 \mu_{t \wedge \tau^{K,R}}^{K,R}(\kappa_{p,\eta}^c) \mathbb{1}_{t \leq \tau^{K,R}},$$

and we deduce that  $\{\mu_{t \wedge \tau^{K,R}}^{K,R}(\kappa_{p,\eta}^c) > \frac{1}{p}\} \cap \{t \leq \tau^{K,R}\} \subset \{M_2(\mu_{t \wedge \tau^{K,R}}^{K,R}) > \frac{a_{p,\eta}^2}{p}\} \cap \{t \leq \tau^{K,R}\}$ . Hence,

$$\begin{aligned} \mathbb{P}\left(\left\{\exists p \in \mathbb{N}^*, \mu_{t \wedge \tau^{K,R}}^{K,R}(\kappa_{p,\eta}^c) > \frac{1}{p}\right\}\right) &\leq \mathbb{P}(\tau^{K,R} < t) + \sum_{p \in \mathbb{N}^*} \mathbb{P}\left(\left\{\mu_{t \wedge \tau^{K,R}}^{K,R}(\kappa_{p,\eta}^c) > \frac{1}{p}\right\} \cap \{t \leq \tau^{K,R}\}\right) \\ &\leq \mathbb{P}(\tau^{K,R} < t) + \sum_{p \in \mathbb{N}^*} \mathbb{P}\left(M_2(\mu_{t \wedge \tau^{K,R}}^{K,R}) \mathbb{1}_{t \leq \tau^{K,R}} > \frac{a_{p,\eta}^2}{p}\right) \\ &\leq \mathbb{P}(\tau^{K,R} < t) + \sum_{p \in \mathbb{N}^*} \frac{\mathbb{E}(M_2(\mu_{t \wedge \tau^{K,R}}^{K,R}) \mathbb{1}_{t \leq \tau^{K,R}})}{\frac{a_{p,\eta}^2}{p}} \\ &\leq \eta. \end{aligned}$$

Therefore, we have proved that

$$\forall \eta > 0, \forall t \in [0, T], \forall K \geq K_0, \quad \mathbb{P}(\mu_{t \wedge \tau^{K,R}}^{K,R} \in \mathcal{K}_\eta) \geq 1 - \eta,$$

where  $\mathcal{K}_\eta := \left\{ \mu \in \mathcal{M}_1(\mathbb{R}) \mid \forall p \in \mathbb{N}^*, \mu(\kappa_{p,\eta}^c) \leq \frac{1}{p} \right\}$  which is compact by PROHOROV's theorem.  $\square$

**Step 3. Tightness of the family of laws of  $(\Gamma^{K,R})_{K \in \mathbb{N}^*}$  on  $\mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ .** Let  $\eta > 0$  be fixed. From Proposition 3.4.8, there exists  $K_0 \in \mathbb{N}^*$  such that for all  $K \geq K_0$ , for all  $t \in [0, T]$ ,  $\mathbb{P}(t \geq \tau^{K,R}) \leq \frac{\eta}{2}$ . Consider  $D_{\varepsilon, \frac{\eta}{2}}^{R,T}$  the compact set in Lemma 3.5.2. It follows that for all  $t \in [0, T]$ ,  $K \geq K_0$ ,

$$\begin{aligned} \mathbb{E}\left(\Gamma^{K,R}\left([0, t] \times D_{\varepsilon, \frac{\eta}{2}}^{R,T}\right)\right) &\geq \mathbb{E}\left(\Gamma^{K,R}\left([0, t] \times D_{\varepsilon, \frac{\eta}{2}}^{R,T}\right) \mathbb{1}_{t < \tau^{K,R}}\right) \\ &= \int_0^t \mathbb{P}\left(\left\{\mu_{s \wedge \tau^{K,R}}^{K,R} \in D_{\varepsilon, \frac{\eta}{2}}^{R,T}\right\} \cap \{t < \tau^{K,R}\}\right) ds \\ &\geq \int_0^t \mathbb{P}\left(\mu_{s \wedge \tau^{K,R}}^{K,R} \in D_{\varepsilon, \frac{\eta}{2}}^{R,T}\right) ds - \int_0^t \mathbb{P}(t \geq \tau^{K,R}) ds \\ &\geq t(1 - \eta). \end{aligned}$$

Therefore, the tightness of the family of laws of  $(\Gamma^{K,R})_{K \in \mathbb{N}^*}$  follows from Lemma 1.2.17 of Chapter 1.

**Step 4. Tightness of the family of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$ .** From Steps 1 and 3 and PROHOROV's theorem, we deduce that the family of laws

of the couple  $\left( \left( z_{\bullet \wedge \tau^{K,R}}^{K,R}, \Gamma^{K,R} \right) \right)_{K \in \mathbb{N}^*}$  is relatively compact in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$ . Hence, there exists a probability measure  $\mathbb{Q}$  on the canonical space  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  and an increasing function  $n : \mathbb{N}^* \rightarrow \mathbb{N}^*$  such that the subsequence of laws of  $\left( \left( z_{\bullet \wedge \tau^{n(K),R}}^{n(K),R}, \Gamma^{n(K),R} \right) \right)_{K \in \mathbb{N}^*}$  converges weakly to the limiting value  $\mathbb{Q}$  when  $K \rightarrow +\infty$ . Thanks to Proposition 3.4.8, we deduce that the family of laws of  $\left( \left( z^{n(K),R}, \Gamma^{n(K),R} \right) \right)_{K \in \mathbb{N}^*}$  converges weakly to  $\mathbb{Q}$  when  $K \rightarrow +\infty$  and therefore that the family of laws of  $\left( \left( z^{K,R}, \Gamma^{K,R} \right) \right)_{K \in \mathbb{N}^*}$  is relatively compact, thus tight on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  by PROHOROV's theorem.

**Step 5. Uniform integrability.** For all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ ,  $K \in \mathbb{N}^*$ , let us consider  $\left( N_t^{f,K,R} \right)_{t \in [0, T]}$  the stochastic process defined by

$$N_t^{f,K,R} := f(z_t^{K,R}) - f(z_0^{K,R}) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(z_s^{K,R}, \mu) \Gamma^{K,R}(ds, d\mu).$$

From (3.9), we have for all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ , there exists a constant  $C > 0$  such that

$$\forall t \in [0, T], \forall K \in \mathbb{N}^*, \quad \left| N_{t \wedge \tau^{K,R}}^{f,K,R} \right| \leq C \left( 1 + \int_0^t M_2(\mu_s^{K,R}) \mathbf{1}_{s \leq \tau^{K,R}} ds \right),$$

Hence, the uniform integrability of  $\left( N_{\bullet \wedge \tau^{K,R}}^{f,K,R} \right)_{K \in \mathbb{N}^*}$  follows from Corollary 3.4.4.

**Step 6. Proof of (3.41) and almost sure convergence of  $\left( \tilde{N}_{\bullet \wedge \tau^{K,R}}^{f,K,R} \right)_{K \in \mathbb{N}^*}$  to  $\tilde{N}^{f,R}$ .** From Step 4, Proposition 3.4.8 and SKOROHOD's representation theorem, there exists an increasing function  $\bar{n} : \mathbb{N}^* \rightarrow \mathbb{N}^*$  and a probability space on which we define, the random variable  $\tilde{\tau}^{\bar{n}(K),R}$ , the families  $\left( \tilde{z}^{\bar{n}(K),R} \right)_{K \in \mathbb{N}^*}$  and  $\left( \tilde{\Gamma}^{\bar{n}(K),R} \right)_{K \in \mathbb{N}^*}$  and  $\tilde{\zeta}^R, \tilde{\Gamma}^R$  copies of  $\left( z^{\bar{n}(K),R} \right)_{K \in \mathbb{N}^*}$ ,  $\left( \Gamma^{\bar{n}(K),R} \right)_{K \in \mathbb{N}^*}, \zeta^R, \Gamma^R$  under  $\mathbb{Q}$  such that the sequence  $\left( \left( \tilde{z}^{\bar{n}(K),R}, \tilde{\Gamma}^{\bar{n}(K),R}, \tilde{\tau}^{\bar{n}(K),R} \right) \right)_{K \in \mathbb{N}^*}$  converges a.s. to  $\left( \tilde{\zeta}^R, \tilde{\Gamma}^R, +\infty \right)$  when  $K \rightarrow +\infty$ . Note that for all  $t \in [0, T]$ ,

$$\tilde{\Gamma}^{\bar{n}(K),R} \left( [0, t \wedge \tilde{\tau}^{\bar{n}(K),R}] \times \mathcal{M}_1(\mathbb{R}) \right) = t \wedge \tilde{\tau}^{\bar{n}(K),R} \xrightarrow[K \rightarrow +\infty]{\text{a.s.}} t = \tilde{\Gamma}^R([0, t] \times \mathcal{M}_1(\mathbb{R})).$$

From Lemma 1.2.18 (b)(c)(d) of Chapter 1 and Corollary 3.4.4, we have

$$\int_0^{t \wedge \tilde{\tau}^{K,R}} \int_{\mathcal{M}_1(\mathbb{R})} M_5(\mu) \tilde{\Gamma}^{K,R}(ds, d\mu) \xrightarrow[K \rightarrow +\infty]{\text{a.s.}} \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} M_5(\mu) \tilde{\Gamma}^R(ds, d\mu)$$

and (3.41) follows. From Lemma 1.2.18 (b)(c)(d), we also deduce that for all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ , for all  $t \in [0, T]$ , a.s.

$$\begin{aligned} \lim_{K \rightarrow +\infty} \int_0^{t \wedge \tilde{\tau}^{\bar{n}(K),R}} \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\tilde{z}_s^{\bar{n}(K),R}, \mu) \tilde{\Gamma}^{\bar{n}(K),R}(ds, d\mu) \\ = \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\tilde{\zeta}_s^R, \mu) \tilde{\Gamma}^R(ds, d\mu) \end{aligned}$$

and thus the sequence  $\left( \tilde{N}_{\bullet \wedge \tilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} \right)_{K \in \mathbb{N}^*}$  converges a.s. to the process  $\left( \tilde{N}_t^{f,R} \right)_{t \in [0, T]}$  where  $\tilde{N}_t^{f,K,R}$  and  $\tilde{N}_t^{f,R}$  are respectively defined for all  $K \in \mathbb{N}^*$  by

$$\begin{aligned} \tilde{N}_t^{f,K,R} &:= f(\tilde{z}_t^{K,R}) - f(\tilde{z}_0^{K,R}) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\tilde{z}_s^{K,R}, \mu) \tilde{\Gamma}^{K,R}(ds, d\mu), \quad t \leq T \wedge \tilde{\tau}^{K,R} \\ \tilde{N}_t^{f,R} &:= f(\tilde{\zeta}_t^R) - f(\tilde{\zeta}_0^R) - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\tilde{\zeta}_s^R, \mu) \tilde{\Gamma}^R(ds, d\mu), \quad t \leq T. \end{aligned}$$

**Step 7.  $N^{f,R}$  is a martingale.** Let us consider the filtration  $(\widetilde{\mathcal{F}}_t^{K,R})_{t \in [0,T]}$  defined by  $\widetilde{\mathcal{F}}_t^{K,R} := \sigma\left(\widetilde{z}_s^{K,R}, \widetilde{\Gamma}^{K,R}([0,s] \times H) \mid s \leq t, H \in \mathcal{B}(\mathcal{M}_1(\mathbb{R}))\right)$ ,  $(f_i)_{1 \leq i \leq p}$ ,  $p \in \mathbb{N}^*$  bounded LIPSCHITZ functions from  $\mathbb{T}_R$  to  $\mathbb{T}_R$  and  $0 < t_1 \leq \dots \leq t_p \leq s < t$ . Let us denote for all  $K \in \mathbb{N}^*$ , for all  $t \leq \widetilde{\tau}^{K,R} \wedge T$ ,  $\widetilde{M}_t^{f,K,R} := \widetilde{N}_t^{f,K,R} + \widetilde{\mathcal{E}}_t^{f,K,R}$  where  $\widetilde{M}_t^{f,K,R}$  is constructed from  $\widetilde{z}^{K,R}$  and  $\widetilde{\mu}^{K,R}$  as in (3.18) and  $\widetilde{\mathcal{E}}_t^{f,K,R}$  is an error term. Note that  $\widetilde{M}_t^{f,K,R}$  is a  $(\widetilde{\mathcal{F}}_t^{K,R})_{t \in [0,T]}$ -martingale as in Proposition 3.3.1. Hence,

$$\mathbb{E}\left(\widetilde{M}_{t \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} \prod_{i=1}^p f_i\left(\widetilde{M}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right)\right) = \mathbb{E}\left(\widetilde{M}_{s \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} \prod_{i=1}^p f_i\left(\widetilde{M}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right)\right).$$

From Proposition 3.3.2 and (3.33),  $\widetilde{\mathcal{E}}_{t \wedge \widetilde{\tau}^{K,R}}^{f,K,R} = \frac{1}{K} O\left(\int_0^t [1 + M_2(\widetilde{\mu}_s^{K,R})] \mathbb{1}_{s \leq \widetilde{\tau}^{K,R}} ds\right)$  and then satisfy the condition  $\lim_{K \rightarrow +\infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} \left|\mathcal{E}_{t \wedge \widetilde{\tau}^{K,R}}^{f,K,R}\right|\right) = 0$ . As for all  $i \in \{1, \dots, p\}$ ,  $f_i$  is LIPSCHITZ, there exists a constant  $C > 0$  such that

$$\begin{aligned} & \mathbb{E}\left(\left|f_i\left(\widetilde{M}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} - \mathcal{E}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right) - f_i\left(\widetilde{M}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right)\right|\right) \\ & \leq \frac{C}{\bar{n}(K)} \mathbb{E}\left(\int_0^{t_i} [1 + M_2(\mu_r^{\bar{n}(K),R})] \mathbb{1}_{r \leq \widetilde{\tau}^{\bar{n}(K),R}} dr\right) \end{aligned}$$

where the term of the right hand side of the previous inequality goes to 0 when  $K \rightarrow +\infty$  according to Lemma 3.4.3. We deduce that for all  $u \in \{s, t\}$ ,

$$\lim_{K \rightarrow +\infty} \mathbb{E}\left(\widetilde{M}_{u \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} \prod_{i=1}^p f_i\left(\widetilde{M}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right)\right) = \lim_{K \rightarrow +\infty} \mathbb{E}\left(\widetilde{N}_{u \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R} \prod_{i=1}^p f_i\left(\widetilde{N}_{t_i \wedge \widetilde{\tau}^{\bar{n}(K),R}}^{f,\bar{n}(K),R}\right)\right).$$

From Steps 5 and 6, we deduce that

$$\mathbb{E}\left(\widetilde{N}_t^{f,R} \prod_{i=1}^p f_i\left(\widetilde{N}_{t_i}^{f,R}\right)\right) = \mathbb{E}\left(\widetilde{N}_s^{f,R} \prod_{i=1}^p f_i\left(\widetilde{N}_{t_i}^{f,R}\right)\right).$$

Hence,

$$\mathbb{E}\left(N_t^{f,R} \prod_{i=1}^p f_i\left(N_{t_i}^{f,R}\right)\right) = \mathbb{E}\left(N_s^{f,R} \prod_{i=1}^p f_i\left(N_{t_i}^{f,R}\right)\right).$$

Since the last property is true for all  $p \in \mathbb{N}^*$ ,  $t_1 \leq \dots \leq t_p \leq s < t$  and for all bounded LIPSCHITZ functions  $(f_i)_{1 \leq i \leq p}$ , the monotone class theorem ensures us that

$$\mathbb{E}\left(N_t^{f,R} \mid \sigma\left(N_u^{f,R} \mid u \leq s\right)\right) = N_s^{f,R}.$$

Hence the announced result.

**Step 8. Nullity of  $N^{f,R}$ .** On the one hand, from ITÔ's formula [200, Theorem 32 of Chapter II], for all  $f \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$  and  $t \in [0, T]$

$$\begin{aligned} f^2(\zeta_t^R) &= f^2(\zeta_0^R) + 2 \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} f(\zeta_s^R) \mathcal{L}_{\text{SLOW}} f(\zeta_s^R, \mu) \Gamma^R(ds, d\mu) \\ &+ 2 \int_0^t f(\zeta_s^R) dN_s^{f,R} + \langle N^{f,R} \rangle_t + \sum_{0 < s \leq t} \left(f(\zeta_s^R) - f(\zeta_{s-}^R)\right)^2. \end{aligned} \tag{3.43}$$

On the other hand, applying (3.40) with  $f^2 \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ , we obtain that for all  $t \in [0, T]$

$$f^2(\zeta_t^R) = f^2(\zeta_0^R) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f^2(\zeta_s^R, \mu) \Gamma^R(ds, d\mu) + N_t^{f^2, R}. \quad (3.44)$$

Comparing (3.43) and (3.44) leads for all  $t \in [0, T]$  to

$$N_t^{f^2, R} - 2 \int_0^t f(\zeta_s^R) dN_s^{f, R} = \langle N^{f, R} \rangle_t + \sum_{0 < s \leq t} (f(\zeta_s^R) - f(\zeta_{s-}^R))^2$$

and thus by [150, Theorem 4.1] that  $\mathbb{Q}$ -a.s. for all  $t \in [0, T]$ ,

$$\langle N^{f, R} \rangle_t = - \sum_{0 < s \leq t} (f(\zeta_s^R) - f(\zeta_{s-}^R))^2 \leq 0$$

so that  $\langle N^{f, R} \rangle_t = 0$ . Hence,  $\mathbb{Q}$ -a.s.  $N^{f, R} = 0$  which completes the proof.  $\square$

### 3.6 Characterisation of the occupation measure limit on the torus

Consider a probability measure  $\mathbb{Q}$  on  $\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  and the canonical process  $(\zeta^R, \Gamma^R)$  as in Theorem 3.5.1. The following lemma gives us a desintegration result of the occupation measure  $\Gamma^R$  that we characterise below.

**Lemma 3.6.1.** *Let  $T \geq 0$  be fixed. With the notations of Theorem 3.5.1, there exists a random probability measure-valued process  $(\gamma_t^R)_{t \in [0, T]}$  that is predictable in  $(\omega, t)$  and such that for all bounded measurable function  $\psi : [0, T] \times \mathcal{M}_1(\mathbb{R}) \rightarrow \mathbb{T}_R$ ,*

$$\int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s, \mu) \Gamma^R(ds, d\mu) = \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \psi(s, \mu) \gamma_s^R(d\mu) ds. \quad (3.45)$$

*Proof.* The desintegration result of  $\Gamma^R$  follows directly from Lemma 1.2.20 of Chapter 1.  $\square$

**Corollary 3.6.2.** *Let  $T \geq 0$  be fixed. With the notations of Theorem 3.5.1, we have that  $\zeta^R \in \mathcal{C}^0([0, T], \mathbb{T}_R)$  is differentiable of derivative in  $L^1(\mathbb{R})$   $\mathbb{Q}$ -a.s.*

*Proof.* Applying (3.40) with  $f = \text{id} \in \mathcal{C}_b^1(\mathbb{T}_R, \mathbb{T}_R)$ , we deduce from Theorem 3.5.1, (3.9) and (3.45),  $\mathbb{Q}$ -a.s., for all  $t \in [0, T]$

$$\zeta_t^R = \zeta_0^R + \int_0^t \left( \int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \gamma_s^R(d\mu) \right) \partial_1 \text{Fit}(\zeta_s^R, \zeta_s^R) ds.$$

From Assumptions **(A)** and (3.41), the integrand of the previous time integral is in  $L^1(\mathbb{R})$   $\mathbb{Q}$ -a.s. Hence, the announced result follows from the fundamental theorem of calculus.  $\square$

We now want to characterise the limiting value  $\Gamma^R(dt, d\mu) = \gamma_t^R(d\mu) dt$  under  $\mathbb{Q}$ .

**Proposition 3.6.3.** *With the notations of Theorem 3.5.1, for a.e.  $t \in [0, T]$ ,  $\mathbb{Q}$ -a.s.,  $\gamma_t^R = \pi^\lambda(\zeta_t^R)$  where  $\pi^\lambda$  is the unique invariant probability measure of the centered FLEMING-VIOT process with resampling rate  $\lambda$  (see Section 2.4).*



It is here that we exploit ergodicity properties for the limit fast component. The proof of Proposition 3.6.3, given in Section 3.6.1 is based on the following technical lemma giving a characterisation of  $\pi^\lambda$  and proved in Section 3.6.2. To state this lemma, let us first recall from (3.21) the definition of polynomials in  $\mu$ :

$$P_{f,n}(\mu) := \langle f, \mu^n \rangle := \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) \mu(dx_1) \cdots \mu(dx_n)$$

with  $n \in \mathbb{N}^*$ ,  $\mu \in \mathcal{M}_1^c(\mathbb{R})$ ,  $f \in \mathcal{C}_b^3(\mathbb{R}^n, \mathbb{R})$ . For all  $n \in \mathbb{N}^*$ , for any function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  whose second derivatives exist, we denote by  $\text{Hess}(f) := \left( \partial_{ij}^2 f \right)_{1 \leq i, j \leq n}$  the Hessian matrix of  $f$ . For all  $n \in \mathbb{N}^*$ , let us denote by  $\mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^n, \mathbb{R})$  the set

$$\left\{ f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}) \mid \exists C > 0, \forall x \in \mathbb{R}^n, |f(x)| + \|\nabla f(x)\|_\infty + \|\text{Hess}(f)(x)\|_\infty \leq C(1 + \|x\|_\infty^2) \right\}.$$

From (3.25), we can see that if  $f \in \mathcal{C}_b^4(\mathbb{R}^n, \mathbb{R})$ ,  $\mathcal{L}_{\text{FVc}}^\lambda P_{f,n}(\mu)$  is a polynomial in  $\mu$  of the form  $P_{J,n+1}(\mu)$  for some function  $J \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^{n+1}, \mathbb{R})$ . We recall from Proposition 2.2.11 of Chapter 2 that if  $\mu \in \mathcal{M}_1^{c,5}(\mathbb{R})$ , then  $\sup_{0 \leq t \leq T} \mathbb{E}_\mu(M_5(X_t)) < \infty$  where  $(X_t)_{t \geq 0}$  denotes the centered FLEMING-VIOT process with resampling rate  $\lambda$ . Since, by (3.25), for all  $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$ ,

$$|\mathcal{L}_{\text{FVc}} P_{J,n+1}(\mu)| \leq C(1 + M_4(\mu)),$$

for some constant  $C > 0$ , we can apply the martingale problem (3.26) to the function  $P_{J,n+1}(\mu)$  using classical localisation techniques to obtain that the process  $(\widehat{M}_t^{P_{J,n+1}})_{t \geq 0}$  defined by

$$\widehat{M}_t^{P_{J,n+1}} := P_{J,n+1}(X_t) - P_{J,n+1}(X_0) - \int_0^t \mathcal{L}_{\text{FVc}}^\lambda P_{J,n+1}(X_s) ds \quad (3.46)$$

is a  $\mathbb{P}_\mu$ -martingale for all  $\mu \in \mathcal{M}_1^{c,5}(\mathbb{R})$ .

**Lemma 3.6.4.** *Let  $\lambda \in \mathbb{R}$  be fixed. If  $\gamma \in \mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$  satisfies  $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu) \gamma(d\mu) < \infty$  and*

$$\forall n \in \mathbb{N}^*, \forall f \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^n, \mathbb{R}), \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^\lambda P_{f,n}(\mu) \gamma(d\mu) = 0,$$

then,  $\gamma = \pi^\lambda$ .

### 3.6.1 Proof of Proposition 3.6.3

Let us define for all  $\ell \in \mathbb{N}$ ,  $\mathcal{C}_K^\ell(\mathbb{R}^n, \mathbb{R})$  the space of real functions of class  $\mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R})$  with compact support. From Lemma 3.3.4 and (3.12), for all  $t \in [0, T]$ , for all  $n \in \mathbb{N}^*$  we have for all  $f \in \mathcal{C}_K^3(\mathbb{R}^n, \mathbb{R})$  that

$$K^2 \sigma^2 M_{t \wedge \tau^{K,R}}^{K, P_{f,n}} = \sum_{i=1}^3 (\mathbf{A})_i^{K,R}(t \wedge \tau^{K,R})$$

where for all  $t \leq \tau^{K,R} \wedge T$

$$\begin{aligned} (\mathbf{A})_1^{K,R}(t) &:= K^2 \sigma^2 \left( P_{f,n}(\mu_t^{K,R}) - P_{f,n}(\mu_0^{K,R}) \right), \\ (\mathbf{A})_2^{K,R}(t) &:= - \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \theta(z_s^{K,R}) m_2(z_s^{K,R}) \mathcal{L}_{\text{FVc}}^{\lambda(z_s^{K,R})} P_{f,n}(\mu) \Gamma^{K,R}(ds, d\mu), \\ (\mathbf{A})_3^{K,R}(t) &:= O \left( \frac{1}{\sqrt{K}} + \sigma K^{\frac{3}{2} + \varepsilon} + \int_0^t \frac{M_3(\mu_s^{K,R})}{K} ds \right), \end{aligned}$$

is a martingale. Since  $\sigma K \rightarrow 0$  by Assumption (3.5),  $(\mathbf{A})_1^{K,R}(t \wedge \tau^{K,R}) \rightarrow 0$  when  $K \rightarrow +\infty$ . From Corollary 3.4.4, the sequence  $\left( (\mathbf{A})_2^{K,R}(t \wedge \tau^K) \right)_{t \in [0, T], K \in \mathbb{N}^*}$  is uniformly integrable and converges in law, when  $K \rightarrow +\infty$ , to

$$M_t^{P_{f,n}} := - \int_0^t \theta(\zeta_s^R) m_2(\zeta_s^R) \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_s^R)} P_{f,n}(\mu) \Gamma^R(ds, d\mu).$$

Note that, from CAUCHY-SCHWARZ's inequality and Lemma 3.4.2, there exists a constant  $C > 0$  such that

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} \int_0^{t \wedge \tau^{K,R}} \frac{M_3(\mu_s^{K,R})}{K} ds \right)^2 \leq \frac{T}{K^2} \mathbb{E} \left( \int_0^{T \wedge \tau^{K,R}} M_6(\mu_s^{K,R}) ds \right) \leq \frac{CT}{K^2}.$$

Hence,  $\lim_{K \rightarrow +\infty} \sup_{0 \leq t \leq T} (\mathbf{A})_3^{K,R}(t \wedge \tau^{K,R}) = 0$  in  $L^2(\mathbb{R})$ . In particular, we deduce that a subsequence of  $\left( \sup_{0 \leq t \leq T} (\mathbf{A})_3^{K,R}(t \wedge \tau^{K,R}) \right)_{K \in \mathbb{N}^*}$  converges almost surely to 0 and the family  $\left( \sup_{0 \leq t \leq T} (\mathbf{A})_3^{K,R}(t \wedge \tau^{K,R}) \right)_{K \in \mathbb{N}^*}$  is uniformly integrable along this subsequence [213, Theorem 13.7]. Using the same method based on SKOROHOD's representation theorem as in the proof of Theorem 3.5.1, we deduce that the process  $\left( M_t^{P_{f,n}} \right)_{t \geq 0}$  is a  $\mathbb{Q}$ -martingale. As it is also a continuous and finite variation process by Lemma 3.6.1, it must hence be  $\mathbb{Q}$ -a.s. null [150, Theorem 4.1]. Hence, using Lemma 3.6.1 again, we have proved that

$$\forall f \in \mathcal{C}_K^3(\mathbb{R}^n, \mathbb{R}), \mathbb{Q} - \text{a.s.}, dt - \text{a.e.}, \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f,n}(\mu) \gamma_t^R(d\mu) = 0.$$

The space  $\mathcal{C}_K^2(\mathbb{R}^n, \mathbb{R})$  equipped with the norm  $\|f\|_{W_0^{2,\infty}} = \|f\|_\infty + \|\nabla f\|_\infty + \|\text{Hess}(f)\|_\infty$  is separable. So, we can choose a dense countable family  $\mathcal{B} \subset \mathcal{C}_K^3(\mathbb{R}^n, \mathbb{R})$  such that

$$\forall f \in \mathcal{C}_K^2(\mathbb{R}^n, \mathbb{R}), \exists (f_p)_{p \in \mathbb{N}^*} \in \mathcal{B}^{\mathbb{N}^*}, \quad f_p \xrightarrow[p \rightarrow +\infty]{\|\cdot\|_{W_0^{2,\infty}}} f.$$

Then,

$$\mathbb{Q} - \text{a.s.}, dt - \text{a.e.}, \forall p \in \mathbb{N}^*, \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f_p,n}(\mu) \gamma_t^R(d\mu) = 0.$$

From (3.41) and Lemma 3.6.1, we have that  $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu) \gamma_t^R(d\mu) < \infty$ . As for all  $p \in \mathbb{N}^*$ ,

$$\mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f_p,n}(\mu) \leq C_1 (1 + M_2(\mu)) \|f_p\|_{W_0^{2,\infty}}.$$

for some constant  $C_1 > 0$ , we obtain by the dominated convergence theorem that

$$\mathbb{Q} - \text{a.s.}, dt - \text{a.e.}, \forall f \in \mathcal{C}_K^2(\mathbb{R}^n, \mathbb{R}), \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f,n}(\mu) \gamma_t^R(d\mu) = 0. \quad (3.47)$$

Let us consider  $f \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^n, \mathbb{R})$  and for all  $q \in \mathbb{N}^*$ ,  $x \mapsto \chi_q(x) = \exp\left(-\frac{1}{q^2 - \|x\|_\infty^2}\right) \mathbf{1}_{\|x\|_\infty < q}$  of class  $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$  with compact support. Then, for all  $q \in \mathbb{N}^*$ ,  $f\chi_q \in \mathcal{C}_K^2(\mathbb{R}^n, \mathbb{R})$ . Noting that,  $\left| \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f\chi_q,n}(\mu) \gamma_t^R(d\mu) \right|$  is dominated by  $C_2 \int_{\mathcal{M}_1(\mathbb{R})} (1 + M_4(\mu)) \gamma_t^R(d\mu)$  for some

constant  $C_2 > 0$ , and since  $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu) \gamma_t^R(d\mu) < \infty$ , we deduce from (3.47) applied to  $f\chi_q$ , by the dominated convergence theorem, that

$$\mathbb{Q} - \text{a.s.}, dt - \text{a.e.}, \forall f \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^n, \mathbb{R}), \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^{\lambda(\zeta_t^R)} P_{f,n}(\mu) \gamma_t^R(d\mu) = 0.$$

Therefore, it follows from Lemma 3.6.4 that  $\mathbb{Q}$ -a.s,  $dt$ -a.e.,  $\gamma_t^R(d\mu) = \pi^{\lambda(\zeta_t^R)}(d\mu)$  which concludes the proof.

### 3.6.2 Proof of Lemma 3.6.4

In Section 3.6.2.1 we extend some duality results for the centered FLEMING-VIOT process, obtained in Chapter 2, which are be useful to prove Lemma 3.6.4 in Section 3.6.2.2.

#### 3.6.2.1 Extension of the duality result for the centered Fleming-Viot process

Let us recall that the dual process  $(\xi_t)_{t \geq 0}$  of the centered FLEMING-VIOT process  $(X_t)_{t \geq 0}$  with resampling rate  $\lambda$ , on the state space  $\bigcup_{n \in \mathbb{N}^*} \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ , obtained in Section 2.3.2 is defined as below.

Let us consider  $M := (M(t))_{t \geq 0}$  a MARKOV's birth and death process in  $\mathbb{N}$  whose transition rates  $q_{i,j}$  from  $i$  to  $j$  are given by:

$$(1) \quad q_{n,n+1} = \lambda n^2 \quad (2) \quad q_{n,n-1} = \lambda n(n-1) \quad (3) \quad q_{i,j} = 0 \text{ otherwise.}$$

For all  $M(0) \in \mathbb{N}^*$ ,  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  and  $\lambda > 0$ , we define

$$\xi_t := T_\lambda^{(M(\tau_n))}(t - \tau_n) \Lambda_n T_\lambda^{(M(\tau_{n-1}))}(\tau_n - \tau_{n-1}) \Lambda_{n-1} \cdots \Lambda_1 T_\lambda^{(M(0))}(\tau_1) \xi_0, \quad \tau_n \leq t < \tau_{n+1}, \quad n \in \mathbb{N}, \quad (3.48)$$

where  $(\tau_n)_{n \in \mathbb{N}}$  is the sequence of jump times of  $M$  with  $\tau_0 = 0$ ,  $(T_\lambda^{(n)}(t))_{t \geq 0}$  is the semi-group of operator associated to the generator  $B_\lambda^{(n)}$  given by (3.22) and where  $(\Lambda_n)_{n \in \mathbb{N}}$  is a sequence of random operators. These are conditionally independent given  $M$  and satisfy for all  $k \in \mathbb{N}$ ,  $n \geq 1$  and  $1 \leq i \neq j \leq n$ ,

$$\mathbb{P}\left(\Lambda_k = \Phi_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n-1\}\right) = \frac{1}{n(n-1)}$$

and for all  $n \geq 1$  and  $1 \leq i, j \leq n$ ,

$$\mathbb{P}\left(\Lambda_k = K_{i,j} \mid \{M(\tau_k^-) = n, M(\tau_k) = n+1\}\right) = \frac{1}{n^2},$$

where  $\Phi_{i,j}$  and  $K_{i,j}$  are respectively defined in (3.23) and (3.24). Note that if  $M(0) = 1$ , the dual process can only jump from  $\xi_0$  to  $K_{ij}\xi_0$ . As for all  $t \geq 0$ ,  $\xi_t \in \mathcal{C}^2(\mathbb{R}^{M(t)}, \mathbb{R})$ ,  $K_{i,j}\xi_t$  is well defined. Moreover, the dual process  $(\xi_t)_{t \geq 0}$  with initial condition  $\xi_0$  is constructed on the same probability space and independently of the centered FLEMING-VIOT process  $(X_t)_{t \geq 0}$  with resampling rate  $\lambda$  and initial condition  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ . We shall denote by  $\mathbb{P}_{(\mu, \xi_0)}$ , the law of the couple  $((X_t, \xi_t))_{t \geq 0}$  on this probability space.

We denote by  $S_t$  the number of jumps of the process  $M$  on  $[0, t]$ . We start with a result giving bounds on the dual process, which is an extension of similar estimates obtained in Lemma 2.6.4 of Chapter 2.

**Lemma 3.6.5.** *For all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  there exists a function  $C_0$  from  $\bigcup_{k \in \mathbb{N}^*} (0, +\infty)^{k-1} \times \{k\}$  to  $\mathbb{R}_+$ , locally bounded, such that*

$$C_0(\tau_1, 1) := \|\xi_0\|_\infty + \frac{3}{M(0)} \|\text{Hess}(\xi_0)\|_\infty$$

and, for all  $(t_j)_{j \in \mathbb{N}} \in (0, +\infty)^\mathbb{N}$ ,  $k \mapsto C_0((t_i)_{0 \leq i \leq k-1}, k)$  is non-decreasing and satisfying

$$\forall k \in \mathbb{N}^*, \forall t \leq \tau_k, \forall x \in \mathbb{R}^{M(t)}, \quad |\xi_t(x)| \leq C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) (1 + \|x\|_\infty^{2S_t}). \quad (3.49)$$

*Proof.* By induction on  $k \in \mathbb{N}^*$ , we prove the property

$$(\mathcal{Q}_k) : \quad \forall x \in \mathbb{R}^{M(\tau_k)}, \quad |\xi_{\tau_k}(x)| \leq C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) (1 + \|x\|_\infty^{2k}).$$

**Step 1. Initial case: computation of  $C_0(\tau_1, 1)$ .** Let us recall some notations of Theorem 2.6.1 of Chapter 2. For all  $n \in \mathbb{N}^*$ , we denote by  $\mathbf{1} \in \mathbb{R}^n$ , the vector whose coordinates are all 1. For all  $n \in \mathbb{N}^*$ ,  $t \geq 0$ ,  $x \in \mathbb{R}^n$  and  $\lambda > 0$ , let us consider  $g_{t,x}^\lambda$  the density of the Gaussian distribution  $\mathcal{N}^{(n)}(m_{t,x}^\lambda, \Sigma_t^\lambda)$  where  $m_{t,x}^\lambda := x - \frac{(1 - \exp(-2\lambda nt))}{n}(x \cdot \mathbf{1})\mathbf{1}$  and  $\Sigma_t^\lambda := P\sigma_t^\lambda P^{-1}$  with

$$\sigma_t^\lambda := \begin{pmatrix} \frac{1 - \exp(-4\lambda nt)}{4\lambda n} & 0 & \dots & \dots & 0 \\ 0 & t & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & t \end{pmatrix} \quad \text{and} \quad P := \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{2}} & \dots & \dots & \frac{1}{\sqrt{n(n-1)}} \\ \vdots & -\sqrt{\frac{1}{2}} & \ddots & \dots & \frac{1}{\sqrt{n(n-1)}} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{\sqrt{n(n-1)}} \\ \frac{1}{\sqrt{n}} & 0 & \dots & 0 & -\sqrt{\frac{n-1}{n}} \end{pmatrix}.$$

Note that for all  $i, j \in \{1, \dots, n\}$ ,  $\partial_{x_i} m_{t,x}^\lambda = \epsilon_i - \frac{1 - \exp(-2\lambda nt)}{n}\mathbf{1}$  and  $\partial_{x_i x_j}^2 m_{t,x}^\lambda = 0$  where  $(\epsilon_1, \dots, \epsilon_n)$  is the canonical basis of  $\mathbb{R}^n$ . For all  $\lambda > 0$ , the key identity for the sequel is

$$\forall f \in L^\infty(\mathbb{R}^n), \forall t \geq 0, \forall x \in \mathbb{R}^n, \quad T_\lambda^{(M(0))}(t)f(x) = (f * g_{t,0}^\lambda)(m_{t,x}^\lambda) \quad (3.50)$$

where  $*$  stands for the convolution product (see (2.46) in Theorem 2.6.1).

On the one hand, note that for all  $t < \tau_1$ ,  $M(t) = M(0)$ . In this case, for all  $x \in \mathbb{R}^{M(0)}$ ,  $\lambda > 0$ ,  $\xi_t(x) = T_\lambda^{(M(0))}(t)\xi_0(x)$  and so, from Theorem 2.6.1,  $|\xi_t(x)| \leq \|\xi_0\|_\infty$ . On the other hand, at time  $\tau_1$ , we make a partition of cases according to whether the dual process loses or gains a variable. Let  $i, j \in \{1, \dots, M(0)\}$  be fixed.

- **Case  $\Lambda_1 = \Phi_{i,j}$ .** In this case,  $M(\tau_1) = M(0) - 1$  and we deduce from (3.48) that

$$\forall x \in \mathbb{R}^{M(0)-1}, \quad \xi_{\tau_1}(x) = \Phi_{i,j} T_\lambda^{(M(0))}(\tau_1) \xi_0(x).$$

By (3.23), we deduce that for all  $x \in \mathbb{R}^{M(0)-1}$ ,  $|\xi_{\tau_1}(x)| \leq \|\xi_0\|_\infty$ .

- **Case  $\Lambda_1 = K_{i,j}$ .** In this case,  $M(\tau_1) = M(0) + 1$  and we deduce from (3.48) and (3.24) then from **(3)** of Theorem 2.6.1, and properties of the convolution product that for all  $x \in \mathbb{R}^{M(0)+1}$ ,

$$\begin{aligned} \xi_{\tau_1}(x) &= K_{i,j} T_\lambda^{(M(0))}(\tau_1) \xi_0(x) = \partial_{x_i x_j} T_\lambda^{(M(0))}(\tau_1) \xi_0(\tilde{x}) x_{M(0)+1}^2 \\ &= \left( \partial_{x_j} m_{\tau_1, \tilde{x}}^\lambda \right)^t \left[ \left( \text{Hess}(\xi_0) * g_{\tau_1, 0}^\lambda \right) \left( m_{\tau_1, \tilde{x}}^\lambda \right) \partial_{x_i} m_{t, \tilde{x}}^\lambda \right] x_{M(0)+1}^2 \end{aligned}$$

where  $\tilde{x} = (x_1, \dots, x_{M(0)}) \in \mathbb{R}^{M(0)}$ . As,

$$\left( \text{Hess}(\xi_0) * g_{\tau_1, 0}^\lambda \right) \left( m_{\tau_1, \tilde{x}}^\lambda \right) = \left( \int_{\mathbb{R}^{M(0)}} \partial_{ij}^2 \xi_0(u) g_{\tau_1, 0}^\lambda \left( m_{\tau_1, \tilde{x}}^\lambda - u \right) du \right)_{1 \leq i, j \leq M(0)}$$

we obtain that  $\left\| \left( \text{Hess}(\xi_0) * g_{\tau_1, 0}^\lambda \right) \left( m_{\tau_1, \tilde{x}}^\lambda \right) \right\|_\infty \leq \left\| \partial_{ij}^2 \xi_0 \right\|_\infty$ . Noting that

$$\left| \left( \partial_{x_j} m_{\tau_1, \tilde{x}}^\lambda \right)^t \partial_{x_i} m_{t, \tilde{x}}^\lambda \right| \leq \frac{3}{M(0)},$$

we deduce that for all  $x \in \mathbb{R}^{M(0)+1}$ ,

$$|\xi_{\tau_1}(x)| \leq \frac{3}{M(0)} \sup_{1 \leq i, j \leq M(0)} \left\| \partial_{ij}^2 \xi_0 \right\|_\infty x_{M(0)+1}^2 \leq \frac{3}{M(0)} \|\text{Hess}(\xi_0)\|_\infty \|x\|_\infty^2$$

and  $(\mathcal{Q}_1)$  follows.

**Step 2. Inductive Step.** We assume that, for  $k \in \mathbb{N} \setminus \{0, 1\}$ ,  $(\mathcal{Q}_{k-1})$  is satisfied and prove that  $(\mathcal{Q}_k)$  is also. We make again a partition of cases according to whether the dual process loses or gains a variable. Let  $i, j \in \{1, \dots, M(\tau_{k-1})\}$  be fixed.

- **Case  $\Lambda_k = \Phi_{i,j}$  at the  $k^{\text{th}}$  jump.** In this case,  $M(\tau_k) = M(\tau_{k-1}) - 1$  and we deduce from (3.48) that for all  $x \in \mathbb{R}^{M(\tau_{k-1})-1}$ ,  $\xi_{\tau_k}(x) = \Phi_{i,j} T_\lambda^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x)$ . By using (3.23) and  $(\mathcal{Q}_{k-1})$ , we deduce from **(1)** of Corollary 2.6.2 of Chapter 2 that for all  $x \in \mathbb{R}^{M(\tau_{k-1})-1}$ ,

$$\begin{aligned} & \left| \Phi_{i,j} T_\lambda^{(M(\tau_{k-1}))}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ & \leq C_2(\tau_k - \tau_{k-1}, M(\tau_{k-1})) C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) \left(1 + \|x\|_\infty^{2(k-1)}\right), \end{aligned}$$

where  $C_2 C_0$  is locally bounded. As  $t \mapsto C_2(t, M(\tau_{k-1}))$  is non-decreasing, we deduce  $(\mathcal{Q}_k)$  in that case.

- **Case  $\Lambda_k = K_{i,j}$  at the  $k^{\text{th}}$  jump.** In this case,  $M(\tau_k) = M(\tau_{k-1}) + 1$ . From (3.48), (3.24) and **(3)** of Theorem 2.6.1, we have for all  $x \in \mathbb{R}^{M(\tau_{k-1})+1}$ ,

$$\begin{aligned} |\xi_{\tau_k}(x)| &= \left| K_{i,j} T_\lambda^{(M(\tau_{k-1})+1)}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ &= \left| \left( \partial_{x_j} m_{\tau_k - \tau_{k-1}, \tilde{x}}^\lambda \right)^t \left[ \left( \xi_{\tau_{k-1}} * \text{Hess} \left( g_{\tau_k - \tau_{k-1}, 0}^\lambda \right) \right) \left( m_{\tau_k - \tau_{k-1}, \tilde{x}}^\lambda \right) \partial_{x_i} m_{\tau_k - \tau_{k-1}, \tilde{x}}^\lambda \right] \right| \\ & \quad \times x_{M(\tau_{k-1})+1}^2, \end{aligned}$$

where  $\tilde{x} = (x_1, \dots, x_{M(\tau_{k-1})})^t \in \mathbb{R}^{M(\tau_{k-1})}$ . From  $(\mathcal{Q}_{k-1})$  and **(2)** of Corollary 2.6.2, we deduce that

$$\begin{aligned} & \left| K_{i,j} T_\lambda^{(M(\tau_{k-1})+1)}(\tau_k - \tau_{k-1}) \xi_{\tau_{k-1}}(x) \right| \\ & \leq C_3(\tau_k - \tau_{k-1}, M(\tau_{k-1})) C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) \left(1 + \|x\|_\infty^{2k}\right), \end{aligned}$$

where  $C_3C_0$  is locally bounded and  $(Q_k)$  follows in that case.

We conclude by the principle of induction.

**Step 3. Proof of (3.49) for  $t < \tau_k$ .** Note that from (3.48), for all  $k \in \mathbb{N}$  and  $t \in (\tau_k, \tau_{k+1})$ ,

$$\forall x \in \mathbb{R}^{M(t)}, \quad \xi_t(x) = T_\lambda^{(M(\tau_k))}(t - \tau_k) \xi_{\tau_k}(x),$$

so the announced result follows from **(1)** of Corollary 2.6.2.  $\square$

Let us consider for all  $k \in \mathbb{N}^*$ ,  $\ell, m \in \mathbb{N}$  the stopping times

$$\begin{aligned} \vartheta_{k,\ell} &:= \inf \left\{ t \geq 0 \mid S_t \geq k \text{ or } \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq \ell \right\}, \\ \vartheta'_{k,m} &:= \inf \left\{ t \geq 0 \mid S_t \geq k \text{ or } C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) \geq m \right\}, \\ \vartheta'_{k,\ell,m} &:= \inf \left\{ t \geq 0 \mid S_t \geq k \text{ or } \exists s \in [0, t], \langle \xi_s, X_{t-s}^{M(s)} \rangle \geq \ell \right. \\ &\quad \left. \text{or } C_0((\tau_{i+1} - \tau_i)_{0 \leq i \leq k-1}, k) \geq m \right\}. \end{aligned}$$

As  $(M(t), \xi_t)_{t \geq 0}$  is independent of  $(X_t)_{t \geq 0}$ , note that for all  $k \in \mathbb{N}^*$ ,  $m \in \mathbb{N}$ ,  $\vartheta'_{k,m}$  is independent of  $(X_t)_{t \geq 0}$ . Recall that for all  $t \geq 0$ ,  $\xi_t \in \mathcal{C}^2(\mathbb{R}^{M(t)}, \mathbb{R})$ . For all  $n \geq M(t)$ , for all  $x \in \mathbb{R}^n$ , we denote by  $\xi_t^{(n)}(x) = \xi_t^{(n)}(x_1, \dots, x_n) := \xi_t(x_1, \dots, x_{M(t)})$  so that for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,

$$\langle \xi_t^{(n)}, \mu^n \rangle = \langle \xi_t, \mu^{M(t)} \rangle.$$

The next lemma is an extension of the duality identity proved in Chapter 2.

**Lemma 3.6.6.** *Given  $(X_t)_{t \geq 0}$ ,  $(\xi_t)_{t \geq 0}$  as above with  $X_0 := \mu \in \mathcal{M}_1^{c,2k}(\mathbb{R})$ ,  $k \in \mathbb{N}^*$  and  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ , we have that  $\mathbb{P}_{(\mu, \xi_0)}$ -a.s., for all  $t \geq 0$ , for all  $m \in \mathbb{N}$ ,*

$$\mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_0, X_{t \wedge \vartheta'_{k,m}}^{M(0)} \rangle \right) = \left\langle \mathbb{E}_{\xi_0} \left( \xi_{t \wedge \vartheta'_{k,m}}^{(M(0)+k)} \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k,m}} M^2(u) du \right) \right), \mu^{M(0)+k} \right\rangle. \quad (3.51)$$

Compared to the duality identity of Chapter 2, the key point of this lemma is that the right-hand side of the last equation, is a polynomial in  $\mu$ .

*Proof.* Let us consider  $t \geq 0$ ,  $k \in \mathbb{N}^*$ ,  $\ell, m \in \mathbb{N}$ . As  $\vartheta'_{k,\ell,m} \leq \vartheta_{k,\ell}$ , the weak duality identity of Theorem 2.3.4 of Chapter 2 implies that:

$$\begin{aligned} \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_0, X_{t \wedge \vartheta'_{k,\ell,m}}^{M(0)} \rangle \right) &= \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_{t \wedge \vartheta'_{k,\ell,m}}, \mu^{M(t \wedge \vartheta'_{k,\ell,m})} \rangle \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k,\ell,m}} M^2(u) du \right) \right) \\ &= \mathbb{E}_{(\mu, \xi_0)} \left( \langle \xi_{t \wedge \vartheta'_{k,\ell,m}}^{(M(0)+k)}, \mu^{M(0)+k} \rangle \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k,\ell,m}} M^2(u) du \right) \right). \end{aligned}$$

Now, from Lemma 3.6.5 and the definition of  $\vartheta'_{k,\ell,m}$ , note that for all  $x \in \mathbb{R}^{M(0)+k}$ ,

$$\mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k,\ell,m}}^{(M(0)+k)}(x) \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k,\ell,m}} M^2(u) du \right) \right) \leq m \left( 1 + \|x\|_\infty^{2k} \right) \exp \left( \lambda t (M(0) + k)^2 \right). \quad (3.52)$$

Since  $\mu \in \mathcal{M}_1^{c,2k}(\mathbb{R})$ , we deduce from FUBINI's theorem that

$$\mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_0, X_{t \wedge \vartheta'_{k, \ell, m}}^{M(0)} \right\rangle \right) = \left\langle \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, \ell, m}}^{(M(0)+k)} \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, \ell, m}} M^2(u) du \right) \right), \mu^{M(0)+k} \right\rangle.$$

On the one hand, from Section 2.6.2.1,  $\lim_{\ell \rightarrow +\infty} \vartheta'_{k, \ell, m} = \vartheta'_{k, m}$   $\mathbb{P}_{(\mu, \xi_0)}$ -a.s., and since  $(X_t)_{t \geq 0}$  has continuous paths for the topology of weak convergence, we have  $\mathbb{P}_{(\mu, \xi_0)}$ -a.s.

$$\lim_{\ell \rightarrow +\infty} \left\langle \xi_0, X_{t \wedge \vartheta'_{k, \ell, m}}^{M(0)} \right\rangle = \left\langle \xi_0, X_{t \wedge \vartheta'_{k, m}}^{M(0)} \right\rangle.$$

Therefore, we deduce from the dominated convergence theorem, that for all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$ ,

$$\lim_{\ell \rightarrow +\infty} \mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_0, X_{t \wedge \vartheta'_{k, \ell, m}}^{M(0)} \right\rangle \right) = \mathbb{E}_{(\mu, \xi_0)} \left( \left\langle \xi_0, X_{t \wedge \vartheta'_{k, m}}^{M(0)} \right\rangle \right).$$

On the other hand, Lemma 3.6.5 and the dominated convergence theorem imply that for all  $x \in \mathbb{R}^{M(0)+k}$ ,

$$\begin{aligned} & \lim_{\ell \rightarrow +\infty} \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, \ell, m}}^{(M(0)+k)}(x) \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, \ell, m}} M^2(u) du \right) \right) \\ &= \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, m}}^{(M(0)+k)}(x) \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, m}} M^2(u) du \right) \right) \end{aligned}$$

and the limit satisfies the inequality (3.52). As  $\mu \in \mathcal{M}_1^{c,2k}(\mathbb{R})$ , we deduce again from the dominated convergence theorem that

$$\begin{aligned} & \lim_{\ell \rightarrow +\infty} \left\langle \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, \ell, m}}^{(M(0)+k)} \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, \ell, m}} M^2(u) du \right) \right), \mu^{M(0)+k} \right\rangle \\ &= \left\langle \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, m}}^{(M(0)+k)} \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, m}} M^2(u) du \right) \right), \mu^{M(0)+k} \right\rangle \end{aligned}$$

and the announced result follows from the fact that  $\vartheta'_{k, m}$  is independent of  $(X_t)_{t \geq 0}$ , which implies

$$\begin{aligned} & \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \vartheta'_{k, m}}^{(M(0)+k)}(x) \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, m}} M^2(u) du \right) \right) \\ &= \mathbb{E}_{\xi_0} \left( \xi_{t \wedge \vartheta'_{k, m}}^{(M(0)+k)}(x) \exp \left( \lambda \int_0^{t \wedge \vartheta'_{k, m}} M^2(u) du \right) \right). \quad \square \end{aligned}$$

From Lemma 3.6.5, we deduce that we can choose  $m \in \mathbb{N}$  large enough such that  $\vartheta'_{1, m} = \tau_1$ . In this case, the function inside the brackets in the left-hand side of (3.51) is defined, for all  $t \geq 0$  and for all  $x \in \mathbb{R}^{M(0)+1}$ , by

$$\psi_t(x) := \mathbb{E}_{(\mu, \xi_0)} \left( \xi_{t \wedge \tau_1}^{(M(0)+1)}(x) \exp \left( \lambda \int_0^{t \wedge \tau_1} M^2(u) du \right) \right).$$

For all  $t \geq 0$ ,  $n \in \mathbb{N}^*$ ,  $f \in \mathcal{C}_b^2(\mathbb{R}^n, \mathbb{R})$  and  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  let us denote

$$P_{\text{FVc}}(t)P_{f, n}(\mu) := \mathbb{E}_\mu(P_{f, n}(X_t))$$

the semi-group of the centered FLEMING-VIOT process  $(X_t)_{t \geq 0}$  with resampling rate  $\lambda$ .

The following result is the main result of this section. It gives an extension of Lemma 3.6.5 to a function which appears naturally in the proof of Lemma 3.6.4.

**Proposition 3.6.7.** For all  $t \geq 0$ , for all  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ , for all  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$

$$P_{\text{FVc}}(t)P_{\xi_0, M(0)}(\mu) = P_{V_t, M(0)+1}(\mu) := \langle V_t, \mu^{M(0)+1} \rangle$$

where

$$V_t := \exp(\alpha_0 t) \psi_t - \alpha_0 \int_0^t \exp(\alpha_0 s) \psi_s ds \quad (3.53)$$

and  $\alpha_0 := \lambda M(0) [2M(0) - 1]$ . In addition, for all  $t \geq 0$ , there exist a constant  $C_t > 0$  such that for all  $x \in \mathbb{R}^{M(0)+1}$ ,

$$|V_t(x)| \leq C_t (1 + \|x\|_\infty^2).$$

*Proof.* Let  $t \geq 0$ ,  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$ ,  $\xi_0 \in \mathcal{C}_b^2(\mathbb{R}^{M(0)}, \mathbb{R})$  be fixed. Note that from Lemma 3.6.5, there exists a constant  $\tilde{C}_t$ , depending only on  $\xi_0$  and  $t$  such that for all  $x \in \mathbb{R}^{M(0)+1}$ ,

$$|\psi_t(x)| \leq C_0(\tau_1, 1) (1 + \|x\|_\infty^2) \mathbb{E}_{(\mu, \xi_0)}(\exp(\lambda t M^2(0))) \leq \tilde{C}_t (1 + \|x\|_\infty^2). \quad (3.54)$$

Now, from Lemma 3.6.6 and the fact that  $\tau_1$  is independent of  $(X_t)_{t \geq 0}$ , we obtain that

$$\langle \psi_t, \mu^{M(0)+1} \rangle = \mathbb{E}_{(\mu, \xi_0)}(P_{\xi_0, M(0)}(X_{t \wedge \tau_1})) = \mathbb{E}_{\xi_0}(P_{\text{FVc}}(t \wedge \tau_1) P_{\xi_0, M(0)}(\mu)).$$

As  $\tau_1$  follows an exponential law of parameter  $\alpha_0$  we have that

$$\langle \psi_t, \mu^{M(0)+1} \rangle = P_{\text{FVc}}(t) P_{\xi_0, M(0)}(\mu) \exp(-\alpha_0 t) + \alpha_0 \int_0^t P_{\text{FVc}}(s) P_{\xi_0, M(0)}(\mu) \exp(-\alpha_0 s) ds. \quad (3.55)$$

Thanks to (3.54) and since  $\mu$  has its moment of order 2 finite, the FUBINI theorem ensures us that

$$\left\langle \int_0^t \exp(\alpha_0 s) \psi_s ds, \mu^{M(0)+1} \right\rangle = \int_0^t \exp(\alpha_0 s) \langle \psi_s, \mu^{M(0)+1} \rangle ds.$$

From (3.55) we deduce that

$$\begin{aligned} & \left\langle \int_0^t \exp(\alpha_0 s) \psi_s ds, \mu^{M(0)+1} \right\rangle \\ &= \alpha_0 \int_0^t \exp(\alpha_0 s) \int_0^s \exp(-\alpha_0 u) P_{\text{FVc}}(u) P_{\xi_0, M(0)} du ds + \int_0^t P_{\text{FVc}}(s) P_{\xi_0, M(0)}(\mu) ds \quad (3.56) \\ &= \exp(\alpha_0 t) \int_0^t \exp(-\alpha_0 u) P_{\text{FVc}}(u) P_{\xi_0, M(0)}(\mu) du. \end{aligned}$$

Then, the first announced result follows from (3.55) and (3.56) and the second one from (3.54).  $\square$

### 3.6.2.2 Proof of Lemma 3.6.4

The key point of this proof is to establish that for all  $\xi_0 \in \mathcal{C}_b^4(\mathbb{R}^{M(0)}, \mathbb{R})$ ,  $t \geq 0$  and  $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$ ,

$$\begin{aligned} P_{\text{FVc}}(t)P_{\xi_0, M(0)}(\mu) &= \langle \xi_0, \mu^{M(0)} \rangle + \int_0^t \mathcal{L}_{\text{FVc}}^\lambda P_{\text{FVc}}(s) P_{\xi_0, M(0)}(\mu) ds \\ &= P_{\xi_0, M(0)}(\mu) + \int_0^t \mathcal{L}_{\text{FVc}}^\lambda P_{V_s, M(0)+1}(\mu) ds. \end{aligned} \quad (3.57)$$



where we use Proposition 3.6.7 in the second equality. To do this, we will first prove that for all  $t \geq 0$ , there exists a constant  $C_t > 0$  such that for all  $x \in \mathbb{R}^{M(0)+1}$ ,  $\|\text{Hess}(V_t)(x)\|_\infty \leq C_t (1 + \|x\|_\infty^2)$ , so that  $V_t \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^{M(0)+1}, \mathbb{R})$ .

For all  $n \in \mathbb{N}^*$ , for any function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $k \in \mathbb{N}$  times differentiable, we denote by  $D^k f$  the differential of order  $k$  of  $f$ . Let  $t \geq 0$ ,  $\lambda > 0$ ,  $\mu \in \mathcal{M}_1^{c,2}(\mathbb{R})$  and  $\xi_0 \in \mathcal{C}_b^4(\mathbb{R}^{M(0)}, \mathbb{R})$  be fixed.

**Step 1. Preliminary bounds.** Let  $x \in \mathbb{R}^{M(0)}$ . It follows from (3.50) that

$$\begin{aligned} & \partial_{x_i x_j x_k x_\ell}^4 T^{(M(0))}(t) \xi_0(x) \\ &= \sum_{m,p,q,r=1}^{M(0)} \partial_{x_i} (m_{t,x}^\lambda)_m \partial_{x_j} (m_{t,x}^\lambda)_p \partial_{x_k} (m_{t,x}^\lambda)_q \partial_{x_\ell} (m_{t,x}^\lambda)_r (g_{t,0}^\lambda * \partial_{y_m y_p y_q y_r}^4 \xi_0) (m_{t,x}^\lambda) \end{aligned}$$

where  $(m_{t,x}^\lambda)_p$  designates the  $p^{\text{th}}$  component of  $m_{t,x}^\lambda$ . Then, for all  $k \in \{0, \dots, 4\}$ , there exists a constant  $C_1 > 0$  independent of  $x$  and  $t$  such that

$$\|D^k T^{(M(0))}(t) \xi_0(x)\|_\infty \leq C_1 (1 + \|D^k \xi_0\|_\infty).$$

From (3.48), (3.23) and (3.24), we deduce that there exists a constant  $C_2 > 0$  independent of  $t$  such that for all  $x \in \mathbb{R}^{M(0)+1}$ , for all  $k \in \{0, 1, 2\}$ ,

$$\begin{aligned} \|D^k \xi_{t \wedge \tau_1}^{(M(0)+1)}(x)\|_\infty &\leq C_1 (1 + \|D^k \xi_0\|_\infty) + \sum_{\substack{i,j=1 \\ i \neq j}}^n \|D^k \Phi_{i,j} T_\lambda^{(M(0))}(\tau_1) \xi_0(x_1, \dots, x_{M(0)-1})\|_\infty \\ &\quad + \sum_{i,j=1}^n \|D^k K_{i,j} T_\lambda^{(M(0))}(\tau_1) \xi_0(x)\|_\infty \\ &\leq C_2 (1 + \|D^k \xi_0\|_\infty + \|D^{k+2} \xi_0\|_\infty x_{M(0)+1}^2). \end{aligned}$$

By the theorem of differentiation under the integral sign, for all  $k \in \{0, 1, 2\}$  and  $x \in \mathbb{R}^{M(0)+1}$ , we deduce that

$$\|D^k \psi_t(x)\|_\infty \leq \exp(\lambda M(0)^2 t) \mathbb{E}_{(\mu, \xi_0)} \left( \|D^k \xi_{t \wedge \tau_1}^{(M(0)+1)}(x)\|_\infty \right) \leq C_2 \exp(\lambda M(0)^2 t) (1 + \|x\|_\infty^2).$$

Then, from the definition of  $V_t$  in (3.53) we deduce that there exists a constant  $C_t$  such that for all  $k \in \{0, 1, 2\}$ , for all  $x \in \mathbb{R}^{M(0)+1}$ ,  $\|D^k V_t(x)\|_\infty \leq C_t (1 + \|x\|_\infty^2)$ , so that  $V_t \in \mathcal{C}_{\|\cdot\|}^2(\mathbb{R}^{M(0)+1}, \mathbb{R})$ .

**Step 2. Proof of the key point (3.57).** Thanks to (3.26), for all  $t \geq 0$ ,

$$P_{\text{FVc}}(t) P_{\xi_0, M(0)}(\mu) = P_{\xi_0, M(0)}(\mu) + \mathbb{E}_\mu \left( \int_0^t \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_s) ds \right).$$

From (3.25), there exists a constant  $C_3 > 0$  such that for all  $t \geq 0$ ,  $\mathbb{E}_\mu (|\mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_t)|) \leq C_3 (1 + \mathbb{E}_\mu (M_2(X_t)))$  which is finite from Proposition 2.2.11 of Chapter 2 since  $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$ . Hence, from FUBINI's theorem we obtain that

$$P_{\text{FVc}}(t) P_{\xi_0, M(0)}(\mu) = P_{\xi_0, M(0)}(\mu) + \int_0^t \mathbb{E}_\mu (\mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_s)) ds. \quad (3.58)$$

As  $\mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(\mu)$  is a polynomial in  $\mu$  and since  $|\mathcal{L}_{\text{FVc}}^\lambda (\mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(\mu))| \leq C_4 (1 + M_4(\mu))$  for some constant  $C_4 > 0$ , we deduce as in (3.46) that

$$\mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_t) \right) - \mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_0) \right) = \mathbb{E}_\mu \left( \int_0^t \mathcal{L}_{\text{FVc}}^\lambda \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)} \right) (X_s) ds \right).$$

In particular,  $t \mapsto \mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_t) \right)$  is continuous, so

$$\begin{aligned} \mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_t) \right) &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( \int_t^{t+h} \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_s) ds \right) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( P_{\xi_0, M(0)}(X_{t+h}) - P_{\xi_0, M(0)}(X_t) \right) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( P_{\text{FVc}}(t) P_{\xi_0, M(0)}(X_h) \right) - P_{\text{FVc}}(t) P_{\xi_0, M(0)}(\mu) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( P_{V_t, M(0)+1}(X_h) - P_{V_t, M(0)+1}(\mu) \right) \end{aligned}$$

where we used MARKOV's property in the third equality and Proposition 3.6.7 in the last equality. From Step 1, we deduce that there exists a constant  $C_5 > 0$  such that

$$\left| \mathcal{L}_{\text{FVc}}^\lambda P_{V_t, M(0)+1}(\mu) \right| \leq C_5 (1 + M_4(\mu)). \quad (3.59)$$

Thus, from the martingale problem (3.46) and the continuity of  $t \mapsto \mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{V_t, M(0)}(\mu) \right)$ , we have that

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( P_{V_t, M(0)+1}(X_h) - P_{V_t, M(0)+1}(\mu) \right) &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_\mu \left( \int_0^h \mathcal{L}_{\text{FVc}}^\lambda P_{V_t, M(0)+1}(X_s) ds \right) \\ &= \mathcal{L}_{\text{FVc}}^\lambda P_{V_t, M(0)+1}(\mu). \end{aligned}$$

Therefore, we have proved that for all  $t \geq 0$ , for all  $\mu \in \mathcal{M}_1^{c,4}(\mathbb{R})$ ,

$$\mathbb{E}_\mu \left( \mathcal{L}_{\text{FVc}}^\lambda P_{\xi_0, M(0)}(X_t) \right) = \mathcal{L}_{\text{FVc}}^\lambda P_{V_t, M(0)+1}(\mu),$$

and so, by (3.58), we obtain (3.57).

**Step 3. Conclusion.** From (3.57) and (3.59) and since  $\int_{\mathcal{M}_1(\mathbb{R})} M_4(\mu) \gamma(d\mu) < \infty$ , we deduce that

$$\begin{aligned} &\int_{\mathcal{M}_1(\mathbb{R})} P_{\text{FVc}}(t) P_{\xi_0, M(0)}(\mu) \gamma(d\mu) \\ &= \int_{\mathcal{M}_1(\mathbb{R})} P_{\xi_0, M(0)}(\mu) \gamma(d\mu) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{FVc}}^\lambda P_{V_s, M(0)+1}(\mu) \gamma(d\mu) ds \\ &= \int_{\mathcal{M}_1(\mathbb{R})} P_{\xi_0, M(0)}(\mu) \gamma(d\mu), \end{aligned}$$

where we used the assumption of Lemma 3.6.4 in the last equality.

As the set of test functions  $\left\{ P_{f,n} \mid f \in \mathcal{C}_b^4(\mathbb{R}^n, \mathbb{R}), n \in \mathbb{N}^* \right\}$  is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -convergence determining [60, Lemma 2.1.2], it is  $\mathcal{M}_1(\mathcal{M}_1(\mathbb{R}))$ -separating [94, Chapter 3, Section 4, p.112], so we have for any bounded continuous function  $\phi$  from  $\mathcal{M}_1(\mathbb{R})$  to  $\mathbb{R}$  that

$$\int_{\mathcal{M}_1(\mathbb{R})} P_{\text{FVc}}(t) \phi(\mu) \gamma(d\mu) = \int_{\mathcal{M}_1(\mathbb{R})} \phi(\mu) \gamma(d\mu).$$

Hence  $\gamma$  is an invariant probability measure for the centered FLEMING-VIOT process with resampling rate  $\lambda$ . Now, from Theorem 2.4.1 of Chapter 2,  $\pi^\lambda$  is its unique invariant probability measure which ends the proof of Lemma 3.6.4.

### 3.7 Characterisation of the limiting values of the slow component

Combining the results of Sections 3.5 and 3.6, we have proved that for all  $R > 0$ , the sequence of laws of  $((z^{K,R}, \Gamma^{K,R}))_{K \in \mathbb{N}^*}$  is tight in  $\mathcal{M}_1(\mathbb{D}([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R})))$  and for any limiting value  $\mathbb{Q}$  of this sequence, the canonical process  $(\zeta^R, \Gamma^R)$  on  $\mathcal{C}^0([0, T], \mathbb{T}_R) \times \mathcal{M}_m^T(\mathcal{M}_1(\mathbb{R}))$  satisfies for all  $f \in \mathcal{C}_b^1(\mathbb{T}_R, \mathbb{T}_R)$ , for all  $t \in [0, T]$ ,  $\mathbb{Q}$ -a.s.,

$$f(\zeta_t^R) = f(\zeta_0^R) + \int_0^t \int_{\mathcal{M}_1(\mathbb{R})} \mathcal{L}_{\text{SLOW}} f(\zeta_s^R, \mu) \pi^\lambda(\zeta_s^R)(d\mu) ds. \quad (3.60)$$

From now on, we will establish in Section 3.7.1 with Lemma 3.7.1 that the sequence of laws of  $(z^{K,R})_{K \in \mathbb{N}^*}$  converges weakly in  $\mathbb{D}([0, T], \mathbb{T}_R)$  to the solution of an ODE in the torus. Finally, Section 3.7.2 allows us to get away from the torus and to prove Theorem 3.2.1.

#### 3.7.1 Convergence of the slow component on the torus

**Lemma 3.7.1.** *The sequence  $(z^{K,R})_{K \in \mathbb{N}^*}$  converges in law in  $\mathbb{D}([0, T], \mathbb{T}_R)$  to the unique solution of*

$$\forall t \in [0, T], \quad z_t = x_0 + \frac{1}{2} \int_0^t \partial_1 \text{Fit}(z_s, z_s) \beta(z_s) m_2(z_s) ds \quad (3.61)$$

on the torus  $\mathbb{T}_R$ .

*Proof.* For any limiting value  $\mathbb{Q}$  of the sequence of laws of  $(z^{K,R})_{K \in \mathbb{N}^*}$ , from (3.9) and applying (3.60) with  $f = \text{id} \in \mathcal{C}_b^2(\mathbb{T}_R, \mathbb{T}_R)$ , we deduce that  $\mathbb{Q}$ -a.s.,

$$\forall t \in [0, T], \quad \zeta_t^R = \zeta_0^R + \int_0^t \left( \int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^\lambda(\zeta_s^R)(d\mu) \right) \partial_1 \text{Fit}(\zeta_s^R, \zeta_s^R) ds.$$

From Corollary 2.4.16 of Chapter 2, for all  $s \in [0, t]$ ,

$$\int_{\mathcal{M}_1(\mathbb{R})} M_2(\mu) \pi^\lambda(\zeta_s^R)(d\mu) = \frac{1}{2\lambda(\zeta_s^R)} = \frac{\beta(\zeta_s^R) m_2(\zeta_s^R)}{2},$$

so  $\zeta^R$  is solution of (3.61)  $\mathbb{Q}$ -a.s. By uniqueness of the solution of (3.61), the announced result follows.  $\square$

#### 3.7.2 End of the proof of Theorem 3.2.1

Let  $T > 0$  be fixed and recall that  $x_0$  is the mean trait value of  $\nu_0^K$ . From Lemma 3.7.1 and since the ODE (3.61) is non-explosive, we can choose  $R > 0$  large enough such that  $\lim_{K \rightarrow +\infty} \mathbb{P}\left(z_t^{K,R} \in [x_0 - \frac{R}{2}, x_0 + \frac{R}{2}], \forall t \in [0, T]\right) = 1$ . From Proposition 3.4.8 and (3.7), we have that,  $\lim_{K \rightarrow +\infty} \mathbb{P}\left(\text{Diam}\left(\text{Supp } \nu_t^{K,R}\right) \leq \frac{1}{K^{1+\varepsilon/2}}, \forall t \in [0, T]\right) = 1$ . Hence,

$$\lim_{K \rightarrow +\infty} \mathbb{P}\left(\text{Supp } \nu_{t/K\sigma^2}^{K,R} \subset [x_0 - R, x_0 + R], \forall t \in [0, T]\right) = 1.$$

Now, on the event  $\{\text{Supp } \nu_{t/K\sigma^2}^{K,R} \subset [x_0 - R, x_0 + R], \forall t \in [0, T]\}$ ,  $\nu_{t/K\sigma^2}^{K,R} = \nu_{t/K\sigma^2}^K$  for all  $t \in [0, T]$ , identifying  $x \in \mathbb{T}_R$  with its unique representant in  $[x_0 - 2R, x_0 + 2R]$ . In particular,

$$\forall t \in [0, T], \quad z_t^{K,R} = z_t^K.$$

Theorem 3.2.1 then follows from Lemma 3.7.1.

## Appendix A

# Technical lemmas involved in the existence proof

### A.1 Approximation lemma

**Lemma A.1.1.** *Let  $p \in \mathbb{N}^*$ ,  $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$ ,  $F \in \mathcal{C}_b^3(\mathbb{R}^p, \mathbb{R})$  and  $g = (g_1, \dots, g_p)$  where for each  $i \in \{1, \dots, p\}$ ,  $g_i \in \mathcal{C}_b^3(\mathbb{R}, \mathbb{R})$ . Then,*

$$\begin{aligned}
& F(\langle g_1 \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle, \dots, \langle g_p \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle) \\
& \quad - F(\langle g_1 \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle, \dots, \langle g_p \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle) \\
& = \sum_{k=1}^p \left\{ \partial_k F(\langle g_1 \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle, \dots, \langle g_p \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle) \left( \langle g_k \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right. \right. \\
& \quad \left. \left. - \langle \text{id}, \mu - \nu \rangle [\langle g'_k \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle + \langle g'_k \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle] \right. \right. \\
& \quad \left. \left. + \frac{1}{2} \langle \text{id}, \mu - \nu \rangle^2 \langle g''_k \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \right) \right\} \\
& + \frac{1}{2} \sum_{i,j=1}^p \left\{ \partial_{ij}^2 F(\langle g_1 \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle, \dots, \langle g_p \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle) \left( \langle g_i \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right. \right. \\
& \quad \times \langle g_j \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle - \langle \text{id}, \mu - \nu \rangle \left[ \langle g'_j \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \langle g_i \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right. \\
& \quad \left. \left. + \langle g'_i \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \langle g_j \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right. \right. \\
& \quad \left. \left. - \langle g'_i \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \langle g'_j \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \langle \text{id}, \mu - \nu \rangle \right] \right) \right\} \\
& + O\left( |\langle \text{id}, \mu - \nu \rangle|^3 + \sum_{k=1}^p \sum_{\ell=0}^2 \left| \langle g_k^{(\ell)} \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right|^3 \right)
\end{aligned}$$

*Proof.* The general case  $p \in \mathbb{N}^*$  can be proved by a straightforward extension of the proof of the case  $p = 1$  which is the only case that we prove. Applying TAYLOR's formula to  $g \circ \tau_{-\langle \text{id}, \mu \rangle} =$

$g(\cdot - \langle \text{id}, \mu \rangle)$ , we obtain that

$$\begin{aligned} & \langle g \circ \tau_{-\langle \text{id}, \mu \rangle} - g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu \rangle \\ &= \int_{\mathbb{R}} \mu(dx) \left[ g'(x - \langle \text{id}, \nu \rangle) \langle \text{id}, \nu - \mu \rangle + \frac{1}{2} g''(x - \langle \text{id}, \nu \rangle) \langle \text{id}, \nu - \mu \rangle^2 + O(|\langle \text{id}, \nu - \mu \rangle|^3) \right] \\ &= -\langle \text{id}, \mu - \nu \rangle \left[ \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle + \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right] \\ &\quad + \frac{1}{2} \langle \text{id}, \mu - \nu \rangle^2 \left[ \langle g'' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle + \langle g'' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right] + O(|\langle \text{id}, \mu - \nu \rangle|^3). \end{aligned}$$

Therefore, we deduce the following approximation:

$$\begin{aligned} & \langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \\ &= \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle - \langle \text{id}, \mu - \nu \rangle \left[ \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle + \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \right] \tag{A.1} \\ &\quad + \frac{1}{2} \langle \text{id}, \mu - \nu \rangle^2 \langle g'' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle + O(|\langle \text{id}, \mu - \nu \rangle|^3 + \langle \text{id}, \mu - \nu \rangle^2 |\langle g'' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle|). \end{aligned}$$

Applying TAYLOR's formula to  $F(\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle)$ , we obtain that

$$\begin{aligned} & F(\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle) - F(\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle) \\ &= F'(\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle) \times [\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle] \\ &\quad + \frac{F''(\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle)}{2} \times [\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle]^2 \\ &\quad + O(|\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle|^3). \end{aligned}$$

Using (A.1), we deduce that

$$\begin{aligned} & [\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle]^2 \\ &= \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle^2 + \langle \text{id}, \mu - \nu \rangle^2 \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle^2 \\ &\quad - 2 \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \langle \text{id}, \mu - \nu \rangle \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle \\ &\quad + O(|\langle \text{id}, \mu - \nu \rangle|^3 + \langle \text{id}, \mu - \nu \rangle^2 \left[ |\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| \right. \\ &\quad \quad \left. + |\langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| + |\langle g'' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| \right] \\ &\quad \left. + |\langle \text{id}, \mu - \nu \rangle \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| \right), \end{aligned}$$

and

$$\begin{aligned} & O(|\langle g \circ \tau_{-\langle \text{id}, \mu \rangle}, \mu \rangle - \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \nu \rangle|^3) \\ &= O(|\langle \text{id}, \mu - \nu \rangle|^3 + \langle \text{id}, \mu - \nu \rangle^2 \left[ |\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| + |\langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle| \right] \\ &\quad + |\langle \text{id}, \mu - \nu \rangle| \left[ \langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle^2 + \langle g' \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle^2 \right] + |\langle g \circ \tau_{-\langle \text{id}, \nu \rangle}, \mu - \nu \rangle|^3). \end{aligned}$$

The announced result follows from YOUNG's inequalities.  $\square$

## A.2 Lemma of convergence

**Lemma A.2.1.** *Let  $k \in \mathbb{N}^*$ . We assume that  $\langle \text{id}^k, \nu \rangle < \infty$ . We consider for  $t \in [0, T]$ , an increasing sequence  $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = T$  of subdivisions of  $[0, T]$  whose mesh tends to 0. Then, for all  $k \in \mathbb{N}^*$ ,  $h_1 \in \{1, \text{id}\}$  and  $h_2 \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ , we obtain that, in  $\mathbb{P}_\nu^{FV}$ -probability,*

$$(1) \quad \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \left| \left\langle h_1 \times \left[ h_2 \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} - h_2 \circ \tau_{-\langle \text{id}, Y_s \rangle} \right], Y_s \right\rangle \right|^k ds = 0,$$

$$(2) \quad \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} |\langle h_1, Y_s \rangle|^k \left| \left\langle h_2 \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} - h_2 \circ \tau_{-\langle \text{id}, Y_s \rangle}, Y_s \right\rangle \right|^k ds = 0.$$

*Proof.* The two properties can be proved similarly. We only prove the first one. Thanks to Lemma 2.5.1 (1)(a) and using that  $h_2$  is LIPSCHITZ, there exists a constant  $C_{\text{Lip}}$  such that

$$\begin{aligned} & \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} \left| \left\langle h_1 \times \left[ h_2 \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} - h_2 \circ \tau_{-\langle \text{id}, Y_s \rangle} \right], Y_s \right\rangle \right|^k ds \\ & \leq \int_{t_i^n \wedge t}^{t_{i+1}^n \wedge t} |\langle h_1, Y_s \rangle|^k \left[ C_{\text{Lip}}^k \left| M_s^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right| \wedge (2 \|h_2\|_\infty)^k \right] ds. \end{aligned}$$

If  $h_1 = 1$ , the dominated convergence theorem allows us to conclude the proof. If  $h_1 = \text{id}$ , using that  $\sup_{t \in [0, T]} |\langle \text{id}, Y_t \rangle| < \infty$   $\mathbb{P}_\nu^{FV}$ -a.s. by Lemma 2.5.1 (1)(b), we can also apply the dominated convergence.  $\square$

## A.3 Control of error terms

**Lemma A.3.1.** *Let  $t > 0$  be fixed and assume that  $\langle \text{id}^2, \nu \rangle < \infty$ . Let  $j \in \{0, 1, 2\}$  and  $g \in \mathcal{C}_b^4(\mathbb{R}, \mathbb{R})$  fixed. The sequences*

$$(1) \quad \left( \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 \right)_{n \in \mathbb{N}}$$

$$(2) \quad \left( \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \right|^3 \right)_{n \in \mathbb{N}}$$

converge to 0 in  $\mathbb{P}_\nu^{FV}$ -probability.

*Proof. Step 1. Proof of (1).* Let  $\varepsilon > 0$  and  $t \geq 0$  fixed. Let  $A > 0$  to be determined later. We introduce the stopping time

$$\tau_A := \inf \left\{ t \geq 0 \mid \left| \langle \text{id}^2, Y_t \rangle - \langle \text{id}, Y_t \rangle^2 \right| \geq A \right\}$$

which satisfies almost surely  $\lim_{A \rightarrow +\infty} \tau_A = +\infty$  by Lemma 2.5.1. Then, using MARKOV's

inequality, we obtain that

$$\begin{aligned} & \mathbb{P}_\nu^{FV} \left( \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 > \varepsilon \right) \\ & \leq \mathbb{P}_\nu^{FV} (\tau_A \leq t) + \mathbb{P}_\nu^{FV} \left( \{\tau_A > t\} \cap \left\{ \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 > \varepsilon \right\} \right) \\ & \leq \mathbb{P}_\nu^{FV} (\tau_A \leq t) + \frac{1}{\varepsilon} \sum_{i=0}^{p_n-1} \mathbb{E} \left( \left| M_{t_{i+1}^n \wedge t \wedge \tau_A}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 \right). \end{aligned}$$

From the definition of  $\tau_A$ , we obtain for all  $s \in [t_i^n \wedge t, t_{i+1}^n \wedge t]$ ,  $\left| M_{s \wedge \tau_A}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3$  is bounded and  $\langle M^{\text{id}}(\text{id}) \rangle_{t \wedge \tau_A}$  is  $2\gamma A^2$ -LIPSCHITZ. Thanks to the BURKOLDER-DAVIS-GUNDY inequality and Lemma 2.5.1 (2), there exists a constant  $C_1$  such that

$$\begin{aligned} \mathbb{E} \left( \left| M_{t_{i+1}^n \wedge t \wedge \tau_A}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 \right) & \leq C_1 \mathbb{E} \left( \left[ \langle M^{\text{id}}(\text{id}) \rangle_{t_{i+1}^n \wedge t \wedge \tau_A} - \langle M^{\text{id}}(\text{id}) \rangle_{t_i^n \wedge t} \right]^{\frac{3}{2}} \right) \\ & \leq C_1 (2\gamma A^2)^{\frac{3}{2}} (t_{i+1}^n \wedge t - t_i^n \wedge t)^{\frac{3}{2}}. \end{aligned}$$

Therefore, if we choose  $A$  such that  $\mathbb{P}_\nu^{FV} (\tau_A \leq t) \leq \frac{\varepsilon}{2}$ , and  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\sqrt{\sup_{i \in [0, p_n-1]} |t_{i+1}^n \wedge t - t_i^n \wedge t|} \leq \frac{\varepsilon^2}{2C_1 (2\gamma A^2)^{\frac{3}{2}} t},$$

we obtain that

$$\begin{aligned} & \mathbb{P}_\nu^{FV} \left( \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}}(\text{id}) - M_{t_i^n \wedge t}^{\text{id}}(\text{id}) \right|^3 > \varepsilon \right) \\ & \leq \mathbb{P}_\nu^{FV} (\tau_A \leq t) + \frac{C_1 (2\gamma A^2)^{\frac{3}{2}}}{\varepsilon} t \sqrt{\sup_{i \in [0, p_n-1]} |t_{i+1}^n \wedge t - t_i^n \wedge t|} \\ & \leq \varepsilon, \end{aligned}$$

and the first announced result follows.

**Step 2. Proof of (2).** In similar way as previously, we obtain that

$$\begin{aligned} & \mathbb{P}_\nu^{FV} \left( \sum_{i=0}^{p_n-1} \left| M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \right|^3 > \varepsilon \right) \\ & \leq \frac{1}{\varepsilon} \sum_{i=0}^{p_n-1} \mathbb{E} \left( \left| M_{t_{i+1}^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) - M_{t_i^n \wedge t}^{\text{id}} \left( g^{(j)} \circ \tau_{-\langle \text{id}, Y_{t_i^n \wedge t} \rangle} \right) \right|^3 \right) \\ & \leq \frac{C_1 \left( 2\gamma \|g^{(j)}\|_\infty^2 \right)^{\frac{3}{2}}}{\varepsilon} t \sqrt{\sup_{i \in [0, p_n-1]} |t_{i+1}^n \wedge t - t_i^n \wedge t|}, \end{aligned}$$

which converges to 0 when  $n \rightarrow +\infty$ . □



# Appendix B

## R code for IBM simulation

Here is the R-source code of the program simulating the versions of KISDI's model [140], firstly for the IBM of Section 1.1.3.2 with parameters (1.13) and secondly for the IBM of Section 1.3.3 with parameters (1.72). This program directly creates one or two `.png` images similar to the figures of Sections 1.1.3.2 and 1.3.3 of the introduction.

Our program is based on the `IBMPopSim` package developed by GIORGI ET AL in [116]. Their package makes it possible to simulate interacting age and trait-structured populations and allows resampling events. In our situation, we have no age-structure, so we set the maximal life length parameter of each individuals to  $+\infty$ . The R-topics of various commands used in our program are given at <https://cran.r-project.org/web/packages/IBMPopSim/IBMPopSim.pdf>.

When there are swap events (individuals can change of characteristics), the dates of swap events and the changes of characteristics following each swap event should be recorded for each individual in the population, which is a memory intensive and computationally costly process. To maintain efficient simulations in the presence of swap events, GIORGI ET AL propose the following solution.

In the presence of swap events, the argument `time` of `popsim` should a vector of dates  $(t_0, \dots, t_n)$ . In this case, `popsim` returns in the object `population` a list of  $n$  population data frames representing the population at time  $t_1, \dots, t_n$  simulated from the initial time  $t_0$ . For  $i \in \{1, \dots, n\}$ , the  $i^{\text{th}}$  data frame describes individuals who lived in the population during the period  $[t_0, t_i]$ , with their characteristics **at time**  $t_i$ .

### B.1 R-code for the IBM for constant population with swap events

```
1 # Current directory
2 getwd() # To know the current directory
3 setwd("/users/hass/Mes_images_simulations") # To choose the new current
  directory
4
5 # Installation of the package IBMPopSim
6 install.packages("IBMPopSim")
7
8 K <- 100 # Number of individuals in the initial population
9
10 # Creation of the trait-structured population
11 pop_init <- data.frame(
```

```

12 "birth" = rep(0,K), # Age of each individual is the same
13 "death" = as.double(NA), # Death dates of individuals
14 "trait" = rep(1,K) # All individuals have initially the same trait 1 (
    Monomorphic population)
15
16 # The two previous commands concerning the birth and death dates (fixed) of
    the individuals are necessary to make the algorithm work but are not
    exploited, only the trait structure matters us.
17
18 # List of parameters
19 param <- list("p" = 0.1, "sigma" = 0.03, "jtrait"=0.)
20
21 # Remarks (1) CUnif is a realisation of a uniform law on [0,1]
22 #           (2) CNorm(0, 1) is a realisation of Gaussian distribution with mean 0
    and variance 1
23
24 # Events creation
25
26 # (1) Event describing the effect of mutations
27 event1 <- mk_event_individual( type = "swap",
28   intensity_code = 'result = 1;', # Rate of mutation : Theta = 1
29   kernel_code = 'if (CUnif() < p)
30     I.trait = min(max(0.,I.trait + sigma*CNorm(0, 1)),4.);') #
    New value of the mutant trait according to the m-law conditional on the
    mutant trait remaining in the trait space [0,4]
31
32 # (2) Resampling event of one individual by another
33 event2 <- mk_event_interaction(type = "swap", name = "swap2",
34   interaction_code = 'result = (4. - J.trait) * ( 1.- 1./
    (1. + 2. * exp(-4. * (I.trait- J.trait))))); # Resampling rate b(x,y) where x
    = I.trait and y = J.trait
35     jtrait = J.trait; ',
36   kernel_code = 'I.trait = jtrait;')
37
38 # Creation of the IBM
39 My_model <- mk_model(
40   characteristics = get_characteristics(pop_init), # Initial model population
41   events = list(event1, event2), # Model events
42   parameters = param # Model parameters
43   )
44
45 # Events bounds
46 b_event1 <- 1 # Bound of the mutation rate (Theta)
47 b_event2 <- 4 # Bounds of the resampling rate (b)
48
49 # Time of simulation
50 T <- 600;
51 STEP <- 1/100;
52 t_sim <- seq(0,T,by = STEP); # Vector of dates (t0, ... , tn) from t0 = 0 up to
    time T with time step of STEP : ti = t0 + i * STEP
53
54 # Simulation of the random evolution of the IBM
55 sim_out <- popsim(
56   model = My_model,
57   population = pop_init,
58   events_bounds = c('swap' = b_event1, 'swap2' = b_event2),
59   parameters = param,
60   time = t_sim)
61

```

```

62 pop_out <- sim_out$population # Data frame containing the date of birth, death,
   and traits of all individuals who lived in the population on each interval [
   t0, ti].
63
64 # Graphical illustration: plot of the evolution of the population trait
   distribution
65
66 # (1) Creation of data frame containing the traits of all individuals who lived
   in the population in [0,T] with time step of STEP by iteration and
   concatenation
67
68 My_YY = pop_out[[1]]$trait
69 for (i in 2:((1/STEP)*T)){
70   My_YY = c(My_YY, pop_out[[i]]$trait)
71 }
72
73 # Creation of the corresponding abscissa (by iteration and concatenation)
74 My_XX = rep(1,K)
75 for (i in 2:((1/STEP)*T)){
76   My_XX = c(My_XX, rep(i*STEP,K))
77 }
78
79 # (2) Creation of .png figure in the current directory : we graphically
   represent all the traits present in the population at each time instant ti
80
81 png(file = "CEAD_Swap_K_100_T_600_sigma_0_03_p_0_1.png")
82 plot(My_XX, My_YY, xlab = "Time", ylab = "Traits", col = "blue", pch = 46)
83 dev.off()

```

## B.2 R-code for the IBM with fluctuate population size

```

1 # Current directory
2 getwd() # To know the current directory
3 setwd("/users/hass/Mes_images_simulations") # To choose the new current
   directory
4
5 # Installation of the package IBMPopSim
6 install.packages("IBMPopSim")
7
8 # Library for graphing
9 library(ggplot2)
10 library(colorspace)
11
12 K <- 300 # Number of individuals in the initial population
13
14 # Creation of the trait-structured population
15 pop_init <- data.frame(
16   "birth" = rep(0,K), # Age of each individual is the same
17   "death" = as.double(NA), # Death dates of individuals
18   "trait" = rep(1,K) # All individuals have initially the same trait 1 (
   Monomorphic population)
19
20 # List of parameters
21
22 # (1) Parameters for birth events
23 params_birth <- list(
24   "p" = 1,

```

```

25 "sigma" = 0.0005)
26
27 # (2) Parameters for death events
28 params_death <- list(
29   "beta" = 2.,
30   "c" = 1.2,
31   "N" = K)
32
33 # Remarks (1) CUnif is a realisation of a uniform law on [0,1]
34 #           (2) CNorm(0, 1) is a realisation of Gaussian distribution with mean 0
35           and variance 1
36
37 # Event creation
38
39 # (1) Birth event
40 birth_event <- mk_event_individual(
41   type = "birth",
42   intensity_code = 'result = 4-I.trait;',
43   kernel_code = 'if (CUnif() < p)
44                 newI.trait = min(max(0., CNorm(I.trait, sigma)), 4.);
45                 else
46                 newI.trait = I.trait;')
47
48 # (2) Death event
49 death_event <- mk_event_interaction(
50   type = "death",
51   interaction_code = "result = (beta / N) * ( 1.- 1./(1. + c * exp(-4. * (I.
52     trait- J.trait)))));"
53 )
54
55 # Creation of the IBM
56 My_model <- mk_model(
57   characteristics = get_characteristics(pop_init), # Intial model population
58   events = list(birth_event, death_event), # Model events
59   parameters = c(params_birth, params_death) # Model parameters obtained by
60     concatenation of params_birth and params_death
61 )
62
63 # Events bounds
64 birth_bound <- 4 # Bound of the birth rate (b)
65 interaction_bound <- 2 # Bound of the competition rate (c)
66
67 # Time of simulation
68 T <- 200000
69
70 # Simulation of the random evolution of the IBM on [0, T]
71 sim_out <- popsim(model = model,
72   population = pop_init,
73   events_bounds = c('birth'=birth_bound, 'death'=interaction_
74     bound),
75   parameters = c(params_birth, params_death),
76   time = T)
77
78 pop_out <- sim_out$population # Data frame containing the date of birth, death,
79   and traits of all individuals who lived in the population over the period
80   [0, T]
81
82 pop_size <- nrow(population_alive(pop_out, t = T)) # Number of individuals alive
83   at final time T

```

```

77
78 # Graphical illustrations
79
80 # (1) Graphical illustration of the evolution of the population size
81 # For a clear illustration, we limit ourselves to about 3,000 points over the
    period [0,T].
82
83 Nb_points_start <- 15
84 Nb_points_middle <- 66
85 Nb_points_final <- 10
86
87 Start_range <- seq(0, 1, 1/Nb_points_start) # Sequence of numbers between 0 and
    1 with a step of 1/Nb_points_start
88 Middle_range <- seq(1, T-100, Nb_points_middle) # Sequence of numbers between 1
    and T-100 with a step of Nb_points_middle
89 Final_range <- seq(T-100, T, Nb_points_final) # Sequence of numbers between T
    -100 and T with a step of Nb_points_final
90
91 My_range <- c(Start_range, Middle_range, Final_range) # Concatenation of the
    three previous ranges
92
93 # (1)(a) Creation of vectors of population size for each date of My_range
94 Pop_size_1 <- c()
95 for(i in 1 : length(Start_range))
96 {
97   Pop_size_1 <- c(Dates_1, nrow(population_alive(pop_out, t = i/Nb_points_start)
    )) # Vector of population size for each date of Start_range
98 }
99
100 Pop_size_2 <- c()
101 for(i in 1 : length(Middle_range))
102 {
103   Pop_size_2 <- c(Pop_size_2, nrow(population_alive(pop_out, t = (i * Nb_points_
    middle)))) # Vector of population size for each date of Middle_range
104 }
105
106 Pop_size_3 <- c()
107 for(i in 1 : length(Final_range))
108 {
109   Pop_size_3 <- c(Pop_size_3, nrow(population_alive(pop_out, t = (max(Middle_
    range) + i)))) # Vector of population size for each date of Final_range
110 }
111
112 My_pop_size <- c(Pop_size_1, Pop_size_2, Pop_size_3) # Concatenation of the
    three previous vectors
113
114 # (1)(b) Plot of the evolution of the population size
115 png(file = "Pop_size_K_300_T_200000_sigma_0_0005_p_1.png") # Creation of .png
    figure in the current directory
116 plot(x = My_range, y = My_pop_size, col = "blue", pch = 19, cex = 0.01, main = "
    Evolution of the population size", ylab = "Number of individuals", xlab = "
    Time")
117 dev.off()
118
119 # (2) Graphical illustration of the evolution of the population trait
    distribution
120 My_draw <- ggplot(pop_out) + geom_segment(
121   aes(x=birth, xend=death, y=trait, yend=trait),
122   na.rm=TRUE, colour = "blue", alpha = 0.1) +

```

*Appendix B. R code for IBM simulation*

---

```
123   labs(title = "Evolution of the trait distribution", x = "Time", y = "Traits")
124
125   ggsave("CEAD_K_300_T_200000_sigma_0_0005_p_1.png", # Creation of .png figure in
126           the current directory with ggplot
127           plot = My_draw,
128           width = 6, height = 4.5
129   )
```

# Bibliography

- [1] P. A. ABRAMS, H. MATSUDA, and Y. HARADA. Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits. *Evolutionary Ecology*, 7(5) :465–487, 1993.
- [2] D. ALDOUS. Stopping times and tightness. *The Annals of Probability*, 6(2) :335–340, 1978.
- [3] C.B. ANFINSEN. Principles that Govern the Folding of Protein Chains. *Science*, 181(4096) :223–230, 1973.
- [4] K. B. ATHREYA and P. E. NEY. *Branching processes*. Number 196 in Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, Berlin Heidelberg, 1972.
- [5] M. BAAR. *Stochastic individual-based models of adaptive dynamics and applications to cancer immunotherapy*. Thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, March 2017.
- [6] M. BAAR, A. BOVIER, and N. CHAMPAGNAT. From stochastic, individual-based models to the canonical equation of adaptive dynamics in one step. *The Annals of Applied Probability*, 27(2), 2017.
- [7] M. BAAR, L. COQUILLE, H. MAYER, M. HÖLZEL, M. ROGAVA, T. TÜTING, and A. BOVIER. A stochastic model for immunotherapy of cancer. *Scientific Reports*, 6(1) :24169, 2016.
- [8] K. BALL, T. G. KURTZ, L. POPOVIC, and G. REMPALA. Asymptotic analysis of multiscale approximations to reaction networks. *The Annals of Applied Probability*, 16(4), 2006.
- [9] G. BALLIF, F. CLÉMENT, and R. YVINEC. Averaging of a Stochastic Slow-Fast Model for Population Dynamics : Application to the Development of Ovarian Follicles. *SIAM Journal on Applied Mathematics*, 82(1) :359–380, 2022.
- [10] V. BANSAYE and S. MÉLÉARD. *Stochastic models for structured populations - Scaling limits and long time behavior*. Springer International Publishing, Cham, 2015.
- [11] G. BARLES, L. C. EVANS, and P. E. SOUGANIDIS. Wavefront propagation for reaction-diffusion systems of PDE. *Duke Mathematical Journal*, 61(3), 1990.
- [12] G. BARLES, S. MIRRAHIMI, and B. PERTHAME. Concentration in Lotka-Volterra parabolic or integral equations : a general convergence result. *Methods and Applications of Analysis*, 16(3) :321–340, 2009.
- [13] G. BARLES and B. PERTHAME. Concentrations and constrained Hamilton-Jacobi equations arising in adaptive dynamics. In D. DANIELLI, editor, *Contemporary Mathematics*, volume 439, pages 57–68. American Mathematical Society, Providence, Rhode Island, 2007.
- [14] G. BEN AROUS, N. CHAMPAGNAT, and R. FERRIÈRE. The canonical equation of adaptive dynamics : a mathematical view. *Selection*, 2 :71–81, 2001.
- [15] S. BENGTTSSON, J. HALLQUIST, A. WERKER, and T. WELANDER. Acidogenic fermentation of industrial wastewaters : Effects of chemostat retention time and pH on volatile fatty acids production. *Biochemical Engineering Journal*, 40(3) :492–499, 2008.
- [16] T. BENOIST, C. BERNARDIN, R. CHETRIT, R. CHHAIBI, J. NAJNUDEL, and C. PELLEGRINI. Emergence of Jumps in Quantum Trajectories via Homogenization. *Communications in Mathematical Physics*, 387(3) :1821–1867, 2021.
- [17] A. BENSOUSSAN, J.-L. LIONS, and G. PAPANICOLAOU. *Asymptotic analysis for periodic structures*. American Mathematical Society, Providence, R.I, 2011.

- [18] S. BILLIARD, P. COLLET, S. MÉLÉARD, and V.C. TRAN. Stochastic dynamics for adaptation and evolution of microorganism. *European Congress of Mathematics V*. MEHRMANN and M. SKUTELLA, editors, pages 525–550, 2018. European Mathematical Society, Zürich.
- [19] S. BILLIARD, R. FERRIÈRE, S. MÉLÉARD, and V.C. TRAN. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *Journal of Mathematical Biology*, 71(5) :1211–1242, 2015.
- [20] P. BILLINGSLEY. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics, Second edition. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1999.
- [21] M. BIRKNER and J. BLATH. Generalised Stable Fleming-Viot Processes as Flickering Random Measures. *Electronic Journal of Probability*, 14 :2418–2437, 2009.
- [22] G. BLANKENSHIP and G. C. PAPANICOLAOU. Stability and Control of Stochastic Systems with Wide-band Noise Disturbances. I. *SIAM Journal on Applied Mathematics*, 34(3) :437–476, 1978.
- [23] J. BLATH, T. PAUL, and A. TÓBIÁS. A stochastic adaptive dynamics model for bacterial populations with mutation, dormancy and transfer. *ALEA, Latin American Journal of Probability and Mathematical Statistics*, 20(1) :313, 2023.
- [24] J. BLATH and A. TÓBIÁS. Invasion and fixation of microbial dormancy traits under competitive pressure. *Stochastic Processes and their Applications*, 130(12) :7363–7395, 2020.
- [25] J. BLATH and A. TÓBIÁS. The interplay of dormancy and transfer in bacterial populations : Invasion, fixation and coexistence regimes. *Theoretical Population Biology*, 139 :18–49, 2021.
- [26] B. BOLKER and S. W. PACALA. Using Moment Equations to Understand Stochastically Driven Spatial Pattern Formation in Ecological Systems. *Theoretical Population Biology*, 52(3) :179–197, 1997.
- [27] B. BOLKER and S. W. PACALA. Spatial Moment Equations for Plant Competition : Understanding Spatial Strategies and the Advantages of Short Dispersal. *The American Naturalist*, 153(6) :575–602, 1999.
- [28] C. BONNET and S. MÉLÉARD. Large fluctuations in multi-scale modeling for rest hematopoiesis. *Journal of Mathematical Biology*, 82(6) :58, 2021.
- [29] A. BOVIER, L. COQUILLE, and C. SMADI. Crossing a fitness valley as a metastable transition in a stochastic population model. *The Annals of Applied Probability*, 29(6), 2019.
- [30] J. S. BROWN and T. L. VINCENT. Coevolution as an evolutionary game. *Evolution*, 41(1) :66–79, 1987.
- [31] J. S. BROWN and T. L. VINCENT. Predator-prey coevolution as an evolutionary game. In S. LEVIN and Y. COHEN, editors, *Applications of Control Theory in Ecology*, volume 73, pages 83–101. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987. Series Title : Lecture Notes in Biomathematics.
- [32] J. S. BROWN and T. L. VINCENT. Organization of predator-prey communities as an evolutionary game. *Evolution*, 46(5) :1269–1283, 1992.
- [33] J. A. CARRILLO, S. CUADRADO, and B. PERTHAME. Adaptive dynamics via Hamilton-Jacobi approach and entropy methods for a juvenile-adult model. *Mathematical Biosciences*, 205(1) :137–161, 2007.
- [34] S. CERRAI. A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *The Annals of Applied Probability*, 19(3), 2009.
- [35] S. CERRAI and M. FREIDLIN. Averaging principle for a class of stochastic reaction-diffusion equations. *Probability Theory and Related Fields*, 144(1-2) :137–177, 2009.
- [36] N. CHAMPAGNAT. *Mathematical study of stochastic models of evolution belonging to the ecological theory of adaptive dynamics*. Theses, Université de Nanterre - Paris X, 2004.
- [37] N. CHAMPAGNAT. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116(8) :1127–1160, 2006.



- 
- [38] N. CHAMPAGNAT. *Approches stochastiques et déterministes en biologie : dynamique adaptative, modélisation pour l'écologie, génétique des populations et dynamique moléculaire ; caractère bien posé d'équations différentielles ordinaires et stochastiques*. HDR en mathématiques appliquées, Université de Lorraine, 2015, 128 p.
- [39] N. CHAMPAGNAT, R. FERRIÈRE, and S. MÉLÉARD. Unifying evolutionary dynamics : From individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69(3) :297–321, 2006.
- [40] N. CHAMPAGNAT and V. HASS. Convergence of individual-based models with small and frequent mutations to the canonical equation of adaptive dynamics. *Preprint (arXiv :2303.09936v1)*, 2023.
- [41] N. CHAMPAGNAT and V. HASS. Existence, uniqueness and ergodicity for the centered Fleming-Viot process. *Stochastic Processes and their Applications*, 166 :104219, 2023.
- [42] N. CHAMPAGNAT and P.-E. JABIN. The evolutionary limit for models of populations interacting competitively via several resources. *Journal of Differential Equations*, 251(1) :176–195, 2011.
- [43] N. CHAMPAGNAT, P.-E. JABIN, and S. MÉLÉARD. Adaptation in a stochastic multi-resources chemostat model. *Journal de Mathématiques Pures et Appliquées*, 101(6) :755–788, 2014.
- [44] N. CHAMPAGNAT and A. LAMBERT. Evolution of discrete populations and the canonical diffusion of adaptive dynamics. *The Annals of Applied Probability*, 17(1), 2007.
- [45] N. CHAMPAGNAT and A. LAMBERT. Adaptive dynamics in logistic branching populations. *Banach Center Publications*, 80(1) :235–244, 2008.
- [46] N. CHAMPAGNAT and S. MÉLÉARD. Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55(2) :147–188, 2007.
- [47] N. CHAMPAGNAT and S. MÉLÉARD. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151(1-2) :45–94, 2011.
- [48] N. CHAMPAGNAT, S. MÉLÉARD, S. MIRRAHIMI, and V.C. TRAN. Filling the gap between individual-based evolutionary models and Hamilton-Jacobi equations, 2022. arXiv :2205.00770 [math, q-bio].
- [49] N. CHAMPAGNAT, S. MÉLÉARD, and V.C. TRAN. Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. *The Annals of Applied Probability*, 31(4), 2021.
- [50] B. CHARLESWORTH. *Evolution in Age-Structured Populations*. Cambridge University Press, 2 edition, 1994.
- [51] P. COLLET, S. MÉLÉARD, and J. A. J. METZ. A rigorous model study of the adaptive dynamics of Mendelian diploids. *Journal of Mathematical Biology*, 67(3) :569–607, 2013.
- [52] L. COQUILLE, A. KRAUT, and C. SMADI. Stochastic individual-based models with power law mutation rate on a general finite trait space. *Electronic Journal of Probability*, 26, 2021.
- [53] C. CORON. Slow-fast stochastic diffusion dynamics and quasi-stationarity for diploid populations with varying size. *Journal of Mathematical Biology*, 72(1-2) :171–202, 2016.
- [54] M. COSTA. A piecewise deterministic model for a prey-predator community. *The Annals of Applied Probability*, 26(6) :3491–3530, 2016.
- [55] M. COSTA, C. HAUZY, N. LOEUILLE, and S. MÉLÉARD. Stochastic eco-evolutionary model of a prey-predator community. *Journal of Mathematical Biology*, 72(3) :573–622, 2016.
- [56] C. COSTANTINI and T. G. KURTZ. Diffusion approximation for transport processes with general reflection boundary conditions. *Mathematical Models and Methods in Applied Sciences*, 16(05) :717–762, 2006.
- [57] C. DARWIN. *On the origin os species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, 1859.
- [58] C. DARWIN and A. WALLACE. On the tendency of species to form varieties ; and on the perpetuation of varieties and species by natural means of selection. *Zoological Journal of the Linnean Society of London*, 3 :45–62, 1858.

- [59] D.A. DAWSON. Geostochastic calculus. *Canadian Journal of Statistics*, 6(2) :143–168, 1978.
- [60] D.A. DAWSON. *Measure-Valued Markov processes*. P.-L. HENNEQUIN (editor) École d'Été de probabilités de Saint-Flour XXI. Lecture Notes in Mathematics, Vol. 1541 : 1-260, 1993.
- [61] D.A. DAWSON and K. HOCHBERG. Wandering random measures in the Fleming-Viot model. *The Annals of Probability*, 10(3) :554–580, 1982.
- [62] D.A. DAWSON and P. MARCH. Resolvent Estimates for Fleming-Viot Operators and Uniqueness of Solutions to Related Martingale Problems. *Journal of Functional Analysis*, 132(2) :417–472, 1995.
- [63] D. L. DEANGELIS. *Individual-Based Models and Approaches in Ecology : Populations, Communities and Ecosystems*. Chapman and Hall/CRC, 1 edition, 2018.
- [64] A. DEMBO and O. ZEITOUNI. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [65] F. DERCOLE, U. DIECKMANN, M. OBERSTEINER, and S. RINALDI. Adaptive dynamics and technological change. *Technovation*, 28(6) :335–348, 2008.
- [66] F. DERCOLE, R. FERRIÈRE, A. GRAGNANI, and S. RINALDI. Coevolution of slow-fast populations : evolutionary sliding, evolutionary pseudo-equilibria and complex Red Queen dynamics. *Proceedings of the Royal Society B : Biological Sciences*, 273(1589) :983–990, 2006.
- [67] F. DERCOLE, R. FERRIÈRE, and S. RINALDI. Ecological bistability and evolutionary reversals under asymmetrical competition. *Evolution*, 56(6) :1081–1090, 2002.
- [68] F. DERCOLE and S. RINALDI. *Analysis of Evolutionary Processes : The Adaptive Dynamics Approach and Its Applications*. Princeton University Press, 2008.
- [69] L. DESVILLETES, P.-E. JABIN, S. MISCHLER, and G. RAOUL. On selection dynamics for continuous structured populations. *Communications in Mathematical Sciences*, 6(3) :729–747, 2008.
- [70] U. DIECKMANN and M. DOEBELI. On the origin of species by sympatric speciation. *Nature*, 400(6742) :354–357, 1999.
- [71] U. DIECKMANN and R. FERRIÈRE. Adaptive dynamics and evolving biodiversity. In R. FERRIÈRE, U. DIECKMANN, and D. COUVET, editors, *Evolutionary Conservation Biology*, pages 188–224. Springer Berlin Heidelberg, Cambridge, 2004. Cambridge University Press.
- [72] U. DIECKMANN and R. LAW. The dynamical theory of coevolution : a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34(5-6) :579–612, 1996.
- [73] U. DIECKMANN and R. LAW. Relaxation projections and the method of moments. pages 412–455. Cambridge University Press, Cambridge, 2000.
- [74] O. DIECKMANN. A beginner's guide to adaptive dynamics. *Banach Center Publications*, 63(1) :47–86, 2003.
- [75] O. DIECKMANN, P.-E. JABIN, S. MISCHLER, and B. PERTHAME. The dynamics of adaptation : An illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67(4) :257–271, 2005.
- [76] M. DOEBELI and U. DIECKMANN. Evolutionary Branching and Sympatric Speciation Caused by Different Types of Ecological Interactions. *The American Naturalist*, 156(S4) :77–101, 2000.
- [77] M. DOEBELI and U. DIECKMANN. Speciation along environmental gradients. *Nature*, 421(6920) :259–264, 2003.
- [78] P. DONNELLY and T.G. KURTZ. A countable representation of the Fleming-Viot measure-valued diffusion. *The Annals of Probability*, 24(2), 1996.
- [79] P. DONNELLY and T.G. KURTZ. Genealogical processes for Fleming-Viot models with selection and recombination. *The Annals of Applied Probability*, 9(4) :1091–1148, 1999.
- [80] P. DONNELLY and T.G. KURTZ. Particle Representations for Measure-Valued Population Models. *The Annals of Probability*, 27(1), 1999.
- [81] R. M. DUDLEY. *Real analysis and probability*. Number 74 in Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge ; New York, 2002.

- 
- [82] M. DURINX, J. A. J. METZ, and G. MESZÉNA. Adaptive dynamics for physiologically structured population models. *Journal of Mathematical Biology*, 56(5) :673–742, 2008.
- [83] R. DURRETT and S. LEVIN. The Importance of Being Discrete (and Spatial). *Theoretical Population Biology*, 46(3) :363–394, 1994.
- [84] B. ERNANDE, U. DIECKMANN, and M. HEINO. Adaptive changes in harvested populations : plasticity and evolution of age and size at maturation. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 271(1537) :415–423, 2004.
- [85] A. ETHERIDGE. *An Introduction to Superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2000.
- [86] A. ETHERIDGE. *Some Mathematical Models from Population Genetics*, volume 2012 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [87] A. ETHERIDGE and T. G. KURTZ. Genealogical constructions of population models. *The Annals of Probability*, 47(4) :1827–1910, 2019.
- [88] S. N. ETHIER and T. G. KURTZ. The infinitely-many-neutral-alleles diffusion model. *Advances in Applied Probability*, 13(3) :429–452, 1981.
- [89] S. N. ETHIER and T. G. KURTZ. The Infinitely-Many-Alleles Model with Selection as a Measure-Valued Diffusion. In *Stochastic Methods in Biology*, volume 70, pages 72–86. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987.
- [90] S. N. ETHIER and T. G. KURTZ. Coupling and ergodic theorems for Fleming-Viot processes. *The Annals of Probability*, 26(2), 1998.
- [91] S. N. ETHIER and T.G. KURTZ. Fleming-Viot Processes in Population Genetics. *SIAM Journal on Control and Optimization*, 31(2) :345–386, 1993.
- [92] S. N. ETHIER and T. SHIGA. A Fleming-Viot Process with Unbounded Selection, II. In Z. HOU, J.A. FILAR, and A. CHEN, editors, *Markov Processes and Controlled Markov Chains*, pages 305–322. Springer US, Boston, MA, 2002.
- [93] S.N. ETHIER and T. G. KURTZ. Convergence to Fleming-Viot processes in the weak atomic topology. *Stochastic Processes and their Applications*, 54(1) :1–27, 1994.
- [94] S.N. ETHIER and T.G. KURTZ. *Markov Processes*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 1986.
- [95] W. J. EWENS. *Mathematical population genetics. I. Theoretical introduction*, volume 27 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, NY, 2004.
- [96] R. FERRIERE, J. L. BRONSTEIN, S. RINALDI, R. LAW, and M. GAUDUCHON. Cheating and the evolutionary stability of mutualisms. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 269(1493) :773–780, 2002.
- [97] R. FERRIÈRE and V. C. TRAN. Stochastic and deterministic models for age-structured populations with genetically variable traits. *ESAIM : Proceedings*, 27 :289–310, 2009.
- [98] R. A. FISHER. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52(2) :399–433, 1919.
- [99] W. FLEMING and M. VIOT. Some measure-valued Markov processes in population genetics theory. *Indiana University Mathematics Journal*, 28(5) :817–843, 1979.
- [100] J.-P. FOUQUE, J. GARNIER, G. PAPANICOLAOU, and K. SØLNA. *Wave Propagation and Time Reversal in Randomly Layered Media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer New York, New York, NY, 2007.
- [101] J. P. FOUQUE, G. PAPANICOLAOU, R. SIRCAR, and K. SØLNA. Singular Perturbations in Option Pricing. *SIAM Journal on Applied Mathematics*, 63(5) :1648–1665, 2003.
- [102] N. FOURNIER and S. MÉLÉARD. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4) : 1880 - 1919, 2004.

- [103] M. FREIDLIN. Limit theorems for large deviations and reaction-diffusion equations. *The Annals of Probability*, 13(3) :639–675, 1985.
- [104] M. I. FREIDLIN and A. D. WENTZELL. *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer US, New York, NY, 1984.
- [105] H. FU, L. WAN, and J. LIU. Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. *Stochastic Processes and their Applications*, 125(8) :3255–3279, 2015.
- [106] M. GALTIER and G. WAINRIB. Multiscale analysis of slow-fast neuronal learning models with noise. *The Journal of Mathematical Neuroscience*, 2(1) :13, 2012.
- [107] H. GAO and Y. SHI. Averaging principle for a stochastic coupled fast-slow atmosphere-ocean model. *Journal of Differential Equations*, 298 :248–297, 2021.
- [108] A. GENADOT. Averaging for some simple constrained Markov processes. *Probability and Mathematical Statistics*, 39(1) :139–158, 2019.
- [109] A. GENADOT and M. THIEULLEN. Averaging for a Fully Coupled Piecewise-Deterministic Markov Process in Infinite Dimensions. *Advances in Applied Probability*, 44(3) :749–773, 2012.
- [110] A. GENADOT and M. THIEULLEN. Multiscale Piecewise Deterministic Markov Process in infinite dimension : central limit theorem and Langevin approximation. *ESAIM : Probability and Statistics*, 18 :541–569, 2014.
- [111] S. A. H. GERITZ. Resident-invader dynamics and the coexistence of similar strategies. *Journal of Mathematical Biology*, 50(1) :67–82, 2005.
- [112] S. A. H. GERITZ, J. A. J. METZ, É. KISDI, and G. MESZÉNA. Dynamics of Adaptation and Evolutionary Branching. *Physical Review Letters*, 78(10) :2024–2027, 1997.
- [113] S.A.H. GERITZ, M. GYLLENBERG, F.J.A. JACOBS, and K. PARVINEN. Invasion dynamics and attractor inheritance. *Journal of Mathematical Biology*, 44(6) :548–560, 2002.
- [114] S.A.H. GERITZ, É. KISDI, G. MESZÉNA, and J.A.J. METZ. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12(1) :35–57, 1998.
- [115] S.A.H. GERITZ, E. VAN DER MEIJDEN, and J.A.J. METZ. Evolutionary Dynamics of Seed Size and Seedling Competitive Ability. *Theoretical Population Biology*, 55(3) :324–343, 1999.
- [116] D. GIORGI, S. KAKAI, and V. LEMAIRE. *IBMPopSim : Individual Based Model Population Simulation*, 2023. <https://github.com/DaphneGiorgi/IBMPopSim>, <https://DaphneGiorgi.github.io/IBMPopSim/>.
- [117] A. GONZALEZ CASANOVA and C. SMADI. On  $\Lambda$ -Fleming-Viot processes with general frequency-dependent selection. *Journal of Applied Probability*, 57(4) :1162–1197, 2020.
- [118] V. GRIMM and S. F. RAILSBACK. *Individual-based Modeling and Ecology*. Princeton University Press, 2005.
- [119] A. GUPTA, J. A. J. METZ, and V. C. TRAN. A New Proof for the Convergence of an Individual Based Model to the Trait Substitution Sequence. *Acta Applicandae Mathematicae*, 131(1) :1–27, 2014.
- [120] M. E. GURTIN and R. C. MACCAMY. Non-linear age-dependent population dynamics. *Archive for Rational Mechanics and Analysis*, 54(3) :281–300, 1974.
- [121] M. GYLLENBERG, K. PARVINEN, and U. DIECKMANN. Evolutionary suicide and evolution of dispersal in structured metapopulations. *Journal of Mathematical Biology*, 45(2) :79–105, 2002.
- [122] J.B.S. HALDANE. A mathematical theory of natural and artificial selection. a series of ten scientific papers published between 1924 and 1934. *Transactions of the Cambridge Philosophical Society, Proceedings of the Cambridge Philological Society and Genetics*, 3, 1924-1934.
- [123] W.D. HAMILTON. The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1) :1–16, 1964.

- 
- [124] H. HE. Fleming-Viot processes in an environment. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 13(03) :489–509, 2010.
- [125] J. HOFBAUER and K. SIGMUND. *The theory of evolution and dynamical systems : mathematical aspects of selection*. Cambridge University Press, New York, 1988.
- [126] J. HOFBAUER and K. SIGMUND. Adaptive dynamics and evolutionary stability. *Applied Mathematics Letters*, 3(4) :75–79, 1990.
- [127] N. IKEDA and S. WATANABE. *Stochastic Differential Equations and Diffusion Processes*, volume 24 of *North-Holland Mathematical Library, Second edition*. North-Holland Pub. Co., Japan, 1989.
- [128] Y. IWASA, A. POMIANKOWSKI, and S. NEE. The evolution of costly mate preferences ii. the “handicap” principle. *Evolution*, 45(6) :1431–1442, 1991.
- [129] P.-E. JABIN. Small populations corrections for selection-mutation models. *Networks & Heterogeneous Media*, 7(4) :805–836, 2012.
- [130] P.-E. JABIN and G. RAOUL. On selection dynamics for competitive interactions. *Journal of Mathematical Biology*, 63(3) :493–517, 2011.
- [131] D. H. JANZEN. When is it coevolution? *Evolution*, 34(3) :611–612, 1980.
- [132] A. JOFFE and M. MÉTIVIER. *Weak convergence of sequences of semimartingales with applications to multitype branching processes*. *Advances in Applied Probability*, 18 : 20–65, 1986.
- [133] H.-W. KANG, T. G. KURTZ, and L. POPOVIC. Central limit theorems and diffusion approximations for multiscale Markov chain models. *The Annals of Applied Probability*, 24(2), 2014.
- [134] I. KARATZAS and S.E. SHREVE. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics, Second edition*. Springer New York, New York, NY, 1998.
- [135] H. KESTEN. The number of alleles in electrophoretic experiments. *Theoretical Population Biology*, 18(2) :290–294, 1980.
- [136] R. KHASHMINSKII. On averaging principle for itô stochastic differential equations. *Kybernitika, Chekhoslovakia*, 4(3) :260–279, 1968.
- [137] R. KHASHMINSKII and N. KRYLOV. On averaging principle for diffusion processes with null-recurrent fast component. *Stochastic Processes and their Applications*, 93(2) :229–240, 2001.
- [138] M. KIMURA. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, 1 edition, 1983.
- [139] J.F.C. KINGMAN. Coherent random walks arising in some genetical models. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 351(1664) :19–31, 1976.
- [140] É. KISDI. Evolutionary Branching under Asymmetric Competition. *Journal of Theoretical Biology*, 197(2) :149–162, 1999.
- [141] É. KISDI. Trade-off geometries and the adaptive dynamics of two co-evolving species. *Evolutionary Ecology Research*, 8(6) :959–973, 2006.
- [142] A. KLENKE. *Probability Theory : A Comprehensive Course*. Universitext. Springer, New York, 2013.
- [143] M. KOT. *Elements of Mathematical Ecology*. Cambridge University Press, 1 edition, 2001.
- [144] T.G. KURTZ. Averaging for martingale problems and stochastic approximation. In I. KARATZAS and D. OCONE, editors, *Applied Stochastic Analysis, Lecture Notes in Control and Information Sciences*, volume 177, pages 186–209. Springer-Verlag, Berlin/Heidelberg, 1992.
- [145] H. J. KUSHNER. Jump-Diffusion Approximations for Ordinary Differential Equations with Wide-Band Random Right Hand Sides. *SIAM Journal on Control and Optimization*, 17(6) :729–744, 1979.
- [146] H. J. KUSHNER. A martingale method for the convergence of a sequence of processes to a jump-diffusion process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 53(2) :207–219, 1980.

- [147] H. J. KUSHNER. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*. Number 6 in The MIT Press series in signal processing, optimization, and control. MIT Press, Cambridge, Mass, 1984.
- [148] R. LANDE. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. *Evolution*, 33(1) :402–416, 1979.
- [149] R. LAW and U. DIECKMANN. Moment approximations of individual-based models. pages 252–270. Cambridge University Press, Cambridge, 2002.
- [150] J.-F. LE GALL. *Mouvement brownien, martingales et calcul stochastique*. Vol 71, Mathématiques et Applications, Springer-Verlag, 2013.
- [151] J.-F. LE GALLIARD, R. FERRIÈRE, and U. DIECKMANN. The adaptive dynamics of altruism in spatially heterogeneous populations. *Evolution*, 57(1) :1–17, 2003.
- [152] O. LEIMAR. Multidimensional convergence stability and the canonical adaptive dynamics. Cambridge University Press, 2001.
- [153] O. LEIMAR. Multidimensional convergence stability. *Evolutionary Ecology Research*, 11 :191–208, 2009.
- [154] H. LEMAN. Convergence of an infinite dimensional stochastic process to a spatially structured trait substitution sequence. *Stochastics and Partial Differential Equations : Analysis and Computations*, 4(4) :791–826, 2016.
- [155] H. LEMAN. *Probabilistic and deterministic analysis of the evolution : influence of a spatial structure and a mating preference*. Theses, Université Paris Saclay (COmUE), June 2016.
- [156] S. A. LEVIN. Dispersion and Population Interactions. *The American Naturalist*, 108(960) :207–228, 1974.
- [157] Z. LI, T. SHIGA, and L. YAO. A Reversibility Problem for Fleming-Viot Processes. *Electronic Communications in Probability*, 4, 1999.
- [158] T. LINDVALL. *Lectures on the Coupling Method*. Dover Publications, Inc, Mineola, N.Y, 2002.
- [159] H. LIU and X. ZHOU. Some support properties for a class of  $\Lambda$ -Fleming-Viot processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 51(3) :1076–1101, 2015.
- [160] A. LORZ, S. MIRRAHIMI, and B. PERTHAME. Dirac Mass Dynamics in Multidimensional Non-local Parabolic Equations. *Communications in Partial Differential Equations, Taylor & Francis*, 36(6) :1071–1098, 2011.
- [161] P. MARROW, R. LAW, and C. CANNINGS. The coevolution of predator-prey interactions : ESSs and red queen dynamics. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 250(1328) :133–141, 1992.
- [162] C. MATESSI and C. DI PASQUALE. Long-term evolution of multilocus traits. *Journal of Mathematical Biology*, 34(5-6) :613–653, 1996.
- [163] J. MAYNARD SMITH. *Evolution and the theory of games*. Cambridge University Press, Cambridge ; New York, 1982.
- [164] J. MAYNARD SMITH and J. HAIGH. The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1) :23–35, 1974.
- [165] J. MAYNARD SMITH and G. R. PRICE. The Logic of Animal Conflict. *Nature*, 246(5427) :15–18, 1973.
- [166] S. MÉLÉARD and V. C. TRAN. Trait Substitution Sequence process and Canonical Equation for age-structured populations. *Journal of Mathematical Biology*, 58(6) :881–921, 2009.
- [167] S. MÉLÉARD and V. C. TRAN. Slow and fast scales for superprocess limits of age-structured populations. *Stochastic Processes and their Applications*, 122(1) :250–276, 2012.
- [168] G. MENDEL. Experiments in plant hybridization. *Journal of the Royal Horticultural Society*, 26 :235–244, 1901.
- [169] G. MESZÉNA, M. GYLLENBERG, F. J. JACOBS, and J. A. J. METZ. Link between Population Dynamics and Dynamics of Darwinian Evolution. *Physical Review Letters*, 95(7) :078105, 2005.

- 
- [170] J. A. J. METZ and C. G. F. DE KOVEL. The canonical equation of adaptive dynamics for Mendelian diploids and haplo-diploids. *Interface Focus*, 3(6) :20130025, 2013.
- [171] J.A.J. METZ, S.A.H. GERITZ, G. MESZENA, F.J.A. JACOBS, and J.S. VAN HEERWAARDEN. Adaptive dynamics : A geometrical study of the consequences of nearly faithful reproduction. *Stochastic and spatial structures of dynamical systems*, KNAW Verhandelingen, Afd. Natuurkunde, Eerste reeks, 45 :183–231, North–Holland, Amsterdam, 1996.
- [172] J.A.J. METZ, R.M. NISBET, and S.A.H. GERITZ. How should we define fitness for general ecological scenarios? *Trends in Ecology & Evolution*, 7(6) :198–202, 1992.
- [173] S. MIRRAHIMI, G. BARLES, B. PERTHAME, and P. E. SOUGANIDIS. A Singular Hamilton–Jacobi Equation Modeling the Tail Problem. *SIAM Journal on Mathematical Analysis*, 44(6) :4297–4319, 2012.
- [174] S. MIRRAHIMI, B. PERTHAME, and J. Y. WAKANO. Evolution of species trait through resource competition. *Journal of Mathematical Biology*, 64(7) :1189–1223, 2012.
- [175] D. MOLLISON. Spatial Contact Models for Ecological and Epidemic Spread. *Journal of the Royal Statistical Society : Series B (Methodological)*, 39(3) :283–313, 1977.
- [176] J. MONOD. La technique de culture continue. *Annales de l’Institut Pasteur. Paris*, 79 :390–410, 1950.
- [177] P.A.P. MORAN. Wandering distributions and the electrophoretic profile. *Theoretical Population Biology*, 8(3) :318–330, 1975.
- [178] P.A.P. MORAN. Wandering distributions and the electrophoretic profile. II. *Theoretical Population Biology*, 10(2) :145–149, 1976.
- [179] J. D. MURRAY. *Mathematical Biology II : Spatial Models and Biomedical Applications*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, NY, 2003.
- [180] J. D. MURRAY. *Mathematical biology I : An introduction*. Number 17 in Interdisciplinary applied mathematics. Springer-Verlag GmbH, Berlin Heidelberg, softcover reprint of the hardcover 3rd edition 2002, corrected second printing edition, 2004.
- [181] S. D. MYLIUS and O. DIEKMANN. The Resident Strikes Back : Invader-Induced Switching of Resident Attractor. *Journal of Theoretical Biology*, 211(4) :297–311, 2001.
- [182] E.G. NEPOMUCENO, R.H.C. TAKAHASHI, and L.A. AGUIRRE. Individual-based model (ibm) : an alternative framework for epidemiological compartment models. *Brazilian Journal of Biometrics*, 34(1) :133–162, 2016.
- [183] A. NOVICK and L. SZILARD. Description of the Chemostat. *Science*, 112(2920) :715–716, 1950.
- [184] A. NOVICK and L. SZILARD. Experiments with the Chemostat on Spontaneous Mutations of Bacteria. *Proceedings of the National Academy of Sciences*, 36(12) :708–719, 1950.
- [185] M. NOWAK and K. SIGMUND. The evolution of stochastic strategies in the Prisoner’s Dilemma. *Acta Applicandae Mathematicae*, 20(3) :247–265, 1990.
- [186] B. OBORNY. Growth Rules in Clonal Plants and Environmental Predictability – A Simulation Study. *The Journal of Ecology*, 82(2) :341–351, 1994.
- [187] K. OELSCHLAGER. Limit theorems for age-structured populations. *The Annals of Probability*, 18(1) :290–318, 1990.
- [188] T. OHTA and M. KIMURA. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research*, 22(2) :201–204, 1973.
- [189] T. OHTA and M. KIMURA. Simulation studies on electrophoretically detectable genetic variability in a finite population. *Genetics*, 76(3) :615–624, 1974.
- [190] L. OVERBECK. Uniqueness of the Fleming-Viot process with selection. In Z.M. MA, M. RÖCKNER, and J.A. YAN, editors, *Dirichlet Forms and Stochastic Processes*. De Gruyter, Berlin, New York, 1995.
- [191] L. OVERBECK, M. ROCKNER, and B. SCHMULAND. An Analytic Approach to Fleming-Viot Processes with Interactive Selection. *The Annals of Probability*, 23(1), 1995.

- [192] G. C. PAPANICOLAOU, D. STROOCK, and S. R. S. VARADHAN. Martingale approach to some limit theorems. In P. L. CHOW, editor, *Duke turbulence conference*, number III in Duke University Mathematics Series, pages ii+120 pp. Duke University, Durham, NC, 1977.
- [193] G. A. PAVLIOTIS and A. M. STUART. *Multiscale methods : averaging and homogenization*. Number 53 in Texts in applied mathematics. Springer, New York, NY, 2008.
- [194] B. PERTHAME and G. BARLES. Dirac Concentrations in Lotka-Volterra Parabolic PDEs. *Indiana University Mathematics Journal*, 57(7) :3275–3301, 2008.
- [195] B. PERTHAME and M. GAUDUCHON. Survival thresholds and mortality rates in adaptive dynamics : conciliating deterministic and stochastic simulations. *Mathematical Medicine and Biology : A Journal of the IMA*, 27(3) :195–210, 2009.
- [196] B. PERTHAME and S. GÉNIEYS. Concentration in the Nonlocal Fisher Equation : the Hamilton-Jacobi Limit. *Mathematical Modelling of Natural Phenomena*, 2(4) :135–151, 2007.
- [197] P. PFAFFELHUBER and L. POPOVIC. How spatial heterogeneity shapes multiscale biochemical reaction network dynamics. *Journal of The Royal Society Interface*, 12(104) :20141106, 2015.
- [198] P. PFAFFELHUBER and L. POPOVIC. Scaling limits of spatial compartment models for chemical reaction networks. *The Annals of Applied Probability*, 25(6), 2015.
- [199] L. POPOVIC, S. A. MCKINLEY, and M. C. REED. A Stochastic Compartmental Model for Fast Axonal Transport. *SIAM Journal on Applied Mathematics*, 71(4) :1531–1556, 2011.
- [200] P. PROTTER. *Stochastic Integration and Differential Equation*. Springer, 2004.
- [201] A. PUGLIESE. On the evolutionary coexistence of parasite strains. *Mathematical Biosciences*, 177-178 :355–375, 2002.
- [202] G. RAOUL. Long Time Evolution of Populations under Selection and Vanishing Mutations. *Acta Applicandae Mathematicae*, 114(1-2) :1–14, 2011.
- [203] J. ROUGHGARDEN. *The theory of coevolution*. Sinauer Associates, Sunderland Massachusetts, 1983.
- [204] W. RUDIN. *Real and complex analysis*. Third edition, McGraw-Hill, New York, 2013.
- [205] T. SHIGA. Wandering phenomena in infinite-allelic diffusion models. *Advances in Applied Probability*, 14(3) :457–483, 1982.
- [206] C. SMADI. *Modèles probabilistes de populations : branchement avec catastrophes et signature génétique de la sélection*. Theses, Université Paris-Est, 2015.
- [207] D. STROOCK and S. VARADHAN. *Multidimensional diffusion processes*. Number 233 in Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, Berlin Heidelberg, reprint of the 1997 edition edition, 2006.
- [208] H. R. THIEME. *Mathematics in population biology*. Princeton series in theoretical and computational biology. Princeton University Press, Princeton, N.J, 2003.
- [209] V. C. TRAN. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM : Probability and Statistics*, 12 :345–386, 2008.
- [210] A. Y. VERETENNIKOV. On the averaging principle for systems of stochastic differential equations. *Mathematics of the USSR-Sbornik*, 69(1) :271–284, 1991.
- [211] T.L. VINCENT, Y. COHEN, and J.S. BROWN. Evolution via Strategy Dynamics. *Theoretical Population Biology*, 44(2) :149–176, 1993.
- [212] D. WAXMAN and S. GAVRILETS. 20 Questions on Adaptive Dynamics : Target review of Adaptive Dynamics. *Journal of Evolutionary Biology*, 18(5) :1139–1154, 2005.
- [213] D. WILLIAMS. *Probability with Martingales*. Cambridge University Press, 1 edition, 1991.
- [214] S. WRIGHT. Evolution in mendelian populations. *Genetics*, 16(2) :97–159, 1931.
- [215] W. R. YOUNG, A. J. ROBERTS, and G. STUHNE. Reproductive pair correlations and the clustering of organisms. *Nature*, 412(6844) :328–331, 2001.



## Résumé

La théorie des dynamiques adaptatives est une branche de la biologie de l'évolution qui étudie les liens entre Écologie et Évolution. Les hypothèses biologiques qui définissent son cadre sont celles de mutations rares et petites et de grande population asexuée. Les modèles de dynamiques adaptatives décrivent la population au niveau des individus, lesquels sont caractérisés par leurs phénotypes, et visent à étudier l'influence des mécanismes d'hérédité, de mutation et de sélection sur l'évolution à long terme de la population. Le succès de cette théorie vient notamment de sa capacité à fournir une description de l'évolution à long terme du phénotype dominant dans la population comme solution de "l'Équation Canonique des Dynamiques Adaptatives" dirigée par un gradient de fitness, où la fitness décrit la possibilité d'invasions mutantes, et est construite à partir de paramètres écologiques. Deux approches mathématiques principales portant sur l'équation canonique ont été développées à ce jour : une approche basée sur des EDP et une approche stochastique. Malgré son succès, l'approche stochastique est critiquée par des biologistes puisqu'elle est basée sur une hypothèse non-réaliste de mutations trop rares. Le but de cette thèse est de corriger cette controverse biologique en proposant des modèles probabilistes plus réalistes. Plus précisément, le but est de s'intéresser mathématiquement, sous une double asymptotique de grande population et de petites mutations, aux conséquences d'une nouvelle hypothèse biologique de mutations fréquentes sur l'équation canonique. Il s'agit de déterminer, à partir d'un modèle stochastique individu-centré, le comportement en temps long du trait phénotypique moyen de la population. La question que l'on se pose se reformule en une analyse asymptotique lent-rapide agissant sur deux échelles de temps éco-évolutives. Une échelle lente correspondant à la dynamique du trait moyen et une rapide correspondant à la dynamique d'évolution de la distribution recentrée et dilatée des traits. Cette analyse asymptotique lent-rapide repose sur des techniques de moyennisation. Cette méthode requiert d'identifier et de caractériser le comportement asymptotique de la composante rapide et que cette dernière possède des propriétés d'ergodicité. Plus précisément, le comportement en temps long de la composante rapide est non-classique et correspond à celui d'une diffusion à valeurs mesures originale qui s'interprète comme un processus de FLEMING-VIOT recentré que l'on caractérise comme l'unique solution d'un certain problème de martingale. Une partie de ces résultats repose sur une relation de dualité portant sur ce processus non-classique et nécessite des conditions de moments sur les données initiales. Au moyen de techniques de couplage et de la correspondance entre les processus particuliers de MORAN et les généalogies de KINGMAN, on établit que le processus de FLEMING-VIOT recentré satisfait une propriété d'ergodicité avec résultat de convergence exponentielle en variation totale. La mise en œuvre des méthodes de moyennisation, inspirée par KURTZ, est fondée sur des arguments de compacité-unicité. L'idée consiste à prouver la compacité des lois du couple constitué de la composante lente et de la mesure d'occupation de la composante rapide puis d'établir un problème de martingale pour tous points d'accumulation de la famille des lois de ce couple. La dernière étape consiste à identifier ces points d'accumulation. Cette méthode requiert notamment l'introduction de temps d'arrêt pour contrôler les moments de la composante rapide et de prouver qu'ils tendent vers l'infini à l'aide d'arguments de grandes déviations, de réduire le problème posé initialement sur la droite réelle au cas du tore afin de prouver la compacité, d'identifier la limite de la composante rapide en adaptant un argument basé sur la dualité de DAWSON, d'identifier la limite de la composante lente puis de passer du tore à la droite réelle.

**Mots-clés:** Dynamiques adaptatives, Équation Canonique, Modèles stochastiques individu-centrés, Analyse asymptotique lent-rapide, Méthode de moyennisation, Processus de diffusions à valeurs mesure, Processus de FLEMING-VIOT, Problème de martingale, Méthode de dualité, Ergodicité, Processus de MORAN, Généalogies de KINGMAN, Look-down modifié de DONNELLY-KURTZ, Couplage.

## Abstract

Adaptive dynamics theory is a branch of evolutionary biology which studies the links between ecology and evolution. The biological assumptions that define its framework are those of rare and small mutations and large asexual populations. Adaptive dynamics models describe the population at the level of individuals, which are characterised by their phenotypes, and aim to study the influence of heredity, mutation and selection mechanisms on the long term evolution of the population. The success of this theory comes in particular from its ability to provide a description of the long term evolution of the dominant phenotype in the population as a solution to the “Canonical Equation of Adaptive Dynamics” driven by a fitness gradient, where fitness describes the possibility of mutant invasions, and is constructed from ecological parameters. Two main mathematical approaches to the canonical equation have been developed so far: an approach based on PDEs and a stochastic approach. Despite its success, the stochastic approach is criticised by biologists as it is based on a non-realistic assumption of too rare mutations. The goal of this thesis is to correct this biological controversy by proposing more realistic probabilistic models. More precisely, the aim is to investigate mathematically, under a double asymptotic of large population and small mutations, the consequences of a new biological assumption of frequent mutations on the canonical equation. The goal is to determine, from a stochastic individual-based model, the long term behaviour of the mean phenotypic trait of the population. The question we ask is reformulated into a slow-fast asymptotic analysis acting on two eco-evolutionary time scales. A slow scale corresponding to the dynamics of the mean trait, and a fast scale corresponding to the evolutionary dynamics of the centred and dilated distribution of traits. This slow-fast asymptotic analysis is based on averaging techniques. This method requires the identification and characterisation of the asymptotic behaviour of the fast component and that the latter has ergodicity properties. More precisely, the long time behaviour of the fast component is non-classical and corresponds to that of an original measure-valued diffusion which is interpreted as a centered FLEMING-VIOT process that is characterised as the unique solution of a certain martingale problem. Part of these results is based on a duality relation on this non-classical process and requires moment conditions on the initial data. Using coupling techniques and the correspondence between MORAN’s particle processes and KINGMAN’s genealogies, we establish that the centered FLEMING-VIOT process satisfies an ergodicity property with exponential convergence result in total variation. The implementation of averaging methods, inspired by KURTZ, is based on compactness-uniqueness arguments. The idea is to prove the compactness of the laws of the couple made up of the slow component and the occupation measure of the fast component and then to establish a martingale problem for all accumulation points of the family of laws of this couple. The last step is to identify these accumulation points. This method requires in particular the introduction of stopping times to control the moments of the fast component and to prove that they tend to infinity using large deviation arguments, to reduce the problem initially posed on the real line to the torus case in order to prove compactness, to identify the limit of the fast component by adapting an argument based on DAWSON duality, to identify the limit of the slow component and then to move from the torus to the real line.

**Keywords:** Adaptive dynamics, Canonical Equation, Stochastic individual-based models, Slow-fast asymptotic analysis, Averaging method, Measure-valued diffusion processes, FLEMING-VIOT’s processes, Martingale problems, Duality method, Ergodicity, MORAN’s processes, KINGMAN’s genealogies, DONNELLY-KURTZ’s modified look-down, Coupling.