



HAL
open science

Study of numerical methods for some stochastic differential equations in finance and modeling of capital distribution in financial market

Houzhi Li

► **To cite this version:**

Houzhi Li. Study of numerical methods for some stochastic differential equations in finance and modeling of capital distribution in financial market. General Mathematics [math.GM]. Université Paris Cité, 2021. English. NNT : 2021UNIP7031 . tel-03550527

HAL Id: tel-03550527

<https://theses.hal.science/tel-03550527>

Submitted on 1 Feb 2022

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 PARIS



École Doctorale de Sciences Mathématiques de Paris Centre (386)
Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR 8001)



THÈSE DE DOCTORAT

en Mathématiques Appliquées

Présentée par

Houzhi LI

**Étude de méthodes numériques pour certaines équations
différentielles stochastiques en finance et modélisation de la
distribution du capital dans le marché financier**

**Study of numerical methods for some stochastic differential
equations in finance and modeling of capital distribution in
financial market**

Sous la direction de **Noufel FRIKHA** et **Peter TANKOV**

Soutenue le *30 Mars 2021* devant le jury composé de :

M. Aurélien ALFONSI	ENPC	Rapporteur
M. Jean-François CHASSAGNEUX	Université de Paris	Examineur
Mme. Christa CUCHIERO	Université de Vienne (Autriche)	Rapporteuse
M. Noufel FRIKHA	Université de Paris	Directeur de thèse
Mme. Zorana GRBAC	Université de Paris	Examinatrice
M. Ahmed KEBAIER	Université Paris 13	Examineur
Mme. Agnès SULEM	INRIA	Examinatrice
M. Peter TANKOV	ENSAE	Directeur de thèse

Remerciements

Mes premiers remerciements s'adressent bien évidemment à mes directeurs de thèse, Noufel Frikha et Peter Tankov, sans qui ce travail n'aurait pas vu le jour. Je vous suis extrêmement reconnaissant pour votre confiance, votre patience à mon égard et bien sûr votre disponibilité. Vous m'avez appris la recherche en mathématiques avec sérieux et rigueur.

Je tiens à remercier sincèrement Aurélien Alfonsi et Christa Cuchiero qui me font l'honneur de rapporter ma thèse. Je remercie également Jean-François Chassagneux, Zorana Grbac, Ahmed Kebaier et Agnès Sulem pour avoir accepté de faire partie de mon jury de thèse.

Je remercie tout le Laboratoire de Probabilités, Statistique et Modélisation, en particulier les doctorants (j'ai peur que la liste soit incomplète donc je m'adresse à vous tous...) et l'équipe de Mathématiques Financières et Actuarielles, Probabilités Numériques pour m'avoir accueilli. Merci à Nathalie et Valérie qui nous accompagnent dans les méandres administratifs que nous rencontrons au quotidien.

Je souhaite remercier les organisateurs du CEMRACS 2017, école d'été durant laquelle la dernière partie de cette thèse a été rédigée, en collaboration avec Andrea Angiuli, Christy V. Graves, Jean-François Chassagneux, François Delarue et René Carmona, que je remercie tous. Un grand merci au CIRM pour l'accueil durant cet été 2017.

Je remercie également Junchao Chen pour la collaboration qui a permis au troisième chapitre de ma thèse de voir le jour.

Je remercie tous les chercheurs et doctorants avec qui j'ai pu avoir des discussions enrichissantes lors de diverses conférences, que ce soit à Shanghai, à Bologne, à Manizales, à Métabief ou à Paris.

Enfin, je remercie ma famille et mes amis qui m'ont accompagné et soutenu, en particulier mes parents et ma cousine.

Résumé

Nous apportons dans cette thèse quelques contributions à la modélisation du marché financier dans le cadre de la théorie stochastique du portefeuille et à l'étude des méthodes numériques pour quelques équations différentielles stochastiques en modélisation financière et en théorie des jeux. Cette thèse se divise en quatre parties indépendantes, bien que des techniques et des méthodes communes soient utilisées dans les parties 2 et 3.

La première partie porte sur la modélisation du portefeuille de marché à travers les capitalisations relatives, aussi appelées poids du marché des actifs. L'évolution au cours du temps des poids du marché est modélisée par un processus de diffusion à valeurs dans le simplexe unité car les poids sont positifs ou nuls avec somme égale 1. Nous étudions d'abord ce processus d'un point de vue probabiliste: nous donnons les conditions sous lesquelles le bord du simplexe unité est atteint, puis montrons l'existence et l'unicité de la distribution stationnaire des poids sous certaines hypothèses bien spécifiées. Nous faisons également le lien avec les modèles existants dans la littérature et notre modèle permet d'avoir plus de flexibilité dans la structure de covariance entre les actifs. Ensuite les propriétés de marché sont étudiées, dont l'importante notion d'arbitrage relatif au marché qui permet de battre le marché à coup sûr et des critères d'existence d'arbitrage relatif sont donnés. Finalement nous considérons le problème de calibration aux données du marché et quelques applications du modèle qui pourraient être des sujets intéressants de recherches futures.

La deuxième partie est dédiée à l'étude d'un schéma probabiliste pour les équations différentielles stochastiques de type McKean-Vlasov. La stratégie principale est l'itération de Picard permettant d'avoir une approximation de la loi marginale de la solution de l'EDS de McKean-Vlasov, qui peut être vue comme un point fixe pour une application définie sur l'espace des flots de mesures de probabilité. La technique de continuation en temps est introduite dans le but de réduire la complexité, qui consiste à couper l'intervalle de temps en des sous-intervalles uniformes et d'effectuer des itérations de Picard sur chacun de ces sous-intervalles. Pour approcher la loi marginale à chaque itéré de Picard, nous utilisons une représentation probabiliste qui sert aussi à la simulation sans biais des processus de diffusion. Plus précisément, nous considérons un schéma d'Euler sur une grille de temps aléatoires bien choisie, les temps de sauts étant de loi exponentielle ou Beta et nous avons un estimateur sans biais de la fonction de densité et ainsi de l'espérance. Les nombres d'itérés de Picard et de sous-intervalles sont choisis pour minimiser la complexité globale de l'algorithme. Nous étudions également le biais causé par l'approximation avec itération de Picard et l'erreur statistique, bien que l'estimation de l'erreur statistique dans le cas général fasse partie du travail en cours. Finalement nous montrons l'efficacité de la méthode à travers des exemples numériques.

Nous nous intéressons, dans la troisième partie, à la représentation probabiliste et

aux formules d'intégrations par parties pour quelques modèles à volatilité stochastique. Notre premier objectif est, comme dans la précédente partie, donner une représentation probabiliste du prix d'un produit financier, défini par une espérance. Le processus considéré ici est celui du couple formé par le prix spot et la volatilité du sous-jacent, dont les dynamiques sont données par deux semi-martingales corrélées. Une particularité de ce modèle est que le drift dans la dynamique de la volatilité est non-borné, typiquement il s'agit d'un terme de retour à la moyenne. Pour que la représentation probabiliste soit bien définie avec des moments L^p bornés, le processus d'approximation de la volatilité sera choisi comme la somme de la solution d'une équation différentielle ordinaire et d'un terme de diffusion. Ensuite nous cherchons à représenter sous forme d'espérances, les dérivées respectives du prix du produit financier par rapport au prix initial et à la volatilité initiale du sous-jacent, aussi appelées *Delta* et *Vega* pour les sensibilités ou "grecques" en finance. Nous utiliserons des techniques de calcul de Malliavin sur mesure appliqué à des chaînes de Markov pour obtenir des formules d'intégration par parties de type Bismut-Elworthy-Li qui seront ensuite simulables. Les exemples numériques seront donnés pour illustrer ces représentations.

Dans la quatrième partie, nous développons deux algorithmes pour la résolution numérique des équations différentielles stochastiques progressives-rétrogrades (EDSPRs) de type McKean-Vlasov issues des jeux à champ moyen (MFG) et montrons quelques exemples numériques. Après avoir rappelé les notions de base sur les jeux à champ moyen, qui consistent à trouver l'équilibre de Nash en présence d'une continuité de joueurs, nous introduisons le problème de contrôle stochastique de type champ moyen équivalent associé à l'équilibre de Nash, qui est décrit par une EDSPR couplée de type McKean-Vlasov dont la solution peut être vue comme un point fixe et cette solution est approchée par un schéma de Picard progressif-rétrograde. Nous utilisons encore une fois la technique de continuation en temps pour que le schéma de Picard converge sur des petits intervalles de temps, cependant la technique est présentée de façon différente que dans la seconde partie. Dans le premier algorithme, nous utilisons l'approximation binaire pour la diffusion Brownienne, qui donnera une structure d'arbre binaire. Afin de réduire la complexité exponentielle en nombre de pas de temps dans l'arbre, nous considérons dans le deuxième algorithme, des grilles d'approximations pour les processus progressif et rétrograde.

Mots-clés : Théorie stochastique du portefeuille; Arbitrage relatif; Poids du marché; Processus ergodiques; EDS de McKean-Vlasov; Équations de Fokker-Planck non-linéaires; Analyse numérique probabiliste; Itération de Picard; Modèles à volatilité stochastique; Formules d'intégration par parties; Jeux à champ moyen; EDSPRs

Abstract

In this thesis, we make some contributions to the modeling of the financial market in the context of stochastic portfolio theory, as well as to the study of numerical methods for some stochastic differential equations in financial modeling and game theory. This thesis is divided into four independent parts, although some common techniques and methods are used in Part 2 and Part 3.

The first part focuses on the modeling of market portfolio through relative capitalizations, also called market weights of assets. The evolution over time of market weights is modeled by a diffusion process with values in the unit simplex since the weights are nonnegative with sum equal to 1. We first study the process from a probabilistic point of view: we give conditions under which the boundary of the unit simplex is attained, then show the existence and uniqueness of the stationary distributions of weights under some well-specified assumptions. We also explain the link with existing models in the literature and show that our model provides more flexibility for the covariance structure between assets. Then market properties are studied, for instance the important notion of relative arbitrage to the market which allows to beat the market for sure and criteria for the existence of relative arbitrage are given. Finally we consider the problem of calibration to market data and some other applications of the model that could be interesting topics for future research.

The second part is dedicated to the study of a probabilistic scheme for stochastic differential equations of McKean-Vlasov type. We propose a Picard iteration algorithm which allows to get an approximation of the marginal law of the solution of McKean-Vlasov SDE, which could be seen as a fixed point of a mapping defined on the space of flows of probability measures. The technique of continuation in time is introduced to reduce the complexity, in fact we split the time interval into uniform subintervals and Picard iterations are performed on each of these subintervals. In order to approximate the marginal law at each Picard iteration, we use a probabilistic representation which is also used for the unbiased simulation of diffusion processes. More precisely, we consider an Euler scheme on a well-chosen random grid, with jump times following exponential or Beta distribution and we get an unbiased estimator of the density function as well as the expectation. The numbers of Picard iterations and of subintervals are chosen to minimize the global complexity of the algorithm. We also study the bias caused by the approximation using Picard iteration and the statistical error, although the estimation of the statistical error in the general case is still work in progress. Finally, the efficiency of the method is illustrated through some numerical examples.

In the third part of the thesis, we are interested in probabilistic representation and integration by parts formulae for some stochastic volatility models. Our first goal is to give a probabilistic representation of the price of a financial product which is defined as an expectation as in the previous part. The process considered here is the pair formed

by the spot price and the volatility of the underlying asset, whose dynamics are given by two correlated semimartingales. One particularity of this model is that the drift coefficient in the dynamics of the volatility is unbounded, typically it corresponds to a term of mean reversion. To have a well-defined probabilistic representation with finite L^p moments, the approximation process of the volatility should be chosen as the sum of the solution to an ordinary differential equation and a diffusion term. Then we seek to represent as expectations, the derivatives of the price of the financial product with respect to initial price and initial volatility, also called *Delta* and *Vega* for sensitivities or "greeks" in finance. We develop a tailor-made Malliavin calculus for the underlying Markov chain to obtain integration by parts formulae of Bismut-Elworthy-Li type that can be simulated without any bias. Numerical examples will be given to illustrate our different formulae.

In the fourth part, we develop two algorithms for the numerical resolution of forward-backward stochastic differential equations (FBSDEs) of McKean-Vlasov type arising from mean-field games (MFG) and we present some numerical examples. After revisiting some basic notions on mean-field games, which consist in finding the Nash equilibrium in the presence of a continuum of players, we introduce the equivalent stochastic control problem of mean-field type associated to Nash equilibrium. This problem is described by a fully coupled FBSDE of McKean-Vlasov type whose solution can be seen as a fixed point and can be approximated by a forward-backward Picard scheme. Again continuation in time is used to have the convergence of Picard scheme on small time intervals, although the technique is presented differently than in the second part. For the first algorithms, the binary approximation is used for the Brownian diffusion, which leads to a structure of a binary tree. In order to reduce the exponential complexity in the number of time steps, we consider approximation grids for the forward and backward processes in the second algorithm.

Keywords : Stochastic Portfolio Theory; Relative Arbitrage, Market Weights; Ergodic Processes; McKean-Vlasov SDEs; Non-linear Fokker-Planck equations; Probabilistic numerical analysis; Picard iteration; Stochastic volatility model; Integration by parts formulae; Mean-field games; FBSDEs

Contents

1	Introduction - Français	1
1.1	Modélisation des capitalisations relatives des actifs	1
1.1.1	Processus et distribution des poids du marché	1
1.1.2	Structure du marché et arbitrage relatif	5
1.1.3	Applications et perspectives	8
1.2	Un schéma probabiliste pour les équations de McKean-Vlasov	10
1.2.1	Itération de Picard et méthode de continuation	10
1.2.2	Représentation probabiliste de la loi marginale	12
1.2.3	Analyse de la convergence et de l'erreur et perspectives	15
1.3	Représentation probabiliste et formules d'intégration par parties pour certains modèles à volatilité stochastique avec drift non-borné	16
1.3.1	Représentation probabiliste du couple spot-volatilité	17
1.3.2	Formules d'intégration par parties	19
1.4	Méthodes numériques pour les EDSPRs issues des jeux à champ moyen .	20
1.4.1	Jeux à champ moyen et approche probabiliste	20
1.4.2	Algorithmes d'arbre et de grille pour les EDSPRs	22
2	Introduction - English	27
2.1	Modeling of relative capitalizations of assets	27
2.1.1	Process and distribution of market weights	27
2.1.2	Structure of the market and relative arbitrage	31
2.1.3	Applications and perspectives	34
2.2	A probabilistic scheme for McKean-Vlasov equations	35
2.2.1	Picard iteration and method of continuation	36
2.2.2	Probabilistic representation of marginal law	37
2.2.3	Convergence and error analysis and perspectives	40
2.3	Probabilistic representation and integration by parts formulae for some stochastic volatility model with unbounded drift	42
2.3.1	Probabilistic representation of spot-volatility couple	42
2.3.2	Integration by parts formulae for sensitivities	44
2.4	Numerical methods for FBSDEs arising from mean-field games	45
2.4.1	Mean-field games and probabilistic approach	45
2.4.2	Tree and grid algorithms for FBSDEs	47

I	Modeling the market by capital distribution	53
3	Modeling the market by capital distribution	55
3.1	Introduction	55
3.2	Equation and process of market weights	58
3.2.1	Market weights equation and its first properties	58
3.2.2	Distribution and transition density of market weights	69
3.2.3	Stationary distribution of market weights	70
3.3	Structure of market and relative arbitrage	76
3.3.1	Market portfolio and growth rate	76
3.3.2	Trading strategy and relative arbitrage	78
3.3.3	Conditions of arbitrage relative to the market	79
3.4	Future works and conclusion	90
3.4.1	Functionally generated portfolio	90
3.4.2	Estimation of model parameters with application to stock market indices	90
3.4.3	Long-term portfolio optimization under market weights equation	92
3.4.4	Conclusion and perspectives	93
3.5	Appendix	93
3.5.1	Attainment of boundary for Theorem 3.2.1	93
3.5.2	Hörmander’s condition and theorem for Proposition 3.2.2	95
3.5.3	Choice of Lyapunov function V for Theorem 3.2.2	95
3.5.4	Generalized local martingale problem and candidate measure for Lemma 3.3.1	97
II	A Probabilistic Scheme for McKean-Vlasov Equations	99
4	A Probabilistic Scheme for McKean-Vlasov Equations	101
4.1	Introduction	101
4.2	Description of the numerical probabilistic scheme	103
4.2.1	Assumptions and well-posedness of the McKean-Vlasov SDE	103
4.2.2	Construction of the Picard iteration scheme	106
4.2.3	Probabilistic representation of the marginal law of the Picard iteration scheme	108
4.2.4	Construction of the Monte Carlo estimator	111
4.3	Convergence Analysis	113
4.3.1	Decomposition of the global error and complexity of the algorithm	113
4.3.2	Convergence analysis of the local Picard iteration schemes	114
4.3.3	Convergence analysis of the global Picard iteration scheme on $[0, T]$	115
4.4	Numerical results	119
4.4.1	Standard linear model	120
4.4.2	Kuramoto model	123
4.4.3	Polynomial drift model	124
4.5	Appendix	126
4.5.1	Proof of Theorem 4.2.1	126
4.5.2	Proof of Proposition 4.3.1	133

III	Probabilistic Representation and Integration by Parts Formulae for some Stochastic Volatility Models	145
5	Probabilistic Representation and Integration by Parts Formulae for some Stochastic Volatility Models	147
5.1	Introduction	147
5.2	Preliminaries: assumptions, definition of the underlying Markov chain and related Malliavin calculus	149
5.2.1	Assumptions	149
5.2.2	Choice of the approximation process	150
5.2.3	Markov chain on random time grid	152
5.2.4	Taylor-made Malliavin calculus for the Markov chain.	153
5.3	Probabilistic representation for the couple (S_T, Y_T) .	159
5.4	Integration by parts formulae	159
5.4.1	The transfer of derivative formula	160
5.4.2	The integration by parts formulae	162
5.5	Numerical Results	171
5.5.1	Black-Scholes Model	172
5.5.2	A Stein-Stein type model	176
5.5.3	A model with a periodic diffusion coefficient function	176
5.6	Proof of Theorem 5.3.1 and Lemma 5.4.1	179
5.6.1	Proof of Theorem 5.3.1	179
5.6.2	Proof of Lemma 5.4.1	190
5.7	Some technical results	194
5.7.1	Emergence of jumps in the renewal process N	194
5.7.2	Formulae for the computations of price, <i>Delta</i> and <i>Vega</i>	196
IV	Probabilistic Numerical Methods for Mean-Field Games	199
6	Probabilistic Numerical Methods for Mean-Field Games	201
6.1	Introduction	201
6.2	Overview of Mean Field Games and FBSDEs	202
6.2.1	N Player Stochastic Differential Games	202
6.2.2	Mean Field Games	203
6.2.3	General System	207
6.3	Algorithms	208
6.3.1	Global Picard Iteration on a Small Time Interval	209
6.3.2	Continuation in Time of the Global Method for Arbitrary Interval/Coupling	212
6.3.3	Tree Algorithm for the Global Method	213
6.3.4	Picard Iteration on the Marginal Laws: a Grid Algorithm	215
6.4	Examples	219
6.4.1	Linear Example	219
6.4.2	Trigonometric Drivers Example	220
6.4.3	Mixed Model	221
6.4.4	Examples: Linear Quadratic Mean Field Games	225

6.5 Conclusion 231

Bibliography **239**

Chapitre 1

Introduction - Français

Dans ce manuscrit nous étudions quatre problèmes issus des mathématiques financières et des probabilités numériques. Nous introduisons les questions abordées, ainsi que les problématiques et les motivations. Nous résumons également les résultats principaux obtenus et les problèmes ouverts qui feraient l'objet de recherches futures.

1.1 Modélisation des capitalisations relatives des actifs

Dans la théorie "classique" du portefeuille, nous modélisons le portefeuille du marché constitué de n actifs risqués par le processus $(\mathbf{S}_t)_{t \geq 0}$, $\mathbf{S}_t = (S_t^i)_{i=1}^n$ avec S_t^i représentant la capitalisation du marché de l'actif i à l'instant t , dont la dynamique est donnée par une semi-martingale dans l'espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ et le taux d'intérêt est supposé nul. La capitalisation totale du marché $(\Sigma_t)_{t \geq 0}$ est la somme des capitalisations des actifs S_t^i , $i = 1, \dots, n$. Pour étudier les propriétés du marché et les stratégies d'investissement relatives au marché dans lequel nous investissons, au lieu de considérer les capitalisations réelles, nous avons aussi la possibilité de décrire le marché par les capitalisations relatives des actifs $(\boldsymbol{\mu}_t)_{t \geq 0}$, $\boldsymbol{\mu}_t = (\mu_t^i)_{i=1}^n$, $\mu_t^i = S_t^i / \Sigma_t$ qui sont définies par leurs parts de marché, ou poids relatifs du marché. Le processus de la capitalisation totale est un processus stochastique qui pourrait être corrélé au processus des poids relatifs. Le couple $(\boldsymbol{\mu}_t, \Sigma_t)_{t \geq 0}$ caractérise donc le portefeuille du marché et souvent nous n'avons pas besoin de connaître la dynamique de la capitalisation totale car nous investissons toujours dans ce même marché et c'est le numéraire sous lequel nous nous plaçons.

1.1.1 Processus et distribution des poids du marché

Nous modélisons le processus des poids relatifs du marché en introduisant l'équation des poids du marché d'exposant $\zeta > 0$ ($MWE(\zeta)$), qui est inspiré de la distribution de probabilité discrète dans problème de filtrage optimal non-linéaire, et aussi de la

dynamique des poids relatifs dans le modèle de marché à volatilité stabilisée :

$$d\mu_t^i = b^i(\boldsymbol{\mu}_t)dt + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t)dW_t^\nu, \quad i \in \{1, \dots, n\}; \quad \boldsymbol{\mu}_0 \in \Delta_+^n$$

$$b^i(\boldsymbol{\mu}) = (B \cdot \boldsymbol{\mu})^i = \sum_{j=1}^n B^{i,j} \mu^j, \quad \sigma^{i,\nu}(\boldsymbol{\mu}) = (\mu^i)^{\zeta} \zeta^{i,\nu} - \mu^i \sum_{k=1}^n (\mu^k)^{\zeta} \zeta^{k,\nu}$$
(1.1.1)

Ici la matrice B de taille $n \times n$ est le transposé d'une matrice de taux de transition d'une chaîne de Markov, c'est-à-dire qu'elle satisfait :

$$B^{i,j} \geq 0 \quad \forall i \neq j \quad \text{et} \quad \sum_{i=1}^n B^{i,j} = 0 \quad \forall j; \quad i, j \in \{1, \dots, n\},$$

et la matrice ζ est de taille $n \times d$, de rang n avec $n \leq d$, en particulier :

$$\forall i \in \{1, \dots, n\}, \quad c^i := \sum_{\nu=1}^d (\zeta^{i,\nu})^2 > 0$$

La matrice de covariance des poids $a(\cdot) = \sigma \sigma^\top(\cdot) = [a^{i,j}(\cdot)]$ est définie par :

$$a^{i,j}(\boldsymbol{\mu}_t) = \frac{d}{dt} \langle \mu^i, \mu^j \rangle_t = \sum_{\nu=1}^d \sigma^{i,\nu} \sigma^{j,\nu}(\boldsymbol{\mu}_t)$$

Notons par Δ^n (resp. Δ_+^n) le simplexe unité de dimension n (resp. son intérieur) et par $\partial\Delta^n := \Delta^n \setminus \Delta_+^n$ le bord du simplexe unité :

$$\Delta^n \text{ (} \Delta_+^n \text{ resp.)} = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^n \mid \mu^i \geq 0 \text{ (} \mu^i > 0 \text{ resp.) } \forall i \in \{1, \dots, n\} \text{ et } \sum_{i=1}^n \mu^i = 1 \right\}$$

Résultat Principal 1.1.1. *Pour tout $\zeta > 0$ et tous poids initiaux strictement positifs i.e. $\forall \boldsymbol{\mu}_0 \in \Delta_+^n$, les poids du marché dont la dynamique est donnée par (1.1.1), restent à tout instant dans le simplexe unité, i.e. $\boldsymbol{\mu}_t \in \Delta^n$ pour tout $t \geq 0$ presque sûrement.*

Il existe une unique solution forte à l'équation des poids du marché et aucun poids du marché ne s'annule sur tout horizon de temps fini $[0, T]$, $T > 0$, i.e. $\forall \boldsymbol{\mu}_0 \in \Delta_+^n$, l'unique solution forte $(\boldsymbol{\mu}_t)_{t \geq 0}$ vérifie $\forall t \geq 0$, $\boldsymbol{\mu}_t \in \Delta_+^n$ p.s., si et seulement si l'une des trois conditions suivantes est satisfaite :

- (i) $\zeta \geq 1$;
- (ii) $\zeta \in (1/2, 1)$ et $\forall i \in \{1, \dots, n\}$, $b^i(\boldsymbol{\mu}) > 0$ sur $\{\mu^i = 0\}$;
- (iii) $\zeta = 1/2$ et $\forall i \in \{1, \dots, n\}$, $2b^i(\boldsymbol{\mu}) \geq c^i > 0$ sur $\{\mu^i = 0\}$.

Dans le cas contraire, si l'une des conditions suivantes est satisfaite, il existe un point $\boldsymbol{\mu}_ \in \Delta_+^n$ et une constante $\epsilon > 0$, si les poids initiaux vérifient $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \epsilon$, alors le bord $\partial\Delta^n$ est atteint avec probabilité sur tout horizon de temps fini $[0, T]$, $T > 0$, définissons le temps d'arrêt $\tau := \inf\{t \geq 0 : \boldsymbol{\mu}_t \in \partial\Delta^n\}$ nous avons donc $\mathbb{P}[\tau \leq T] > 0$. Plus précisément il existe un indice $i \in \{1, \dots, n\}$ tel que $\mathbb{P}[\exists t \in (0, T) \mid \mu_t^i = 0] > 0$ et la solution forte existe et est unique jusqu'au temps d'arrêt $\tau \wedge T$:*

(iv) $\zeta \in (0, 1/2)$;

(v) $\zeta = 1/2$ et $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $2b^i(\boldsymbol{\mu}_*) < c^i$;

(vi) $\zeta \in (1/2, 1)$ et $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $b^i(\boldsymbol{\mu}_*) = 0$.

L'exposant $\zeta > 0$ est un paramètre du modèle à choisir ou à estimer. Lorsque $\zeta = 1$ nous retrouvons la dynamique de la probabilité discrète dans le problème de filtrage optimal non-linéaire.

Il est possible de retrouver la dynamique des poids du marché dans le modèle du marché à volatilité stabilisée. Ce modèle et ses généralisations sont largement étudiés depuis une vingtaine d'années dans le cadre de la théorie stochastique des portefeuilles, voir [FK05] et [Pic14], tout comme son processus des poids, appelé aussi processus de Jacobi ou Wright-Fisher, voir [GJ06] et [Pal+11]. En fait il suffit de prendre dans notre modèle $\zeta = 1/2$, la matrice B vérifiant $B^{i,j} = \alpha \xi^i > 0$ pour $i \neq j$ et $B^{i,i} = -\alpha \sum_{k=1, k \neq i}^n \xi^k < 0$, $\alpha > 0$ avec le niveau de retour à la moyenne $\boldsymbol{\xi} = (\xi^i)_{i=1}^n \in \Delta_+^n$ et la matrice ς diagonale, $d = n$ et $\varsigma = \sqrt{c}I^n$, $c > 0$.

Notre modèle a le même drift que les modèles polynomiaux étudiés dans [Cuc19], mais la structure de volatilité est différente, le choix de l'exposant ζ pourrait nous fournir plus de flexibilité pour caractériser la structure de covariance, contrairement aux modèles polynomiaux où l'exposant est $1/2$ comme dans le modèle du marché à volatilité stabilisée.

Maintenant que nous sommes sûrs que les poids vont rester dans le simplexe unité, nous pouvons considérer la *forme réduite* de l'équation des poids du marché en prenant simplement les poids des $(n-1)$ premiers actifs, le simplexe unité et son intérieur seront donc vus comme des ensembles équivalents dans \mathbb{R}^{n-1} et la matrice de covariance $a(\cdot)$ sera de dimension $(n-1)$:

$$\Delta^n (\Delta_+^n \text{ resp.}) = \{ \boldsymbol{\mu} = (\mu^i)_{i=1}^{n-1} \mid \mu^i \geq 0 (\mu^i > 0 \text{ resp.}) \forall i \text{ et } \sum_{i=1}^{n-1} \mu^i \geq 1 (< 1 \text{ resp.}) \}$$

Résultat Principal 1.1.2. *La matrice de covariance réduite $a(\boldsymbol{\mu})$ est définie positive si et seulement si $\boldsymbol{\mu} \in \Delta_+^n$, ce qui s'écrit aussi :*

$$\begin{aligned} \forall \boldsymbol{\mu} \in \Delta_+^n; \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} > 0 \\ \forall \boldsymbol{\mu} \in \partial \Delta^n = \Delta^n \setminus \Delta_+^n; \exists \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0 \end{aligned}$$

Définissons pour $\epsilon \in (0, 1/n)$, le sous-ensemble Δ_ϵ^n du simplexe unité constitué des poids $\geq \epsilon$ et nous avons dans ce sous-ensemble :

$$\begin{aligned} \Delta_\epsilon^n = \{ \boldsymbol{\mu} = (\mu^i)_{i=1}^{n-1} \mid \mu^i \geq \epsilon \forall i \in \{1, \dots, n-1\} \text{ et } \sum_{i=1}^n \mu^i \leq 1 - \epsilon \} \\ \exists \lambda_\epsilon > 0, \forall \boldsymbol{\mu} \in \Delta_\epsilon^n; \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} \geq \lambda_\epsilon |\mathbf{x}|^2 \end{aligned}$$

De manière équivalente, la matrice de volatilité réduite $\sigma(\boldsymbol{\mu})$ de taille $(n-1) \times d$ est de plein rang si et seulement si $\boldsymbol{\mu} \in \Delta_+^n$. Pour tout $\boldsymbol{\mu} \in \partial \Delta^n$, la matrice de covariance

et de volatilité ont le même rang, qui est égal à $n - 1 - m(\boldsymbol{\mu})$, où $m(\boldsymbol{\mu})$ est le nombre d'indices $i \in \{1, \dots, n\}$ tels que $\mu^i = 0$.

Par conséquent, le processus des poids du marché défini comme solution de l'équation des poids du marché ne vérifie pas la condition d'uniforme ellipticité, mais uniquement l'ellipticité locale sur les ensembles Δ_ϵ^n , $\epsilon \in (0, 1/n)$.

La distribution des poids du marché, vue comme une distribution n -dimensionnelle ne pourrait pas être absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R}^n car les poids sont à valeurs dans Δ^n qui est de mesure de Lebesgue nulle dans \mathbb{R}^n . Nous avons donc besoin de considérer la forme réduite des poids qui est à valeurs dans \mathbb{R}^{n-1} . Lorsque la volatilité $\sigma(\cdot)$ est polynomiale, *i.e.* l'exposant ζ est un entier naturel strictement positif, le drift et la volatilité sont toutes les deux des fonctions bornées et infiniment différentiables avec dérivées bornées sur \mathbb{R}^{n-1} . En vérifiant l'hypothèse de Hörmander, voir par exemple [Nua06], avec les poids initiaux strictement positifs, *i.e.* $\boldsymbol{\mu}_0 \in \Delta_+^n \subset \mathbb{R}^{n-1}$, à tout instant $t > 0$, la distribution des poids $m_t(\cdot|\boldsymbol{\mu}_0)$ est absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R}^{n-1} et la densité de probabilité associée $p_t(\cdot|\boldsymbol{\mu}_0)$ existe est infiniment différentiable par rapport à sa variable d'espace :

$$\begin{aligned} m_t(d\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) &= p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) \cdot d\boldsymbol{\mu}_t \\ p_t(\cdot|\boldsymbol{\mu}_0) &\in \mathcal{C}^\infty(\mathbb{R}^{n-1}, \mathbb{R}_+) \end{aligned}$$

Une distribution $m(\cdot)$ à support dans Δ^n est *stationnaire* (ou c'est une mesure invariante) par rapport au processus $(\boldsymbol{\mu}_t)_{t \geq 0}$ si pour toute fonction f continue bornée et tout $t > 0$:

$$\langle \mathcal{P}_t f, m \rangle = \int_{\Delta^n} \left(\int_{\Delta^n} f(\boldsymbol{\mu}_t) p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) d\boldsymbol{\mu}_t \right) m(d\boldsymbol{\mu}_0) = \langle f, m \rangle = \int_{\Delta^n} f(\boldsymbol{\mu}_0) m(d\boldsymbol{\mu}_0)$$

Résultat Principal 1.1.3. *Partant de $\boldsymbol{\mu}_0 \in \Delta_+^n$ et avec volatilité polynomiale, *i.e.* $\zeta \in \mathbb{N}^*$, la distribution stationnaire des poids du marché, notée $m_\infty(\cdot)$ existe, est unique et est absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R}^{n-1} .*

De plus, avec la notation d_{TV} pour la distance en variation totale entre deux mesures, la vitesse de convergence de la distribution des poids du marché $m_t(\cdot|\boldsymbol{\mu}_0)$ vers l'unique distribution stationnaire $m_\infty(\cdot)$ quand $t \uparrow \infty$ est sous-géométrique, c'est-à-dire pour tout $\gamma > 0$ il existe une constante $C(\gamma) > 0$ indépendante de t et de $\boldsymbol{\mu}_0$ telle que :

$$\begin{aligned} d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_\infty(\cdot)) &\leq C(\gamma) \cdot (V(\boldsymbol{\mu}_0)/t^{\gamma+1} + 1/t^\gamma) \\ V(\boldsymbol{\mu}) &= - \sum_{i=1}^{n-1} \log(\mu^i) - \log \left(1 - \sum_{j=1}^{n-1} \mu^j \right) = - \sum_{i=1}^n \log(\mu^i) \end{aligned}$$

Supposons qu'il existe une solution non-négative $p(\cdot)$ définie sur Δ_+^n à l'équation de Fokker-Planck avec dérivée temporelle nulle, avec conditions aux bords $p(\boldsymbol{\mu}) = 0 \forall \boldsymbol{\mu} \in \partial\Delta^n$:

$$\mathcal{L}^* p = \sum_{i=1}^{n-1} \partial_i [b^i(\boldsymbol{\mu}) p(\boldsymbol{\mu})] - \frac{1}{2} \sum_{i,j=1}^{n-1} \partial_{ij}^2 [a^{i,j}(\boldsymbol{\mu}) p(\boldsymbol{\mu})] = 0$$

Si de plus l'intégrale de $p(\cdot)$ sur Δ_+^n est égale à 1, $p(\cdot)$ est donc une densité de probabilité, alors la mesure ainsi définie $m(\boldsymbol{\mu}) = p(\boldsymbol{\mu})d\boldsymbol{\mu}$ est stationnaire.

Cette équation peut être résolue analytiquement dans certains cas, notamment lorsque le nombre d'actifs n est égal à 2. Par exemple pour le modèle du marché à volatilité stabilisée, la solution correspond à la distribution de Dirichlet.

1.1.2 Structure du marché et arbitrage relatif

Il est possible de décrire la structure complète du marché par le couple $(\boldsymbol{\mu}_t, \Sigma_t)$ en introduisant le processus de la capitalisation totale du marché, ou la taille du portefeuille de marché $(\Sigma_t)_{t \geq 0}$, qui pourrait être corrélé au processus des poids du marché à travers les coefficients de volatilité $(\sigma^{\Sigma, \nu}(\cdot))_{\nu=1}^d$:

$$d\Sigma_t = b^\Sigma(\Sigma_t)dt + \sum_{\nu=0}^d \sigma^{\Sigma, \nu}(\Sigma_t)dW_t^\nu, \quad \Sigma_0 > 0$$

La capitalisation de l'actif i est simplement $S_t^i = \mu_t^i \cdot \Sigma_t$. En passant au logarithme, nous obtenons la *décomposition du taux de croissance* suivante :

$$d \log S_t^i = d \log \mu_t^i + d \log \Sigma_t = (\gamma^i(\boldsymbol{\mu}) + \gamma^\Sigma(\Sigma))dt + \text{martingale}$$

Le taux de croissance se décompose en $\gamma^i(\cdot)$, la partie associée au poids de l'actif i et en $\gamma^\Sigma(\cdot)$ qui est le taux de croissance de la capitalisation totale du marché commun à tous les actifs.

En prenant la capitalisation totale égale à constante, ce qui revient à utiliser le portefeuille de marché comme numéraire, nous retrouvons la dynamique du logarithme du poids du marché de l'actif i :

$$d \log \mu_t^i = \gamma^i(\boldsymbol{\mu})dt + \sum_{\nu=1}^d (\mu_t^i)^{-1} \sigma^{i, \nu}(\boldsymbol{\mu}_t) dW_t^\nu$$

Regardons de plus près le terme de volatilité et définissons la volatilité totale $\sigma^i(\boldsymbol{\mu}_t)$ de l'actif i :

$$\begin{aligned} \sigma^i(\boldsymbol{\mu}_t)^2 &= \sum_{\nu=1}^d (\mu_t^i)^{-2} \sigma^{i, \nu}(\boldsymbol{\mu}_t)^2 = (\mu_t^i)^{2(\zeta-1)} (1 - 2\mu_t^i) \sum_{\nu=1}^d (\zeta^{i, \nu})^2 + \sum_{\nu=1}^d \left(\sum_{k=1}^n (\mu^k)^\zeta \zeta^{k, \nu} \right)^2 \\ &\quad - 2(\mu_t^i)^{(\zeta-1)} \sum_{\nu=1}^d \zeta^{i, \nu} \sum_{k=1, k \neq i}^n (\mu^k)^\zeta \zeta^{k, \nu} \end{aligned}$$

Rappelons que la matrice ζ est de rang n , donc $c^i = \sum_{\nu=1}^d (\zeta^{i, \nu})^2 > 0$. Pour $\zeta \in (0, 1)$, quand le poids μ_t^i de l'actif i est petit, le premier terme après la dernière égalité $(\mu_t^i)^{2(\zeta-1)}(1 - 2\mu_t^i)c^i$ est prédominant et $\sigma^i(\boldsymbol{\mu}_t) \rightarrow \infty$ lorsque $\mu_t^i \rightarrow 0$, les actifs avec plus petits poids ont tendance à avoir plus grandes volatilités comme dans le modèle du marché à volatilité stabilisée [Pal+11]. On dit que la *stabilisation par volatilité* est *asymptotique*. Plus $\zeta \in (0, 1)$ est petit, plus le premier terme est grand, plus l'effet de stabilisation par volatilité est important. Lorsque $\zeta \geq 1$, la volatilité totale $\sigma^i(\boldsymbol{\mu}_t)$ est

bornée pour $\boldsymbol{\mu}_t \in \Delta_n$, la stabilisation par volatilité asymptotique n'est en général pas observée.

En théorie stochastique du portefeuille classique, nous définissons l'ensemble des stratégies de trading auto-finançantes $\mathcal{T}(\mathbf{S})$ comme des processus à valeurs dans \mathbb{R}^n , prévisibles et intégrales par rapport au processus des capitalisations. Ici en nous plaçant dans le numéraire du portefeuille de marché $(\Sigma_t)_{t \geq 0}$, la valeur du portefeuille d'investissement $(V_t^\pi)_{t \geq 0}$, aussi appelée le processus de richesse, peut s'exprimer en termes des poids du marché $(\boldsymbol{\mu}_t)_{t \geq 0}$, avec la stratégie $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t \geq 0} \in \mathcal{T}(\boldsymbol{\mu})$. En fait selon [KR17], l'ensemble des stratégies de trading auto-finançantes est le même par rapport aux capitalisations et aux poids relatifs des actifs, *i.e.* $\mathcal{T}(\boldsymbol{\mu}) = \mathcal{T}(\mathbf{S})$. Pour simplifier, en normalisant nous pouvons supposer que la valeur initiale du portefeuille est égale à 1.

$$V_t^\pi = \boldsymbol{\pi}_t \cdot \boldsymbol{\mu}_t = \sum_{i=1}^n \pi_t^i \mu_t^i, \quad V_0^\pi = 1$$

Pour deux stratégies de trading $\boldsymbol{\pi}, \boldsymbol{\rho} \in \mathcal{T}(\boldsymbol{\mu})$, on dit que $\boldsymbol{\pi}$ réalise un arbitrage relatif par rapport à $\boldsymbol{\rho}$ sur l'horizon de temps $[0, T]$ si :

$$\mathbb{P}[V_T^\pi \geq V_T^\rho] = 1 \text{ et } \mathbb{P}[V_T^\pi > V_T^\rho] > 0$$

L'arbitrage relatif est fort si :

$$\mathbb{P}[V_T^\pi > V_T^\rho] = 1$$

Un arbitrage relatif qui nous intéresse en particulier est celui par rapport au marché. La valeur du portefeuille de marché étant toujours égale à 1 dans son propre numéraire *i.e.* $V_t^\mu = 1 \forall t \geq 0$, nous écrivons simplement :

$$\begin{aligned} \mathbb{P}[V_T^\pi \geq 1] &= 1 \text{ et } \mathbb{P}[V_T^\pi > 1] > 0 \\ \mathbb{P}[V_T^\pi > 1] &= 1 \text{ pour l'arbitrage relatif fort} \end{aligned}$$

Le taux de croissance $\gamma^\pi(\cdot)$ associée à la stratégie de trading $\boldsymbol{\pi}$ est défini à partir de sa forme logarithmique :

$$d \log V_t^\pi = \gamma^\pi(\boldsymbol{\mu}_t) dt + \text{martingale}$$

Définissons le taux croissance en excès $\gamma_*^\pi(\cdot)$ et la même quantité associée au marché $\gamma_*(\cdot)$:

$$\begin{aligned} \gamma_*^\pi(\boldsymbol{\mu}_t) &= \gamma^\pi(\boldsymbol{\mu}_t) - \sum_{i=1}^n \pi_t^i \gamma^i(\boldsymbol{\mu}_t) \\ \gamma_*(\boldsymbol{\mu}_t) &= \frac{1}{2} \sum_{i=1}^n (\mu_t^i)^{-1} a^{i,i}(\boldsymbol{\mu}_t) \end{aligned}$$

Le taux de croissance en excès cumulée (du marché) est définie comme l'intégrale de la croissance en excès le long de la trajectoire des poids du marché :

$$\Gamma_*(t) = \int_0^t \gamma_*(\boldsymbol{\mu}_u) du = \frac{1}{2} \sum_{i=1}^n \int_0^t (\mu_u^i)^{-1} a^{i,i}(\boldsymbol{\mu}_u) du$$

Une condition suffisante pour l'existence d'arbitrage relatif (au marché) obtenue dans [FK05] porte sur le taux de croissance en excès du marché, si $\gamma_*(\cdot)$ est minoré par une constante strictement positive, i.e. $\exists \underline{\gamma} > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma}, \forall t \in [0, T]$ p.s., alors il existe un seuil $T^* > 0$ dépendant des paramètres du modèles, tel que l'arbitrage relatif existe sur l'horizon du temps $[0, T], \forall T > T^*$.

Résultat Principal 1.1.4. *Selon les valeurs de l'exposant $\zeta > 0$, nous avons les estimations suivantes du taux de croissance en excès :*

1. *Lorsque $\zeta \leq 1/2$ et sous l'hypothèse $\zeta \boldsymbol{\varsigma}^\top = \text{diag}(c^1, \dots, c^n)$, $c^i > 0 \forall i \in \{1, \dots, n\}$, il existe une constante $\underline{\gamma} > 0$ telle que :*

$$\forall t > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma} \text{ a.s.}$$

La condition suffisante d'existence d'arbitrage relatif ci-dessus est donc vérifiée, au moins pour un horizon de temps suffisamment long.

2. *Lorsque $\zeta \in N^*$ nous avons :*

$$\forall \epsilon > 0 : \forall t > 0 : \mathbb{P}[\gamma_*(\boldsymbol{\mu}_t) < \epsilon] > 0$$

C'est donc le cas opposé de la condition suffisante pour l'arbitrage relatif. Pour tout $\epsilon > 0$, le taux de croissance en excès a une probabilité strictement positive d'être inférieure à ϵ .

La prime de risque du marché associée aux poids $\boldsymbol{\theta}(\cdot) = (\theta^\nu(\cdot))_{\nu=1}^d$ satisfait la relation :

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, \forall t \geq 0, \mathbf{b}(\boldsymbol{\mu}_t) = \boldsymbol{\sigma}(\boldsymbol{\mu}_t) \cdot \boldsymbol{\theta}(\boldsymbol{\mu}_t) \text{ et } \forall T \geq 0, \int_0^T |\boldsymbol{\theta}(\boldsymbol{\mu}_t)|^2 dt < \infty$$

Le déflateur de risque du marché $(Z_t)_{t \geq 0}$ est défini à partir de la prime de risque par une martingale exponentielle avec $Z_0 = 1$, c'est une martingale locale :

$$Z_t = \mathcal{E}\left(-(\boldsymbol{\theta}(\boldsymbol{\mu}_u))_{u \in [0, t]}\right) = \exp\left(-\int_0^t \boldsymbol{\theta}(\boldsymbol{\mu}_u) \cdot d\mathbf{W}_u - \frac{1}{2} \int_0^t |\boldsymbol{\theta}(\boldsymbol{\mu}_u)|^2 du\right)$$

Quand le déflateur de risque $(Z_t)_{t \geq 0}$ existe, on dit que la condition de "Pas de profit non borné avec risque borné" ou *NUPBR* introduite dans [DS94] est vérifiée. Comme son nom l'indique, cette condition est équivalente au fait que l'ensemble des processus de richesse $\{V_T^\pi \mid \boldsymbol{\mu} \in \mathcal{T}(\boldsymbol{\mu})\}$ est borné en probabilité. La condition de "Pas de déjeuner gratuit avec risque qui s'annule", ou *NFLVR* est satisfaite lorsque nous avons (*NUPBR*) et il n'y a pas d'arbitrage relatif au marché.

Résultat Principal 1.1.5. *Pour $\boldsymbol{\mu}_0 \in \Delta_+^n$ et sous l'hypothèse $d \geq n$ et $\boldsymbol{\varsigma}$ de plein rang, alors la prime de risque du marché $\boldsymbol{\theta}(\cdot)$, ainsi que le déflateur de risque du marché $(Z_t)_{t \geq 0}$ existent, la condition (*NUPBR*) est donc vérifiée sur l'horizon de temps $[0, T], T > 0$, si et seulement si l'une des conditions suivantes est satisfaite :*

- (1) *L'une des conditions (i) - (iii) dans Résultat Principal 1.1.1 est satisfaite, rappelons que dans ce cas $\forall t \geq 0, \boldsymbol{\mu}_t \in \Delta_+^n$ a.s..*

(2) L'une des conditions (iv) - (vi) dans Résultat Principal 1.1.1 est satisfaite, et $b^i(\boldsymbol{\mu}) = 0$ sur $\{\mu^i = 0\}$ pour tous les indices i tels que $\mathbb{P}[\exists t \in (0, T] \mid \mu_t^i = 0] > 0$.

De plus, la prime de risque du marché peut s'écrire comme suit :

$$\boldsymbol{\theta}(\boldsymbol{\mu}_t) = (\boldsymbol{\sigma}^\top \cdot (\mathbf{a}^+ \cdot \mathbf{b}))(\boldsymbol{\mu}_t)$$

Où $\mathbf{a}^+(\cdot)$ est le pseudo-inverse de Moore-Penrose de la matrice de covariance.

Dans le cas contraire, lorsqu'aucune des conditions (1) et (2) ci-dessus n'est satisfaite, i.e. pour un certain indice $i \in \{1, \dots, n\}$, $(\mu_t^i)_{t \geq 0}$ s'annule avec probabilité non nulle et il existe $\boldsymbol{\mu} \in \{\mu^i = 0\}$ tel que $b^i(\boldsymbol{\mu}) > 0$, alors la condition (NUPBR) n'est pas satisfaite, en particulier la prime $\boldsymbol{\theta}(\cdot)$ et le déflateur $(Z_t)_{t \geq 0}$ de risque du marché n'existent pas.

Le lien entre la martingalité du déflateur de risque du marché et la condition d'existence d'arbitrage relatif au marché est établi dans le résultat suivant.

Résultat Principal 1.1.6. *Il n'y a pas d'arbitrage relatif au marché sur l'horizon de temps $[0, T]$ si et seulement si le déflateur de risque du marché $(Z_t)_{t \geq 0}$ associé aux poids du marché est une vraie martingale sur $[0, T]$, i.e. $\mathbb{E}[Z_T] = 1$. Ou de manière équivalente, l'arbitrage relatif existe sur $[0, T]$ si et seulement si la martingale locale $(Z_t)_{t \geq 0}$ est stricte sur $[0, T]$ avec $\mathbb{E}[Z_T] < 1$.*

Le résultat précédent nous donne une condition nécessaire et suffisante pour l'existence d'arbitrage relatif par rapport au marché, mais la difficulté technique consiste à vérifier si le déflateur de risque du marché $(Z_t)_{t \geq 0}$ est une vraie martingale sur $[0, T]$.

Résultat Principal 1.1.7. *Lorsque le nombre n d'actifs est égal à 2, avec la matrice $\varsigma = [\varsigma^1 \ \varsigma^2]^\top$ vérifiant $\varsigma^1 \geq 0 \geq \varsigma^2$, $(\varsigma^1, \varsigma^2) \neq (0, 0)$, alors selon la valeur de l'exposant $\zeta > 0$ dans l'équation des poids du marché, nous avons :*

1. *Quand $\zeta \in (0, 1)$: L'arbitrage relatif par rapport au marché existe, au moins pour des horizons de temps suffisamment longs.*
2. *Quand $\zeta \geq 1$: Nous avons l'absence d'arbitrage relatif pour tout horizon de temps $[0, T]$, $T > 0$.*

Pour le cas général où le nombre d'actifs $n \geq 2$ est arbitraire, nous avons pour $\zeta \geq 1$, l'absence d'arbitrage relatif pour tout horizon de temps $[0, T]$, $T > 0$.

1.1.3 Applications et perspectives

Pour construire d'autres portefeuilles d'investissement dans ce même marché, la notion de portefeuilles générées par des fonctions a été introduite dans [Fer99]. Cette fonction a pour argument les poids relatifs des actifs, qui sont donnés immédiatement par notre modèle. Pour une fonction G définie sur le simplexe unité Δ^n à valeurs strictement positives, deux fois différentiables avec dérivées continues, un nouveau portefeuille avec poids relatifs $(\boldsymbol{\pi}_t)_{t \geq 0}$ est construit à partir du portefeuille de marché $(\boldsymbol{\mu}_t)_{t \geq 0}$ par :

$$\pi_t^i = \mu_t^i \cdot (1 + \partial_i \log G(\boldsymbol{\mu}_t) - \sum_{j=1}^n \mu_t^j \cdot \partial_j \log G(\boldsymbol{\mu}_t))$$

Lorsque l'arbitrage relatif par rapport au marché existe sur $[0, T]$, $\mathbb{E}[Z_T] < 1$, définissons la fonction $u(t, \boldsymbol{\mu}_t) = \mathbb{E}[Z_T | \mathcal{F}_t] / Z_t$ et si u est régulière en prenant $G(\boldsymbol{\mu}_t) = u(t, \boldsymbol{\mu}_t)$ nous construisons l'arbitrage relatif au marché, optimal en terme d'investissement initial avec delta hedging.

Pour calibrer notre modèle aux données du marché, ou simplement estimer les paramètres du modèle, supposons par exemple que $\zeta = 1$. Notons $z_t^i = \log(\mu_t^i)$, nous avons l'expression de variation quadratique suivante :

$$\frac{d\langle \tilde{z}^i, \tilde{z}^j \rangle_t}{dt} = \sum_{\nu=1}^d \zeta^{i,\nu} \zeta^{j,\nu} = (\zeta \zeta^\top)^{i,j}; \quad \tilde{z}_t^i = z_t^i - \frac{1}{n} \sum_{k=1}^n z_t^k$$

Cette matrice de variations quadratiques peut être estimée sur une grille de temps \mathcal{T} :

$$\begin{aligned} \mathcal{T} &= \{t_m = m \cdot \Delta t, \Delta t = T/M_T; m = 0, \dots, M_T\} \\ (\zeta \zeta^\top)^{i,j} &\approx \frac{1}{T} \sum_{m=1}^{M_T} (\tilde{z}_{t_m}^i - \tilde{z}_{t_{m-1}}^i)(\tilde{z}_{t_m}^j - \tilde{z}_{t_{m-1}}^j), \quad i \leq j \end{aligned}$$

La matrice ζ peut donc être définie comme racine carrée de la matrice de variation quadratique.

Les coefficients de la matrice B sont obtenus par des régressions linéaires multiples pour chaque actif $i \in \{1, \dots, n\}$:

$$\mu_{t_m}^i \approx \mu_{t_{m-1}}^i + \left[\sum_{j=1}^{n-1} (B^{i,j} - B^{i,n}) \mu_{t_{m-1}}^j + B^{i,n} \right] \Delta t + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_{t_{m-1}}) (W_m^\nu - W_{m-1}^\nu)$$

L'exposant $\zeta > 0$, qui mesure l'effet de stabilisation par volatilité, peut lui-même être estimé à partir des données. L'estimation des matrices B et ζ dans le cas général reste donc à étudier.

Avec l'ergodicité du processus des poids du marché démontrée dans Résultat Principal 1.1.3, nous étudions l'optimisation de la richesse sur le long terme, en définissant le *taux de croissance espéré sensibilisé par le risque* $J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T)$ pour l'horizon de temps T et sa version long terme $J(\boldsymbol{\mu}, \boldsymbol{\pi})$, avec $\theta > -2$, $\theta \neq 0$ un paramètre de sensibilité au risque :

$$\begin{aligned} J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T) &= -\frac{2}{\theta} \log \mathbb{E}[e^{-\frac{\theta}{2} \log V_T^\pi}] \\ J(\boldsymbol{\mu}, \boldsymbol{\pi}) &= \limsup_{T \rightarrow \infty} \frac{1}{T} J(\boldsymbol{\mu}, \boldsymbol{\pi}; T) \end{aligned}$$

Il est possible d'établir l'équation de Bellman ou sa version ergodique associée au problème d'optimisation de $J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T)$ ou de $J(\boldsymbol{\mu}, \boldsymbol{\pi})$. Cependant, pour établir l'existence et l'unicité de la solution de l'équation de Bellman, nous avons besoin de l'uniforme ellipticité, ce qui n'est pas le cas pour le processus des poids du marché qui n'admet que l'ellipticité locale. Ce problème de contrôle stochastique ergodique peut donc faire l'objet de recherches futures.

Typiquement nous appliquons notre modèle à des portefeuilles de marché avec un nombre d'actifs n important. Nous nous intéresserons naturellement au comportement

des poids du marché lorsque $n \rightarrow \infty$. L'équation des poids du marché, qui donne des distributions de probabilités discrètes, devient à valeurs mesures. Pour le modèle de marché à volatilité stabilisée, la dynamique des poids correspond au processus de Jacobi ou Wright-Fisher, a pour limite le processus de Fleming-Viot, qui est à valeurs mesures comme indiqué dans [Pal+11]. Les processus polynomiaux à valeurs mesures sont également étudiés dans le papier récent [CLSF+19]. Le lien avec les modèles basés sur les rangs, voir [JR15] pourrait aussi être établi, en prenant en compte l'interaction de type champ moyen entre les poids des actifs.

1.2 Un schéma probabiliste pour les équations de McKean-Vlasov

La seconde partie est dédiée à l'approximation numérique de la loi marginale de l'équation différentielle stochastique de McKean-Vlasov suivante :

$$X_t^{s,\xi} = \xi + \int_s^t b(X_r^{s,\xi}, [X_r^{s,\xi}]) dr + \int_s^t \sigma(X_r^{s,\xi}, [X_r^{s,\xi}]) dW_r, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d) \quad (1.2.1)$$

où $\mathcal{P}(\mathbb{R}^d)$ désigne l'ensemble des mesures de probabilités sur \mathbb{R}^d . Nous nous plaçons dans un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, W est un mouvement Brownien de dimension q , les coefficients b, σ sont définis sur \mathbb{R}^d à valeurs dans \mathbb{R}^d et $\mathbb{R}^{d \times q}$ respectivement. Ici $[\theta]$ désigne la loi de la variable aléatoire θ . Sous des hypothèses de faible régularité sur les fonctions b et $a = \sigma\sigma^*$, *i.e.* si elles sont bornées, Hölderiennes en la variable d'espace (uniformément en la variable mesure), si elles admettent des dérivées linéaires fonctionnelles (ou dérivées plates) par rapport à la variable mesure notées $\delta b/\delta m$ et $\delta a/\delta m$, avec $\delta b/\delta m$ bornée, $(x, y) \mapsto [\delta a(x, m)/\delta m](y)$ Hölderienne (uniformément en la variable m) et enfin si le coefficient de diffusion a est uniformément elliptique : $\exists \lambda \geq 1, \forall (x, u, m) \in (\mathbb{R}^d)^2 \times \mathcal{P}(\mathbb{R}^d), \lambda^{-1}|u|^2 \leq \langle a(x, m)u, u \rangle \leq \lambda|u|^2$ alors l'EDS de McKean-Vlasov ci-dessus admet une unique solution faible pour toute condition initiale $(s, \mu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$, c.f. Chaudru de Raynal et Frikha [RF19].

1.2.1 Itération de Picard et méthode de continuation

Le flot stochastique découplé associée à l'EDS de McKean-Vlasov est défini comme une EDS standard avec des coefficients inhomogènes en temps, en considérant la loi μ de la variable aléatoire ξ :

$$X_t^{s,x,\mu} = x + \int_s^t b(X_r^{s,x,\mu}, [X_r^{s,\xi}]) dr + \int_s^t \sigma(X_r^{s,x,\mu}, [X_r^{s,\xi}]) dW_r$$

Notre stratégie est d'approcher la loi marginale de la solution de l'EDS de McKean-Vlasov en utilisant *l'itération de Picard* et le taux de convergence est géométrique selon [Szn91]. Pour réduire la complexité notamment lorsque le nombre d'itérations augmente, nous utilisons la *méthode de continuation en temps* sur l'horizon $[0, T]$ en introduisant la grille de temps uniforme $\mathcal{S} = \{t_0 = 0 \leq t_1 \leq \dots \leq t_N = T\}$, $N \geq 1$ avec $t_k = k\delta$, $\delta = T/N$. Des itérations de Picard locales sont effectuées sur chaque sous-intervalle de temps $[t_k, t_{k+1}]$, $k = 0, \dots, N-1$ en utilisant l'approximation obtenue sur les sous-intervalles précédents.

Définissons l'ensemble $\mathcal{A}_{t_k, t_{k+1}, \mu_k}$ comme un sous-espace fermé de $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$ et la métrique uniforme $d_{t_k, t_{k+1}}$:

$$\mathcal{A}_{t_k, t_{k+1}, \mu_k} = \left\{ \mathbf{P} \in \mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d)) : \mathbf{P}(t_k) = \mu_k \right\},$$

$$d_{t_k, t_{k+1}}(\mathbf{P}, \mathbf{Q}) = \sup_{s \in [t_k, t_{k+1}]} |d_{TV}(\mathbf{P}(s), \mathbf{Q}(s))|$$

où d_{TV} est la distance en variation totale sur $\mathcal{P}(\mathbb{R}^d)$. Remarquons que l'espace $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$ muni de la métrique $d_{t_k, t_{k+1}}$ est complet. Par ailleurs, $\mathcal{A}_{t_k, t_{k+1}, \mu_k}$ étant un sous-espace fermé de $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$, il est donc également complet muni de la métrique $d_{t_k, t_{k+1}}$. Définissons l'application $\mathcal{T}_k : \mathcal{A}_{t_k, t_{k+1}, \mu_k} \rightarrow \mathcal{A}_{t_k, t_{k+1}, \mu_k}$, qui a $\mathbf{Q} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ associe $\mathcal{T}_k(\mathbf{Q})(t) = [X_t^{\mathbf{Q}}]$, $t \in [t_k, t_{k+1}]$, où $X^{\mathbf{Q}}$ est donnée par l'unique solution faible de l'EDS suivante :

$$X_t^{\mathbf{Q}} = \xi_k + \int_{t_k}^t b(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) ds + \int_{t_k}^t \sigma(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) dW_s, \quad t \in [t_k, t_{k+1}], \quad [\xi_k] = \mu_k.$$

L'application \mathcal{T}_k est bien définie sous les hypothèses de régularité et d'uniforme ellipticité précédemment mentionnées. Le point fixe de \mathcal{T}_k est l'unique solution de l'EDS de McKean-Vlasov sur $[t_k, t_{k+1}]$ avec la condition initiale (t_k, μ_k) .

La condition initiale $\mathbf{P}_k^{(0)} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ du schéma d'itération de Picard est donnée par le processus \bar{X}_t^k en gelant les coefficients en (t_k, ξ_k, μ_k) :

$$\bar{X}_t^k = \xi_k + b(\xi_k, \mu_k)(t - t_k) + \sigma(\xi_k, \mu_k)(W_t - W_{t_k})$$

avec $\mathbf{P}_k^{(0)}(t) = [\bar{X}_t^k]$, $t \in [t_k, t_{k+1}]$ qui s'écrit donc :

$$\mathbf{P}_k^{(0)}(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) \mu_k(dx) dz$$

où pour une matrice réelle a de dimension $d \times d$ symétrique et inversible, $x \mapsto g(a, x)$ désigne la fonction de densité la loi normale centrée de matrice de variance covariance a . Le flot de mesures de probabilité $\left\{ \mathbf{P}_k^{(L)}(t) = \mathcal{T}_k^{(L)}(\mathbf{P}_k^{(0)})(t), t \in [t_k, t_{k+1}] \right\}$ fournit une approximation de $([X_t])_{t \in [t_k, t_{k+1}]}$, où $\mathcal{T}_k^{(L)}$ représente l'application \mathcal{T}_k itérée L fois, avec L le nombre d'itérations de Picard, aussi appelé le nombre de niveaux. Pour $k = 0$, $\mathbf{P}_0^{(0)}(t_0) = \mu_0 = [\xi]$ et $\mathbf{P}_k^{(L)}(t_k) = \mu_k^L := \mathbf{P}_{k-1}^{(L)}(t_k)$ pour $k = 1, \dots, N-1$.

Sur chaque sous-intervalle de temps $[t_k, t_{k+1}]$ et pour tout niveau $\ell = 1, \dots, L$, $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ correspond au flot de lois marginales de l'unique solution $X^{(\ell)}$ de l'EDS *découplée* suivante qui a pour paramètre le flot des mesures au niveau précédent $(\mathbf{P}_k^{(\ell-1)}(t))_{t \in [t_k, t_{k+1}]}$:

$$X_t^{(\ell)} = \xi_k^{(\ell)} + \int_{t_k}^t b(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) ds + \int_{t_k}^t \sigma(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) dW_s \quad (1.2.2)$$

avec $[X_0^{(\ell)}] = \mu_k^{(\ell)} = \mathbf{P}_k^{(\ell)}(t_k)$.

D'après [Fri08], sous les hypothèses précédentes, la variable aléatoire $X_t^{(\ell)}$ admet une densité $z \mapsto p_\ell(t_k, t, z)$ vérifiant :

$$\mathbf{P}_k^{(\ell)}(t)(dz) = p_\ell(t_k, t, z) dz = \int_{\mathbb{R}^d} p_\ell(t_k, t, x, z) \mu_k^{(\ell)}(dx) dz.$$

Nous avons la borne supérieure Gaussienne suivante pour tout $t \in (t_k, t_{k+1}]$, $(x, z) \in (\mathbb{R}^d)^2$ et $\ell \geq 0$, avec deux constantes $c, C > 0$ indépendantes de ℓ :

$$p_\ell(t_k, t, x, z) \leq E_{1/2,1}(C(|b|_\infty + 1)) g(c(t - t_k), z - x)$$

où pour $c > 0$, $x \mapsto g(c, x)$ désigne la fonction densité de la loi normale centrée, de matrice de variance-covariance cI_d .

1.2.2 Représentation probabiliste de la loi marginale

Nous cherchons maintenant une approximation de la suite de mesures de probabilité $(\mathbf{P}^{(\ell)}(t_k))_{1 \leq \ell \leq L}$, $k = 1, \dots, N$ à l'aide d'une représentation probabiliste pour les lois marginales des processus de diffusion standard, introduite dans [BKH+15] puis étudiée dans [AKH+17] et [HLTT+17].

Soit une suite de variables aléatoires $\tau := (\tau_n)_{n \geq 1}$, $\tau_0 = 0$ avec $(\tau_n - \tau_{n-1})_{n \geq 1}$ i.i.d. de densité f et de fonction de répartition $F(t) = \int_{-\infty}^t f(s) ds$. Le processus de renouvellement $N := (N_t)_{t \geq 0}$ avec temps de sauts τ est défini par $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$.

Pour toute fonction $\Phi : \Delta_n([s, t]) \rightarrow \mathbb{R}$ vérifiant $\mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}} |\Phi(\tau_1, \dots, \tau_n)] < \infty$ où l'ensemble $\Delta_n([s, t]) = \{\mathbf{s}_n \in [s, t]^n : s_0 := s \leq s_1 \leq \dots \leq s_n \leq t =: s_{n+1}\}$:

$$\begin{aligned} & \mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}} \Phi(\tau_1, \dots, \tau_n)] \\ &= \int_{\Delta_n([s, t])} \Phi(r_1 - s, \dots, r_n - s) (1 - F(t - r_n)) \prod_{j=0}^{n-1} f(r_{j+1} - r_j) d\mathbf{r}_n. \end{aligned}$$

En prenant certaines fonctions de densité, nous reconnaissons les processus de renouvellement suivants :

1. Pour $f(t) = \lambda e^{-\lambda t} \mathbf{1}_{[0, \infty)}(t)$, $\lambda > 0$, N est un processus de Poisson d'intensité λ .
2. Pour $f(t) = \frac{1-\alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{t^\alpha} \mathbf{1}_{[0, \bar{\tau}]}(t)$, $(\alpha, \bar{\tau}) \in (0, 1) \times (0, \infty)$, les temps de sauts de N suivent la loi $Beta(1 - \alpha, 1)$ à valeurs dans $[0, \bar{\tau}]$.

Soient $t \in [t_k, t_{k+1}]$, $\ell \geq 0$, une mesure initiale μ_k^ℓ , un mouvement Brownien W^ℓ et un processus de renouvellement N^ℓ avec temps de sauts $\tau^\ell = (\tau_n^\ell)_{n \geq 0}$, le schéma d'Euler $\bar{X}^\ell = (\bar{X}_{\zeta_n^\ell}^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$ partant de $\bar{X}_{t_k}^\ell$, $[\bar{X}_{t_k}^\ell] = \mu_k^\ell$ est défini sur la grille de temps aléatoire $\zeta^\ell := (\zeta_n^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$, $\zeta_n^\ell := t_k + \tau_n^\ell \wedge (t - t_k)$, $\zeta_0^\ell := t_k$:

$$\begin{aligned} \bar{X}_{\zeta_{n+1}^\ell}^\ell &= \bar{X}_{\zeta_n^\ell}^\ell + b_n^\ell (\zeta_{n+1}^\ell - \zeta_n^\ell) + \sigma_n^\ell (W_{\zeta_{n+1}^\ell}^\ell - W_{\zeta_n^\ell}^\ell), \quad 0 \leq n \leq N_{t-t_k}^\ell \\ \sigma_n^\ell &= \sigma(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell)), \quad b_n^\ell = b(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell)), \quad a_n^\ell = \sigma_n^\ell(\sigma_n^\ell)^* \end{aligned}$$

Résultat Principal 1.2.1. *Supposons $x \mapsto b(x, m) \in C_b^1(\mathbb{R}^d, \mathbb{R}^d)$, $x \mapsto a(x, m) \in C_b^2(\mathbb{R}^d, \mathbb{R}^{d \times d})$ et la matrice a uniformément elliptique. Soit h une fonction vérifiant $\forall z \in \mathbb{R}^d$, $|h(z)| \leq C \exp\left(\alpha \frac{|z|^2}{2T}\right)$ pour des constantes $C, \alpha > 0$, $\alpha < c^{-1}$ avec c, C apparues dans la borne supérieure Gaussienne de la densité. Alors il existe une constante*

$\kappa := \kappa(c, \alpha)$ telle que si $\int_{\mathbb{R}^d} \exp(\kappa|x|^2) \mu_k^\ell(dx) < \infty$, l'espérance $\mathbb{E}[h(X_t^{(\ell)})]$ admet la représentation probabiliste suivante :

$$\forall t \in [t_k, t_{k+1}], \mathbb{E}[h(X_t^{(\ell)})] = \mathbb{E}\left[h\left(\bar{X}_{\zeta_{N_{t-t_k}^\ell}^\ell}^\ell\right) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^\ell\right], \ell \geq 1$$

où $\theta_r^\ell := \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}(\bar{X}_{\zeta_r^\ell}^\ell, \bar{X}_{\zeta_{r+1}^\ell}^\ell, \mathbf{P}_k^{(\ell-1)})$ pour $r = 0, \dots, N_{t-t_k}^\ell - 1$ et $\theta_{N_{t-t_k}^\ell}^\ell = (1 - F(t - \zeta_{N_{t-t_k}^\ell}^\ell))^{-1}$ avec :

$$\theta_{s,t}(x, y, \mathbf{P}) := (f(t-s))^{-1} \left[\sum_{i,j=1}^d \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) - \sum_{i=1}^d \rho_{s,t}^i(x, y, \mathbf{P}) \right]$$

et les fonctions $\kappa_{s,t}^{i,j}$, $\rho_{s,t}^i$, pour $i, j = 1, \dots, d$, sont définies par :

$$\begin{aligned} \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) &:= \frac{1}{2} [a_{i,j}(y, \mathbf{P}(t)) - a_{i,j}(x, \mathbf{P}(s))] H_2^{i,j}(a(x, \mathbf{P}(s))(t-s), y-x) \\ &\quad + \partial_{y_i} a_{i,j}(y, \mathbf{P}(t)) H_1^i(a(x, \mathbf{P}(s))(t-s), y-x) + \frac{1}{2} \partial_{y_i, y_j}^2 a_{i,j}(y, \mathbf{P}(t)), \\ \rho_{s,t}^i(x, y, \mathbf{P}) &:= [b_i(y, \mathbf{P}(t)) - b_i(x, \mathbf{P}(s))] H_1^i(a(x, \mathbf{P}(s))(t-s), y-x) + \partial_{y_i} b_i(y, \mathbf{P}(t)) \end{aligned}$$

De plus, pour tout $t \in (t_k, t_{k+1}]$, $\ell \geq 1$, les fonctions de densité $z \mapsto p_\ell(t_k, t, z)$, $p_\ell(t_k, t, x, z)$ admettent aussi des formulations probabilistes, par exemple pour $p_\ell(t_k, t, x, z)$:

$$p_\ell(t_k, t, x, z) = \mathbb{E}\left[g\left(a_{N_{t-t_k}^\ell}^\ell(t - \zeta_{N_{t-t_k}^\ell}^\ell), z - \bar{X}_{N_{t-t_k}^\ell}^{\ell,x}\right) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^{\ell,x}\right]$$

avec poids $\theta_r^{\ell,x} = \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}(\bar{X}_{\zeta_r^\ell}^{\ell,x}, \bar{X}_{\zeta_{r+1}^\ell}^{\ell,x}, \mathbf{P}_k^{(\ell-1)})$ et le schéma d'Euler $(\bar{X}_{\zeta_r^\ell}^{\ell,x})_{0 \leq r \leq N_{t-t_k}^\ell}$ est défini de la même façon que précédemment sauf que sa condition initiale est x .

Pour tout $k \in \{0, \dots, N-1\}$, en appliquant la représentation probabiliste k fois, nous avons pour $t \in [t_k, t_{k+1}]$:

$$\begin{aligned} \mathbb{E}[h(X_t^{(\ell)})] &= \mathbb{E}\left[h\left(\bar{X}_{\zeta_{N_{t-t_k}^\ell}^\ell}^\ell\right) \prod_{r=0}^{N_{t-t_k}^{\ell,k+1}} \theta_r^{\ell,k+1} \times \Gamma_k^\ell\right], \quad \ell \geq 1 \\ \Gamma_k^\ell &:= \Gamma_{k-1}^\ell \prod_{r=0}^{N_\delta^{\ell,k}} \theta_r^{\ell,k} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell,j}} \theta_r^{\ell,j}, \quad \text{avec } \Gamma_0^\ell = 1. \end{aligned}$$

Nous posons $\hat{\mu}_0^L = \frac{1}{M_0} \sum_{i=1}^{M_0} \delta_{\xi^{(i)}}$, où les variables aléatoires $(\xi^{(i)})_{1 \leq i \leq M_0}$ sont i.i.d. de loi μ . Nous construisons pas à pas de $k = 0$ jusqu'à $N-1$, les flots de mesures signées

$$\begin{cases} \hat{\mu}^0(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \hat{\mu}_k^L)(t-t_k), z - b(x, \hat{\mu}_k^L)(t-t_k) - x) \hat{\mu}_k^L(dx) dz, t \in [t_k, t_{k+1}], \\ k = 0, \dots, N-1 \end{cases}$$

en posant $\widehat{\mu}_k^L = \widehat{\mu}^L(t_k)$, $k = 1, \dots, N$ et

$$\forall t \in [t_k, t_{k+1}], \quad \widehat{\mu}^\ell(t)(dz) := \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Gamma_k^{\ell,(m)} \prod_{r=0}^{N_{t-t_k}^{\ell,k+1,(m)}} \theta_r^{\ell,k+1,(m)} \delta_{\bar{X}_{N_{t-t_k}^{\ell,k+1,(m)+1}}^{\ell,(m)}}(dz),$$

$$\Gamma_k^{\ell,(m)} := \Gamma_{k-1}^{\ell,(m)} \prod_{r=0}^{N_\delta^{\ell,k,(m)}} \theta_r^{\ell,k,(m)} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell,j,(m)}} \theta_r^{\ell,j,(m)}, \quad \text{avec } \Gamma_0^{\ell,(m)} := 1,$$

et pour $j \in \{1, \dots, k+1\}$ et pour tout $r \in \{0, \dots, N_\delta^{\ell,j,(m)} - 1\}$ (si $1 \leq j \leq k$) ou $r \in \{0, \dots, N_{t-t_k}^{\ell,j,(m)}\}$ (si $j = k+1$)

$$\begin{aligned} \theta_r^{\ell,j,(m)} &:= \theta_{\zeta_r^{\ell,(m)}, \zeta_{r+1}^{\ell,(m)}}^{\ell}(\bar{X}_r^{\ell,(m)}, \bar{X}_{r+1}^{\ell,(m)}, \widehat{\mu}^{\ell-1}), \\ \theta_{N_\delta^{\ell,j}}^{\ell,j,(m)} &:= (1 - F(t_j - \zeta_{N_\delta^{\ell,j,(m)}}^{\ell,(m)}))^{-1}, \\ \theta_{N_{t-t_k}^{\ell,k+1}}^{\ell,k+1,(m)} &:= (1 - F(t - \zeta_{N_{t-t_k}^{\ell,k+1,(m)}}^{\ell,(m)}))^{-1}, \\ \bar{X}_{\zeta_{r+1}^{\ell,(m)}}^{\ell,(m)} &= \bar{X}_{\zeta_r^{\ell,(m)}}^{\ell,(m)} + b_r^\ell(\zeta_{r+1}^{\ell,(m)} - \zeta_r^{\ell,(m)}) + \sigma_r^{\ell,(m)}(W_{\zeta_{r+1}^{\ell,(m)}}^{\ell,(m)} - W_{\zeta_r^{\ell,(m)}}^{\ell,(m)}) \end{aligned}$$

avec les notations $b_r^{\ell,(m)} := b(\bar{X}_r^{\ell,(m)}, \widehat{\mu}^{\ell-1}(\zeta_r^{\ell,(m)}))$, $\sigma_r^{\ell,(m)} := \sigma(\bar{X}_r^{\ell,(m)}, \widehat{\mu}^{\ell-1}(\zeta_r^{\ell,(m)}))$, $a_r^{\ell,(m)} = \sigma_r^{\ell,(m)}(\sigma_r^{\ell,(m)})^T$. Ici, les variables aléatoires $(W^{\ell,(m)}, N^{\ell,(m)}, \bar{X}_0^{\ell,(m)})$, $1 \leq m \leq M_\ell$, $\ell \geq 1$, sont i.i.d. de loi (W, N, X_0) . De plus pour tout $m \in \{1, \dots, M_\ell\}$ fixé, les variables aléatoires $\{N^{\ell,j,(m)}, j = 1, \dots, k+1\}$ sont des processus de renouvellement i.i.d. avec temps de sauts $\tau^{\ell,j,(m)} = (\tau_r^{\ell,j,(m)})_{r \geq 0}$ et nous fixons $\zeta_r^{\ell,(m)} := t_j + \tau_r^{\ell,j,(m)} \wedge \delta$, pour $r = 0, \dots, N_\delta^{\ell,j,(m)} + 1$, $j = 1, \dots, k$ et $\zeta_r^{\ell,(m)} := t_k + \tau_r^{\ell,k+1,(m)} \wedge (t - t_k)$, pour $r = 0, \dots, N_{t-t_k}^{\ell,k+1,(m)} + 1$.

L'estimateur Monte-Carlo de l'espérance $\mathbb{E}[h(X_T)]$ est donné par :

$$\langle h, \widehat{\mu}^L(T) \rangle = \frac{1}{M_L} \sum_{m=1}^{M_L} h(\bar{X}_{N_\delta^{L,N,(m)+1}}^L) \times \Gamma_N^{L,(m)}.$$

La mesure signée $\widehat{\mu}^L(T)$ ayant pour fonction de densité $z \mapsto \widehat{p}_L(0, T, z)$ définie par

$$\widehat{p}_L(0, T, z) := \frac{1}{M_L} \sum_{m=1}^{M_L} g \left(a_{N_\delta^{L,N,(m)}}^{L,(m)}(T - \zeta_{N_\delta^{L,N,(m)}}^{L,(m)}), z - \bar{X}_{N_\delta^{L,N,(m)}}^{L,(m)} \right) \times \Gamma_N^{L,(m)}$$

est une approximation de la densité $z \mapsto p(0, T, z)$ à l'instant T de la solution $X_T^{0,\xi}$ de l'EDS de McKean-Vlasov.

Remarquons que pour estimer $\widehat{\mu}^L(t)$, $t \in [0, t_1]$ nous avons besoin de connaître $\widehat{\mu}^{L-1}$ puis $\widehat{\mu}^{L-2}$ et ainsi de suite, la complexité de la méthode est de l'ordre $M_0 \times \dots \times M_L$. Pour $\widehat{\mu}^L(t)$, $t \in [t_{k-1}, t_k]$, $k = 2, \dots, N$, nous avons besoin de connaître $\widehat{\mu}^L(t_k)$ qui est calculé sur $[t_{k-2}, t_{k-1}]$, la complexité est donc de l'ordre $M_L \times M_0 \times \dots \times M_L$. En additionnant les N terme, la complexité de notre algorithme est donc de l'ordre $M_0 \times \dots \times M_L + (N-1)M_L \times M_1 \times \dots \times M_L$.

1.2.3 Analyse de la convergence et de l'erreur et perspectives

Notre estimateur Monte-Carlo de la quantité $\mathbb{E}[h(X_T)]$ et aussi de la fonction de densité $z \mapsto p(0, T, z)$ contient deux sources d'erreur :

$$\begin{aligned}\mathbb{E}[h(X_T)] - \langle h, \widehat{\mu}^L(T) \rangle &= \mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^L)] + \mathbb{E}[h(X_T^L)] - \langle h, \widehat{\mu}^L(T) \rangle \\ p(0, T, z) - \widehat{p}_L(0, T, z) &= p(0, T, z) - p_L(0, T, z) + p_L(0, T, z) - \widehat{p}_L(0, T, z).\end{aligned}$$

La première erreur correspond au biais introduit par l'approximation de la loi marginale de l'équation de McKean-Vlasov par l'itération de Picard avec L niveaux. Sa convergence est géométrique en L , d'ordre $(C\delta)^L/\Gamma(1 + (L - 1)/2) \sim (C\delta)^L/(\lfloor L/2 \rfloor)!$.

La seconde erreur est une erreur statistique due à l'approximation à chaque niveau $\ell = 1, \dots, L$, du flot de mesures de probabilité μ^ℓ par le flot de mesures signées $\widehat{\mu}^\ell$ obtenues par la représentation probabiliste. Le moment $L^2(\mathbb{P})$ de cette erreur devrait être (au mieux) de l'ordre $M^{-\frac{1}{2}}$, en supposant que $M_\ell = M$, $\ell = 0, \dots, L$, pour simplifier la discussion.

Pour avoir une erreur moyenne quadratique globale de l'ordre ε , N (M resp.) doit être de l'ordre $\varepsilon^{-\frac{1}{L}}$ (resp. ε^{-2}), le coût computationnel NM^{L+1} est donc de l'ordre $\varepsilon^{-\frac{1}{L}}\varepsilon^{-2L-2}$. Il est donc optimal de choisir simplement $L = 1$ pour avoir la complexité optimale ε^{-5} .

Nous donnons d'abord un résultat concernant l'estimation du biais du schéma d'itération de Picard local.

Résultat Principal 1.2.2. *Supposons que les hypothèses de régularité du Résultat Principal 1.2.1 soient satisfaites. Supposons que pour tout $(i, j) \in \{1, \dots, d\}^2$, les fonctions $(x, y) \mapsto [\delta b_i(x, m)/\delta m](y)$ et $(x, y) \mapsto [\delta a_{i,j}(x, m)/\delta m](y)$ appartiennent aux espaces $\mathcal{C}_b^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$ et $\mathcal{C}_b^{2,2}(\mathbb{R}^d \times \mathbb{R}^d)$ respectivement. Alors, pour tout $k = 0, \dots, n - 1$, la suite de flots de mesures $\left\{ \mathbf{P}_k^{(\ell)} = (\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}, \ell \geq 0 \right\}$, où $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ est la loi marginale de la solution l'EDS découplée (1.2.2) avec distribution initiale $\mu_k \in \mathcal{P}_2(\mathbb{R}^d)$ à l'instant t_k , converge vers le flot de mesures de probabilité $\mathbf{P}_k^\infty = (\mathbf{P}_k^\infty(t))_{t \in [t_k, t_{k+1}]}$ lorsque $\ell \uparrow \infty$ donné par l'unique solution faible de l'EDS champ moyen (1.2.1) avec condition initiale (t_k, μ_k) . De plus, pour tout $h \in \mathcal{C}^2(\mathbb{R}^d)$, h et ses dérivées étant à croissance au plus quadratique, nous avons pour tout $L \geq 0$ et $t \in [t_k, t_{k+1}]$:*

$$|\mathbb{E}[h(X_t)] - \mathbb{E}[h(X_t^{(L)})]| \leq \frac{(C\delta)^L}{\Gamma(1 + (L - 1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)})$$

avec les notations :

$$\begin{aligned}\mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) &:= \sup_{t_k \leq r \leq t_{k+1}} \mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r)), \\ \mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r)) &:= \sup_{|h|_{\mathcal{C}_b^2} \leq 1} \left(\int_{\mathbb{R}^d} h(z) \mathbf{P}_k^\infty(r)(dz) - \int_{\mathbb{R}^d} h(z) \mathbf{P}_k^{(0)}(r)(dz) \right).\end{aligned}$$

Nous avons aussi la majoration pour le biais de la densité, pour tout $(t, x, z) \in (t_k, t_{k+1}] \times$

$(\mathbb{R}^d)^2$ et des constantes positives $c > \lambda$, C non-décroissante par rapport à T :

$$\begin{aligned} & |p(t_k, t, x, z) - p_L(t_k, t, x, z)| \\ & \leq \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) g(c(t-t_k), z-x) \end{aligned}$$

Finalemnt, nous avons la majoration suivante avec une constante $C' > 0$, pour tout $k = 0, \dots, n-1$:

$$\mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \leq C\delta.$$

Pour le biais du schéma d'itération de Picard global, le but de notre travail en cours est d'obtenir l'estimation suivante :

$$\begin{aligned} & |p(0, T, z) - p_L(0, T, z)| \\ & \leq CT^{-1} \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \int_{\mathbb{R}^d} g(cT, z-z_0) \mu(dz_0) \end{aligned}$$

en utilisant les deux relations suivantes :

$$\begin{aligned} p_L(0, T, z_N) &= \int_{(\mathbb{R}^d)^N} \prod_{i=1}^N p_L(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}-1} \mu(dz_0), \\ p(0, T, z_N) &= \int_{(\mathbb{R}^d)^N} \prod_{i=1}^N p(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}-1} \mu(dz_0) \end{aligned}$$

avec la notation $dz_{\mathbf{N}-1} = dz_1 \cdots dz_{N-1}$. Puis il suffit d'intégrer contre la fonction h pour avoir le biais de l'espérance :

$$\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^L)] = \int_{(\mathbb{R}^d)^2} h(z_N) (p(0, T, z_0, z_N) - p_L(0, T, z_0, z_N)) dz_N \mu(dz_0).$$

L'analyse de l'erreur statistique fait également partie du travail en cours. Quelques difficultés sont présentes, notamment lorsque nous remplaçons pour chaque instant sur la grille $k = 1, \dots, N$, $\mu_{t_k}^L$ par son estimateur empirique $\hat{\mu}^L(t_k)$, l'erreur s'accumule jusqu'à l'instant final $T = t_N$.

1.3 Représentation probabiliste et formules d'intégration par parties pour certains modèles à volatilité stochastique avec drift non-borné

Le modèle à volatilité stochastique qui nous intéresse dans la troisième partie est décrit par le couple $(S_t, Y_t)_{t \in [0, T]}$, où S_t représente le prix (spot) du sous-jacent à l'instant t et Y_t est la volatilité, la corrélation entre le spot et la volatilité est ρ :

$$\begin{aligned} dS_t &= rS_t dt + \sigma_S(Y_t) S_t dW_t \\ dY_t &= b_Y(Y_t) dt + \sigma_Y(Y_t) dB_t \\ d\langle B, W \rangle_t &= \rho dt, \rho \in (-1, 1) \end{aligned}$$

1.3. Représentation probabiliste et formules d'intégration par parties pour certains modèles à volatilité stochastique avec drift non-borné

Nous supposons que les coefficients de diffusion de $a_S = \sigma_S^2$, $a_Y = \sigma_Y^2$ et de drift b_Y sont infiniment différentiables à dérivées bornées, a_S et a_Y sont bornés mais b_Y n'est pas nécessairement borné. Typiquement, le drift correspond à un terme de retour à la moyenne, *i.e.* $b_Y(y) = \lambda(\mu - y)$, $\lambda > 0$ comme dans le modèle de Stein-Stein [SS91]. De plus, nous supposons l'uniforme ellipticité, *i.e.* $\exists \kappa \geq 1, \forall x \in \mathbb{R}, \kappa^{-1} \leq a_S(x) \leq \kappa, \kappa^{-1} \leq a_Y(x) \leq \kappa$.

Au lieu de considérer le prix spot, nous prenons plutôt son logarithme $X_t = \ln(S_t)$ et le couple qui nous intéresse dans la suite devient donc $(X_t, Y_t)_{t \in [0, T]}$ avec conditions initiales $X_0 = x_0, Y_0 = y_0 = \ln(s_0)$:

$$dX_t = \left(r - \frac{1}{2} a_S(Y_t) \right) dt + \sigma_S(Y_t) dW_t$$

1.3.1 Représentation probabiliste du couple spot-volatilité

La technique principale pour la représentation probabiliste de la loi marginale du processus de diffusion standard est la même que dans la partie précédente pour les EDS de McKean-Vlasov, toujours dans l'esprit de la méthode de simulation sans biais introduite dans [BKH+15]. La difficulté majeure par rapport aux travaux utilisant de cette méthode, par exemple comme dans [BKH+15] pour les processus de diffusion multi-dimensionnelle ou [FKHL+19] pour les processus de diffusion tués, nous devons prendre en compte dans ce modèle le fait que le drift b_Y du processus de volatilité n'est pas borné. Une application naïve de la méthodologie présentée dans les travaux mentionnés précédemment ne fonctionne pas. Pour cela, il faut changer le processus d'approximation autour duquel l'argument de perturbation est effectué. Introduisons la solution $(m_t(y_0))_{t \in [0, T]}$ de l'équation différentielle ordinaire $dm_t = b_Y(m_t) dt$, $m_0 = y_0$ obtenue en enlevant le bruit dans la dynamique de Y . Nous introduisons alors le processus d'approximation (\bar{X}, \bar{Y}) obtenu en gelant les coefficients de la dynamique du couple (X, Y) le long de la trajectoire de m :

$$\begin{aligned} \bar{X}_t^{x_0} &= x_0 + \int_0^t \left(r - \frac{1}{2} a_S(m_s) \right) ds + \int_0^t \sigma_S(m_s) dW_s, \\ \bar{Y}_t^{y_0} &= y_0 + \int_0^t b_Y(m_s) ds + \int_0^t \sigma_Y(m_s) dB_s. \end{aligned}$$

Comme dans la partie précédente, nous considérons une suite d'instants de sauts $\tau = (\tau_n)_{n \geq 1}$ et le processus de renouvellement associé $N = (N_t)_{t \geq 0}$, $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$, N est indépendant du mouvement Brownien bi-dimensionnel (W, B) . Le schéma d'Euler

pour (\bar{X}, \bar{Y}) est défini sur la grille de temps aléatoire $(\zeta_i)_{i \geq 0}$ avec $\zeta_i = \tau_i \wedge T$, $\zeta_0 = 0$:

$$\begin{aligned}\bar{X}_{i+1} &= \bar{X}_i + \left(r(\zeta_{i+1} - \zeta_i) - \frac{1}{2} a_{S,i} \right) + \sigma_{S,i} Z_{i+1}^1, \\ \bar{Y}_{i+1} &= m_i + \sigma_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right), \\ a_{S,i} &= \sigma_{S,i}^2 = \int_0^{\zeta_{i+1} - \zeta_i} a_S(m_s(\bar{Y}_i)) ds, \\ a_{Y,i} &= \sigma_{Y,i}^2 = \int_0^{\zeta_{i+1} - \zeta_i} a_Y(m_s(\bar{Y}_i)) ds, \\ \sigma_{S,Y,i} &= \int_0^{\zeta_{i+1} - \zeta_i} (\sigma_S \sigma_Y)(m_s(\bar{Y}_i)) ds, \quad \rho_i = \rho \frac{\sigma_{S,Y,i}}{\sigma_{S,i} \sigma_{Y,i}}.\end{aligned}$$

Nous utilisons la notation $m_i = m_{\zeta_{i+1} - \zeta_i}(\bar{Y}_i)$ pour la solution à l'instant $\zeta_{i+1} - \zeta_i$ de l'EDO avec condition initiale \bar{Y}_i , $Z = (Z_n^1, Z_n^2)_{n \geq 1}$ est une suite de variables aléatoires i.i.d. de lois $\mathcal{N}(0, I_2)$ et $\sigma'_{S,i}$, $\sigma'_{Y,i}$, $\sigma'_{S,Y,i}$, ρ'_i et m'_i sont des dérivées par rapport à \bar{Y}_i .

Le schéma d'Euler (\bar{X}, \bar{Y}) que nous venons de définir peut être vu comme une chaîne de Markov par rapport à la filtration $\mathcal{G} = (\mathcal{G}_i)_{i \geq 0}$ avec $\mathcal{G}_i = \sigma(Z_j, 0 \leq j \leq i)$. Nous développons un calcul de Malliavin adapté à la chaîne de Markov (\bar{X}, \bar{Y}) . Commençons par définir l'ensemble $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, $n \in \mathbb{N}$, $i \in \{0, \dots, n\}$ comme l'espace des variables aléatoires H vérifiant :

- $H = h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1})$ sur $\{N_T = n\}$, où $\zeta^{n+1} := (0 = \zeta_0, \zeta_1, \dots, \zeta_n, \zeta_{n+1} = T)$
- Pour tout $\sigma_{n+1} \in \Delta_{n+1}(T) := \{\sigma_{n+1} \in [0, T]^n : 0 \leq \sigma_1 < \dots < \sigma_{n+1} \leq T\}$, la fonction $h(\cdot, \sigma_{n+1}) \in \mathcal{C}_p^\infty(\mathbb{R}^4)$.

Définissons d'abord l'opérateur de dérivée $\mathcal{D}_{i+1}^{(\alpha)} H$, $\alpha \in \{1, 2\}$ pour $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$:

$$\mathcal{D}_{i+1}^{(1)} H = \partial_{\bar{X}_{i+1}} H \text{ et } \mathcal{D}_{i+1}^{(2)} H = \partial_{\bar{Y}_{i+1}} H.$$

L'opérateur d'intégrale satisfait l'identité de dualité suivante, avec la notation pour l'espérance conditionnelle $\mathbb{E}_{i,n}[\cdot] = \mathbb{E}[\cdot | \mathcal{G}_i, \tau^{n+1}, N_T = n]$, pour tout $\alpha \in \{1, 2\}$:

$$\mathbb{E}_{i,n} \left[\mathcal{D}_{i+1}^{(\alpha)} f(\bar{X}_{i+1}, \bar{Y}_{i+1}) H \right] = \mathbb{E}_{i,n} \left[f(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(\alpha)}(H) \right].$$

L'identité précédente se généralise au cas où α est un multi-indice $\alpha = (\alpha_1, \dots, \alpha_p)$, avec $\alpha_i \in \{1, 2\}$ en itérant l'opérateur de dérivée et d'intégrale. Pour un indice $\alpha_{p+1} \in \{1, 2\}$ nous avons la relation de récurrence :

$$\mathcal{I}_{i+1}^{(\alpha, \alpha_{p+1})}(H) = \mathcal{I}_{i+1}^{(\alpha_{p+1})}(\mathcal{I}_{i+1}^{(\alpha)}(H)), \quad \mathcal{D}_{i+1}^{(\alpha, \alpha_{p+1})} H = \mathcal{D}_{i+1}^{(\alpha_{p+1})}(\mathcal{D}_{i+1}^{(\alpha)} H).$$

Définissons pour $\gamma > 0$, $\mathcal{B}_\gamma(\mathbb{R}^2)$ l'ensemble de fonctions boréliennes sur \mathbb{R}^2 qui admettent une croissance exponentielle quadratique à l'infini :

$$\exists C > 0, \forall (x, y) \in \mathbb{R}^2, |h(x, y)| \leq C \exp(\gamma(|x|^2 + |y|^2)).$$

Résultat Principal 1.3.1. *Sous les hypothèses de régularité et d'uniforme ellipticité précédentes, la loi du couple (X_T, Y_T) à l'instant T vérifie une formule de représentation probabiliste, il existe $\gamma > 0$ tel que pour tout $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, l'espérance (ou le prix d'un produit qui paie $h(X_T, Y_T)$ à la maturité T), s'écrit comme suit :*

$$\mathbb{E}[h(X_T, Y_T)] = \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i\right]$$

où, sur l'ensemble $\{N_T = n\}$, les variables aléatoires $\theta_i \in \mathbb{S}_{i-1, n}(\bar{X}, \bar{Y})$, pour $i = 1, \dots, n$, sont définies par :

$$\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{I}_i^{(1)}(c_S^i) + \mathcal{I}_i^{(2,2)}(c_Y^i) + \mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{I}_i^{(1,2)}(c_{Y,S}^i) \right]$$

et pour $i = n + 1$, $\theta_{n+1} = (1 - F(T - \zeta_n))^{-1}$.

Si par exemple N est un processus de renouvellement avec temps de sauts de loi Beta(1/2, 1), alors la variable aléatoire dans l'espérance de droite $h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i$ admet un moment $L^p(\mathbb{P})$ fini pour tout $p \geq 1$.

1.3.2 Formules d'intégration par parties

Delta (Vega respectivement) est défini par la dérivée par rapport au prix spot (par rapport à la volatilité respectivement) du sous-jacent. Nous cherchons donc à donner des formules de Bismut-Elworthy-Li aux deux quantités suivantes qui peuvent être vues comme Delta et Vega à l'instant initial $t = 0$, avec $s_0 = \exp(X_0)$, $y_0 = Y_0$:

$$\partial_{s_0} \mathbb{E}[h(X_T, Y_T)] \text{ et } \partial_{y_0} \mathbb{E}[h(X_T, Y_T)]$$

L'objectif est donc d'écrire les dérivées d'espérances ci-dessus sous forme d'espérances de certaines fonctions. D'abord nous utilisons la représentation probabiliste pour l'espérance. L'étape suivante est donc d'effectuer le "transfert de la dérivée" à l'intérieur de l'espérance à l'aide d'une formule d'intégration par parties, mais pour des raisons d'intégrabilité, nous appliquons plutôt des formules d'IPP locales sur chacun des intervalles de temps aléatoires $[\zeta_i, \zeta_{i+1}]$, $i = 0, \dots, N_T$. Finalement, les dérivées de l'espérance, ou Delta et Vega, s'écriront comme des sommes pondérées des formules d'IPP locales, avec les poids égaux aux longueurs des intervalles aléatoires.

Résultat Principal 1.3.2. *Toujours sous les hypothèses de régularité et d'uniforme ellipticité, la loi du couple (X_T, Y_T) vérifient les formules de Bismut-Elworthy-Li suivantes. Il existe $\gamma > 0$ tel que pour tout $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ et pour tout $(s_0, y_0) \in \mathbb{R}^2$, les dérivées de l'espérance par rapport à s_0 et y_0 , ou Delta et Vega s'écrivent comme suit :*

$$s_0 T \partial_{s_0} \mathbb{E}[h(X_T, Y_T)] = \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \overrightarrow{\theta}^{\mathcal{I}_k^{(1), N_T+1}}\right],$$

et

$$\begin{aligned} & T \partial_{y_0} \mathbb{E}[h(X_T, Y_T)] \\ &= \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\overrightarrow{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \overrightarrow{\theta}^{C_j^{N_T+1}} + \overrightarrow{\theta}^{\mathcal{I}_k^{(1), N_T+1}} \right)\right]. \end{aligned}$$

où $\vec{\theta}^{\mathcal{I}_k^{(1),n+1}}$, $\vec{\theta}^{C_j^{n+1}}$, $\vec{\theta}^{\mathcal{I}_k^{(2),n+1}}$ et $\vec{\theta}^{\mathcal{I}_j^{(1),n+1}}$ avec $n \geq 0$ sur $\{N_T = n\}$, $k \in \{1, \dots, n+1\}$ et $1 \leq j \leq k$, sont des fonctions explicites des poids θ_i et des autres paramètres du modèle.

Si par exemple N est encore une fois un processus de renouvellement avec temps de sauts de loi $\text{Beta}(1/2, 1)$, alors les variables aléatoires dans les espérances de droite ci-dessus admettent des moments $L^p(\mathbb{P})$ finis pour tout $p \geq 1$.

L'intérêt majeur des deux formules d'intégration par partie précédente réside dans le fait qu'elles conduisent à une méthode de Monte-Carlo sans biais et donc de complexité optimale. En effet, les variables aléatoires apparaissant sous l'espérance du terme de droite est parfaitement simulable. Ceci fait contraste avec les formules d'intégration par partie que l'on peut obtenir à l'aide du calcul de Malliavin usuel c.f. [Nua06].

1.4 Méthodes numériques pour les EDSPRs issues des jeux à champ moyen

Dans cette dernière partie du manuscrit, nous nous intéressons aux solutions numériques des jeux à champ moyen, notamment dans le cadre des jeux linéaires-quadratiques. Il y a deux principales approches possibles, la première est analytique et nous utilisons des systèmes d'équation aux dérivées partielles. La seconde est celle étudiée ici, qui consiste à écrire le problème de contrôle stochastique de type champ moyen associé sous forme d'une équation différentielle stochastique progressive-rétrograde entièrement couplée de type McKean-Vlasov.

1.4.1 Jeux à champ moyen et approche probabiliste

Commençons par les jeux différentiels stochastiques à N joueurs dans un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$, où l'état X_t^i de chaque joueur $i \in \{1, \dots, N\}$ est donné par l'équation différentielle stochastique suivante :

$$dX_t^i = b^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i)dt + \sigma^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i)dW_t^i.$$

Chaque joueur choisit son contrôle $\alpha^i = (\alpha_t^i)_{0 \leq t \leq T}$ dans un ensemble de contrôles admissibles \mathbb{A} et $\bar{\mu}_t = \frac{1}{N} \sum_{j=1}^N \delta_{X_t^j} \in \mathcal{P}_2(\mathbb{R}^d)$ est la distribution empirique des états des joueurs où $\mathcal{P}_2(\mathbb{R}^d)$ est l'espace des mesures de probabilité avec moment de second ordre fini et W_2 est la distance de Wasserstein associée. Le drift b^i et la volatilité σ^i sont Lipschitziennes et continue par rapport à tous leurs arguments, notamment nous utilisons la distance de Wasserstein W_2 pour la variable mesure.

La fonction de coût J^i que le joueur i cherche à minimiser est définie comme la somme des coûts courant f^i et terminal g^i :

$$J^i(\alpha) = \mathbb{E} \left[\int_0^T f^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i)dt + g^i(X_T^i, \bar{\mu}_T) \right].$$

Les joueurs sont en *équilibre de Nash* si chaque joueur ne choisirait pas une autre stratégie s'il ou elle considère que les stratégies des autres joueurs sont fixées. L'ensemble des stratégies des joueurs est un équilibre de Nash si :

$$J^i(\alpha) \leq J^i(\alpha^1, \dots, \alpha^{i-1}, \alpha, \alpha^{i+1}, \dots, \alpha^N), \forall \alpha \in \mathbb{A}, \forall i \in \{1, \dots, N\}.$$

Calculer l'équilibre de Nash pour N joueurs revient en fait à résoudre un système à N équations, ce qui est compliqué d'un point de vue computationnel. L'idée des jeux à champ moyen est d'étudier le système quand N tend vers l'infini et nous devons supposer tous les joueurs identiques, *i.e.* $b = b^i$, $\sigma = \sigma^i$, $f = f^i$, $g = g^i$ et $\mathbb{A} = A$, c.f. [LL07] et [CD+18].

Au lieu de considérer le système à N joueurs, nous avons une continuité de joueurs et il suffit de regarder le joueur représentatif. Le passage à la limite est justifié par exemple dans [Szn91] dans le cadre de la propagation du chaos pour un système de particules. Dans ce cas, l'équilibre de Nash peut être vu comme un point fixe, qui est résolu en deux étapes :

1. Pour un flot fixé de mesures de probabilité $\mu = (\mu_t)_{0 \leq t \leq T} \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$, nous résolvons le problème de contrôle stochastique classique :

$$\inf_{\alpha \in A} J^\mu(\alpha) = \mathbb{E} \left[\int_0^T f(t, X_t^\alpha, \mu_t, \alpha_t) dt + g(X_T^\alpha, \mu_T) \right] \quad (1.4.1)$$

où la dynamique de $(X_t^\alpha)_{0 \leq t \leq T}$ est donnée par

$$\begin{aligned} dX_t^\alpha &= b(t, X_t^\alpha, \mu_t, \alpha_t) dt + \sigma(t, X_t^\alpha, \mu_t, \alpha_t) dW_t, \\ X_0^\alpha &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d). \end{aligned}$$

2. Trouver un point fixe, μ tel que $\mathcal{L}(X_t^\alpha) = \mu_t$ pour tout $0 \leq t \leq T$.

Ce problème peut être résolu par une approche analytique, c'est-à-dire avec des équations aux dérivées partielles de Jacobi-Bellman et de Fokker-Planck. Ici nous nous concentrons sur les deux approches probabilistes qui nous permettront d'écrire des équations différentielles stochastiques progressives-rétrogrades entièrement couplées de type champ moyen (EDSPRs ou FBSDEs).

La première approche est appelée *approche faible*, pour un flot de mesures fixé $\mu = (\mu_t)_{0 \leq t \leq T}$, nous définissons la fonction valeur u :

$$u(t, x) = \inf_{(\alpha_s)_{t \leq s \leq T} \in A} \mathbb{E} \left[\int_t^T f(s, X_s, \mu_s, \alpha_s) ds + g(X_T, \mu_T) \mid X_t = x \right]$$

Le processus $(Y_t)_{0 \leq t \leq T}$, $Y_t = u(t, X_t)$ est défini comme la fonction valeur le long du processus d'état $(X_t)_{0 \leq t \leq T}$. Pour simplifier, supposons que la volatilité $\sigma > 0$ est constante. Sous des conditions de régularité et de bornitude, le couple $(X_t, Y_t)_{0 \leq t \leq T}$ vérifie l'EDSPR suivante :

$$\begin{aligned} dX_t &= b(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1} Z_t)) dt + \sigma dW_t, \quad X_0 = \xi \\ dY_t &= -f(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1} Z_t)) dt + Z_t dW_t, \quad Y_T = g(X_T, \mu_T) \end{aligned}$$

La fonction $\hat{\alpha}$ est l'unique minimiseur de l'Hamiltonien H :

$$\begin{aligned} H(t, x, \mu, \alpha, y) &= b(t, x, \mu, \alpha) \cdot y + f(t, x, \mu, \alpha) \\ \hat{\alpha}(t, x, \mu, y) &= \arg \inf_{\alpha \in A} H(t, x, \mu, \alpha, y) \end{aligned}$$

Selon [CD+18], le flot de mesures $\mu = (\mu_t)_{0 \leq t \leq T}$ est un équilibre de Nash du jeu à champ moyen si et seulement si μ correspond à la loi de X , *i.e.* $\mu_t = \mathcal{L}(X_t), \forall t \in [0, T]$. Nous obtenons finalement l'EDSPR de l'approche faible en $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$:

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1}Z_t)) dt + \sigma dW_t, \quad X_0 = \xi, \\ dY_t &= -f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1}Z_t)) dt + Z_t dW_t, \quad Y_T = g(X_T, \mathcal{L}(X_T)). \end{aligned}$$

La seconde approche, appelée *approche de Pontryagin*, est basée sur le principe du maximum stochastique de Pontryagin. Pour le processus Y , à la place de la fonction valeur u , nous prenons sa dérivée par rapport à la variable spatiale, *i.e.* $Y_t = \nabla_x u(t, X_t)$. Les hypothèses sur b, f et g sont plus fortes, notamment la fonction du coût courant f est continûment différentiable avec des dérivées Lipschitz et bornées, de plus f est strictement convexe, pour tout $(t, x, x', \mu, \alpha, \alpha') \in [0, T] \times (\mathbb{R}^d)^2 \times \mathcal{P}_2(\mathbb{R}^d) \times A^2$:

$$f(t, x', \mu, \alpha') - f(t, x, \mu, \alpha) - (x - x', \alpha - \alpha') \cdot \partial_{(x, \alpha)} f(t, x, \mu, \alpha) \geq C_1 |\alpha' - \alpha|^2.$$

De façon similaire, l'EDSPR de l'approche de Pontryagin avec solution (X, Y, Z) s'écrit :

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t)) dt + \sigma dW_t, \quad X_0 = \xi, \\ dY_t &= -[\nabla_x b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t))] \cdot Y_t \\ &\quad + \nabla_x f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t))] dt + Z_t dW_t, \quad Y_T = \nabla_x g(X_T, \mathcal{L}(X_T)). \end{aligned}$$

Un avantage de l'approche de Pontryagin est que le processus Y est plus étroitement lié au contrôle optimal, notamment dans le cas des jeux linéaires-quadratiques que nous verrons dans les exmples numériques.

Le système général d'EDSPR entièrement couplée dans les deux approches peut s'écrire comme suit, avec $[X]$ représentant la loi de la variable aléatoire X et nous nous plaçons dans le cas uni-dimensionnel pour simplifier :

$$\begin{aligned} dX_t &= B(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + \sigma dW_t, \quad X_0 = \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}), \\ dY_t &= -F(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + Z_t dW_t, \quad Y_T = G(X_T, [X_T]). \end{aligned}$$

1.4.2 Algorithmes d'arbre et de grille pour les EDSPRs

Nous définissons le schéma de Picard découplé $\Phi_{\xi, G}$, qui à l'itéré de Picard $j-1$ associe l'itéré suivant j :

$$\Phi_{\xi, G} : (X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}]) \mapsto (X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$$

où ξ et G correspondent aux données initiale et finale du problème respectivement. En fait $\Phi_{\xi, G}$ est naturellement définie à partir de l'EDSPR générale :

1. Résoudre pour $(X^j, [X^j])$:

$$\begin{aligned} dX_t^j &= B(t, X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}, [X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}]) dt + \sigma dW_t \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}) \end{aligned}$$

2. Ensuite résoudre pour $(Y^j, Z^j, [Y^j], [Z^j])$:

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^{j-1}, Z_t^{j-1}, [X_t^j, Y_t^{j-1}, Z_t^{j-1}]) dt + Z_t^j dW_t \\ Y_T^j &= G(X_T^j, [X_T^j]) \end{aligned}$$

3. Retourner $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$.

Avec l'initialisation par $(X^0, Y^0, Z^0, [X^0, Y^0, Z^0])$, il est clair que $\Phi_{\xi, G}$ définit une suite $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])_{j \geq 0}$. Si cette suite converge vers $(X, Y, Z, [X, Y, Z])$, c'est donc un point fixe du schéma $\Phi_{\xi, G}$ et $(X, Y, Z, [X, Y, Z])$ résout donc l'EDSPR général. Le problème est que, seulement pour des petites valeurs d'horizon de temps T , le schéma $\Phi_{\xi, G}$ est contractant, ce qui nous assurerait l'existence d'un point fixe.

L'idée de *continuation en temps*, proposée dans [CCD+19], consiste à découper l'intervalle de temps en des petits intervalles, appelés des levels, que nous supposons uniformes avec $h_\ell = T/N_\ell$:

$$\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}$$

Ensuite nous faisons des itérations de Picard de façon récursive entre les levels. Commençons par définir le schéma *picard* $[k](X, Y_{T_{k+1}})$ pour level $[T_k, T_{k+1}]$:

1. Initialiser $Y_t = 0$ et $Z_t = 0, \forall t \in [T_k, T_{k+1})$.
2. Pour $1 \leq j \leq J_p$ (nombre d'itérations de Picard) :

$$(X, Y, Z, [X, Y, Z]) = \Phi_{X_{T_k}, Y_{T_{k+1}}}(X, Y, Z, [X, Y, Z])$$

3. Retourner $(X, Y, Z, [X, Y, Z])$.

Le solveur global, noté *solver*, applique la continuation en temps et est défini de façon récursive et rétrograde par un entier $J_s > 0$. Nous avons pour chaque level k :

1. Initialiser $X_t = X_{T_k}, Y_t = 0$ et $Z_t = 0, \forall t \in [T_k, T_{k+1})$.
2. Pour $1 \leq j \leq J_s$
 - (a) $(Y_{T_{k+1}}, [Y_{T_{k+1}}]) = \text{solver}[k+1](X_{T_{k+1}}, [X_{T_{k+1}}])$
 - (b) $(X, Y, Z, [X, Y, Z]) = \text{picard}[k](X, Y_{T_{k+1}})$
3. Retourner $(Y_{T_k}, [Y_{T_k}])$.

Avant de présenter les algorithmes, écrivons la discrétisation des EDS dans le cadre du schéma de Picard $\Phi_{\xi, G}$ sur une grille de temps $\{0 = t_0 < t_1, \dots, t_i = ih, \dots, t_{N_t-1} < t_{N_t} = T\}$ avec $h = T/N_t$ qui est un raffinement de la grille correspondant aux levels $\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}$:

$$\begin{cases} X_{t_{i+1}}^j = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) + \sigma \Delta W_i, & X_0^j = \xi \\ Y_{t_i}^j = \mathbb{E}_{t_i}(Y_{t_{i+1}}^j) + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]), & Y_T^j = G(X_T^{j-1}, [X_T^{j-1}]) \\ Z_{t_i}^j = h^{-1} \mathbb{E}_{t_i}(Y_{t_{i+1}}^j \Delta W_i), & Z_T^j = 0. \end{cases}$$

Remarquons que ce schéma discrétisé est découplé, contrairement à l'EDSPR originale. Nous pouvons donc résoudre le schéma rétrograde puis le schéma progressif pour obtenir $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$ à partir de $(X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}])$ où $X^{j-1} = (X_{t_i}^{j-1})_{0 \leq i \leq N_t}$.

Notre premier algorithme, *l'algorithme d'arbre* a été présenté dans [CCD+19]. La structure d'arbre binaire est obtenue par l'approximation binaire des incréments du mouvement Brownien : $\Delta W_i = \pm\sqrt{h}$ avec probabilité 1/2. Chaque noeud à la profondeur i représente une valeur de X_{t_i} et a deux enfants qui correspondent aux valeurs de $X_{t_{i+1}}$ ("haut \uparrow " et "bas \downarrow ") qui sont calculées comme suit :

$$X_{t_{i+1}}^j(\uparrow\downarrow) = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j, [X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j]) \pm \sigma\sqrt{h}.$$

Si nous utilisons M points pour l'approximation de la loi initiale ξ du processus X , nous aurons M arbres binaires parallèles. Pour connaître la loi marginale de X à l'instant t_i , il suffit de regarder les valeurs des noeuds à la profondeur i et il y en a $M \times 2^i$. La complexité de l'algorithme d'arbre est donc exponentielle en le nombre de pas de temps.

Pour le schéma rétrograde, nous réutilisons l'arbre binaire, à l'instant final $T = t_{N_t}$, nous avons $Y_T^j = G(X_T^j, [X_T^j])$ pour chacun des $M \times 2^{N_t}$ noeuds. Puis l'espérance conditionnelle correspond simplement à la moyenne des branches "hautes" et "basses". L'itération de Picard est initialisée en $j = 0$ par $X_{t_i} = \xi, \forall i \in \{0, \dots, N_t\}$.

La contribution principale de cette partie est *l'algorithme de grille* sur les lois marginales, inspiré de [DM+06], nous utilisons la notion du champ de découplage. En fait il existe des fonctions feedback déterministes $(u, v) : [0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}$, u étant solution d'une certaine EDP, voir par exemple [CCD14], telles que la solution (X, Y, Z) du système général d'EDSPR satisfait :

$$Y_t = u(t, X_t, [X_t]) \quad \text{et} \quad Z_t = v(t, X_t, [X_t])$$

La loi de la solution $([X_t, Y_t, Z_t])_{0 \leq t \leq T}$ peut donc être caractérisée par $(\mu_t, u(t, \cdot), v(t, \cdot))_{0 \leq t \leq T}$ avec $[X_t] = \mu_t$. Nous nous concentrons donc sur l'approximation de μ et (u, v) . Dans cet algorithme de grille, nous calculons donc seulement les lois marginales à chaque pas de temps, à la place de lois trajectorielles comme précédemment.

Notre schéma de Picard découplé $\Psi_{[\xi], G}$ est ensuite redéfini par :

$$\Psi_{[\xi], G} : (\mu_t^{j-1}, u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))_{0 \leq t \leq T} \mapsto (\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T}, \quad \mu_0^{j-1} = \mu_0^j = [\xi]$$

1. Résoudre pour X^j l'EDS suivante et définir $\mu_t^j := [X_t^j]$:

$$\begin{aligned} dX_t^j &= B(t, X_t^j, u^{j-1}(t, X_t^j), v^{j-1}(t, X_t^j), \mu_t^{j-1}, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))) \# \mu_t^{j-1} dt + \sigma dW_t \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}) \end{aligned}$$

Avec $\varphi \# \nu$ la mesure push-forward de la mesure ν par une fonction φ .

2. Trouver $(u^j, v^j)(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^2$ par $(u^j, v^j)(t, X_t^j) := (Y_t^j, Z_t^j)$ en résolvant :

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^j, Z_t^j, \mu_t^j, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))) \# \mu_t^j dt + Z_t^j dW_t, \\ Y_T^j &= G(X_T^j, \mu_T^j). \end{aligned}$$

3. Retourner $((\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T})$.

Pour rendre notre algorithme simulable (pouvoir calculer (μ_t^j, u^j, v^j) explicitement), nous avons recours à une grille de discrétisation spatiale pour le processus progressif X . Par exemple une grille uniforme Γ de pas Δx , avec la fonction de projection Π associée :

$$\Gamma = \{x_k = x_1 + (k - 1)\Delta x, k = 1, \dots, N_x\}.$$

Le processus X est défini sur la grille Γ :

$$\begin{aligned} X_{t_{i+1}}^j &= \Pi \left(X_{t_i}^j + h B(X_{t_i}^j, (u_i^{j-1}, v_i^{j-1})(X_{t_i}^j), \mu_i^{j-1}, (u_i^{j-1}, v_i^{j-1})\sharp\mu_i^{j-1}) + \sigma \Delta W_i \right) \\ [X_{t_i}^j] &= \mu_i^j(\cdot) \text{ et } \mu_{t_{i+1}}^j(\cdot) = [X_{t_{i+1}}^j]. \end{aligned}$$

La loi marginale du processus X à l'instant t_i de l'itéré de Picard j est défini par une moyenne pondérée des mesures de Dirac sur Γ :

$$\mu_i^j(\cdot) = \sum_{k=1}^{N_x} p_{i,k}^j \delta_{x_k}(\cdot), p_{i,k}^j \geq 0 \forall k \in \{1, \dots, N_x\} \text{ et } \sum_{k=1}^{N_x} p_{i,k}^j = 1.$$

La loi marginale $\mu_{t_{i+1}}^j(\cdot)$, ou plutôt les poids $(p_{i+1,n}^j, n \in \{1, \dots, N_x\})$ à l'instant t_{i+1} est la convolution de $\mu_i^j(\cdot)$ avec la probabilité de transition $\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k), k, n \in \{1, \dots, N_x\}$:

$$p_{i+1,n}^j = \sum_{k=1}^{N_x} p_{i,k}^j \times \mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k).$$

Pour connaître explicitement les probabilités de transition, par exemple nous utilisons l'approximation binaire pour les incréments Browniens, même s'il est aisé d'utiliser d'autres approximations avec plus de points :

$$\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k) = \frac{1}{2} \left(\mathbf{1}(X_{t_{i+1}}^j(\uparrow) = x_n | X_{t_i}^j = x_k) + \mathbf{1}(X_{t_{i+1}}^j(\downarrow) = x_n | X_{t_i}^j = x_k) \right).$$

Le schéma progressif nous donne donc pour chaque itéré de Picard j le flot de mesures $(\mu_i^j)_{i=0}^{N_t}$ aux instants $(t_i)_{i=0}^{N_t}$. D'autre part, le schéma rétrograde donne $u_i^j(x)$ et $v_i^j(x)$ pour $x \in \Gamma$ sur la grille. Avec la condition terminale $T = t_{N_t}, (u_{N_t}^j, v_{N_t}^j) = (G, 0)$, nous avons pour $i \leq N_t - 1$:

$$\begin{aligned} u_i^j(x) &= \mathbb{E} \left(u_{t_{i+1}}^j(X_{t_{i+1}}^j, \mu_{t_{i+1}}^j) + h \cdot F(X_{t_i}^j, u_i^{j-1}(X_{t_i}^j), v_i^j(X_{t_i}^j), \mu_i^j, (u_i^{j-1}, v_i^{j-1})\sharp\mu_i^j) | X_{t_i}^j = x \right), \\ v_i^j(x) &= \mathbb{E} \left(u_{t_{i+1}}^j(X_{t_{i+1}}^j) \cdot \Delta W_i / h | X_{t_i}^j = x \right). \end{aligned}$$

La variable $X_{t_{i+1}}^j$ à valeurs dans Γ et sa loi $[X_{t_{i+1}}^j] = \mu_{t_{i+1}}^j$ à support dans Γ , sont calculées par le schéma progressif partant de $X_{t_i}^j = x \in \Gamma$. Nous avons une expression simplifiée en utilisant encore une fois, l'approximation binaire pour ΔW_i .

Finalement, en remplaçant le schéma de Picard découpé précédent $\Phi_{\xi, G}$ dans l'algorithme d'arbre, par le schéma progressif-rétrograde pour les lois marginales sur la grille $\Psi_{[\xi], G}$ que nous venons de définir, nous pouvons appliquer la continuation en temps avec des schémas similaires à *picard[k]* et *solver* pour compléter l'implémentation numérique.

Chapter 2

Introduction - English

In this manuscript we study four problems from financial mathematics and numerical probabilities. We start by introducing the main issues and motivations then we summarize the main results obtained and the open questions that could be part of future research.

2.1 Modeling of relative capitalizations of assets

In the classical portfolio theory, the market portfolio made up of n assets is modeled by the process $(\mathbf{S}_t)_{t \geq 0}$, $\mathbf{S}_t = (S_t^i)_{i=1}^n$ with S_t^i representing the market capitalization of the i -th asset at time t , whose dynamics are given by a semimartingale in the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The total capitalization of the market $(\Sigma_t)_{t \geq 0}$ is the sum of capitalizations of assets S_t^i , $i = 1, \dots, n$. To study the properties of the market and the investment strategies relative to the market in which we are investing, instead of considering the real capitalizations, we also have the possibility to describe the market by the relative capitalizations of assets $(\boldsymbol{\mu}_t)_{t \geq 0}$, $\boldsymbol{\mu}_t = (\mu_t^i)_{i=1}^n$, $\mu_t^i = S_t^i / \Sigma_t$ defined by their market shares, or relative market weights. The process of total capitalization is a stochastic process that could be correlated to the process of relative weights. The couple $(\boldsymbol{\mu}_t, \Sigma_t)_{t \geq 0}$ then characterizes the market portfolio and we often do not need the knowledge of the dynamics of total market capitalization since we always invest in this same market and we use total capitalization as the numéraire.

2.1.1 Process and distribution of market weights

We model the process of relative market weights by introducing the *market weights equation* of exponent $\zeta > 0$ ($MWE(\zeta)$), which is inspired by the discrete probability distribution in the problem of nonlinear optimal filtering, as well as from the dynamics

of relative weights in the volatility-stabilized market model:

$$d\mu_t^i = b^i(\boldsymbol{\mu}_t)dt + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t)dW_t^\nu, \quad i \in \{1, \dots, n\}; \quad \boldsymbol{\mu}_0 \in \Delta_+^n$$

$$b^i(\boldsymbol{\mu}) = (B \cdot \boldsymbol{\mu})^i = \sum_{j=1}^n B^{i,j} \mu^j, \quad \sigma^{i,\nu}(\boldsymbol{\mu}) = (\mu^i)^\zeta \zeta^{i,\nu} - \mu^i \sum_{k=1}^n (\mu^k)^\zeta \zeta^{k,\nu}$$

Here the matrix B of dimension $n \times n$ is the transpose of a transition rate matrix:

$$B^{i,j} \geq 0 \quad \forall i \neq j \quad \text{and} \quad \sum_{i=1}^n B^{i,j} = 0 \quad \forall j; \quad i, j \in \{1, \dots, n\},$$

and the matrix ζ is of dimension $n \times d$, of rank n with $n \leq d$, in particular:

$$\forall i \in \{1, \dots, n\}, \quad c^i = \sum_{\nu=1}^d (\zeta^{i,\nu})^2 > 0$$

The covariance matrix between the weights $a(\cdot) = \sigma \sigma^\top(\cdot) = [a^{i,j}(\cdot)]$ is defined as:

$$a^{i,j}(\boldsymbol{\mu}_t) = \frac{d}{dt} \langle \mu^i, \mu^j \rangle_t = \sum_{\nu=1}^d \sigma^{i,\nu} \sigma^{j,\nu}(\boldsymbol{\mu}_t)$$

Denote Δ_+^n (resp. Δ^n) for the unit simplex of dimension n (resp. its closure) and its boundary $\partial\Delta^n = \Delta^n \setminus \Delta_+^n$:

$$\Delta_+^n \text{ (resp. } \Delta^n) = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^n \mid \mu^i > 0 \text{ (resp. } \mu^i \geq 0) \forall i \in \{1, \dots, n\} \text{ and } \sum_{i=1}^n \mu^i = 1 \right\}$$

Main Result 2.1.1. *For all $\zeta > 0$ and starting from any positive initial weights i.e. $\boldsymbol{\mu}_0 \in \Delta_+^n$, there exists a unique strong solution $(\boldsymbol{\mu}_t)_{t \geq 0}$ to the market weights equation, satisfying $\boldsymbol{\mu}_t \in \Delta^n$ for $t \geq 0$ almost surely.*

There exists a unique strong solution to market weights equation and none of the market weights vanishes over any finite time horizon $[0, T]$, $T > 0$, i.e. $\forall \boldsymbol{\mu}_0 \in \Delta_+^n$, the unique strong solution $(\boldsymbol{\mu}_t)_{t \geq 0}$ satisfies $\forall t \geq 0$, $\boldsymbol{\mu}_t \in \Delta_+^n$ a.s., if and only if one of the following conditions holds:

- (i) $\zeta \geq 1$;
- (ii) $\zeta \in (1/2, 1)$ and $\forall i \in \{1, \dots, n\}$, $b^i(\boldsymbol{\mu}) > 0$ on $\{\mu^i = 0\}$;
- (iii) $\zeta = 1/2$ and $\forall i \in \{1, \dots, n\}$, $2b^i(\boldsymbol{\mu}) \geq c^i > 0$ on $\{\mu^i = 0\}$.

Otherwise, if one of the following conditions is satisfied, there exists a point $\forall \boldsymbol{\mu}_ \in \Delta_+^n$ and a constant $\epsilon > 0$, if the initial weights satisfy $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \epsilon$, the boundary $\partial\Delta^n$ is attained with positive probability over any finite time horizon $[0, T]$, $T > 0$, define the stopping time $\tau := \inf\{t \geq 0 : \boldsymbol{\mu}_t \in \partial\Delta^n\}$ we have $\mathbb{P}[\tau \leq T] > 0$. More precisely there exists an index $i \in \{1, \dots, n\}$ such that $\mathbb{P}[\exists t \in (0, T) \mid \mu_t^i = 0] > 0$ and the strong solution exists and is unique until the stopping time $\tau \wedge T$:*

(iv) $\zeta \in (0, 1/2)$;

(v) $\zeta = 1/2$ and $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $2b^i(\boldsymbol{\mu}_*) < c^i$;

(vi) $\zeta \in (1/2, 1)$ and i.e. $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $b^i(\boldsymbol{\mu}_*) = 0$.

The exponent $\zeta > 0$ is a model parameter to be chosen or estimated. When $\zeta = 1$ we recognize the dynamics of the discrete probability distribution in the optimal non-linear filtering problem. It is also possible to recognize the dynamics of weights in the volatility-stabilized market model as a special case. This model and its generalizations have been widely studied, see [FK05] and [Pic14], as well as its process of weights, also called Jacobi or Wright-Fisher process, see [GJ06] and [Pal+11]. In fact to get back to the classical volatility-stabilized model, we just choose in our model $\zeta = 1/2$, the matrix B satisfying $B^{i,j} = \alpha \xi^i > 0$ for $i \neq j$ and $B^{i,i} = -\alpha \sum_{k=1, k \neq i}^n \xi^k < 0$, $\alpha > 0$ with mean reversion level $\boldsymbol{\xi} = (\xi^i)_{i=1}^n \in \Delta_+^n$ and the matrix ς diagonal, $d = n$ and $\varsigma = \sqrt{c}I^n$, $c > 0$.

Our model has the same drift coefficient as polynomial models studied in [Cuc19], but the volatility structure is different, the choice of the exponent ζ could provide more flexibility to characterise the covariance structure, as opposed to polynomial models where the exponent is $1/2$ as in volatility-stabilized market models.

As we are now certain that the weights will stay in the unit simplex, it is legitimate to consider the *reduced form* of market weights equation by simply taking the weights of the $(n - 1)$ first assets. From now on, the unit simplex and its closure are regarded as equivalent subsets of \mathbb{R}^{n-1} and the covariance matrix $a(\cdot)$ is of dimension $(n - 1)$:

$$\Delta_+^n \text{ (resp. } \Delta^n) = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^{n-1} \mid \mu^i > 0 \text{ (resp. } \mu^i \geq 0) \forall i \text{ and } \sum_{i=1}^{n-1} \mu^i \leq 1 \text{ (resp. } < 1) \right\}$$

Main Result 2.1.2. *The reduced covariance matrix $a(\boldsymbol{\mu})$ is positive definite if and only if $\boldsymbol{\mu} \in \Delta_+^n$, which can be also written as:*

$$\begin{aligned} & \forall \boldsymbol{\mu} \in \Delta_+^n; \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} > 0 \\ & \forall \boldsymbol{\mu} \in \partial \Delta^n = \Delta^n \setminus \Delta_+^n; \exists \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0 \end{aligned}$$

Equivalently, the reduced volatility matrix $\sigma(\boldsymbol{\mu})$ of dimension $(n - 1) \times d$ is of full rank if and only if $\boldsymbol{\mu} \in \Delta_+^n$. For all $\boldsymbol{\mu} \in \partial \Delta^n$, the volatility and covariance matrices have the same rank, equal to $n - 1 - m(\boldsymbol{\mu})$, where $m(\boldsymbol{\mu})$ is the number of indices $i \in \{1, \dots, n\}$ such that $\mu^i = 0$.

Define for $\epsilon \in (0, 1/n)$, the subset Δ_ϵ^n of unit simplex only containing weights with values greater than ϵ :

$$\Delta_\epsilon^n = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^{n-1} \mid \mu^i \geq \epsilon \forall i \in \{1, \dots, n - 1\} \text{ and } \sum_{i=1}^n \mu^i \leq 1 - \epsilon \right\}$$

We have for this subset:

$$\exists \lambda_\epsilon > 0, \forall \boldsymbol{\mu} \in \Delta_\epsilon^n; \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} \geq \lambda_\epsilon |\mathbf{x}|^2$$

Therefore, the market weights process defined given by market weights equation does not have uniform ellipticity, but only local ellipticity on the sets Δ_ϵ^n , $\epsilon \in (0, 1/n)$.

The distribution of market weights, seen as an n -dimensional distribution, could not be absolutely continuous with respect to Lebesgue measure in \mathbb{R}^n since the weights take value in the n -dimensional unit simplex, whose Lebesgue measure is zero in \mathbb{R}^n so we need to consider the reduced form of weights that are \mathbb{R}^{n-1} -valued. When the volatility $\sigma(\cdot)$ is polynomial, *i.e.* the exponent ζ is a positive integer, the drift and the volatility are both bounded and infinitely differentiable functions defined on \mathbb{R}^{n-1} . By checking Hörmander's condition, see for example [Nua06], with positive initial weights, *i.e.* $\boldsymbol{\mu}_0 \in \Delta_+^n \subset \mathbb{R}^{n-1}$, for all $t > 0$, the distribution $m_t(\cdot|\boldsymbol{\mu}_0)$ is absolutely continuous with respect to Lebesgue measure in \mathbb{R}^{n-1} and the associated probability density function $p_t(\cdot|\boldsymbol{\mu}_0)$ exists and is infinitely differentiable with respect to its spatial variable:

$$\begin{aligned} m_t(d\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) &= p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) \cdot d\boldsymbol{\mu}_t \\ p_t(\cdot|\boldsymbol{\mu}_0) &\in \mathcal{C}^\infty(\mathbb{R}^{n-1}, \mathbb{R}_+) \end{aligned}$$

A distribution $m(\cdot)$ of support Δ^n is *stationary* (is an invariant measure) with respect to the process $(\boldsymbol{\mu}_t)_{t \geq 0}$ if for all continuous and bounded function f and all $t > 0$:

$$\langle \mathcal{P}_t f, m \rangle = \int_{\Delta^n} \left(\int_{\Delta^n} f(\boldsymbol{\mu}_t) p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) d\boldsymbol{\mu}_t \right) f(\boldsymbol{\mu}_0) m(d\boldsymbol{\mu}_0) = \langle f, m \rangle = \int_{\Delta^n} f(\boldsymbol{\mu}_0) m(d\boldsymbol{\mu}_0)$$

Main Result 2.1.3. *Starting from $\boldsymbol{\mu}_0 \in \Delta_+^n$ and with polynomial volatility, *i.e.* $\zeta \in \mathbb{N}^*$, the stationary distribution of market weights, noted $m_\infty(\cdot)$ exists, is unique and is absolutely continuous with respect to Lebesgue measure in \mathbb{R}^{n-1} .*

Furthermore, with the notation d_{TV} for the total variation distance between two measures, the convergence speed of the distribution of market weights $m_t(\cdot|\boldsymbol{\mu}_0)$ to the unique stationary distribution $m_\infty(\cdot)$ when $t \uparrow \infty$ is subgeometric, that is for all $\gamma > 0$, there exists a constant $C(\gamma) > 0$ independent of t and of $\boldsymbol{\mu}_0$ such that:

$$\begin{aligned} d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_\infty(\cdot)) &\leq C(\gamma) \cdot (V(\boldsymbol{\mu}_0)/t^{\gamma+1} + 1/t^\gamma) \\ V(\boldsymbol{\mu}) &= - \sum_{i=1}^{n-1} \log(\mu^i) - \log \left(1 - \sum_{j=1}^{n-1} \mu^j \right) = - \sum_{i=1}^n \log(\mu^i) \end{aligned}$$

Suppose that there exists a nonnegative solution $p(\cdot)$ defined on Δ_+^n to Fokker-Planck equation with temporal derivative equal to zero, with conditions at the boundary $p(\boldsymbol{\mu}) = 0 \forall \boldsymbol{\mu} \in \partial\Delta^n$:

$$\mathcal{L}^* p = \sum_{i=1}^{n-1} \partial_i [b^i(\boldsymbol{\mu}) p(\boldsymbol{\mu})] - \frac{1}{2} \sum_{i,j=1}^{n-1} \partial_{ij}^2 [a^{i,j}(\boldsymbol{\mu}) p(\boldsymbol{\mu})] = 0$$

If additionally the integral of $p(\cdot)$ on Δ_+^n is equal to 1, meaning that $p(\cdot)$ is a probability density function, then the measure defined by $m(\boldsymbol{\mu}) = p(\boldsymbol{\mu}) d\boldsymbol{\mu}$ is stationary.

This equation can be solved analytically in some cases, especially when the number of assets n is equal to 2. For example for the volatility-stabilized market model, the solution corresponds to the Dirichlet distribution.

2.1.2 Structure of the market and relative arbitrage

It is possible to describe the complete structure of the market by using the couple $(\boldsymbol{\mu}_t, \Sigma_t)$ and introducing the process of total market capitalization, or the size of the market portfolio $(\Sigma_t)_{t \geq 0}$, which could be correlated to the market weights process through the volatility coefficients $(\sigma^{\Sigma, \nu}(\cdot))_{\nu=1}^d$:

$$d\Sigma_t = b^\Sigma(\Sigma_t)dt + \sum_{\nu=0}^d \sigma^{\Sigma, \nu}(\Sigma_t) dW_t^\nu, \quad \Sigma_0 > 0$$

The capitalization of the i -th asset is simply $S_t^i = \mu_t^i \cdot \Sigma_t$. By using logarithm representation, the *decomposition of growth rate* is then obtained:

$$d \log S_t^i = d \log \mu_t^i + d \log \Sigma_t = (\gamma^i(\boldsymbol{\mu}) + \gamma^\Sigma(\Sigma))dt + \text{martingale}$$

The growth rate is decomposed into $\gamma^i(\cdot)$, the part associated to the weight of the i -th asset and into $\gamma^\Sigma(\cdot)$ which is the growth of total market capitalization common to all assets.

By taking the total capitalization equal to constant, we obtain the dynamics of the logarithm of market weights of the i -th asset:

$$d \log \mu_t^i = \gamma^i(\boldsymbol{\mu})dt + \sum_{\nu=1}^d (\mu_t^i)^{-1} \sigma^{i, \nu}(\boldsymbol{\mu}_t) dW_t^\nu$$

Let us have a closer look at the volatility term and define the total volatility $\sigma^i(\boldsymbol{\mu}_t)$ of the i -th asset:

$$\begin{aligned} \sigma^i(\boldsymbol{\mu}_t)^2 &= \sum_{\nu=1}^d (\mu_t^i)^{-2} \sigma^{i, \nu}(\boldsymbol{\mu}_t)^2 = (\mu_t^i)^{2(\zeta-1)} (1 - 2\mu_t^i) \sum_{\nu=1}^d (\zeta^{i, \nu})^2 + \sum_{\nu=1}^d \left(\sum_{k=1}^n (\mu^k)^\zeta \zeta_{\zeta^{k, \nu}} \right)^2 \\ &\quad - 2(\mu_t^i)^{(\zeta-1)} \sum_{\nu=1}^d \zeta^{i, \nu} \sum_{k=1, k \neq i}^n (\mu^k)^\zeta \zeta_{\zeta^{k, \nu}} \end{aligned}$$

Recall that the rank of the matrix ζ is n so $c^i = \sum_{\nu=1}^d (\zeta^{i, \nu})^2 > 0$. For $\zeta \in (0, 1)$, when the weight μ_t^i of the i -th weight is small, the first term after the last equality $(\mu_t^i)^{2(\zeta-1)} (1 - 2\mu_t^i) c^i$ is the dominant one and $\sigma^i(\boldsymbol{\mu}_t) \rightarrow \infty$ when $\mu_t^i \rightarrow 0$, assets with smaller weights tend to have greater volatilities as in volatility-stabilized market models [Pic14], *stabilization by volatility* is then *asymptotic*. The smaller the exponent $\zeta \in (0, 1)$, the larger the first term, the higher the effect of volatility stabilization. When $\zeta \geq 1$, the total volatility $\sigma^i(\boldsymbol{\mu}_t)$ is bounded for $\boldsymbol{\mu}_t \in \Delta_n$, the asymptotic stabilization by volatility is not observed in general.

In classical stochastic portfolio theory, we define the set of self-financing trading strategies $\mathcal{F}(\mathbf{S})$ as \mathbb{R}^n -valued processes, which are predictable and integrable with respect to the process of capitalizations. Here by using the numéraire of market portfolio $(\Sigma_t)_{t \geq 0}$, the value of the portfolio $(V_t^\pi)_{t \geq 0}$, also called wealth process, can be written in terms of market weights $(\boldsymbol{\mu}_t)_{t \geq 0}$, with strategy $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t \geq 0} \in \mathcal{F}(\boldsymbol{\mu})$. According to [KR17], the set of self-financing trading strategies is the same with respect to market capitalizations and relative weights of the assets, *i.e.* $\mathcal{F}(\boldsymbol{\mu}) = \mathcal{F}(\mathbf{S})$. For the sake of

simplicity, by normalization we can suppose the initial value of the portfolio equal to 1.

$$V_t^\pi = \boldsymbol{\pi}_t \cdot \boldsymbol{\mu}_t = \sum_{i=1}^n \pi_t^i \mu_t^i, \quad V_0^\pi = 1$$

For two trading strategies $\boldsymbol{\pi}, \boldsymbol{\rho} \in \mathcal{T}(\boldsymbol{\mu})$, we say that $\boldsymbol{\pi}$ realizes an arbitrage with respect to $\boldsymbol{\rho}$ over time horizon $[0, T]$ if:

$$\mathbb{P}[V_T^\pi \geq V_T^\rho] = 1 \text{ and } \mathbb{P}[V_T^\pi > V_T^\rho] > 0$$

The relative arbitrage is strong if:

$$\mathbb{P}[V_T^\pi > V_T^\rho] = 1$$

One particular relative arbitrage that we are interested in is that with respect to the market. The value of the market portfolio being always equal to 1 in its own numéraire *i.e.* $V_t^\mu = 1 \forall t \geq 0$, we can simply write:

$$\begin{aligned} \mathbb{P}[V_T^\pi \geq 1] &= 1 \text{ and } \mathbb{P}[V_T^\pi > 1] > 0 \\ \mathbb{P}[V_T^\pi > 1] &= 1 \text{ for strong relative arbitrage} \end{aligned}$$

The growth rate $\gamma^\pi(\cdot)$ associated to the trading strategy $\boldsymbol{\pi}$ is defined from the logarithm representation:

$$d \log V_t^\pi = \gamma^\pi(\boldsymbol{\mu}_t) dt + \text{martingale}$$

Define the excess growth rate $\gamma_*^\pi(\cdot)$ and the excess growth rate associated to the market $\gamma_*(\cdot)$:

$$\begin{aligned} \gamma_*^\pi(\boldsymbol{\mu}_t) &= \gamma^\pi(\boldsymbol{\mu}_t) - \sum_{i=1}^n \pi_t^i \gamma^i(\boldsymbol{\mu}_t) \\ \gamma_*(\boldsymbol{\mu}_t) &= \frac{1}{2} \sum_{i=1}^n (\mu_t^i)^{-1} a^{i,i}(\boldsymbol{\mu}_t) \end{aligned}$$

The cumulative excess growth rate (of the market) is defined as the integral of excess growth rate along the path of market weights:

$$\Gamma_*(t) = \int_0^t \gamma_*(\boldsymbol{\mu}_u) du = \frac{1}{2} \sum_{i=1}^n \int_0^t (\mu_u^i)^{-1} a^{i,i}(\boldsymbol{\mu}_u) du$$

One sufficient condition for the existence of relative arbitrage (with respect to the market) obtained in [FK05] deals with the excess growth rate of the market, if $\gamma_*(\cdot)$ is bounded from below by a positive constant, *i.e.* $\exists \underline{\gamma} > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma}, \forall t \in [0, T]$ *a.s.*, then there exists a threshold $T^* > 0$ depending on model parameters, so that relative arbitrage exists over the time horizon $[0, T]$, $\forall T > T^*$.

Main Result 2.1.4. *Depending on the value of the exponent $\zeta > 0$, we have the following estimates of excess growth rate:*

1. When $\zeta \leq 1/2$ and under the assumption $\zeta \zeta^\top = \text{diag}(c^1, \dots, c^n)$, $c^i > 0 \forall i \in \{1, \dots, n\}$, there exists a constant $\underline{\gamma} > 0$ such that:

$$\forall t > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma} \text{ a.s.}$$

The sufficient condition for the existence of relative arbitrage is then satisfied, at least for a long enough time horizon.

2. When $\zeta \in \mathbb{N}^*$ we have:

$$\forall \epsilon > 0 : \forall t > 0 : \mathbb{P}[\gamma_*(\boldsymbol{\mu}_t) < \epsilon] > 0$$

It is then the opposite case to the sufficient condition for relative arbitrage. For all $\epsilon > 0$, the excess growth rate has positive probability to be smaller than ϵ .

The market risk premium $\boldsymbol{\theta}(\cdot) = (\theta^\nu(\cdot))_{\nu=1}^d$ satisfies the relation:

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, \forall t \geq 0, \mathbf{b}(\boldsymbol{\mu}_t) = \sigma(\boldsymbol{\mu}_t) \cdot \boldsymbol{\theta}(\boldsymbol{\mu}_t) \text{ and } \forall T \geq 0, \int_0^T |\boldsymbol{\theta}(\boldsymbol{\mu}_t)|^2 dt < \infty$$

The market risk deflator $(Z_t)_{t \geq 0}$ is defined from the risk premium by an exponential martingale with $Z_0 = 1$, it is a local martingale:

$$Z_t = \mathcal{E}\left(-(\boldsymbol{\theta}(\boldsymbol{\mu}_u))_{u \in [0, t]}\right) = \exp\left(-\int_0^t \boldsymbol{\theta}(\boldsymbol{\mu}_u) \cdot d\mathbf{W}_u - \frac{1}{2} \int_0^t |\boldsymbol{\theta}(\boldsymbol{\mu}_u)|^2 du\right)$$

When the market risk premium $(Z_t)_{t \geq 0}$ exists, we say that the condition ‘‘No unbounded profit with bounded risk’’ or *NUPBR* introduced in [DS94] is satisfied. As its name suggests, this condition is equivalent to the fact that the set of wealth processes $\{V_T^\pi \mid \boldsymbol{\mu} \in \mathcal{T}(\boldsymbol{\mu})\}$ is bounded in probability. The condition ‘‘No free lunch with vanishing risk’’, ou *NFLVR* is satisfied when we have (NUPBR) and there is no relative arbitrage with respect to the market.

Main Result 2.1.5. For $\boldsymbol{\mu}_0 \in \Delta_+^n$ and under the assumption $d \geq n$ and ζ of full rank, the market risk premium $\boldsymbol{\theta}(\cdot)$, as well as the market risk deflator $(Z_t)_{t \geq 0}$ exist, so the (NUPBR) condition is satisfied over the time horizon $[0, T]$, $T > 0$, if and only if one of the following conditions is satisfied:

- (1) One of the conditions (i) - (iii) in Main Result 2.1.1 holds, recall that in this case $\forall t \geq 0, \boldsymbol{\mu}_t \in \Delta_+^n$ a.s..
- (2) One of the conditions (iv) - (vi) in Main Result 2.1.1 holds, and $b^i(\boldsymbol{\mu}) = 0$ on $\{\mu^i = 0\}$ for all indices i such that $\mathbb{P}[\exists t \in (0, T) \mid \mu_t^i = 0] > 0$.

In addition, the market risk premium can be written as follows:

$$\boldsymbol{\theta}(\boldsymbol{\mu}_t) = (\sigma^\top \cdot (a^+ \cdot \mathbf{b}))(\boldsymbol{\mu}_t),$$

where $a^+(\cdot)$ is the pseudo-inverse of Moore-Penrose of the covariance matrix.

In the opposite case, when none of the conditions (1) and (2) above holds, i.e. for a certain index $i \in \{1, \dots, n\}$, $(\mu_t^i)_{t \geq 0}$ vanishes on $[0, T]$ with positive probability and there exists $\boldsymbol{\mu} \in \{\mu^i = 0\}$ such that $b^i(\boldsymbol{\mu}) > 0$, then the (NUPBR) condition is not satisfied, in particular the market risk premium $\boldsymbol{\theta}(\cdot)$ and deflator $(Z_t)_{t \geq 0}$ do not exist.

The martingale property of the market risk deflator is related to the condition of existence of relative arbitrage with respect to the market by the following result.

Main Result 2.1.6. *There is no relative arbitrage with respect to the market over the time horizon $[0, T]$ if and only if the market risk deflator $(Z_t)_{t \geq 0}$ associated to market weights is a true martingale on $[0, T]$, i.e. $\mathbb{E}[Z_T] = 1$. Or equivalently, relative arbitrage exists over $[0, T]$ if and only if the local martingale $(Z_t)_{t \geq 0}$ is strict on $[0, T]$ with $\mathbb{E}[Z_T] < 1$.*

The previous result provides a necessary and sufficient condition for the existence of relative arbitrage with respect to the market, although the technical difficulty is to check whether the market risk deflator $(Z_t)_{t \geq 0}$ is a true martingale over the time horizon $[0, T]$.

Main Result 2.1.7. *When the number n of asset is equal to 2, with the matrix $\varsigma = [\varsigma^1 \ \varsigma^2]^\top$ satisfying $\varsigma^1 \geq 0 \geq \varsigma^2$, $(\varsigma^1, \varsigma^2) \neq (0, 0)$, then according to the value of the exponent $\zeta > 0$ in market weights equation, we have:*

1. *When $\zeta \in (0, 1)$: Relative arbitrage with respect to the market exists, at least for long enough time horizons.*
2. *When $\zeta \geq 1$: We have absence of relative arbitrage for any time horizon $[0, T]$, $T > 0$.*

For the general case where the number of assets $n \geq 2$ is arbitrary, we have for $\zeta \geq 1$, absence of relative arbitrage for any time horizon $[0, T]$, $T > 0$.

2.1.3 Applications and perspectives

To construct more investment portfolios in this same market, the notion of functionally generated portfolios has been introduced in [Fer99]. This function takes the relative market weights as argument, which are directly given by our model. For a positive function G defined on the unit simplex Δ^n twice differentiable with continuous derivatives, a new portfolio with relative weights $(\pi_t)_{t \geq 0}$ is constructed from the market portfolio $(\mu_t)_{t \geq 0}$ by:

$$\pi_t^i = \mu_t^i \cdot (1 + \partial_i \log G(\mu_t) - \sum_{j=1}^n \mu_t^j \cdot \partial_j \log G(\mu_t))$$

When relative arbitrage with respect to the market exists on $[0, T]$, $\mathbb{E}[Z_T] < 1$, define the function $u(t, \mu_t) = \mathbb{E}[Z_T | \mathcal{F}_t] / Z_t$ and if u is regular, by taking $G(\mu_t) = u(t, \mu_t)$ we construct the relative arbitrage relative to the market which is optimal in terms of initial investment using delta hedging.

To calibrate our model to market data, or simply estimate the model parameters, suppose for example $\zeta = 1$. Noting that $z_t^i = \log(\mu_t^i)$, we get the expression of quadratic variation as follows:

$$\frac{d\langle \tilde{z}^i, \tilde{z}^j \rangle_t}{dt} = \sum_{\nu=1}^d \varsigma^{i,\nu} \varsigma^{j,\nu} = (\varsigma \varsigma^\top)^{i,j}; \quad \tilde{z}_t^i = z_t^i - \frac{1}{n} \sum_{k=1}^n z_t^k$$

This matrix of quadratic variations can be estimated on a time grid \mathcal{T} :

$$\mathcal{T} = \{t_m = m \cdot \Delta t, \Delta t = T/M_T; m = 0, \dots, M_T\}$$

$$(\zeta \zeta^\top)^{i,j} \approx \frac{1}{T} \sum_{m=1}^{M_T} (\tilde{z}_{t_m}^i - \tilde{z}_{t_{m-1}}^i)(\tilde{z}_{t_m}^j - \tilde{z}_{t_{m-1}}^j), \quad i \leq j$$

The matrix ζ can then be defined as the square root of the matrix of quadratic variations.

The coefficients of the matrix B are obtained by multiple linear regressions of each asset $i \in \{1, \dots, n\}$:

$$\mu_{t_m}^i \approx \mu_{t_{m-1}}^i + \left[\sum_{j=1}^{n-1} (B^{i,j} - B^{i,n}) \mu_{t_{m-1}}^j + B^{i,n} \right] \Delta t + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_{t_{m-1}}) (W_m^\nu - W_{m-1}^\nu)$$

The exponent $\zeta > 0$, which measures the effect of stabilization by volatility, can be also estimated from market data. The estimation of matrices B and ζ in the general case thus deserves further investigation.

With the ergodicity of market weights process shown in Main Result 2.1.3, we study the optimization of wealth over the long term, by defining the *expected risk-sensitive growth rate* $J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T)$ for time horizon T and its long term version $J(\boldsymbol{\mu}, \boldsymbol{\pi})$:

$$J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T) = -\frac{2}{\theta} \log \mathbb{E}[e^{-\frac{\theta}{2} \log V_T^\pi}]$$

$$J(\boldsymbol{\mu}, \boldsymbol{\pi}) = \limsup_{T \rightarrow \infty} \frac{1}{T} J(\boldsymbol{\mu}, \boldsymbol{\pi}; T),$$

where $\theta > -2$, $\theta \neq 0$ is the parameter of risk sensitivity. It is possible to write the Bellman equation or its ergodic version associated to the problem of maximizing $J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T)$ or $J(\boldsymbol{\mu}, \boldsymbol{\pi}_0)$. However, to establish the existence and uniqueness of solution to the Bellman equation, we need uniform ellipticity which is not satisfied by market weights equation, in fact we only have local ellipticity. This ergodic stochastic control problem could be a topic for future research.

Typically our model is applied to market portfolios with a large number n of assets. We are naturally interested in the behavior of market weights when $n \rightarrow \infty$. The market weights equation, which gives discrete probability distributions, becomes measure-valued. For volatility-stabilized market model, where the dynamics of weights correspond to Jacobi or Wright-Fisher process, its measure-valued limit is called Fleming-Viot process as indicated in [Pal+11]. The measure-valued polynomial processes are also studied, see the recent paper [CLSF+19]. The link with rank-based models, see [JR15], could also be established, by taking into account the interaction of mean-field type between the weights of assets.

2.2 A probabilistic scheme for McKean-Vlasov equations

The second part is dedicated to the numerical approximation of the marginal law of the following *McKean-Vlasov differential equation*:

$$X_t^{s,\xi} = \xi + \int_s^t b(X_r^{s,\xi}, [X_r^{s,\xi}]) dr + \int_s^t \sigma(X_r^{s,\xi}, [X_r^{s,\xi}]) dW_r, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d). \quad (2.2.1)$$

On the probability space $(\Omega, \mathcal{A}, \mathcal{P}(\mathbb{R}^d))$, W is a Brownian motion of dimension q , the coefficients b, σ are defined on \mathbb{R}^d and take values in \mathbb{R}^d and $\mathbb{R}^{d \times q}$ respectively. Here $[\theta]$ is the law of a random variable θ . Under weak regularity assumptions on b and $a = \sigma\sigma^*$, *i.e.* if they are bounded, Hölder in the space variable and have functional linear derivatives with respect to the measure variable, denoted as $\delta b/\delta m$ and $\delta a/\delta m$, such that $\delta b/\delta m$ is bounded, $(x, y) \mapsto [\delta a(x, m)/\delta m](y)$ is Hölder (uniformly with respect to variable m), and finally if the diffusion coefficient a is uniformly elliptic: $\exists \lambda \geq 1, \forall (x, u, m) \in (\mathbb{R}^d)^2 \times \mathcal{P}(\mathbb{R}^d), \lambda^{-1}|u|^2 \leq \langle a(x, m)u, u \rangle \leq \lambda|u|^2$, then the McKean-Vlasov SDE (2.2.1) has a unique weak solution for every initial condition $(s, \mu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$, see e.g. Chaudru de Raynal and Frikha [RF19].

2.2.1 Picard iteration and method of continuation

The decoupled stochastic flow associated to the McKean-Vlasov SDE is defined as a standard SDE with time inhomogeneous coefficients by considering the law μ of the random variable ξ :

$$X_t^{s,x,\mu} = x + \int_s^t b(X_r^{s,x,\mu}, [X_r^{s,\xi}]) dr + \int_s^t \sigma(X_r^{s,x,\mu}, [X_r^{s,\xi}]) dW_r$$

Our strategy is to approximate the marginal law of the solution of McKean-Vlasov SDE by using *Picard iterations* and the convergence rate is geometric according to Sznitman [Szn91]. To reduce the complexity especially when the number of iterations increases, we apply the *method of continuation in time* for horizon $[0, T]$ by introducing the uniform time grid $\mathcal{S} = \{t_0 = 0 \leq t_1 \leq \dots \leq t_N = T\}$, $N \geq 1$ with $t_k = k\delta$, $\delta = T/N$. Local Picard iterations are made on each time subinterval $[t_k, t_{k+1}]$, $k = 0, \dots, N-1$ while using the approximation obtained on previous subintervals.

Let us define the set $\mathcal{A}_{t_k, t_{k+1}, \mu_k}$ as a closed subspace of $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$ and the uniform metric $d_{t_k, t_{k+1}}$:

$$\mathcal{A}_{t_k, t_{k+1}, \mu_k} = \left\{ \mathbf{P} \in \mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d)) : \mathbf{P}(t_k) = \mu_k \right\},$$

$$d_{t_k, t_{k+1}}(\mathbf{P}, \mathbf{Q}) = \sup_{s \in [t_k, t_{k+1}]} |d_{TV}(\mathbf{P}(s), \mathbf{Q}(s))|$$

where d_{TV} stands for the total variation distance on the space of probability measures. Remark that the space $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$ equipped with the metric $d_{t_k, t_{k+1}}$ is complete. Moreover, $\mathcal{A}_{t_k, t_{k+1}, \mu_k}$ being a closed subspace of $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$, it is also a complete metric space. Define the mapping $\mathcal{T}_k : \mathcal{A}_{t_k, t_{k+1}, \mu_k} \rightarrow \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ which maps $\mathbf{Q} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ to $\mathcal{T}_k(\mathbf{Q})(t) = [X_t^{\mathbf{Q}}]$, $t \in [t_k, t_{k+1}]$, where $X^{\mathbf{Q}}$ is given by the unique weak solution to the following SDE:

$$X_t^{\mathbf{Q}} = \xi_k + \int_{t_k}^t b(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) ds + \int_{t_k}^t \sigma(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) dW_s, \quad t \in [t_k, t_{k+1}], \quad [\xi_k] = \mu_k.$$

The mapping \mathcal{T}_k is well-defined under the aforementioned regularity and uniform ellipticity assumptions. The fixed point of \mathcal{T}_k is the unique solution of McKean-Vlasov SDE on $[t_k, t_{k+1}]$ with initial condition (t_k, μ_k) .

The initial condition $\mathbf{P}_k^{(0)} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ of Picard iteration scheme is given by the process \bar{X}_t^k with coefficients frozen at (t_k, ξ_k, μ_k) :

$$\bar{X}_t^k = \xi_k + b(\xi_k, \mu_k)(t - t_k) + \sigma(\xi_k, \mu_k)(W_t - W_{t_k})$$

with $\mathbf{P}_k^{(0)}(t) = [\bar{X}_t^k]$, $t \in [t_k, t_{k+1}]$ written as:

$$\mathbf{P}_k^{(0)}(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) \mu_k(dx) dz$$

where for a symmetric and invertible real-valued matrix a of dimension $d \times d$, the map $x \mapsto g(a, x)$ is the density function of a centered normal distribution with covariance matrix a . The flow of probability measures $\left\{ \mathbf{P}_k^{(L)}(t) = \mathcal{T}_k^{(L)}(\mathbf{P}_k^{(0)}(t)), t \in [t_k, t_{k+1}] \right\}$ provides an approximation for $([X_t])_{t \in [t_k, t_{k+1}]}$, where $\mathcal{T}_k^{(L)}$ represents L times iteration of the mapping \mathcal{T}_k , L being the number of Picard iterations, also called number of levels. For $k = 0$, $\mathbf{P}_0^{(0)}(t_0) = \mu_0 = [\xi]$ and $\mathbf{P}_k^{(L)}(t_k) = \mu_k^L := \mathbf{P}_{k-1}^{(L)}(t_k)$ for $k = 1, \dots, N - 1$.

On each time subinterval $[t_k, t_{k+1}]$ and for each level $\ell = 1, \dots, L$, $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ corresponds to the flow of marginal laws of the unique solution $X^{(\ell)}$ of the following *decoupled SDE* which has the flow of measures from previous level as parameter $(\mathbf{P}_k^{(\ell-1)}(t))_{t \in [t_k, t_{k+1}]}$:

$$X_t^{(\ell)} = \xi_k^{(\ell)} + \int_{t_k}^t b(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) ds + \int_{t_k}^t \sigma(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) dW_s, [X_0^{(\ell)}] = \mu_k^{(\ell)} = \mathbf{P}_k^{(\ell)}(t_k) \quad (2.2.2)$$

where $[X_0^{(\ell)}] = \mu_k^{(\ell)} = \mathbf{P}_k^{(\ell)}(t_k)$.

According to [Fri08], under the aforementioned assumptions, the random variable $X_t^{(\ell)}$ has a density $z \mapsto p_\ell(t_k, t, z)$ satisfying:

$$\mathbf{P}_k^{(\ell)}(t)(dz) = p_\ell(t_k, t, z) dz = \int_{\mathbb{R}^d} p_\ell(t_k, t, x, z) \mu_k^{(\ell)}(dx) dz.$$

We have the following Gaussian upper bound: for all $t \in (t_k, t_{k+1}]$, $(x, z) \in (\mathbb{R}^d)^2$ and $\ell \geq 0$, with two constants $c, C > 0$ independent on ℓ :

$$p_\ell(t_k, t, x, z) \leq E_{1/2,1}(C(|b|_\infty + 1)) g(c(t - t_k), z - x)$$

where for $c > 0$, $x \mapsto g(c, x)$ stands for the density function of the centered normal distribution with covariance matrix cI_d .

2.2.2 Probabilistic representation of marginal law

We now look for an approximation of the sequence of probability measures $(\mathbf{P}^{(\ell)}(t_k))_{1 \leq \ell \leq L}$, $k = 1, \dots, N$ based on a probabilistic representation for the marginal laws of standard diffusion processes introduced [BKH+15] and studied further in [AKH+17] and [HLTT+17].

Consider a sequence of random variables $\tau := (\tau_n)_{n \geq 1}$, $\tau_0 = 0$ with $(\tau_n - \tau_{n-1})_{n \geq 1}$ i.i.d. with density f and of cumulative distribution function $F(t) = \int_{-\infty}^t f(s) ds$. The renewal process $N := (N_t)_{t \geq 0}$ with jump times τ is defined by $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$.

For all functions $\Phi : \Delta_n([s, t]) \rightarrow \mathbb{R}$ satisfying $\mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}}|\Phi(\tau_1, \dots, \tau_n)] < \infty$ where the set $\Delta_n([s, t]) = \{\mathbf{s}_n \in [s, t]^n : s_0 := s \leq s_1 \leq \dots \leq s_n \leq t =: s_{n+1}\}$:

$$\begin{aligned} & \mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}}\Phi(\tau_1, \dots, \tau_n)] \\ &= \int_{\Delta_n([s, t])} \Phi(r_1 - s, \dots, r_n - s)(1 - F(t - r_n)) \prod_{j=0}^{n-1} f(r_{j+1} - r_j) d\mathbf{r}_n. \end{aligned}$$

By taking some particular density functions, we recognize the following renewal processes:

1. For $f(t) = \lambda e^{-\lambda t} \mathbf{1}_{[0, \infty)}(t)$, $\lambda > 0$, N is a Poisson process with intensity λ .
2. For $f(t) = \frac{1-\alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{t^\alpha} \mathbf{1}_{[0, \bar{\tau}]}(t)$, $(\alpha, \bar{\tau}) \in (0, 1) \times (0, \infty)$, the jump times of N follow the $Beta(1 - \alpha, 1)$ distribution with values in $[0, \bar{\tau}]$.

For $t \in [t_k, t_{k+1}]$, $\ell \geq 0$, an initial measure μ_k^ℓ , a Brownian motion W^ℓ and a renewal process N^ℓ with jump times $\tau^\ell = (\tau_n^\ell)_{n \geq 0}$, the Euler scheme $\bar{X}^\ell = (\bar{X}_{\zeta_n^\ell}^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$ starting from $\bar{X}_{t_k}^\ell$, $[\bar{X}_{t_k}^\ell] = \mu_k^\ell$ is defined on the random time grid $\zeta^\ell := (\zeta_n^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$, $\zeta_n^\ell := t_k + \tau_n^\ell \wedge (t - t_k)$, $\zeta_0^\ell := t_k$:

$$\begin{aligned} \bar{X}_{\zeta_{n+1}^\ell}^\ell &= \bar{X}_{\zeta_n^\ell}^\ell + b_n^\ell(\zeta_{n+1}^\ell - \zeta_n^\ell) + \sigma_n^\ell(W_{\zeta_{n+1}^\ell}^\ell - W_{\zeta_n^\ell}^\ell), \quad 0 \leq n \leq N_{t-t_k}^\ell, \\ \sigma_n^\ell &= \sigma(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell)), \quad b_n^\ell = b(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell)), \quad a_n^\ell = \sigma_n^\ell(\sigma_n^\ell)^*. \end{aligned}$$

Main Result 2.2.1. Suppose $x \mapsto b(x, m) \in \mathcal{C}_b^1(\mathbb{R}^d, \mathbb{R}^d)$, $x \mapsto a(x, m) \in \mathcal{C}_b^2(\mathbb{R}^d, \mathbb{R}^{d \times q})$ and that the matrix a is uniformly elliptic. Let h be a function satisfying $\forall z \in \mathbb{R}^d$, $|h(z)| \leq C \exp\left(\alpha \frac{|z|^2}{2T}\right)$ for constants $C, \alpha > 0$, $\alpha < c^{-1}$ with c, C appeared in the Gaussian upper bound. Then there exists a constant $\kappa := \kappa(c, \alpha)$ such that if $\int_{\mathbb{R}^d} \exp(\kappa|x|^2) \mu_k^\ell(dx) < \infty$, the expectation $\mathbb{E}[h(X_t^{(\ell)})]$ has the following probabilistic representation:

$$\forall t \in [t_k, t_{k+1}], \quad \mathbb{E}[h(X_t^{(\ell)})] = \mathbb{E}\left[h(\bar{X}_{\zeta_{N_{t-t_k}^\ell + 1}^\ell}^\ell) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^\ell\right], \quad \ell \geq 1$$

where $\theta_r^\ell := \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}(\bar{X}_{\zeta_r^\ell}^\ell, \bar{X}_{\zeta_{r+1}^\ell}^\ell, \mathbf{P}_k^{(\ell-1)})$ pour $r = 0, \dots, N_{t-t_k}^\ell - 1$ and $\theta_{N_{t-t_k}^\ell}^\ell = (1 - F(t - \zeta_{N_{t-t_k}^\ell}^\ell))^{-1}$ with:

$$\theta_{s,t}(x, y, \mathbf{P}) := (f(t - s))^{-1} \left[\sum_{i,j=1}^d \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) - \sum_{i=1}^d \rho_{s,t}^i(x, y, \mathbf{P}) \right]$$

and the functions $\kappa_{s,t}^{i,j}$, $\rho_{s,t}^i$, $i, j = 1, \dots, d$ are defined as:

$$\begin{aligned} \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) &:= \frac{1}{2} [a_{i,j}(y, \mathbf{P}(t)) - a_{i,j}(x, \mathbf{P}(s))] H_2^{i,j}(a(x, \mathbf{P}(s))(t - s), y - x) \\ &\quad + \partial_{y_i} a_{i,j}(y, \mathbf{P}(t)) H_1^i(a(x, \mathbf{P}(s))(t - s), y - x) + \frac{1}{2} \partial_{y_i y_j}^2 a_{i,j}(y, \mathbf{P}(t)), \\ \rho_{s,t}^i(x, y, \mathbf{P}) &:= [b_i(y, \mathbf{P}(t)) - b_i(x, \mathbf{P}(s))] H_1^i(a(x, \mathbf{P}(s))(t - s), y - x) + \partial_{y_i} b_i(y, \mathbf{P}(t)). \end{aligned}$$

2.2. A probabilistic scheme for McKean-Vlasov equations

In addition, for all $t \in (t_k, t_{k+1}]$, $\ell \geq 1$, the density functions $z \mapsto p_\ell(t_k, t, z)$, $p_\ell(t_k, t, x, z)$ also have probabilistic representations, for example for $p_\ell(t_k, t, x, z)$:

$$p_\ell(t_k, t, x, z) = \mathbb{E} \left[g \left(a_{N_{t-t_k}^\ell}^\ell \left(t - \zeta_{N_{t-t_k}^\ell}^\ell \right), z - \bar{X}_{N_{t-t_k}^\ell}^{\ell, x} \right) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^{\ell, x} \right]$$

with weights $\theta_r^{\ell, x} = \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}(\bar{X}_{\zeta_r^\ell}^{\ell, x}, \bar{X}_{\zeta_{r+1}^\ell}^{\ell, x}, \mathbf{P}_k^{(\ell-1)})$ and the Euler scheme $(\bar{X}_{\zeta_r^\ell}^{\ell, x})_{0 \leq r \leq N_{t-t_k}^\ell}$ is defined similarly as previously except its initial condition is x

For each $k \in \{0, \dots, N-1\}$, by applying the probabilistic representation k times, we have for $t \in [t_k, t_{k+1}]$:

$$\begin{aligned} \mathbb{E}[h(X_t^{(\ell)})] &= \mathbb{E} \left[h \left(\bar{X}_{\zeta_{N_{t-t_k}^\ell}^\ell}^\ell \right) \prod_{r=0}^{N_{t-t_k}^{\ell, k+1}} \theta_r^{\ell, k+1} \times \Gamma_k^\ell \right], \quad \ell \geq 1, \\ \Gamma_k^\ell &:= \Gamma_{k-1}^\ell \prod_{r=0}^{N_\delta^{\ell, k}} \theta_r^{\ell, k} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell, j}} \theta_r^{\ell, j}, \quad \text{with } \Gamma_0^\ell = 1. \end{aligned}$$

We set $\hat{\mu}_0^L = \frac{1}{M_0} \sum_{i=1}^{M_0} \delta_{\xi^{(i)}}$, where the random variables $(\xi^{(i)})_{1 \leq i \leq M_0}$ are i.i.d. of law μ . We construct step by step from $k=0$ until $N-1$, the flows of signed measures:

$$\begin{aligned} \left\{ \hat{\mu}^0(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \hat{\mu}_k^L)(t-t_k), z - b(x, \hat{\mu}_k^L)(t-t_k) - x) \hat{\mu}_k^L(dx) dz, t \in [t_k, t_{k+1}], \right. \\ \left. k = 0, \dots, N-1 \right\} \end{aligned}$$

by setting $\hat{\mu}_k^L = \hat{\mu}^L(t_k)$, $k = 1, \dots, N$ and

$$\begin{aligned} \forall t \in [t_k, t_{k+1}], \quad \hat{\mu}^\ell(t)(dz) &:= \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Gamma_k^{\ell, (m)} \prod_{r=0}^{N_{t-t_k}^{\ell, k+1, (m)}} \theta_r^{\ell, k+1, (m)} \delta_{\bar{X}_{N_{t-t_k}^{\ell, k+1, (m)}+1}^{\ell, (m)}}(dz), \\ \Gamma_k^{\ell, (m)} &:= \Gamma_{k-1}^{\ell, (m)} \prod_{r=0}^{N_\delta^{\ell, k, (m)}} \theta_r^{\ell, k, (m)} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell, j, (m)}} \theta_r^{\ell, j, (m)}, \quad \text{with } \Gamma_0^{\ell, (m)} := 1, \end{aligned}$$

and for $j \in \{1, \dots, k+1\}$ and for all $r \in \{0, \dots, N_\delta^{\ell, j, (m)} - 1\}$ (if $1 \leq j \leq k$) or $r \in \{0, \dots, N_{t-t_k}^{\ell, j, (m)}\}$ (if $j = k+1$)

$$\begin{aligned} \theta_r^{\ell, j, (m)} &:= \theta_{\zeta_r^{\ell, (m)}, \zeta_{r+1}^{\ell, (m)}}(\bar{X}_r^{\ell, (m)}, \bar{X}_{r+1}^{\ell, (m)}, \hat{\mu}^{\ell-1}), \\ \theta_{N_\delta^{\ell, j}}^{\ell, j, (m)} &:= (1 - F(t_j - \zeta_{N_\delta^{\ell, j, (m)}}^{\ell, (m)}))^{-1}, \\ \theta_{N_{t-t_k}^{\ell, k+1}}^{\ell, k+1, (m)} &:= (1 - F(t - \zeta_{N_{t-t_k}^{\ell, k+1, (m)}}^{\ell, (m)}))^{-1}, \\ \bar{X}_{\zeta_{r+1}^{\ell, (m)}}^{\ell, (m)} &= \bar{X}_{\zeta_r^{\ell, (m)}}^{\ell, (m)} + b_r^\ell(\zeta_{r+1}^{\ell, (m)} - \zeta_r^{\ell, (m)}) + \sigma_r^{\ell, (m)} (W_{\zeta_{r+1}^{\ell, (m)}}^{\ell, (m)} - W_{\zeta_r^{\ell, (m)}}^{\ell, (m)}) \end{aligned}$$

with the notations $b_r^{\ell,(m)} := b(\bar{X}_r^{\ell,(m)}, \hat{\mu}^{\ell-1}(\zeta_r^{\ell,(m)}))$, $\sigma_r^{\ell,(m)} := \sigma(\bar{X}_r^{\ell,(m)}, \hat{\mu}^{\ell-1}(\zeta_r^{\ell,(m)}))$, $a_r^{\ell,(m)} = \sigma_r^{\ell,(m)}(\sigma_r^{\ell,(m)})^T$. Here, we require that $(W^{\ell,(m)}, N^{\ell,(m)}, \bar{X}_0^{\ell,(m)})$, $1 \leq m \leq M_\ell$, $\ell \geq 1$, is an i.i.d. sequence of random variables with the same law as (W, N, X_0) . Moreover, for any fixed $m \in \{1, \dots, M_\ell\}$, $\{N^{\ell,j,(m)}, j = 1, \dots, k+1\}$ is a sequence of i.i.d. renewal processes with respective jump times $\tau_r^{\ell,j,(m)} = (\tau_r^{\ell,j,(m)})_{r \geq 0}$ so that $\zeta_r^{\ell,(m)} := t_j + \tau_r^{\ell,j,(m)} \wedge \delta$, for $r = 0, \dots, N_\delta^{\ell,j,(m)} + 1$, $j = 1, \dots, k$ and $\zeta_r^{\ell,(m)} := t_k + \tau_r^{\ell,k+1,(m)} \wedge (t - t_k)$, for $r = 0, \dots, N_{t-t_k}^{\ell,k+1,(m)} + 1$.

The Monte-Carlo estimator of the expectation $\mathbb{E}[h(X_T)]$ is:

$$\langle h, \hat{\mu}^L(T) \rangle = \frac{1}{M_L} \sum_{m=1}^{M_L} h(\bar{X}_{N_\delta^L, N, (m)+1}^L) \times \Gamma_N^{L,(m)}$$

The signed measure $\hat{\mu}^L(T)$ with density function $z \mapsto \hat{p}_L(0, T, z)$:

$$\hat{p}_L(0, T, z) := \frac{1}{M_L} \sum_{m=1}^{M_L} g(a_{N_\delta^L, N, (m)}^{L,(m)}(T - \zeta_{N_\delta^L}^{L,(m)}), z - \bar{X}_{N_\delta^L}^{L,(m)}) \times \Gamma_N^{L,(m)}$$

is an approximation of the density $z \mapsto p(0, T, z)$ at time T of the solution $X_T^{0,\xi}$ to the McKean-Vlasov SDE.

Observe that in order to construct $\hat{\mu}^L(t)$, for a given time $t \in [0, t_1]$ and a given number of level L , one needs to have access to the signed measure μ^{L-1} on the time grid given by the jump times of the renewal processes $N^{L,1,(m)}$, $1 \leq m \leq M_L$. Having this measure at hand, one may then compute the weights $(\theta_r^{L,1,(m)})_{0 \leq r \leq N_\delta^{L,1,(m)}}$ and then the quantity $\Gamma_1^{L,(m)}$. However, the same procedure appears at level $L-1$ since in order to compute the measure $\hat{\mu}^{L-1}(\zeta_r^{L,1,(m)})$, one needs to have access to the signed measure $\hat{\mu}^{L-2}$ on the time grid of $N^{L-1,1,(m)}$, $1 \leq m \leq M_{L-1}$, so on and so forth. We thus observe that the (averaged) computational cost for constructing the measure $\hat{\mu}^L(t_1)$ is of order $M_0 \times \dots \times M_L$. We then continue by constructing the measure $\hat{\mu}^L(t)$ for a given time $t \in [t_1, t_2]$, the sequence of weights $\Gamma_1^{\ell,(m)}$, $1 \leq \ell \leq L$, $1 \leq m \leq M_L$, being already computed at the preceding step. The same lines of reasoning apply so that the computation cost for constructing the measure $\hat{\mu}^L(t_2)$ is of order $M_L \times M_1 \times \dots \times M_L$. We proceed accordingly up to the last time interval $[t_{N-1}, T]$. Hence, the (averaged) total complexity of our algorithm is of order $M_0 \times \dots \times M_L + (N-1)M_L \times M_1 \times \dots \times M_L$.

2.2.3 Convergence and error analysis and perspectives

Our Monte-Carlo estimator for the quantities $\mathbb{E}[h(X_T)]$ and $z \mapsto p(0, T, z)$ contains two sources of error:

$$\begin{aligned} \mathbb{E}[h(X_T)] - \langle h, \hat{\mu}^L(T) \rangle &= \mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^L)] + \mathbb{E}[h(X_T^L)] - \langle h, \hat{\mu}^L(T) \rangle, \\ p(0, T, z) - \hat{p}_L(0, T, z) &= p(0, T, z) - p_L(0, T, z) + p_L(0, T, z) - \hat{p}_L(0, T, z). \end{aligned}$$

The first error term corresponds to the bias introduced by the approximation of the marginal law of McKean-Vlasov equation using Picard iteration with L levels. Its convergence is geometric in L , of order $(C\delta)^L / \Gamma(1 + (L-1)/2) \sim (C\delta)^L / ([L/2])!$.

The second error is a statistical error due to the approximation at each level $\ell = 1, \dots, L$, the flow of probability measures μ^ℓ by the flow of signed measures $\widehat{\mu}^\ell$ obtained from the probabilistic representation. The $L^2(\mathbb{P})$ -moment of this error should be (at best) of order $M^{-\frac{1}{2}}$, assuming that $M_\ell = M$, $\ell = 0, \dots, L$, to simplify the discussions.

In order to achieve a root mean squared error of order ϵ , N (resp. M) have to be of order $\epsilon^{-\frac{1}{L}}$ (resp. ϵ^{-2}), the computational cost NM^{L+1} is then of order $\epsilon^{-\frac{1}{L}} \epsilon^{-2L-2}$. Thus it is optimal to simply choose $L = 1$ to have the optimal complexity ϵ^{-5} .

We first investigate the bias of local Picard iteration scheme.

Main Result 2.2.2. *Suppose that the regularity assumptions of Main Result 2.2.1 are satisfied. Suppose that for all $(i, j) \in \{1, \dots, d\}^2$, the functions $(x, y) \mapsto [\delta b_i(x, m)/\delta m](y)$ and $(x, y) \mapsto [\delta a_{i,j}(x, m)/\delta m](y)$ belong to the space $\mathcal{C}_b^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$ and $\mathcal{C}_b^{2,2}(\mathbb{R}^d \times \mathbb{R}^d)$ respectively. Then for all $k = 0, \dots, n-1$, the sequence of flows of measures $\{\mathbf{P}_k^{(\ell)} = (\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}, \ell \geq 0\}$, where $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ is the marginal law of the solution of the decoupled SDE (2.2.2) with initial distribution $\mu_k \in \mathcal{P}_2(\mathbb{R}^d)$ at time t_k , converges to the flow of probability measures $\mathbf{P}_k^\infty = (\mathbf{P}_k^\infty(t))_{t \in [t_k, t_{k+1}]}$ when $\ell \uparrow \infty$ given by the unique weak solution of mean-field SDE (2.2.1) with initial condition (t_k, μ_k) . In addition, for a function $h \in \mathcal{C}^2(\mathbb{R}^d)$, h and its derivative are of at most quadratic growth, we have for all $L \geq 0$ and $t \in [t_k, t_{k+1}]$:*

$$|\mathbb{E}[h(X_t)] - \mathbb{E}[h(X_t^{(L)})]| \leq \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)})$$

with the notations:

$$\begin{aligned} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) &:= \sup_{t_k \leq r \leq t_{k+1}} \mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r)), \\ \mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r)) &:= \sup_{|h|_{\mathcal{C}_b^2} \leq 1} \left(\int_{\mathbb{R}^d} h(z) \mathbf{P}_k^\infty(r)(dz) - \int_{\mathbb{R}^d} h(z) \mathbf{P}_k^{(0)}(r)(dz) \right). \end{aligned}$$

We also have the upper bound for the bias of the density: for all $(t, x, z) \in (t_k, t_{k+1}] \times (\mathbb{R}^d)^2$ and positive constants $c > \lambda$, C that are non-decreasing with respect to T :

$$\begin{aligned} &|p(t_k, t, x, z) - p_L(t_k, t, x, z)| \\ &\leq \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) g(c(t-t_k), z-x). \end{aligned}$$

Finally, we have the following upper bound with a constant $C' > 0$, for all $k = 0, \dots, n-1$:

$$\mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \leq C\delta.$$

For the bias of global Picard iteration, the goal of our current work in progress is to obtain the following estimation:

$$\begin{aligned} &|p(0, T, z) - p_L(0, T, z)| \\ &\leq CT^{-1} \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \int_{\mathbb{R}^d} g(cT, z-z_0) \mu(dz_0) \end{aligned}$$

by using the following relations:

$$p_L(0, T, z_N) = \int_{(\mathbb{R}^d)^N} \prod_{i=1}^N p_L(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}-1} \mu(dz_0),$$

$$p(0, T, z_N) = \int_{(\mathbb{R}^d)^N} \prod_{i=1}^N p(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}-1} \mu(dz_0),$$

where we used the notation $dz_{\mathbf{N}-1} = dz_1 \cdots dz_{N-1}$. Then we only need to integrate against the function h to get the bias of expectation:

$$\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^L)] = \int_{(\mathbb{R}^d)^2} h(z_N)(p(0, T, z_0, z_N) - p_L(0, T, z_0, z_N)) dz_N \mu(dz_0).$$

The analysis of statistical error makes part of our current work. The main difficulty is when we replace for each time on the grid $k = 1, \dots, N$, $\mu_{t_k}^L$ by its empirical estimator $\widehat{\mu}^L(t_k)$, we have the cumulative errors until final time $T = t_N$.

2.3 Probabilistic representation and integration by parts formulae for some stochastic volatility model with unbounded drift

In the third part, we are interested in the stochastic volatility model described by the couple $(S_t, Y_t)_{t \in [0, T]}$, where S_t represents the spot price of the underlying asset at time t and Y_t is the volatility, the correlation between spot and volatility is ρ :

$$\begin{aligned} dS_t &= rS_t dt + \sigma_S(Y_t)S_t dW_t, \\ dY_t &= b_Y(Y_t) dt + \sigma_Y(Y_t)dB_t, \\ d\langle B, W \rangle_t &= \rho dt, \rho \in (-1, 1). \end{aligned}$$

We suppose that the diffusion $a_S = \sigma_S^2$, $a_Y = \sigma_Y^2$ and drift b_Y coefficients are infinitely differentiable with bounded derivatives, a_S and a_Y are bounded but b_Y is not necessarily bounded. Typically the drift corresponds to a mean reversion term, *i.e.* $b_Y(y) = \lambda(\mu - y)$, $\lambda > 0$ as in Stein-Stein model [SS91]. In addition we suppose uniform ellipticity, *i.e.* $\exists \kappa \geq 1$, $\forall x \in \mathbb{R}$, $\kappa^{-1} \leq a_S(x) \leq \kappa$, $\kappa^{-1} \leq a_Y(x) \leq \kappa$. Rather than use the spot price, we instead take its logarithm $X_t = \ln(S_t)$ and the couple $(X_t, Y_t)_{t \in [0, T]}$ with initial conditions $X_0 = x_0, Y_0 = y_0$ is considered:

$$dX_t = \left(r - \frac{1}{2}a_S(Y_t) \right) dt + \sigma_S(Y_t) dW_t.$$

2.3.1 Probabilistic representation of spot-volatility couple

The principal technique for the probabilistic representation of the marginal law is the same as in the previous part for McKean-Vlasov equation, in the spirit of the unbiased simulation method introduced [BKH+15]. The major difference compared to other works using this method, see for example in [FKHL+19] for killed processes, is that we have to take account the unboundedness of the drift b_Y of the volatility process. In

2.3. Probabilistic representation and integration by parts formulae for some stochastic volatility model with unbounded drift

order to circumvent this difficulty, introduce the solution $(m_t(y_0))_{t \in [0, T]}$ to the ordinary differential equation $dm_t = b_Y(m_t) dt$, $m_0 = y_0$ obtained by removing the diffusion term in the dynamics of Y . The approximation process (\bar{X}, \bar{Y}) is obtained by freezing the coefficients in the dynamics of couple (X, Y) along the path of m :

$$\begin{aligned}\bar{X}_t^{x_0} &= x_0 + \int_0^t \left(r - \frac{1}{2} a_S(m_s)\right) ds + \int_0^t \sigma_S(m_s) dW_s, \\ \bar{Y}_t^{y_0} &= y_0 + \int_0^t b_Y(m_s) ds + \int_0^t \sigma_Y(m_s) dB_s.\end{aligned}$$

As in the previous part, the sequence of jump times $\tau = (\tau_n)_{n \geq 1}$ and its associated renewal process $N = (N_t)_{t \geq 0}$, $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$ are considered, N is independent of the bidimensional Brownian motion (W, B) . The Euler scheme for (\bar{X}, \bar{Y}) is defined on the random time grid $(\zeta_i)_{i \geq 0}$ with $\zeta_i = \tau_i \wedge T$, $\zeta_0 = 0$:

$$\begin{aligned}\bar{X}_{i+1} &= \bar{X}_i + \left(r(\zeta_{i+1} - \zeta_i) - \frac{1}{2} a_{S,i}\right) + \sigma_{S,i} Z_{i+1}^1, \\ \bar{Y}_{i+1} &= m_i + \sigma_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2\right), \\ a_{S,i} &= \sigma_{S,i}^2 = \int_0^{\zeta_{i+1} - \zeta_i} a_S(m_s(\bar{Y}_i)) ds, \\ a_{Y,i} &= \sigma_{Y,i}^2 = \int_0^{\zeta_{i+1} - \zeta_i} a_Y(m_s(\bar{Y}_i)) ds, \\ \sigma_{S,Y,i} &= \int_0^{\zeta_{i+1} - \zeta_i} (\sigma_S \sigma_Y)(m_s(\bar{Y}_i)) ds, \quad \rho_i = \rho \frac{\sigma_{S,Y,i}}{\sigma_{S,i} \sigma_{Y,i}}.\end{aligned}$$

We use the notation $m_i = m_{\zeta_{i+1} - \zeta_i}(\bar{Y}_i)$ for the solution at time $\zeta_{i+1} - \zeta_i$ of the ODE with initial condition \bar{Y}_i , $Z = (Z_n^1, Z_n^2)_{n \geq 1}$ is a sequence of i.i.d. random variables of law $\mathcal{N}(0, I_2)$ and $\sigma'_{S,i}$, $\sigma'_{Y,i}$, $\sigma'_{S,Y,i}$, ρ'_i and m'_i are derivatives with respect to \bar{Y}_i .

The Euler scheme (\bar{X}, \bar{Y}) that we just defined can be seen as a Markov chain with respect to the filtration $\mathcal{G} = (\mathcal{G}_i)_{i \geq 0}$ where $\mathcal{G}_i = \sigma(Z_j, 0 \leq j \leq i)$. The tailor-made Malliavin calculus for Markov chain (\bar{X}, \bar{Y}) can be applied, we begin by defining the set $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, $n \in \mathbb{N}$, $i \in \{0, \dots, n\}$ as the space of random variables H satisfying:

- $H = h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1})$ on $\{N_T = n\}$, where $\zeta^{n+1} := (0 = \zeta_0, \zeta_1, \dots, \zeta_n, \zeta_{n+1} = T)$
- For all $\sigma_{n+1} \in \Delta_{n+1}(T) := \{\sigma_{n+1} \in [0, T]^n : 0 \leq \sigma_1 < \dots < \sigma_{n+1} \leq T\}$, the function $h(\cdot, \sigma_{n+1}) \in \mathcal{C}_p^\infty(\mathbb{R}^4)$.

First, let us define the derivative operator $\mathcal{D}_{i+1}^{(\alpha)} H$, $\alpha \in \{1, 2\}$ for $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$:

$$\mathcal{D}_{i+1}^{(1)} H = \partial_{\bar{X}_{i+1}} H \text{ and } \mathcal{D}_{i+1}^{(2)} H = \partial_{\bar{Y}_{i+1}} H.$$

The integral operator satisfies the following duality identity, with notation for conditional expectation $\mathbb{E}_{i,n}[\cdot] = \mathbb{E}[\cdot | \mathcal{G}_i, \tau^{n+1}, N_T = n]$:

$$\mathbb{E}_{i,n} \left[\mathcal{D}_{i+1}^{(\alpha)} f(\bar{X}_{i+1}, \bar{Y}_{i+1}) H \right] = \mathbb{E}_{i,n} \left[f(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(\alpha)}(H) \right].$$

The previous identity can be generalized to the case where α is a multi-index $\alpha = (\alpha_1, \dots, \alpha_p)$, $\alpha_i \in \{1, 2\}$ by iterating the previous derivative and integral operators. For an index $\alpha_{p+1} \in \{1, 2\}$, we have the following relations:

$$\mathcal{I}_{i+1}^{(\alpha, \alpha_{p+1})}(H) = \mathcal{I}_{i+1}^{(\alpha_{p+1})}(\mathcal{I}_{i+1}^{(\alpha)}(H)), \quad \mathcal{D}_{i+1}^{(\alpha, \alpha_{p+1})}H = \mathcal{D}_{i+1}^{(\alpha_{p+1})}(\mathcal{D}_{i+1}^{(\alpha)}H).$$

Define for $\gamma > 0$, $\mathcal{B}_\gamma(\mathbb{R}^2)$ the set of Borel functions defined on \mathbb{R}^2 with exponential growth at infinity:

$$\exists C > 0, \forall (x, y) \in \mathbb{R}^2, |h(x, y)| \leq C \exp(\gamma(|x|^2 + |y|^2)).$$

Main Result 2.3.1. *Under the previous regularity and uniform ellipticity assumptions, the law of the couple (X_T, Y_T) at time T satisfies the following probabilistic representation: there exists $\gamma > 0$ such that for all $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, the expectation (or the price of a product paying $h(X_T, Y_T)$ at maturity T) satisfies:*

$$\mathbb{E}[h(X_T, Y_T)] = \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i\right]$$

where, on the set $\{N_T = n\}$, the random variables $\theta_i \in \mathbb{S}_{i-1, n}(\bar{X}, \bar{Y})$, for $i = 1, \dots, N_T$, are defined by:

$$\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{I}_i^{(1)}(c_S^i) + \mathcal{I}_i^{(2,2)}(c_Y^i) + \mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{I}_i^{(1,2)}(c_{Y,S}^i) \right]$$

and for $i = N_T + 1$, we have: $\theta_{N_T+1} = (1 - F(T - \zeta_{N_T}))^{-1}$.

If for example N is a renewal process with jump times of distribution $\text{Beta}(1/2, 1)$, then the random variable $h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i$ has finite $L^p(\mathbb{P})$ moment for all $p \geq 1$.

2.3.2 Integration by parts formulae for sensitivities

Delta (respectively Vega) is defined by the derivative with respect to the spot price (respectively the volatility) of the underlying asset. We establish Bismut-Elworthy-Li type formulae for the following quantities which could be seen as Delta and Vega at initial time $t = 0$, with $s_0 = \exp(X_0)$, $y_0 = Y_0$:

$$\partial_{s_0} \mathbb{E}[h(X_T, Y_T)] \text{ and } \partial_{y_0} \mathbb{E}[h(X_T, Y_T)].$$

The goal is then to write the derivatives of expectations above as expectations of some functions. First we use the probabilistic representation for the expectation. The next step is to do the "transfer of the derivative" inside the expectation using an integration by parts formula, but due to some integrability issues, we rather apply local IBP formulae to each of the random time intervals $[\zeta_i, \zeta_{i+1}]$, $i = 0, \dots, N_T$. Finally, the derivatives of expectations, or Delta and Vega, can be written as weighted sums of local IBP formulae, the weights being equal to the lengths of random intervals.

Main Result 2.3.2. *Under the previous assumptions of regularity and uniform ellipticity, the law of the couple (X_T, Y_T) at time T satisfies the following Bismut-Elworthy-Li*

type formulae. There exists $\gamma > 0$ such that for all $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ and all $(s_0, y_0) \in \mathbb{R}^2$:

$$s_0 T \partial_{s_0} \mathbb{E} \left[h(X_T, Y_T) \right] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \bar{\theta}^{\mathcal{I}_k^{(1), N_T+1}} \right].$$

and

$$\begin{aligned} & T \partial_{y_0} \mathbb{E} \left[h(X_T, Y_T) \right] \\ &= \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\bar{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \bar{\theta}^{C_j^{N_T+1}} + \bar{\theta}_j^{\mathcal{I}_k^{(1), N_T+1}} \right) \right] \end{aligned}$$

where $\bar{\theta}^{\mathcal{I}_k^{(1), n+1}}$, $\bar{\theta}^{C_j^{n+1}}$, $\bar{\theta}^{\mathcal{I}_k^{(2), n+1}}$ and $\bar{\theta}_j^{\mathcal{I}_k^{(1), n+1}}$ with $n \geq 0$ on $\{N_T = n\}$, $k \in \{1, \dots, n+1\}$ and $1 \leq j \leq k$, are explicit functions of the weights θ_i and the parameters of the model.

If N is again a renewal process with jump times of distribution $\text{Beta}(1/2, 1)$, the random variables inside the expectations appearing on the right-hand side above have finite $L^p(\mathbb{P})$ moments for all $p \geq 1$.

The major interest of the above integration by parts formulae lies in the fact that they lead to an unbiased Monte-Carlo method with optimal complexity. In fact, the random variables that appear inside the expectations on the right-hand side can perfectly be simulated. This is in contrast with the classical integration by parts formulae that could be obtained in this context using standard Malliavin calculus tools as in Nualart [Nua06].

2.4 Numerical methods for FBSDEs arising from mean-field games

In this last part of the manuscript, we investigate the numerical solutions of mean-field games, especially for linear-quadratic games. There are two main approaches, the first one is analytic and we use systems of partial differential equations. The second approach is studied here, we write the associated stochastic control problem of mean-field type as a fully coupled forward-backward stochastic differential equation of McKean-Vlasov type.

2.4.1 Mean-field games and probabilistic approach

Let us start with stochastic differential games with N players in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$, the state X_t^i of each player $i \in \{1, \dots, N\}$ is given by the following stochastic differential equation:

$$dX_t^i = b^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i) dt + \sigma^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i) dW_t^i.$$

Each player chooses their control $\alpha^i = (\alpha_t^i)_{0 \leq t \leq T}$ in a set of admissible controls \mathbb{A} and $\bar{\mu}_t = \frac{1}{N} \sum_{j=0}^N \delta_{X_t^j} \in \mathcal{P}_2(\mathbb{R}^d)$ is the empirical distribution of the states of players where $\mathcal{P}_2(\mathbb{R}^d)$ is the space of probability measures with finite second order moment and

W_2 is the associated Wasserstein distance. The drift b^i and volatility σ^i are Lipschitz continuous with respect to all their arguments and the W_2 distance is used between measure variables.

The cost function J^i that the player i tries to minimize is defined as the sum of running f^i and terminal g^i costs that depend on the controls of all players $\alpha = (\alpha^1, \dots, \alpha^N)$:

$$J^i(\alpha) = \mathbb{E} \left[\int_0^T f^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i) dt + g^i(X_T^i, \bar{\mu}_T) \right].$$

The players are in *Nash equilibrium* if each player is no better off for switching their strategy when they consider the other players' strategies to be fixed. The set of strategies of players is a Nash equilibrium if:

$$J^i(\alpha) \leq J^i(\alpha^1, \dots, \alpha^{i-1}, \alpha, \alpha^{i+1}, \dots, \alpha^N), \forall \alpha \in \mathbb{A}, \forall i \in \{1, \dots, N\}.$$

To solve the Nash equilibrium for N players, we actually have a system of N equations which is complicated to solve. The idea of mean-field games is to solve the system when N goes to infinity and we should suppose identical players, *i.e.* $b = b^i$, $\sigma = \sigma^i$, $f = f^i$, $g = g^i$ and $\mathbb{A} = A$, see [LL07] and [CD+18].

Instead of considering the system of N players, we rather have an infinity of players and we just need to focus on a representative player. The limit is justified for example in [Szn91] for the propagation of chaos for a particle system. In this case, the Nash equilibrium can be seen as a fixed point, which is solved in two steps:

1. For a fixed flow of probability measures $\mu = (\mu_t)_{0 \leq t \leq T} \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$, we solve the classical stochastic control problem:

$$\inf_{\alpha \in A} J^\mu(\alpha) = \mathbb{E} \left[\int_0^T f(t, X_t^\alpha, \mu_t, \alpha_t) dt + g(X_T^\alpha, \mu_T) \right] \quad (2.4.1)$$

with the dynamics of $(X_t^\alpha)_{0 \leq t \leq T}$:

$$\begin{aligned} dX_t^\alpha &= b(t, X_t^\alpha, \mu_t, \alpha_t) dt + \sigma(t, X_t^\alpha, \mu_t, \alpha_t) dW_t, \\ X_0^\alpha &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d). \end{aligned}$$

2. Find a fixed point, μ such that $\mathcal{L}(X_t^\alpha) = \mu_t$ for all $0 \leq t \leq T$.

This problem can be solved by an analytical approach with Jacobi-Bellman and Fokker-Planck partial differential equations. Here we focus on two probabilistic approaches which enable us to write fully coupled forward-backward stochastic differential equations of mean-field type (FBSDEs).

The first approach is called *weak approach*, for a fixed flow of measures, $\mu = (\mu_t)_{0 \leq t \leq T}$, we define the value function u :

$$u(t, x) = \inf_{(\alpha_s)_{t \leq s \leq T} \in \mathbb{A}} \mathbb{E} \left[\int_t^T f(s, X_s, \mu_s, \alpha_s) ds + g(X_T, \mu_T) \mid X_t = x \right].$$

The process $(Y_t)_{0 \leq t \leq T}$, $Y_t = u(t, X_t)$ is defined as the value function along the state process $(X_t)_{0 \leq t \leq T}$. For the sake of simplicity, we suppose that $\sigma > 0$ is constant. Under

regularity and boundedness conditions, the couple $(X_t, Y_t)_{0 \leq t \leq T}$ satisfy the following FBSDE:

$$\begin{aligned} dX_t &= b(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1}Z_t)) dt + \sigma dW_t, \quad X_0 = \xi, \\ dY_t &= -f(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1}Z_t)) dt + Z_t dW_t, \quad Y_T = g(X_T, \mu_T). \end{aligned}$$

The function $\hat{\alpha}$ is the unique minimiser of the Hamiltonian H :

$$\begin{aligned} H(t, x, \mu, \alpha, y) &= b(t, x, \mu, \alpha) \cdot y + f(t, x, \mu, \alpha), \\ \hat{\alpha}(t, x, \mu, y) &= \arg \inf_{\alpha \in A} H(t, x, \mu, \alpha, y). \end{aligned}$$

From [CD+18], the flow of measures $\mu = (\mu_t)_{0 \leq t \leq T}$ is a Nash equilibrium of the mean-field game if and only if μ corresponds to the law of X , *i.e.* $\mu_t = \mathcal{L}(X_t), \forall t \in [0, T]$. We finally obtain the FBSDE of the weak approach in $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$:

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1}Z_t)) dt + \sigma dW_t, \quad X_0 = \xi, \\ dY_t &= -f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1}Z_t)) dt + Z_t dW_t, \quad Y_T = g(X_T, \mathcal{L}(X_T)). \end{aligned}$$

The second approach, called *Pontryagin approach*, is based on the Pontryagin stochastic principle. For the process Y , instead of the value function u , we take the derivative with respect to its space variable, *i.e.* $Y_t = \nabla_x u(t, X_t)$. The assumptions on b, f and g are stronger, especially the function of running cost f is continuously differentiable with Lipschitz and bounded derivatives, moreover f is strictly convex, for all $(t, x, x', \mu, \alpha, \alpha') \in [0, T] \times (\mathbb{R}^d)^2 \times \mathcal{P}_2(\mathbb{R}^d) \times A^2$:

$$f(t, x', \mu, \alpha') - f(t, x, \mu, \alpha) - (x - x', \alpha - \alpha') \cdot \partial_{(x, \alpha)} f(t, x, \mu, \alpha) \geq C_1 |\alpha' - \alpha|^2.$$

Similarly, the FBSDE of Pontryagin approach with solution (X, Y, Z) is written as follows:

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t)) dt + \sigma dW_t, \quad X_0 = \xi, \\ dY_t &= -[\nabla_x b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t))] \cdot Y_t \\ &\quad + \nabla_x f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t)) dt + Z_t dW_t, \quad Y_T = \nabla_x g(X_T, \mathcal{L}(X_T)). \end{aligned}$$

An advantage of Pontryagin approach is that the process Y is more closely related to the optimal control, especially in the case of linear-quadratic games as we will see in numerical examples.

The general system of fully coupled FBSDE in both approaches can be written as follows, with $[X_t]$ representing the law of the random variable X_t and we are in the unidimensional case to simplify the discussions:

$$\begin{aligned} dX_t &= B(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + \sigma dW_t, \quad X_0 = \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}), \\ dY_t &= -F(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + Z_t dW_t, \quad Y_T = G(X_T, [X_T]). \end{aligned}$$

2.4.2 Tree and grid algorithms for FBSDEs

Let us define the decoupled Picard scheme $\Phi_{\xi, G}$, which takes the $j - 1$ Picard iterate and produces the j Picard iterate:

$$\Phi_{\xi, G} : (X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}]) \mapsto (X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$$

where ξ and G respectively correspond to the initial and final data of the problem. Actually, $\Phi_{\xi,G}$ is naturally defined from the general FBSDE:

1. Solve for $(X^j, [X^j])$:

$$\begin{aligned} dX_t^j &= B(t, X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}, [X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}])dt + \sigma dW_t, \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}). \end{aligned}$$

2. Then solve for $(Y^j, Z^j, [Y^j], [Z^j])$:

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^{j-1}, Z_t^{j-1}, [X_t^j, Y_t^{j-1}, Z_t^{j-1}])dt + Z_t^j dW_t, \\ Y_T^j &= G(X_T^j, [X_T^j]). \end{aligned}$$

3. Return $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$.

For a given initialization $(X^0, Y^0, Z^0, [X^0, Y^0, Z^0])$, it is clear that $\Phi_{\xi,G}$ defines a sequence $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])_{j \geq 0}$. If this sequence converges to $(X, Y, Z, [X, Y, Z])$, it is then a fixed point of the scheme $\Phi_{\xi,G}$ and $(X, Y, Z, [X, Y, Z])$ solves the general FBSDE. The problem is that only for small values of time horizon T , the scheme $\Phi_{\xi,G}$ is contracting, which may guarantee the existence of a fixed point.

The idea of *continuation in time*, proposed in [CCD+19], consists of dividing the time intervals to small intervals, called levels that we suppose uniform with size $h_\ell = T/N_\ell$:

$$\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}.$$

Then we use Picard iterations recursively between the levels. Let us begin by defining the scheme $picard[k](X, Y_{T_{k+1}})$ pour level $[T_k, T_{k+1}]$:

1. Initialize $Y_t = 0$ and $Z_t = 0, \forall t \in [T_k, T_{k+1}]$.
2. For $1 \leq j \leq J_p$ (number of Picard iterations):

$$(X, Y, Z, [X, Y, Z]) = \Phi_{X_{T_k}, Y_{T_{k+1}}}(X, Y, Z, [X, Y, Z])$$

3. Return $(X, Y, Z, [X, Y, Z])$.

The global solver, denoted as *solver*, applies the continuation in time and is defined recursively and backwards for a prescribed integer $J_s > 0$. We have, for each level k :

1. Initialize $X_t = X_{T_k}, Y_t = 0$ and $Z_t = 0, \forall t \in [T_k, T_{k+1}]$.
2. For $1 \leq j \leq J_s$
 - (a) $(Y_{T_{k+1}}, [Y_{T_{k+1}}]) = solver[k+1](X_{T_{k+1}}, [X_{T_{k+1}}])$
 - (b) $(X, Y, Z, [X, Y, Z]) = picard[k](X, Y_{T_{k+1}})$
3. Return $(Y_{T_k}, [Y_{T_k}])$.

Before presenting the algorithms, let us write the discretization of the SDEs in the decoupled Picard scheme $\Phi_{\xi, G}$ on a time grid $\{0 = t_0 < t_1, \dots, t_i = ih, \dots, t_{N_t-1} < t_{N_t} = T\}$ with $h = T/N_t$ which is supposed to be a refinement of the grid corresponding to the levels $\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}$:

$$\begin{cases} X_{t_{i+1}}^j = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) + \sigma \Delta W_i, & X_0^j = \xi, \\ Y_{t_i}^j = \mathbb{E}_{t_i}(Y_{t_{i+1}}^j) + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]), & Y_T^j = G(X_T^{j-1}, [X_T^{j-1}]), \\ Z_{t_i}^j = h^{-1} \mathbb{E}_{t_i}(Y_{t_{i+1}}^j \Delta W_i), & Z_T^j = 0. \end{cases}$$

Notice that this discretized scheme is decoupled, as opposed to the original FBSDE. We are then able to solve the backward scheme then the forward scheme to obtain $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$ from $(X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}])$ where $X^{j-1} = (X_{t_i}^{j-1})_{0 \leq i \leq N_t}$.

Our first algorithm, the *tree algorithm*, has been presented in [CCD+19]. The structure of binary tree comes from the binary approximation of increments of Brownian motion: $\Delta W_i = \pm\sqrt{h}$ with probability 1/2. Each node at depth i represents a value of X_{t_i} and has two children corresponding to values $X_{t_{i+1}}$ ("up \uparrow " and "down \downarrow ") which are calculated as follows:

$$X_{t_{i+1}}^j(\uparrow\downarrow) = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j, [X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j]) \pm \sigma\sqrt{h}.$$

If M points are used for the approximation of the initial distribution ξ of process X , we get M parallel binary trees. In order to know the marginal law of X at time t_i , we just look at the values at depth i and there are $M \times 2^i$ values. The complexity of tree algorithm is then exponential in number of time steps.

For the backward scheme we reuse the binary tree, at final time $T = t_{N_t}$, we have $Y_T^j = G(X_T^j, [X_T^j])$ for each of the $M \times 2^{N_t}$ nodes. Then the conditional expectation corresponds simply to the average of "up" and "down" branches. The Picard iteration is initialized at $j = 0$ by $X_{t_i} = \xi, \forall i \in \{0, \dots, N_t\}$.

The main contribution of this part is the *grid algorithm* on the marginal laws, inspired by [DM+06]. We use the notion of decoupling field. More precisely, there exist deterministic feedback functions $(u, v) : [0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}$ with u being solution of a certain PDE, see for example [CCD14], such that the solution of the general system of FBSDE satisfies:

$$Y_t = u(t, X_t, [X_t]) \quad \text{and} \quad Z_t = v(t, X_t, [X_t]).$$

The law of the solution $([X_t, Y_t, Z_t])_{0 \leq t \leq T}$ can be characterized by $(\mu_t, u(t, \cdot), v(t, \cdot))_{0 \leq t \leq T}$ with $[X_t] = \mu_t$. We then focus on the approximation of μ and (u, v) . In this grid algorithm, only the marginal laws at each time step are calculated, instead of pathwise laws previously.

Our decoupled Picard scheme $\Psi_{[\xi], G}$ is then redefined by:

$$\Psi_{[\xi], G} : (\mu_t^{j-1}, u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))_{0 \leq t \leq T} \mapsto (\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T}, \quad \mu_0^{j-1} = \mu_0^j = [\xi]$$

1. Solve for X^j the following SDE and define $\mu_t^j := [X_t^j]$:

$$\begin{aligned} dX_t^j &= B(t, X_t^j, u^{j-1}(t, X_t^j), v^{j-1}(t, X_t^j), \mu_t^{j-1}, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))) \# \mu_t^{j-1}) dt + \sigma dW_t, \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}). \end{aligned}$$

with $\varphi\#\nu$ the push-forward measure of the measure ν by a function φ .

2. Find $(u^j, v^j)(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^2$ by $(u^j, v^j)(t, X_t^j) := (Y_t^j, Z_t^j)$ by solving:

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^j, Z_t^j, \mu_t^j, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))\#\mu_t^j)dt + Z_t^j dW_t, \\ Y_T^j &= G(X_T^j, \mu_T^j). \end{aligned}$$

3. Return $((\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T})$.

In order to implement our algorithm, we need to be able to calculate (μ_t^j, u^j, v^j) explicitly, we use a spatial discretization grid for the forward process X . For example a uniform grid Γ of step size Δx , with associated projection function Π :

$$\Gamma = \{x_k = x_1 + (k-1)\Delta x, k = 1, \dots, N_x\}.$$

The process X is defined on the grid Γ :

$$\begin{aligned} X_{t_{i+1}}^j &= \Pi \left(X_{t_i}^j + h B(X_{t_i}^j, (u_i^{j-1}, v_i^{j-1})(X_{t_i}^j), \mu_i^{j-1}, (u_i^{j-1}, v_i^{j-1})\#\mu_i^{j-1}) + \sigma \Delta W_i \right) \\ [X_{t_i}^j] &= \mu_i^j(\cdot) \text{ and } \mu_{t_{i+1}}^j(\cdot) = [X_{t_{i+1}}^j]. \end{aligned}$$

The marginal law of the process X at time t_i of Picard iterate j is defined by a weighted average of Dirac measure on the grid Γ :

$$\mu_i^j(\cdot) = \sum_{k=1}^{N_x} p_{i,k}^j \delta_{x_k}(\cdot), \quad p_{i,k}^j \geq 0 \quad \forall k \in \{1, \dots, N_x\} \text{ and } \sum_{k=1}^{N_x} p_{i,k}^j = 1.$$

The marginal law $\mu_{t_{i+1}}^j(\cdot)$, or rather the weights $(p_{i+1,n}^j, n \in \{1, \dots, N_x\})$ at time t_{i+1} is the convolution of $\mu_i^j(\cdot)$ with the transition probability $\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k)$, $k, n \in \{1, \dots, N_x\}$:

$$p_{i+1,n}^j = \sum_{k=1}^{N_x} p_{i,k}^j \times \mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k)$$

In order to calculate the transition probabilities, we use for example binary approximation of Brownian increments again, although it is easy to use other approximations with more than two points:

$$\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k) = \frac{1}{2} \left(\mathbf{1}(X_{t_{i+1}}^j(\uparrow) = x_n | X_{t_i}^j = x_k) + \mathbf{1}(X_{t_{i+1}}^j(\downarrow) = x_n | X_{t_i}^j = x_k) \right)$$

At each Picard iterate j , the forward scheme provides us the flow of measures $(\mu_i^j)_{i=0}^{N_t}$ at times $(t_i)_{i=0}^{N_t}$. On the other hand, the backward scheme gives $u_i^j(x)$ and $v_i^j(x)$ for $x \in \Gamma$ on the grid. With the terminal condition $T = t_{N_t}$, $(u_{N_t}^j, v_{N_t}^j) = (G, 0)$ we have, for $i \leq N_t - 1$:

$$\begin{aligned} u_i^j(x) &= \mathbb{E} \left(u_{i+1}^j(X_{t_{i+1}}^j, \mu_{i+1}^j) + h \cdot F(X_{t_i}^j, u_i^{j-1}(X_{t_i}^j), v_i^j(X_{t_i}^j), \mu_i^j, (u_i^{j-1}, v_i^{j-1})\#\mu_i^j) \mid X_{t_i}^j = x \right) \\ v_i^j(x) &= \mathbb{E} \left(u_{i+1}^j(X_{t_{i+1}}^j) \cdot \Delta W_i / h \mid X_{t_i}^j = x \right) \end{aligned}$$

The variable $X_{t_{i+1}}^j$ taking values in Γ and its law $[X_{t_{i+1}}^j] = \mu_{i+1}^j$ of support Γ , are calculated by the forward scheme starting from $X_{t_i}^j = x \in \Gamma$. We have a simplified expression by using binary approximation again for ΔW_i .

Finally, the previous decoupled Picard scheme $\Phi_{\xi, G}$ in tree algorithm is replaced by the forward-backward scheme for marginal laws on the grid $\Psi_{[\xi], G}$ that we just defined, we can apply continuation in time with schemes similar to *picard*[k] and *solver* to complete the numerical implementation.

Part I

Modeling the market by capital distribution

Chapter 3

Modeling the market by capital distribution

The content of this chapter is from a work in progress, in collaboration with my PhD supervisor Peter Tankov.

3.1 Introduction

Since it has been introduced by Fernholz in [FS82] then in [Fer02], a considerable number of models have been developed in *stochastic portfolio theory*, or SPT for short, with the aim of studying equity market structure, as well as analyzing and optimizing portfolio performance. The portfolio of natural interest is the market portfolio, as it defines the market structure, and other portfolios are simply constructed by investing different quantities in the assets of the market portfolio. The components of market portfolios are companies' market capitalizations, or companies' share prices. In classical SPT, with zero interest rate for simplicity, the portfolio components, *i.e.* capitalizations or prices of assets are denoted as a vector of dimension n : $(S_t)_{t \geq 0}$, $S_t = (S_t^i)_{i=1}^n$ and each $(S_t^i)_{t \geq 0}$ has its dynamics for example given by the following stochastic differential equation, or EDS for short:

$$dS_t^i = S_t^i (b^i(t)dt + \sum_{\nu=1}^d \sigma^{i,\nu}(t)dW_t^\nu), \quad S_0^i = s^i > 0$$

The total capitalization of the market $(\Sigma_t)_{t \geq 0}$ is simply the sum of all capitalizations $\Sigma_t = \sum_{i=1}^n S_t^i$.

More specifically, most of the models in SPT, based on different assumptions on market behavior, see [Ver15], can be categorized as follows:

1. Rank-based models, with the Atlas model [BFK+05] and the mean-field Atlas model [JR15].
2. Diverse models [FKK05], where no asset can dominate the market.
3. Intrinsically volatile models, with volatility-stabilized model [FK05], its generalized version [Pic14] and polynomial SPT models [Cuc19].

In rank-based models, as their name suggests, we focus on the ranks of the assets' capitalizations (or prices). In Atlas model only the asset with the lowest capitalization has non-zero positive growth and the capitalizations are uncorrelated. In the more recent mean-field Atlas model, the capitalizations interact through their ranks and we naturally study the behavior of an interacting particle system. An advantage of Atlas model (as well as for its mean-field version) is that the capital distribution curves observed in the market can be matched. However, it is in general difficult to calibrate rank-based models to real market data because of limited tractability.

When the market is diverse, that is, the total market capitalization is not concentrated in one asset, together with the additional assumption of non-degeneracy of the instantaneous covariance matrix of log capitalizations, it is possible to generate a relative arbitrage with respect to the market, or equivalently, outperform the market portfolio systematically on arbitrary time horizon, see [FKK05]. Alternatively, when the cumulative growth rate has its slope bounded away from 0, it is possible to find an arbitrage relative to the market, at least for sufficiently long time horizons.

For intrinsically volatile models, the total market capitalization follows a classical Black-Scholes model, more importantly, the assets with lower capitalizations tend to have larger volatility which is coherent with market observations. The main limitation of the volatility-stabilized model in [FK05] is the absence of correlations between assets' capitalizations and with the total market capitalization. In the more recent polynomial SPT models introduced in [Cuc19], the assets' capitalizations and the total market capitalization are modeled by polynomial processes and there is more flexibility for the correlation structure.

The properties that are frequently investigated in the SPT literature such as capital distribution, relative arbitrage, diversity, and functional generation of portfolios, concern only the relative weights of assets in the market portfolio. It is then natural to directly consider the stochastic process of relative weights, defined as the percentage of the asset's capitalization with respect to total market capitalization *i.e.* $(\boldsymbol{\mu}_t)_{t \geq 0}$, $\boldsymbol{\mu}_t = (\mu_t^i)_{i=1}^n$, $\mu_t^i = S_t^i / \Sigma_t$, which leads to a class of processes with values in the unit simplex, whose components are nonnegative and of sum 1. In this chapter, the process of relative weights in the market, also called *markets weights process* representing the evolution through time of relative weights of assets, is given by the unique strong solution of the *market weights equation* and the process of total market capitalization $(\Sigma_t)_{t \geq 0}$ is defined as another process, which could be correlated to the market weights process. The couple $(\boldsymbol{\mu}_t, \Sigma_t)_{t \geq 0}$ provides an equivalent description of the market to that given by the vector of capitalizations $(\boldsymbol{S}_t)_{t \geq 0}$. Although the process of relative weights in the volatility-stabilized model, also known as Jacobi or Wright-Fisher process has been studied in [Pal+11] and the direct modeling of market weights using polynomial processes has been done in [Cuc19], the market weights process studied in this chapter has some particular interests:

1. The market weights process has regular density and has unique stationary distribution under adequate assumptions, which ensures the long-term behavior, or ergodicity of capital distribution in the market.
2. We introduce a parameter which can be seen as a measure of stabilization of volatility, thus our model has more flexibility and the covariance structure be-

tween weights is richer compared to the classical volatility-stabilized model as well as the polynomial process model.

3. Like for polynomial process models, volatility-stabilized market model can be seen as a special case of our general model.
4. The model remains tractable, with possibility to calibrate to real market data. The ergodicity of market weights process also opens the door for some ergodic control problems with the aim of maximizing long-term wealth.

The chapter is structured as follows: First we introduce the stochastic differential equation depending on an exponent parameter ζ with values in the unit simplex, satisfied by the process of relative weights, called market weights equation. Under standard assumptions, this equation has unique strong solution until the boundary attainment time of the unit simplex. The boundary of the unit simplex is attained if and only if some conditions specified in Theorem 3.2.1 are fulfilled. The process of relative weights under some other models, especially the well-known volatility-stabilized market model, can be seen as a special case of our model. Market weights equation, being a degenerate diffusion, has properties such as local ellipticity, see Proposition 3.2.1. The long-term behavior of the market is ensured by the ergodicity of market weights process, in fact the stationary distribution (or invariant measure) of market weights exists and is unique under appropriate assumptions specified in Theorem 3.2.2.

The second part of the chapter is more market-related, the growth rate of the asset is decomposed into the part associated to total capitalization and the part specific to its weight. By changing the numéraire to that of market portfolio, trading strategy and relative arbitrage are redefined in terms of market weights. Next, the conditions for the existence of relative arbitrage are studied. The first condition is related to an estimate of the excess growth rate, see Proposition 3.3.1. The second one is a necessary and sufficient condition involving the martingality of the market risk deflator, and is stated in Proposition 3.3.2. For a portfolio with two assets, this latter condition is translated into the value of the exponent ζ , which can be seen as a measure of volatility stabilization, see Theorem 3.3.2. In the general case of arbitrary number of assets, it is harder to check whether the market risk deflator is a true martingale. We show in Corollary 3.3.1 that there is no relative arbitrage to the market for $\zeta \geq 1$, when the effect of volatility stabilization is absent as pointed out in Remark 3.3.3.

Finally we discuss the applications of our model and some new questions that could be addressed in future research. The notion of functionally generated portfolio allowing to construct new portfolios from market weights could be explored in our model. The calibration of model to real market data can be explored, as shown for the case $\zeta = 1$. The ergodic control for long-term optimization is an interesting problem for future research, which appears more more difficult to address since most results for existence and uniqueness of the solution require uniform ellipticity of the process, which is not the case for the market weights process in our model.

Remarks on notation We denote by Δ^n (resp. Δ_+^n), the unit simplex (resp. its interior):

$$\Delta^n \text{ (resp. } \Delta_+^n) = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^n \mid \mu^i \geq 0 \text{ (resp. } \mu^i > 0) \forall i \in \{1, \dots, n\} \text{ and } \sum_{i=1}^n \mu^i = 1 \right\}$$

We also introduce the boundary of the unit simplex $\partial\Delta^n = \Delta^n \setminus \Delta_+^n$ and for $i \in \{1, \dots, n\}$, the subset of the boundary $\{\mu^i = 0\} = \{\boldsymbol{\mu} \in \partial\Delta^n \mid \mu^i = 0\}$. The unit vector \mathbf{e}^i has all its coordinates equal to 0 except the i -th which is equal to 1.

For $\epsilon \in (0, 1/n)$, set Δ_ϵ^n is defined as the closed subset of Δ_+^n where all components are no less than ϵ :

$$\Delta_\epsilon^n = \left\{ \boldsymbol{\mu} = (\mu^i)_{i=1}^n \mid \mu^i \geq \epsilon \forall i \in \{1, \dots, n\} \text{ and } \sum_{i=1}^n \mu^i = 1 \right\} \quad (3.1.1)$$

The set of real-valued matrices of dimension $n \times d$ is denoted as $\mathbb{M}^{n,d}(\mathbb{R})$, when it is a square matrix of dimension $n \times n$ we just denote $\mathbb{M}^n(\mathbb{R})$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space with a right continuous filtration.

3.2 Equation and process of market weights

In this chapter, we consider a financial market with n risky assets and we use the vector $\mathbf{S}_t = (S_t^i)_{i=1}^n$ to denote either their prices at time t , or more frequently their capitalizations defined the product of price and number of shares outstanding. The *market weights process* is defined by $\mu_t^i := S_t^i / \Sigma_t$, where $\Sigma_t = \sum_{i=1}^n S_t^i$ is the *total market capitalization process*. By construction, the market weights process takes values in the unit simplex: $\boldsymbol{\mu}_t \in \Delta^n$ for all $t \geq 0$. The couple $((\boldsymbol{\mu}_t, \Sigma_t))_{t \geq 0}$ thus provides an equivalent description of the market prices. Instead of modeling the prices directly, our approach is to model separately the process of market weights process and the process of total market capitalization, and in this section we focus on the market weights process $(\boldsymbol{\mu}_t)_{t \geq 0}$ and propose a model for it.

3.2.1 Market weights equation and its first properties

In this chapter, the dynamics of market weights $(\boldsymbol{\mu}_t)_{t \geq 0}$ is modeled by the following *market weights equation of exponent $\zeta > 0$* , or *MWE*(ζ) for short:

$$d\mu_t^i = b^i(\boldsymbol{\mu}_t)dt + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t) dW_t^\nu, \quad i \in \{1, \dots, n\}; \quad \boldsymbol{\mu}_0 \in \Delta_+^n \quad (3.2.1)$$

$$b^i(\boldsymbol{\mu}) = (B \cdot \boldsymbol{\mu})^i = \sum_{j=1}^n B^{i,j} \mu^j, \quad \sigma^{i,\nu}(\boldsymbol{\mu}) = (\mu^i)^\zeta \zeta^{i,\nu} - \mu^i \sum_{k=1}^n (\mu^k)^\zeta \zeta^{k,\nu}$$

The volatility matrix $\sigma(\cdot) = [\sigma^{i,\nu}(\cdot)] \in \mathbb{M}^{n,d}(\mathbb{R})$ is defined from matrix $\zeta = [\zeta^{i,\nu}] \in \mathbb{M}^{n,d}(\mathbb{R})$, and the drift vector $\mathbf{b}(\cdot) = (b^i(\cdot))_{i=1}^n$ is defined from $B = [B^{i,j}] \in \mathbb{M}^n(\mathbb{R})$ which be seen as the transpose of a transition rate matrix and satisfies the following properties:

$$B^{i,j} \geq 0 \quad \forall i \neq j \text{ and } \sum_{i=1}^n B^{i,j} = 0 \quad \forall j; \quad i, j \in \{1, \dots, n\} \quad (3.2.2)$$

Furthermore, the covariance matrix $a(\cdot) = \sigma\sigma^\top(\cdot) = [a^{i,j}(\cdot)] \in \mathbb{M}^n(\mathbb{R})$ is written as:

$$\begin{aligned} a^{i,j}(\boldsymbol{\mu}_t) &= \frac{d}{dt} \langle \mu^i, \mu^j \rangle_t = \sum_{\nu=1}^d \sigma^{i,\nu} \sigma^{j,\nu}(\boldsymbol{\mu}_t) \\ &= \sum_{\nu=1}^d \left((\mu_t^i)^{\zeta_{\nu}^{i,\nu}} - \mu_t^i \sum_{k=1}^n (\mu_t^k)^{\zeta_{\nu}^{k,\nu}} \right) \left((\mu_t^j)^{\zeta_{\nu}^{j,\nu}} - \mu_t^j \sum_{k=1}^n (\mu_t^k)^{\zeta_{\nu}^{k,\nu}} \right) \end{aligned}$$

For a function $f \in \mathcal{C}_b^2(\mathbb{R}^n)$, we define the infinitesimal operator \mathcal{L} and its adjoint \mathcal{L}^* associated to the market weights equation, as well as $\nabla(f)(\boldsymbol{\mu}) = (\nabla_i f(\boldsymbol{\mu}))_{i=1}^n$ the gradient vector and $\nabla^2(f)(\boldsymbol{\mu}) = [\nabla_{i,j}^2 f(\boldsymbol{\mu})] \in \mathbb{M}^n(\mathbb{R})$ the Hessian matrix:

$$\begin{aligned} \nabla_i f(\boldsymbol{\mu}) &= \partial_i f(\boldsymbol{\mu}) = \frac{\partial}{\partial \mu^i} f(\boldsymbol{\mu}), \quad \nabla_{i,j}^2 f(\boldsymbol{\mu}) = \partial_{i,j}^2 f(\boldsymbol{\mu}) = \frac{\partial^2}{\partial \mu^i \partial \mu^j} f(\boldsymbol{\mu}) \\ \mathcal{L}f(\boldsymbol{\mu}) &= \frac{1}{2} \text{tr}[a(\boldsymbol{\mu}) \cdot \nabla^2(f)(\boldsymbol{\mu})] + \mathbf{b}(\boldsymbol{\mu})^\top \cdot \nabla(f)(\boldsymbol{\mu}) \\ \mathcal{L}^*f(\boldsymbol{\mu}) &= \frac{1}{2} \text{tr}[\nabla^2(a \cdot f)(\boldsymbol{\mu})] + \nabla(\mathbf{b} \cdot f)(\boldsymbol{\mu})^\top \cdot \mathbf{1} \end{aligned} \quad (3.2.3)$$

Let us first make the assumption on the rank of volatility matrix $\varsigma \in \mathbb{M}^{n,d}(\mathbb{R})$ that *will always hold* in this chapter, unless otherwise stated.

Assumption 3.2.1 (Rank of volatility matrix). *Suppose that $d \geq n$, i.e. the dimension d of Brownian motion is greater than or equal to the number n of assets. Moreover, the matrix $\varsigma \in \mathbb{M}^{n,d}(\mathbb{R})$ associated to volatility matrix σ is nondegenerate, i.e. ς is of rank n . In particular, by denoting (c^1, \dots, c^n) the diagonal coefficients of matrix $\varsigma\varsigma^\top$, we have $\forall i \in \{1, \dots, n\}$, $c^i = \sum_{\nu=1}^d (\varsigma^{i,\nu})^2 > 0$. We always suppose that this assumption holds, except otherwise stated.*

Assumption 3.2.2 (Dominance of drift vector). *Throughout the chapter, we will make the following successive assumptions on the matrix B and with respect to the matrix ς :*

- (1) *For all $i \in \{1, \dots, n\}$, there exists $j \in \{1, \dots, n\}$, $j \neq i$ such that $B^{i,j} > 0$. Or equivalently, there exists $\boldsymbol{\mu}^* \in \{\mu^i = 0\}$ such that $b^i(\boldsymbol{\mu}^*) = \sum_{j=1}^n B^{i,j} \mu^j > 0$.*
- (2) *For all $i, j \in \{1, \dots, n\}$, $i \neq j$, $B^{i,j} > 0$, or for all $i \in \{1, \dots, n\}$, $\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} > 0$. We can write equivalently, $b^i(\boldsymbol{\mu}) > 0$ on $\{\mu^i = 0\}$. From the property (3.2.2) of B , this assumption implies $B^{i,i} < 0$ for all $i \in \{1, \dots, n\}$.*
- (3) *For all $i \in \{1, \dots, n\}$, with (c^1, \dots, c^n) the diagonal of matrix $\varsigma\varsigma^\top$:*

$$\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} \geq \frac{c^i}{2} = \frac{1}{2} \sum_{\nu=1}^d (\varsigma^{i,\nu})^2 \geq 0 \quad (3.2.4)$$

As in (2) above, this is equivalent to $2b^i(\boldsymbol{\mu}) \geq c^i \geq 0$ on $\{\mu^i = 0\}$.

Under Assumption 3.2.1, the last inequality in (3.2.4) is strict and it is then clear that (3) \Rightarrow (2) \Rightarrow (1). Let us write the negation of (1):

(0) There exists $i \in \{1, \dots, n\}$ such that for all $j \in \{1, \dots, n\}$, $i \neq j$, we have $B^{i,j} = 0$, or equivalently $b^i(\boldsymbol{\mu}) = 0$ on $\{\mu^i = 0\}$.

Theorem 3.2.1 (Market weights as solution of MWE and boundary attainment). *There exists a weak solution $(\boldsymbol{\mu}_t)_{t \geq 0}$ to the market weights equation (3.2.1) with initial condition $\boldsymbol{\mu}_0 \in \Delta^n$ that remains all times in the unit simplex almost surely, i.e. $\forall t \geq 0$, $\boldsymbol{\mu}_t \in \Delta^n$ a.s. and we will be considering this weak solution in the rest of the chapter.*

Under Assumption 3.2.1 and starting from any positive initial weights, i.e. $\boldsymbol{\mu}_0 \in \Delta_+^n$, for any finite time horizon $[0, T]$, $T > 0$ there exists a unique strong solution to the market weights equation (3.2.1) and the market weights do not vanish on $[0, T]$, i.e. the strong solution $(\boldsymbol{\mu}_t)_{t \geq 0}$ satisfies $\boldsymbol{\mu}_t \in \Delta_+^n$ a.s. for all $t \geq 0$ if and only if we are in one of the following cases:

- (i) $\zeta \geq 1$;
- (ii) $\zeta \in (1/2, 1)$ and Assumption 3.2.2.(2) holds true, i.e. $\forall i \in \{1, \dots, n\}$, $b^i(\boldsymbol{\mu}) > 0$ on $\{\mu^i = 0\}$;
- (iii) $\zeta = 1/2$ and Assumption 3.2.2.(3) holds true, i.e. $\forall i \in \{1, \dots, n\}$, $2b^i(\boldsymbol{\mu}) \geq c^i > 0$ on $\{\mu^i = 0\}$.

Otherwise in the following cases, for any finite time horizon $[0, T]$, $T > 0$, there exists a point $\boldsymbol{\mu}_* \in \Delta^n$ and a constant $\epsilon > 0$, such that if the initial weights $\boldsymbol{\mu}_0 \in \Delta_+^n$ satisfy $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \epsilon$, then the boundary $\partial\Delta^n$ is attained with positive probability on $[0, T]$, define the stopping time $\tau := \inf\{t \geq 0 : \boldsymbol{\mu}_t \in \partial\Delta^n\}$ then we have $\mathbb{P}[\tau \leq T] > 0$. More precisely, there exist an index $i \in \{1, \dots, n\}$ such that for every time horizon $T > 0$, $\mathbb{P}[\exists t \in (0, T] \mid \mu_t^i = 0] > 0$ and the strong solution of market weights equation exists and is unique until the stopping $\tau \wedge T$:

- (iv) $\zeta \in (0, 1/2)$;
- (v) $\zeta = 1/2$ and Assumption 3.2.2.(3) does not hold, i.e. $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $2b^i(\boldsymbol{\mu}_*) < c^i$;
- (vi) $\zeta \in (1/2, 1)$ and Assumption 3.2.2.(2) does not hold, i.e. $\exists i \in \{1, \dots, n\}$, $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $b^i(\boldsymbol{\mu}_*) = 0$.

Proof. (1) We apply Theorem 2.2 in [DPF07] for the stochastic viability of the unit simplex Δ^n as a convex set. However to be able to apply the result, we should make sure that the drift and volatility have at most linear growth, i.e. $\exists M > 0$, $|\mathbf{b}(\boldsymbol{\mu})|, |\boldsymbol{\sigma}(\boldsymbol{\mu})| \leq M(1 + |\boldsymbol{\mu}|) \forall \boldsymbol{\mu} \in \mathbb{R}^n$, which does not hold in \mathbb{R}^n given the forms of drift and volatility. Introduce the *localized version* of market weights equation (3.2.1) with the same initial weights, and $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ a nonnegative, infinitely differentiable localizing function of compact support:

$$\begin{aligned}
 d\tilde{\mu}_t^i &= \tilde{b}^i(\tilde{\boldsymbol{\mu}}_t)dt + \sum_{\nu=1}^d \tilde{\sigma}^{i,\nu}(\tilde{\boldsymbol{\mu}}_t)dW_t^\nu, \quad i \in \{1, \dots, n\}; \quad \tilde{\boldsymbol{\mu}}_0 = \boldsymbol{\mu}_0 \in \Delta_+^n \\
 \tilde{b}^i(\boldsymbol{\mu}) &= \psi(\boldsymbol{\mu})b^i(\boldsymbol{\mu}), \quad \tilde{\sigma}^{i,\nu}(\boldsymbol{\mu}) = \psi(\boldsymbol{\mu})\sigma^{i,\nu}(\boldsymbol{\mu}), \quad \tilde{a} = \tilde{\sigma}\tilde{\sigma}^\top \\
 \psi(\boldsymbol{\mu}) &= 1 \text{ if } \boldsymbol{\mu} \in [0, 1]^n, \quad \psi(\boldsymbol{\mu}) = 0 \text{ if } \boldsymbol{\mu} \in [-1, 2]^n
 \end{aligned} \tag{3.2.5}$$

The drift $\tilde{\mathbf{b}}$ and volatility $\tilde{\sigma}$ of localized equation, being of compact support, clearly have at most linear growth. It remains to show that the following condition holds:

$$\begin{aligned} \tilde{\sigma}(\Pi_\Delta(\boldsymbol{\mu})) \cdot \nabla(d)(\boldsymbol{\mu}) &= 0, \quad \mathcal{L}_\Delta d(\boldsymbol{\mu}) \leq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^n \setminus \Delta^n \\ \mathcal{L}_\Delta f(\boldsymbol{\mu}) &= \frac{1}{2} \text{tr}[\tilde{a}(\Pi_\Delta(\boldsymbol{\mu})) \cdot \nabla^2(f)(\boldsymbol{\mu})] + \tilde{\mathbf{b}}(\Pi_\Delta(\boldsymbol{\mu}))^\top \cdot \nabla(f)(\boldsymbol{\mu}) \end{aligned} \quad (3.2.6)$$

Where $\Pi_\Delta(\cdot)$ represents the projection on the unit simplex Δ^n and $d(\cdot)$ is the distance to Δ^n . With $\nabla(d)(\boldsymbol{\mu}) = \pm \sum_{i=1}^n \mathbf{e}^i / \sqrt{n} = \pm(1, \dots, 1) / \sqrt{n}$, $\nabla^2(d)(\boldsymbol{\mu}) \quad \forall \boldsymbol{\mu} \in \mathbb{R}^n \setminus \Delta^n$, the condition (3.2.6) above is equivalent to:

$$\sum_{i=1}^n \tilde{b}^i(\Pi_\Delta(\boldsymbol{\mu})) = 0, \quad \sum_{i=1}^n \tilde{\sigma}^{i,\nu}(\Pi_\Delta(\boldsymbol{\mu})) = 0 \quad \forall \nu \in \{1, \dots, d\} \quad \forall \boldsymbol{\mu} \in \mathbb{R}^n \setminus \Delta^n$$

Which is automatically satisfied since from the definition of drift and volatility in (3.2.1), we have for all $\boldsymbol{\mu} \in \Delta^n$:

$$\sum_{i=1}^n \tilde{b}^i(\boldsymbol{\mu}) = \sum_{i=1}^n b^i(\boldsymbol{\mu}) = 0, \quad \sum_{i=1}^n \sigma^{i,\nu}(\boldsymbol{\mu}) = \sum_{i=1}^n \tilde{\sigma}^{i,\nu}(\boldsymbol{\mu}) = 0 \quad \forall \nu \in \{1, \dots, d\}.$$

By using Theorem 2.2 in [DPF07], the localized market weights equation (3.2.5) has a weak solution $(\tilde{\boldsymbol{\mu}}_t)_{t \geq 0}$ that remains in the unit simplex Δ^n . Notice that this solution also satisfies the original market weights equation since their drift and volatility coefficient coincide on Δ^n , it is also a weak solution of the original market weights (3.2.1) that remains in the unit simplex.

- (2) For each $\epsilon \in (0, 1/n)$ such that $\boldsymbol{\mu}_0 \in \Delta_\epsilon^n$, define the stopping time $\tau^\epsilon = \inf\{t \geq 0 : \boldsymbol{\mu}_t \notin \Delta_\epsilon^n\}$ with Δ_ϵ^n defined in (3.1.1) and consider the stopped process $(\boldsymbol{\mu}_{t \wedge \tau^\epsilon})_{t \geq 0}$. The localized market weights equation defined by $\tilde{\mathbf{b}}(\cdot) = \mathbf{b}\psi^\epsilon(\cdot)$ and $\tilde{\sigma}(\cdot) = \sigma\psi^\epsilon(\cdot)$ with $\psi^\epsilon(\cdot) = 1$ on Δ_ϵ^n and $\psi^\epsilon(\cdot) = 0$ on $\Delta^n \setminus \Delta_{\epsilon/2}^n$ clearly has globally Lipschitz continuous coefficients thus has unique strong solution $(\tilde{\boldsymbol{\mu}}_{t \in [0, T]})$ on Δ^n over any finite horizon $[0, T]$, $T > 0$. Notice this strong solution and a solution of the original equation are equal until the stopping time $\tau^\epsilon \wedge T$ almost surely since both drift and volatility coefficients of the original and the localized equations coincide on Δ_ϵ^n , their corresponding stopped processes are identical, *i.e.* $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau^\epsilon})_{t \in [0, T]} = (\boldsymbol{\mu}_{t \wedge \tau^\epsilon})_{t \in [0, T]}$ and the equality holds for any $\epsilon \in (0, \epsilon_0]$ with $\epsilon_0 > 0$ satisfying $\boldsymbol{\mu}_0 \in \Delta_{\epsilon_0}^n$. Observe that almost surely, $\lim_{\epsilon \rightarrow 0^+} \tau^\epsilon \wedge T = \tau \wedge T = \inf\{t \in [0, T] : \boldsymbol{\mu}_t \in \partial \Delta^n\}$ and $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau^\epsilon})_{t \in [0, T]}$, $(\boldsymbol{\mu}_{t \wedge \tau^\epsilon})_{t \in [0, T]}$ both have continuous paths, we are able to set $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau})_{t \in [0, T]}$ and $(\boldsymbol{\mu}_{t \wedge \tau})_{t \in [0, T]}$ with $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau})_{t \in [0, T]} = (\boldsymbol{\mu}_{t \wedge \tau \wedge T})_{t \geq 0}$. That is, every solution of the original weights equation $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau})_{t \in [0, T]}$ stopped at $\tau \wedge T$ is equal almost surely to $(\boldsymbol{\mu}_{t \wedge \tau})_{t \in [0, T]}$ which is defined as the limit when $\epsilon \rightarrow 0^+$ of the unique strong solution $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau^\epsilon})_{t \in [0, T]}$ of the localized equation, then $(\tilde{\boldsymbol{\mu}}_{t \wedge \tau})_{t \in [0, T]}$ is necessarily the unique strong solution of the original MWE(ζ) until the stopping time $\tau \wedge T$.

(3) For $i \in \{1, \dots, n\}$ and $t \geq 0$, we write the dynamic for $\log(\mu_t^i)$:

$$\begin{aligned} d \log(\mu_t^i) &= \gamma^i(\boldsymbol{\mu}_t) dt + (\mu_t^i)^{-1} \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t) dW_t^\nu \\ &= \gamma^i(\boldsymbol{\mu}_t) dt + \sum_{\nu=1}^d [(\mu_t^i)^{\zeta-1} \zeta^{i,\nu} - \sum_{j=1}^n (\mu_t^j)^\zeta \zeta^{j,\nu}] dW_t^\nu \end{aligned}$$

where $\gamma^i(\boldsymbol{\mu}_t)$ is the *growth rate* associated to the weight of the i -th asset:

$$\begin{aligned} \gamma^i(\boldsymbol{\mu}_t) &= (\mu_t^i)^{-1} b^i(\boldsymbol{\mu}_t) - \frac{1}{2} (\mu_t^i)^{-2} a^{i,i}(\boldsymbol{\mu}_t) \\ &= (\mu_t^i)^{-1} \sum_{j=1}^n B^{i,j} \mu_t^j - \frac{1}{2} \sum_{\nu=1}^d ((\mu_t^i)^{\zeta-1} \zeta^{i,\nu} - \sum_{j=1}^n (\mu_t^j)^\zeta \zeta^{j,\nu})^2 \end{aligned} \quad (3.2.7)$$

It is clear that for all $\zeta > 0$, $\gamma^i(\boldsymbol{\mu}_t) \rightarrow B^{i,i}$ when $\mu_t^i \rightarrow 1$. Let us study the behavior of the $\gamma^i(\boldsymbol{\mu}_t)$ when $\mu_t^i \rightarrow 0$ in following cases according to the value of exponent ζ :

- When $\zeta \geq 1$, we have the following lower bound for $\gamma^i(\boldsymbol{\mu}_t)$:

$$\gamma^i(\boldsymbol{\mu}_t) \geq B^{i,i} - \frac{1}{2} \sum_{\nu=1}^d (|\zeta^{i,\nu}| + \sum_{j=1}^n |\zeta^{j,\nu}|)^2 =: \underline{\gamma}^i$$

- When $\zeta \in (1/2, 1)$ and $\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} > 0$:

$$\begin{aligned} \gamma^i(\boldsymbol{\mu}_t) &= (\mu_t^i)^{-1} \left[\sum_{\substack{j=1 \\ j \neq i}}^n B^{i,j} \mu_t^j - \frac{1}{2} (\mu_t^i)^{2\zeta-1} (1 - \mu_t^i)^2 \sum_{\nu=1}^d (\zeta^{i,\nu})^2 \right. \\ &\quad \left. + (\mu_t^i)^\zeta (1 - \mu_t^i) \sum_{\nu=1}^d \zeta^{i,\nu} \sum_{\substack{j=1 \\ j \neq i}}^n (\mu_t^j)^\zeta \zeta^{j,\nu} \right] + B^{i,i} - \frac{1}{2} \sum_{\nu=1}^d \left[\sum_{\substack{j=1 \\ j \neq i}}^n (\mu_t^j)^\zeta \zeta^{j,\nu} \right]^2 \\ &\geq (\mu_t^i)^{-1} (1 - \mu_t^i) \left[\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} - \frac{1}{2} (\mu_t^i)^{2\zeta-1} (1 - \mu_t^i) \sum_{\nu=1}^d (\zeta^{i,\nu})^2 \right. \\ &\quad \left. - (\mu_t^i)^\zeta \sum_{\nu=1}^d |\zeta^{i,\nu}| \sum_{j=1}^n |\zeta^{j,\nu}| \right] + B^{i,i} - \frac{1}{2} \sum_{\nu=1}^d \left(\sum_{j=1}^n |\zeta^{j,\nu}| \right)^2 \geq \underline{\gamma}^i \end{aligned}$$

the constant $\underline{\gamma}^i$ is defined as:

$$\underline{\gamma}^i := \tilde{\gamma}^i + B^{i,i} - \frac{1}{2} \sum_{\nu=1}^d \left(\sum_{j=1}^n |\zeta^{j,\nu}| \right)^2$$

where $\tilde{\gamma}^i$ is the minimum on $[0, 1]$ of the function f defined as follows:

$$f(x) = x^{-1} (1-x) \left[\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} - \frac{1}{2} x^{2\zeta-1} (1-x) \sum_{\nu=1}^d (\zeta^{i,\nu})^2 + x^\zeta \sum_{\nu=1}^d |\zeta^{i,\nu}| \sum_{j=1}^n |\zeta^{j,\nu}| \right]$$

The minimum is finite and is reached at a point in $(0, 1)$ since f is continuous on $(0, 1)$, $f(1) = 0$ and $\lim_{x \rightarrow 0^+} f(x) = \infty$ from the fact that $\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} > 0$ and $\zeta \in (1/2, 1)$.

- When $\zeta = 1/2$ and $\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} \geq \sum_{\nu=1}^d (\zeta^{i,\nu})^2 / 2 > 0$:

$$\begin{aligned}
 \gamma^i(\boldsymbol{\mu}_t) &= B^{i,i} + (\mu_t^i)^{-1} \left[\sum_{\substack{j=1 \\ j \neq i}}^n B^{i,j} \mu_t^j - \frac{1}{2} (1 - \mu_t^i)^2 \sum_{\nu=1}^d (\zeta^{i,\nu})^2 \right] \\
 &+ (\mu_t^i)^{-1/2} (1 - \mu_t^i) \sum_{\nu=1}^d \zeta^{i,\nu} \sum_{\substack{j=1 \\ j \neq i}}^n (\mu_t^j)^{1/2} \zeta^{j,\nu} - \frac{1}{2} \sum_{\nu=1}^d \left[\sum_{\substack{j=1 \\ j \neq i}}^n (\mu_t^j)^{1/2} \zeta^{j,\nu} \right]^2 \\
 &\geq (\mu_t^i)^{-1} (1 - \mu_t^i) \left[\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} - (1 - \mu_t^i) \min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} \right] \\
 &+ B^{i,i} - \sum_{\nu=1}^d |\zeta^{i,\nu}| \sum_{j=1}^n |\zeta^{j,\nu}| - \frac{1}{2} \sum_{\nu=1}^d \left(\sum_{j=1}^n |\zeta^{j,\nu}| \right)^2 \\
 &\geq (1 - \mu_t^i) \min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} + B^{i,i} - \sum_{\nu=1}^d |\zeta^{i,\nu}| \sum_{j=1}^n |\zeta^{j,\nu}| - \frac{1}{2} \sum_{\nu=1}^d \left(\sum_{j=1}^n |\zeta^{j,\nu}| \right)^2 \\
 &\geq B^{i,i} - \sum_{\nu=1}^d |\zeta^{i,\nu}| \sum_{j=1}^n |\zeta^{j,\nu}| - \frac{1}{2} \sum_{\nu=1}^d \left(\sum_{j=1}^n |\zeta^{j,\nu}| \right)^2 =: \underline{\gamma}^i
 \end{aligned}$$

For the 3 cases above, there exists $\underline{\gamma}^i \in \mathbb{R}$ such that for all $\boldsymbol{\mu}_t \in \Delta_n^+$, $\gamma^i(\boldsymbol{\mu}_t) \geq \underline{\gamma}^i$ as $\tilde{\mu}_t^i \rightarrow 0$. Define the stopping time $\tau^i := \inf\{t \geq 0 : \mu_t^i = 0\}$ and for any $T > 0$, we are able to write:

$$\inf_{t' \in [0, \tau^i \wedge T]} \int_0^{t'} \gamma^i(\boldsymbol{\mu}_t) dt \geq \underline{\gamma}^i T > -\infty$$

By using McKean's argument from Proposition 4.3 in [MPS11], we conclude in these 3 cases: $\zeta \geq 1$; $\zeta \in (1/2, 1)$ and Assumption 3.2.2.(2); $\zeta = 1/2$ and Assumption 3.2.2.(3), we have $\forall i \in \{1, \dots, n\}, \tau^i = \infty$ then $\forall t \geq 0, \boldsymbol{\mu}_t \in \Delta_n^+$. Additionally, the stopping time satisfies $\tau = \inf\{t \geq 0 : \boldsymbol{\mu}_t \in \partial\Delta^n\} = \min_{i \in \{1, \dots, n\}} \tau^i = \infty$ almost surely and from the conclusion of (2) the market weights equation has a unique strong solution over any finite time horizon $[0, T]$ since $\mathbb{P}[\tau > T] = 1 \forall T > 0$, thus the unique solution exists for all $t \geq 0$ and is denoted correctly as $(\boldsymbol{\mu}_t)_{t \geq 0}$.

- (4) To show that the boundary $\{\mu^i = 0\} \subset \partial\Delta^n$ is attained with positive probability, let us show that under the conditions (iv)-(vi), there exists $\boldsymbol{\mu}_* \in \Delta^n$ such that:

$$2b^i(\boldsymbol{\mu}_*) - (\mu_*^i)^{-1} a^{i,i}(\boldsymbol{\mu}_*) < 0 \quad \text{and} \quad (\mu_*^i)^{-1} a^{i,i}(\boldsymbol{\mu}_*) > 0$$

We recognized the growth rate defined (3.2.7) multiplied by the i -th weight:

$$\begin{aligned}
 2b^i(\boldsymbol{\mu}) - (\mu^i)^{-1} a^{i,i}(\boldsymbol{\mu}) &= 2\mu^i \cdot \gamma^i(\boldsymbol{\mu}) \\
 &= 2b^i(\boldsymbol{\mu}) - \sum_{\nu=1}^d \left((\mu^i)^{\zeta-1/2} \zeta^{i,\nu} - (\mu^i)^{1/2} \sum_{j=1}^n (\mu^j)^{\zeta} \zeta^{j,\nu} \right)^2
 \end{aligned}$$

- When $\zeta \in (0, 1/2)$, we have the equivalence for the second term when $\mu^i \rightarrow 0$:

$$\sum_{\nu=1}^d ((\mu^i)^{\zeta-1/2} \zeta^{i,\nu} - (\mu^i)^{1/2} \sum_{j=1}^n (\mu^j)^{\zeta} \zeta^{j,\nu})^2 \sim (\mu^i)^{2\zeta-1} \sum_{\nu=1}^d (\zeta^{i,\nu})^2 \rightarrow \infty$$

The drift $b^i(\boldsymbol{\mu})$ being bounded for $\boldsymbol{\mu} \in \Delta^n$, we can find $\boldsymbol{\mu}_* \in \Delta^n$ so that $2b^i(\boldsymbol{\mu}_*) - (\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) < 0$ and $(\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) > 0$.

- When $\zeta = 1/2$ and under the assumption $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $2b^i(\boldsymbol{\mu}_*) < \sum_{\nu=1}^d (\zeta^{i,\nu})^2$ we have:

$$2b^i(\boldsymbol{\mu}_*) - (\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) = 2b^i(\boldsymbol{\mu}_*) - \sum_{\nu=1}^d (\zeta^{i,\nu})^2 < 0 \quad \text{and} \quad (\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) = (\zeta^{i,\nu})^2 > 0$$

- When $\zeta \in (1/2, 1)$ and under the assumption $\exists \boldsymbol{\mu}_* \in \{\mu^i = 0\}$, $b^i(\boldsymbol{\mu}_*) = 0$ or equivalently $\exists j \in \{1, \dots, n\}$, $B^{i,j} = 0$, we choose $\mu_*^i = \mu \in [0, 1]$ and $\mu_*^j = 1 - \mu$ and $\mu_*^k = 0$ for $k \neq i, j$ and write:

$$\begin{aligned} 2b^i(\boldsymbol{\mu}_*) - (\mu)^{-1} a^{i,i}(\boldsymbol{\mu}_*) &= 2B^{i,i} \mu - (\mu)^{2\zeta-1} \sum_{\nu=1}^d ((1 - \mu) \zeta^{i,\nu} - (\mu)^{1-\zeta} (1 - \mu) \zeta^{j,\nu})^2 \\ &\leq -(\mu)^{2\zeta-1} \sum_{\nu=1}^d ((1 - \mu) \zeta^{i,\nu} - (\mu)^{1-\zeta} (1 - \mu) \zeta^{j,\nu})^2 \end{aligned}$$

Moreover, since $\zeta < 1$, when $\mu \rightarrow 0$ we have:

$$\sum_{\nu=1}^d ((1 - \mu) \zeta^{i,\nu} - (\mu)^{1-\zeta} (1 - \mu) \zeta^{j,\nu})^2 \rightarrow \sum_{\nu=1}^d (\zeta^{i,\nu})^2 > 0$$

It is then possible to choose $\mu_*^i = \mu > 0$ and close to 0 such that:

$$\sum_{\nu=1}^d ((1 - \mu) \zeta^{i,\nu} - (\mu)^{1-\zeta} (1 - \mu) \zeta^{j,\nu})^2 < 0$$

With this choice of $\boldsymbol{\mu}_*$, we have $2b^i(\boldsymbol{\mu}_*) - (\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) < 0$ and $(\mu_*^i)^{-1} a^{i,j}(\boldsymbol{\mu}_*) > 0$.

Then we can use similar argument as Theorem 5.7.(iii) in [FL16] to show in Appendix 3.5.1 that for any $T > 0$, there exists $\epsilon > 0$ such that if $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \epsilon$ almost surely, then $\mu_t^i = 0$ for some $t \leq T$ with positive probability. In terms of stopping time, $\mathbb{P}[\tau = \min_{i \in \{1, \dots, n\}} \tau^i \leq T] > 0 \forall T > 0$ and using again the conclusion of (2), under conditions (iv)-(vi) the solution to market weights equation exists and is unique until $\tau \wedge T$.

□

Assumption 3.2.3 (Volatility matrix). We suppose for the matrix ς associated to volatility matrix σ :

$$\varsigma\varsigma^\top = \text{diag}(c^1, \dots, c^n)$$

This is obviously the case when ζ just has nonzero main diagonal. The covariance matrix $a(\cdot)$ then has following simplified form:

$$a^{i,j}(\boldsymbol{\mu}) = \mu^i \mu^j \sum_{k=1}^n (\mu^k)^{2\zeta} c^k - \mu^i (\mu^j)^{2\zeta} c^j - \mu^j (\mu^i)^{2\zeta} c^i, \quad i \neq j$$

$$a^{i,i}(\boldsymbol{\mu}) = (1 - 2\mu^i) (\mu^i)^{2\zeta} c^i + (\mu^i)^2 \sum_{k=1}^n (\mu^k)^{2\zeta} c^k$$

Although this assumption is of particular interest in some of the cases that we will be considering in this chapter, it does not hold in general in the rest of the chapter.

Remark 3.2.1 (Link with existing models). From our general $MWE(\zeta)$, we may recognize some well-known models. In addition to financial mathematics, these models could also be applied to in population genetics, or dynamics of continuous time switching regimes:

1. The process of market weights given by $MWE(\zeta)$ can be naturally regarded as dynamics of a *discrete probability distribution*. For instance by taking the exponent $\zeta = 1$, $MWE(1)$ gives the dynamics of the discrete probabilistic distribution in optimal nonlinear filtering problem, see [Lip+01].
2. *Volatility-stabilized market (VSM)* models are very popular in modeling the market capitalization. They are able to reproduce some stylized fact in the market, such as the fact that smaller stocks have larger volatility, see [FK05] and [FK09]. The process of market weights associated to SVM model is a Wright-Fisher diffusion process, studied thoroughly in [Pal+11], also called a multivariate Jacobi diffusion as in [GJ06]. It can be seen as a special case of $MWE(\zeta)$, by taking $\zeta = 1/2$ and $\alpha, c > 0$:

$$d\mu_t^i = \alpha(\xi^i - \mu_t^i)dt + \sqrt{c\mu_t^i}dW_t^i - \mu_t^i \sum_{j=1}^n \sqrt{c\mu_t^j}dW_t^j, \quad i \in \{1, \dots, n\},$$

with the initial condition $\boldsymbol{\mu}_0 \in \Delta_+^n$ and the parameter $\boldsymbol{\xi} = (\xi^i)_{i=1}^n \in \Delta_+^n$. For the drift vector, we chose the matrix B as $B^{i,j} = \alpha\xi^i > 0$ for $i \neq j$ and $B^{i,i} = -\alpha \sum_{k=1, k \neq i}^n \xi^k < 0$ and for the volatility matrix, we chose $d = n$ and $\varsigma = \sqrt{c}I^n$ where $I^n \in \mathbb{M}^n(\mathbb{R})$ is the identity matrix. It is also possible to recognize the volatility in generalized volatility-stabilized processes of parameter β and $K(\cdot) = 1$, studied in [Pic14] by taking $\zeta = 1 - \beta$.

When $n \rightarrow \infty$, the limit process of Wright-Fisher process called Fleming-Viot process, becomes measure-valued and its applications in population genetics are studied in [EK93].

3. The market weights given by *polynomial processes* introduced in [Cuc19] are an extension of SVM model, the drift vector and covariance matrix are defined as follows:

$$b^i(\boldsymbol{\mu}) = \beta^i + \sum_{j=1}^n \tilde{\beta}^{i,j} \mu^j, \quad a^{i,i}(\boldsymbol{\mu}) = \mu^i \sum_{j \neq i} \gamma^{i,j} \mu^j, \quad a^{i,j}(\boldsymbol{\mu}) = \gamma^{i,j} \mu^i \mu^j, \quad i \neq j$$

with $\forall j \in \{1, \dots, n\}$, $\sum_{i=1}^n (\beta^i + \tilde{\beta}^{i,j}) = 0$ and $\beta^i + \tilde{\beta}^{i,j} \geq 0$, $\forall i \neq j$ and $\gamma = [\gamma^{i,j}] \in \mathbb{M}^{n,n}(\mathbb{R})$ symmetric with nonnegative off-diagonal elements and its diagonal elements are equal to 0. Then it is easy to see that the drift of this model can be matched by simply taking $B^{i,j} = \beta^i + \tilde{\beta}^{i,j}$. For the volatility coefficient, by taking $n = d$ and under Assumption 3.2.3, *i.e.* $\zeta \zeta^\top = \text{diag}(c^1, \dots, c^n)$, $c^i > 0$, we compute for $i, j \in \{1, \dots, n\}$:

$$a^{i,i}(\boldsymbol{\mu}) = \mu^i \sum_{j \neq i} \gamma^{i,j}(\boldsymbol{\mu}) \mu^j, \quad a^{i,j}(\boldsymbol{\mu}) = \gamma^{i,j}(\boldsymbol{\mu}) \mu^i \mu^j, \quad i \neq j$$

$$\gamma^{i,i}(\boldsymbol{\mu}) = 0, \quad \gamma^{i,j}(\boldsymbol{\mu}) = c^i + c^j - \sum_{k=1}^n c^k \mu^k, \quad i \neq j$$

The matrix $\gamma(\boldsymbol{\mu}) = [\gamma^{i,j}(\boldsymbol{\mu})]$ is symmetric and has diagonal equal to 0 but contrary to the polynomial model, the off-diagonal coefficients depend on the weights. However when the inequality $\max_{i \in \{1, \dots, n\}} c^i \leq 2 \cdot \min_{i \in \{1, \dots, n\}} c^i$ is satisfied, $\forall i, j \in \{1, \dots, n\}$, $\forall \boldsymbol{\mu} \in \Delta^n$, we make sure that $\gamma^{i,j}(\boldsymbol{\mu}) \geq 0$ so the matrix $\gamma(\boldsymbol{\mu})$ has nonnegative off-diagonal coefficients for all weights. The limit when $n \uparrow \infty$ of polynomial process, called measure-valued polynomial process, is studied in the recent paper [CLSF+19].

Remark 3.2.2 (Mean-reversion level of market weights). Similarly to mean-reverting processes such as Ornstein-Uhlenbeck process, a market weights vector $\boldsymbol{\mu}_{\mathcal{M}} = (\mu_{\mathcal{M}}^i)_{i=1}^n \in \Delta_+^n$ is called *mean-reversion level* of market weights if the drift vanishes, *i.e.* $\mathbf{b}(\boldsymbol{\mu}_{\mathcal{M}}) = \mathbf{0}$, or equivalently $\boldsymbol{\mu}_{\mathcal{M}}$ solves the linear equation $B \cdot \boldsymbol{\mu}_{\mathcal{M}} = \mathbf{0}$. In the above example of volatility-stabilized market models, the corresponding mean-reversion level is $\boldsymbol{\mu}_{\mathcal{M}} = \boldsymbol{\xi}$.

Reduced form of market weights equation From Theorem 3.2.1, the sum of market weights $(\mu_t^i)_{i=1}^n$ given by market weights equation (3.2.1) is equal to 1 at all times. Since the last weight is given by $\mu_t^n = 1 - \sum_{i=1}^{n-1} \mu_t^i$, it is legitimate to just consider the vector of $(n-1)$ first market weights $\bar{\boldsymbol{\mu}}_t = (\mu_t^i)_{i=1}^{n-1}$, called reduced weights $\bar{\boldsymbol{\mu}}_t = (\mu_t^i)_{i=1}^{n-1}$, which is in natural bijection with the vector of complete weights $\boldsymbol{\mu}_t = (\mu_t^i)_{i=1}^n$. This leads to the reduced form of the market weights equation:

$$d\mu_t^i = \bar{b}^i(\bar{\boldsymbol{\mu}}_t) dt + \sum_{\nu=1}^d \bar{\sigma}^{i,\nu}(\bar{\boldsymbol{\mu}}_t) dW_t^\nu, \quad \mu_0^i > 0, \quad i = 1, \dots, n-1, \quad (3.2.8)$$

where the reduced drift vector $\bar{\mathbf{b}}(\cdot) = (\bar{b}^i(\cdot))_{i=1}^{n-1}$ and volatility matrix $\bar{\sigma}(\cdot) = [\bar{\sigma}^{i,\nu}(\cdot)] \in \mathbb{M}^{n-1,d}(\mathbb{R})$ are obtained in the straightforward way from \mathbf{b} and σ .

From now on, by abuse of notation, when considering the reduced weights vector, we will interpret Δ_+^n , Δ^n and the boundary $\partial\Delta^n = \Delta^n \setminus \Delta_+^n$ as subsets of \mathbb{R}^{n-1} defined in the natural way:

$$\Delta_+^n = \left\{ \bar{\boldsymbol{\mu}} = (\mu^i)_{i=1}^{n-1} \mid \mu^i > 0 \ \forall i \in \{1, \dots, n-1\} \text{ and } \sum_{i=1}^{n-1} \mu^i < 1 \right\}$$

From now on, depending on the context, we use the unified and equivalent notation $\boldsymbol{\mu}, \mathbf{b}, \sigma, \varsigma, a$ for the both complete (3.2.1) and reduced (3.2.8) forms, $MWE(\zeta)$ is now written as follows:

$$d\boldsymbol{\mu}_t = \mathbf{b}(\boldsymbol{\mu}_t) \cdot dt + \sigma(\boldsymbol{\mu}_t) \cdot d\mathbf{W}_t; \ \boldsymbol{\mu}_0 \in \Delta_+^n \subset \mathbb{R}^n \text{ or } \mathbb{R}^{n-1} \quad (3.2.9)$$

Proposition 3.2.1 (Local ellipticity of market weights process). *Under Assumption 3.2.1, i.e. $d \geq n$, $\varsigma \in \mathbb{M}^{n,d}(\mathbb{R})$ of rank n , the reduced covariance matrix $a(\boldsymbol{\mu}) = \sigma\sigma^\top(\boldsymbol{\mu}) \in \mathbb{M}^{n-1}(\mathbb{R})$ is positive definite if and only if $\boldsymbol{\mu} \in \Delta_+^n$, that is:*

$$\begin{aligned} \forall \boldsymbol{\mu} \in \Delta_+^n; \ \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} > 0 \\ \forall \boldsymbol{\mu} \in \partial\Delta^n = \Delta^n \setminus \Delta_+^n; \ \exists \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0 \end{aligned}$$

Equivalently, the reduced volatility matrix $\sigma(\boldsymbol{\mu}) \in \mathbb{M}^{n-1,d}(\mathbb{R})$ defined in (3.2.8) is of full rank, i.e. equal to $n-1$, if and only if $\boldsymbol{\mu} \in \Delta_+^n$.

For $\forall \boldsymbol{\mu} \in \partial\Delta^n$, the rank of the reduced covariance matrix is the same as that of the reduced volatility matrix and is equal to $n-1-m(\boldsymbol{\mu})$, with $m(\boldsymbol{\mu})$ being the number of indices $i \in \{1, \dots, n\}$ such that $\mu^i = 0$.

Moreover for each $\epsilon \in (0, 1/n)$, there exists $\lambda_\epsilon > 0$ such that:

$$\forall \boldsymbol{\mu} \in \Delta_\epsilon^n; \ \forall \mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\} : \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} \geq \lambda_\epsilon |\mathbf{x}|^2$$

As the reduced covariance matrix $a(\cdot)$ degenerates at the boundary $\partial\Delta^n$, the market weights process $(\boldsymbol{\mu}_t)_{t \geq 0}$ as the unique strong solution of $MWE(\zeta)$ (3.2.9) does not have uniform ellipticity, but only local ellipticity holds on the sets of type Δ_ϵ^n where $\epsilon \in (0, 1/n)$.

Proof. For the reduced covariance matrix $a(\boldsymbol{\mu}) \in \mathbb{M}^{n-1}(\mathbb{R})$, we check when the quantity $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} \geq 0$ vanishes for $\boldsymbol{\mu} \in \Delta^n$ and $\mathbf{x} = (x^i)_{i=1}^{n-1} \in \mathbb{R}^{n-1}$. It is equivalent to consider $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x}$ with the complete covariance matrix $a(\boldsymbol{\mu}) \in \mathbb{M}^n(\mathbb{R})$ and $\mathbf{x} = (x^i)_{i=1}^n \in \mathbb{R}^n$, $x^n = 0$:

$$\begin{aligned} \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} &= |\sigma^\top(\boldsymbol{\mu}) \mathbf{x}|^2 = 0 \Leftrightarrow \mathbf{x} \in \text{Ker}(\sigma^\top(\boldsymbol{\mu})); \ \sigma(\boldsymbol{\mu}) \in \mathbb{M}^{n,d}(\mathbb{R}), \ \mathbf{x} \in \mathbb{R}^n, \ x_n = 0 \\ &\Leftrightarrow \sum_{i=1}^n x^i ((\mu^i)^\zeta \varsigma^{i,\nu} - \mu^i \sum_{k=1}^n (\mu^k)^\zeta \varsigma^{k,\nu}) = 0; \ \forall \nu \in \{1, \dots, d\}, \ x_n = 0 \\ &\Leftrightarrow \sum_{i=1}^n x^i (\mu^i)^\zeta \varsigma^{i,\nu} = \sum_{i=1}^{n-1} x^i \mu^i \cdot \sum_{k=1}^n (\mu^k)^\zeta \varsigma^{k,\nu}; \ \forall \nu \in \{1, \dots, d\}, \ x_n = 0 \\ &\Leftrightarrow \varsigma^\top(\tilde{\mathbf{x}} - C \cdot \boldsymbol{\mu}^\zeta) = \mathbf{0} \\ &\Leftrightarrow \tilde{\mathbf{x}} - C \cdot \boldsymbol{\mu}^\zeta \in \text{Ker}(\varsigma^\top) \text{ with } \tilde{x}^i = x^i (\mu^i)^\zeta, \ x^n = 0, \text{ and } C = \sum_{k=1}^{n-1} x^k \boldsymbol{\mu}^k. \end{aligned} \quad (3.2.10)$$

Under Assumption 3.2.1, ς is nondegenerate and ς^\top can be regarded as an linear operator from \mathbb{R}^n to \mathbb{R}^d with $n \geq d$, using rank theorem we have $\dim \text{Ker}(\varsigma^\top) = 0$. The equivalence (3.2.10) above then becomes:

$$\tilde{\mathbf{x}} = C \cdot \boldsymbol{\mu}^\varsigma \Leftrightarrow x^i (\mu^i)^\varsigma = C \cdot (\mu^i)^\varsigma; \forall i \in \{1, \dots, n-1\} \text{ and } x^n (\mu^n)^\varsigma = 0 = C \cdot (\mu^n)^\varsigma \quad (3.2.11)$$

Then we discuss the following cases, again for $a(\boldsymbol{\mu}) \in \mathbb{M}^{n-1}(\mathbb{R})$:

- (i) If $\boldsymbol{\mu} \in \Delta_+^n$, *i.e.* $\mu^i > 0 \forall i \in \{1, \dots, n\}$, from $0 = C \cdot (\mu^n)^\varsigma$ on the r.h.s. of the equivalence above we get $C = \sum_{k=1}^{n-1} x^k \mu^k = 0$, then $\forall i \in \{1, \dots, n-1\}$, $x^i (\mu^i)^\varsigma = C \cdot (\mu^i)^\varsigma = 0$ so $x^i = 0$, in order to get $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0$, it is necessary to have $\mathbf{x} = \mathbf{0}$.
- (ii) For $\boldsymbol{\mu} \in \partial\Delta^n = \Delta^n \setminus \Delta_+^n$, first we suppose $\mu^n = 0$, then $\sum_{k=1}^{n-1} \mu^k = 1$. Consider the vector \mathbf{x} with $x^i = C \neq 0$, $i \in \{1, \dots, n-1\}$, then $\sum_{k=1}^{n-1} x^k \mu^k = x^i \sum_{k=1}^{n-1} \mu^k = x^i = C$, $\forall i \in \{1, \dots, n-1\}$, the r.h.s. of (3.2.11) above is clearly satisfied. With this choice of $\mathbf{x} \neq \mathbf{0}$ we have $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0$.
- (iii) The remaining case is $\boldsymbol{\mu} \in \partial\Delta^n$ and $\mu^n \neq 0$, so $\exists j \in \{1, \dots, n-1\}$ such that $\mu^j = 0$. Consider the vector with $x^j \neq 0$ and $x^i = 0$ for $i \neq j$, then $C = \sum_{k=1}^{n-1} x^k \mu^k = 0$ and (3.2.11) is satisfied.

From the discussions above, for $\boldsymbol{\mu} \in \Delta_+^n$, the fact that $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0$ implies $\mathbf{x} = \mathbf{0}$, while for $\boldsymbol{\mu} \in \partial\Delta^n$, there exists $\mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$ such that $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0$. More precisely for all $\boldsymbol{\mu} \in \partial\Delta^n$:

$$\begin{aligned} \text{Ker}(\sigma^\top(\boldsymbol{\mu})) &= \{x \in \mathbb{R}^{n-1} \mid \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = 0\} \\ &= \text{Span}\left\{ \mathbf{e}^i \text{ (if } \mu^i = 0), i \in \{1, \dots, n-1\}; \sum_{i=1}^{n-1} \mathbf{e}^i \text{ (if } \mu^n = 0) \right\} \end{aligned}$$

The vectors in the linear span above are linearly independent since \mathbf{e}^i , $i \in \{1, \dots, n-1\}$ and $\sum_{i=1}^{n-1} \mathbf{e}^i$ cannot be included simultaneously in $\text{Ker}(\sigma^\top(\boldsymbol{\mu}))$ so the dimension of $\text{Ker}(\sigma^\top(\boldsymbol{\mu}))$ is equal to the number of indices $i \in \{1, \dots, n\}$ such that $\mu^i = 0$, that we denote $m(\boldsymbol{\mu})$. Then the rank of $\sigma^\top(\boldsymbol{\mu}) \in \mathbb{M}^{n-1,d}(\mathbb{R})$, $d \geq n$ or $a(\boldsymbol{\mu})$ is equal to $n-1-m(\boldsymbol{\mu})$.

For $\boldsymbol{\mu} \in \Delta_\epsilon^n \subsetneq \Delta_+^n$ and $\mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$, it holds that $\mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} / |\mathbf{x}|^2 > 0$. By homogeneity of the expression we can suppose $|\mathbf{x}| = 1$ and using the compactness of the unit sphere of \mathbb{R}^{n-1} , there exists $\mathbf{x}_\epsilon \in \mathbb{R}^{n-1}$, $|\mathbf{x}_\epsilon| = 1$ and $\lambda_\epsilon > 0$ such that $\inf_{|\mathbf{x}|=1} \mathbf{x}^\top a(\boldsymbol{\mu}) \mathbf{x} = \mathbf{x}_\epsilon^\top a(\boldsymbol{\mu}) \mathbf{x}_\epsilon = \lambda_\epsilon > 0$. \square

Remark 3.2.3 (Degeneracy of covariance matrix). When Assumption 3.2.1 is not satisfied, *i.e.* $d \leq n-1$ or ς is degenerate, in general the reduced covariance matrix $a(\boldsymbol{\mu})$ is not positive definite for all $\boldsymbol{\mu} \in \Delta_+^n$, which means that there might exist $\boldsymbol{\mu} \in \Delta_+^n$ such that $a(\boldsymbol{\mu})$ is degenerate.

- (1) For $d < n-1$ whether ς is degenerate or not, for all $\boldsymbol{\mu} \in \Delta_+^n$, the rank of $a(\boldsymbol{\mu}) = \sigma\sigma^\top(\boldsymbol{\mu}) \in \mathbb{M}^{n-1}(\mathbb{R})$ is less than or equal to the rank of $\sigma(\boldsymbol{\mu}) \in \mathbb{M}^{n-1,d}(\mathbb{R})$, which itself is less than or equal to $d < n-1$, so the rank of the reduced covariance matrix $a(\boldsymbol{\mu}) \in \mathbb{M}^{n-1}(\mathbb{R})$ is strictly less than $(n-1)$ and it cannot be positive definite.

(2) For $d = n - 1$, from the last line in equivalence (3.2.10) we have:

$$x^i(\mu^i)^\zeta = \left(\sum_{k=1}^{n-1} x^k \mu^k \right) \cdot (\mu^i)^\zeta + y^i; \quad \forall i \in \{1, \dots, n-1\}, \quad x^n = 0, \quad (3.2.12)$$

where $\mathbf{y} = (y^i)_{i=1}^n \in \text{Ker}(\zeta^\top) \setminus \{\mathbf{0}\} \neq \emptyset$, as $\dim \text{Ker}(\zeta^\top) \geq 1$. If we are able to find $\mathbf{x}' \in \mathbb{R}^{n-1} \setminus \{0\}$ and $\boldsymbol{\mu}' \in \Delta_+^n$ such that the identity (3.2.12) above is satisfied, then the reduced covariance matrix $a(\cdot)$ is not positive definite for this $\boldsymbol{\mu}' \in \Delta_+^n$. Let us illustrate this with an example taking $n = 2$, $d = 1$ and the matrix $\zeta = [\zeta^1 \ \zeta^2]^\top$. From Theorem 3.2.1 we know that $\forall t \geq 0$, $\mu_t^1, \mu_t^2 \geq 0$, $\mu_t^1 + \mu_t^2 = 1$. Using the notation $\mu = \mu_t^1$, the matrix B , the drift vector $b(\mu)$ and the volatility matrix $\sigma(\mu)$ are written as follows:

$$\begin{aligned} B^{2,1} &= \alpha = -B^{1,1}, \quad B^{1,2} = \beta = -B^{2,2}; \quad \alpha, \beta \geq 0 \\ b(\mu) &= \beta - (\alpha + \beta)\mu, \quad \sigma(\mu) = \zeta^1 \mu^\zeta (1 - \mu) - \zeta^2 \mu (1 - \mu)^\zeta \end{aligned} \quad (3.2.13)$$

When $\zeta \neq 1$ and ζ^1, ζ^2 are either both strictly positive or both strictly negative, we have $\sigma(\mu') = 0$ for the following values of $\mu' \in \Delta_+^2 = (0, 1)$:

$$\mu' = \frac{(\zeta^1)^{1/\zeta}}{(\zeta^1)^{1/\zeta} + (\zeta^2)^{1/\zeta}} \quad \text{or} \quad \mu' = \frac{(-\zeta^1)^{1/\zeta}}{(-\zeta^1)^{1/\zeta} + (-\zeta^2)^{1/\zeta}}.$$

For ζ^1 and ζ^2 with different signs, for example $\zeta^1 > 0 \geq \zeta^2$, we have $\forall \mu \in (0, 1)$, $\sigma(\mu) \neq 0$.

3.2.2 Distribution and transition density of market weights

Given the initial weights $\boldsymbol{\mu}_0 \in \Delta_+^n$ we are interested in the probability distribution of the process of market weights of assets $(\boldsymbol{\mu}_t)_{t \geq 0}$ as the solution of $MWE(\zeta)$ at time $t \geq 0$, denoted as $m_t(\cdot | \boldsymbol{\mu}_0)$ and defined as follows:

$$A \in \mathcal{B}(\mathbb{R}^n), \quad m_t(A | \boldsymbol{\mu}_0) = \mathbb{P}[\boldsymbol{\mu}_t \in A | \boldsymbol{\mu}_0]$$

The first question to be addressed is whether the probability distribution $m_t(\cdot | \boldsymbol{\mu}_0)$ of market weights admits a density, or equivalently whether $m_t(\cdot | \boldsymbol{\mu}_0)$ is absolutely continuous with respect to Lebesgue measure. From Theorem 3.2.1 we know that $\forall t \geq 0$, $\mathbb{P}[\boldsymbol{\mu}_t \in \Delta^n | \boldsymbol{\mu}_0] = 1$ so the support of $m_t(\cdot | \boldsymbol{\mu}_0)$ is included in Δ^n , but since the Lebesgue measure of Δ^n in \mathbb{R}^n is zero, the distribution $m_t(\cdot | \boldsymbol{\mu}_0)$ cannot be continuous with respect to the Lebesgue measure on \mathbb{R}^n . This naturally leads to the reduced form of market weights $\boldsymbol{\mu}_t \in \Delta^n \subset \mathbb{R}^{n-1}$ and the absolute continuity with respect to the Lebesgue measure $\lambda(\cdot)$ on \mathbb{R}^{n-1} .

Proposition 3.2.2 (Absolute continuity of distribution and regularity of density for polynomial volatility). *With the exponent $\zeta \in \mathbb{N}^*$, i.e. the volatility coefficient $\sigma(\cdot)$ is polynomial, starting from any positive initial weights $\boldsymbol{\mu}_0 \in \Delta_+^n$ and under Assumption 3.2.1 for local ellipticity, for all $t > 0$, the probability distribution of market weights of assets $m_t(\cdot | \boldsymbol{\mu}_0)$ is absolutely continuous with respect to the Lebesgue measure*

$\lambda(\cdot)$ on \mathbb{R}^{n-1} , or equivalently, the probability density function $p_t(\cdot|\boldsymbol{\mu}_0)$ associated to $m_t(\cdot|\boldsymbol{\mu}_0)$ exists and satisfies:

$$m_t(d\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) = p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) \cdot d\boldsymbol{\mu}_t$$

Moreover, the density function $p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0)$ is infinitely differentiable with respect to its space variable, i.e. $\boldsymbol{\mu}_t \mapsto p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) \in \mathcal{C}^\infty(\mathbb{R}^{n-1}, \mathbb{R}_+)$.

Proof. For $\zeta \in N^*$, the drift and volatility of $MWE(\zeta)$ (3.2.9) are polynomials on \mathbb{R}^{n-1} , the boundedness in Hörmander's condition stated in Appendix 3.5.2 is not satisfied. We use again the technique of localization with the localizing function $\psi(\cdot) \in \mathcal{C}_c^\infty(\mathbb{R}^{n-1})$ satisfying for example $\psi(\cdot) = 1$ on $[0, 1]^{n-1}$ and $\psi(\cdot) = 1$ on $\mathbb{R}^n \setminus [-1, 2]^{n-1}$, the localized drift vector $\tilde{b}(\cdot) = b\psi(\cdot) \in \mathbb{R}^{n-1}$ and volatility matrix $\tilde{\sigma}(\cdot) = \sigma\psi(\cdot) \in \mathbb{M}^{n-1,d}(\mathbb{R})$ with its d columns denoted as $\{\tilde{\sigma}^1(\cdot), \dots, \tilde{\sigma}^d(\cdot)\}$, are functions in $\mathcal{C}_c^\infty(\mathbb{R}^{n-1})$ with bounded derivatives, also notice that the boundedness does not hold if the volatility $\sigma(\cdot)$ is not polynomial or ζ is not a positive integer. We are then able to check the Hörmander's hypothesis ($H_k^{\mu_0}$) in (3.5.2) with $k = 0$ for the localized equation, which is written explicitly as:

$$(H_0^{\mu_0}) : \text{Span}\{\tilde{\sigma}^1(\boldsymbol{\mu}_0), \dots, \tilde{\sigma}^d(\boldsymbol{\mu}_0)\} = \mathbb{R}^{n-1}$$

The covariance matrix $\tilde{a}(\boldsymbol{\mu}_0) = \tilde{\sigma}\tilde{\sigma}^\top(\boldsymbol{\mu}_0)$ is invertible, necessarily the rank of the matrix of diffusion for the initial distribution $\tilde{\sigma}(\boldsymbol{\mu}_0)$ equal to $n-1$, thus $\text{Span}\{\tilde{\sigma}^1(\boldsymbol{\mu}_0), \dots, \tilde{\sigma}^d(\boldsymbol{\mu}_0)\} = \mathbb{R}^{n-1}$ and ($H_0^{\mu_0}$) is satisfied, finally we conclude with Hörmander's theorem stated in Appendix 3.5.2 for both absolute continuity of distribution $m_t(\cdot|\boldsymbol{\mu}_0)$ with respect to the Lebesgue measure on \mathbb{R}^{n-1} and infinite differentiability of the density function with respect to the space variable $\boldsymbol{\mu}_t \mapsto p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0)$. \square

3.2.3 Stationary distribution of market weights

We are now interested in the long-term relative performance of the assets in the market, which amounts to studying the long-term behavior of the market weights given by $MWE(\zeta)$ (3.2.9). It is then natural to study the stationary probability distribution of market weights.

Definition 3.2.1 (Feller semigroup and stationary distribution). The semigroup associated to the process of market weights $(\boldsymbol{\mu}_t)_{t \geq 0}$ given by $MWE(\zeta)$, applied to a continuous bounded function $f \in \mathcal{C}_b(\Delta^n, \mathbb{R})$ is denoted as $(\mathcal{P}_t f)_{t \geq 0}$. For a probability distribution $m(\cdot)$ on Δ^n , we define the action $\langle \mathcal{P}_t f, m \rangle$ of the semigroup \mathcal{P}_t on $m(\cdot)$:

$$\begin{aligned} \mathcal{P}_t f(\boldsymbol{\mu}_0) &= \mathbb{E}^{\boldsymbol{\mu}_0} f(\boldsymbol{\mu}_t) = \int_{\Delta^n} f(\boldsymbol{\mu}_t) p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) d\boldsymbol{\mu}_t \\ \langle \mathcal{P}_t f, m \rangle &= \mathbb{E}^m f(\boldsymbol{\mu}_t) = \int_{\Delta^n} \mathcal{P}_t f(\boldsymbol{\mu}_0) m(d\boldsymbol{\mu}_0) \end{aligned}$$

We say the semigroup $(\mathcal{P}_t)_{t \geq 0}$ has *Feller* (resp. *strong Feller*) property when:

$$\forall f \text{ continuous bounded (resp. bounded), } \forall t > 0, \boldsymbol{\mu}_0 \mapsto \mathcal{P}_t f(\boldsymbol{\mu}_0) \text{ is continuous.}$$

The probability distribution $m(\cdot)$ of support Δ^n is *stationary* (or *invariant*) with respect to the process $(\boldsymbol{\mu}_t)_{t \geq 0}$ when for all continuous bounded function i.e. $f \in \mathcal{C}_b(\Delta^n, \mathbb{R})$ we have:

$$\forall f \in \mathcal{C}_b(\Delta^n, \mathbb{R}), \forall t > 0, \langle \mathcal{P}_t f, m \rangle = \langle f, m \rangle = \int_{\Delta^n} f(\boldsymbol{\mu}_0) m(d\boldsymbol{\mu}_0)$$

It is easy to see that the semigroup $(\mathcal{P}_t)_{t \geq 0}$ associated to the process of market weights has Feller property. In fact, the process of market weights given by $MWE(\zeta)$ (3.2.9) has continuous paths $(\boldsymbol{\mu}_t)_{t \geq 0}$ so $\boldsymbol{\mu}_0 \mapsto \mathbb{E}^{\boldsymbol{\mu}_0} f(\boldsymbol{\mu}_t) = \mathcal{P}_t f(\boldsymbol{\mu}_0)$ is continuous for all $f \in \mathcal{C}_b(\Delta^n, \mathbb{R})$. However for strong Feller property, which implies that there exists at most one stationary distribution, the proof is less straightforward and relies on the regularity of the density function with respect to the starting point of the process: $\boldsymbol{\mu}_0 \mapsto p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0)$, $\boldsymbol{\mu}_0 \in \Delta^n$. Although there are some results of regularity for $\boldsymbol{\mu}_0 \in \Delta_+^n$ using techniques from Malliavin calculus, especially under the assumption of uniform ellipticity, it seems difficult to prove the continuity for $\boldsymbol{\mu}_0 \in \partial\Delta^n$ since we see in Proposition 3.2.1 that the covariance matrix $a(\cdot)$ degenerates on the boundary $\partial\Delta^n$.

Theorem 3.2.2 (Existence and uniqueness of stationary distribution of weights for polynomial volatility). *Under the same assumption as in Proposition 3.2.2, i.e. $\varsigma \in \mathbb{M}^{n,d}(\mathbb{R})$, $d \geq n$ is of full rank and $\zeta \in \mathbb{N}^*$ so the volatility $\sigma(\cdot)$ is polynomial, the stationary distribution of market weights, denoted as $m_\infty(\cdot)$ exists and is unique and absolutely continuous with respect to the Lebesgue measure in \mathbb{R}^{n-1} . Moreover, with the notation d_{TV} for the total variation distance between two probability distributions, the speed of convergence of the distribution of market weights $m_t(\cdot | \boldsymbol{\mu}_0)$ to the unique stationary distribution $m_\infty(\cdot)$ when $t \uparrow \infty$ is subgeometric, more precisely for each $\gamma > 0$, there exists a constant $C(\gamma) > 0$ independent of t and $\boldsymbol{\mu}_0$, such that the total variation distance between them can be written as follows:*

$$d_{TV}(m_t(\cdot | \boldsymbol{\mu}_0), m_\infty(\cdot)) \leq C(\gamma) \cdot (V(\boldsymbol{\mu}_0)/t^{\gamma+1} + 1/t^\gamma)$$

$$V(\boldsymbol{\mu}) = - \sum_{i=1}^n \log(\mu^i)$$

Proof. (i) Since the support Δ^n of the probability distribution of market weights given by market weights equation is compact and the semigroup has Feller property, a handful of results can be applied to show the existence of stationary distribution. For example can apply the continuous version of theorem of Krylov-Bogolioubov. Let us define the sequence of probability distributions $(m_N(\cdot))_{N \in \mathbb{N}^*}$, with $(t_N)_{N \in \mathbb{N}^*}$ a positive and strictly increasing sequence such that $t_N \uparrow \infty$ as $N \uparrow \infty$:

$$m_N(A) := \frac{1}{t_N} \int_0^{t_N} \int_A p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) d\boldsymbol{\mu}_t dt, \quad A \in \mathcal{B}(\Delta^n).$$

It is clear that this sequence is tight because the support Δ^n is compact. By theorem of Prokhorov, by extraction we have a subsequence $(m_{N'}(\cdot))_{N' \in \mathbb{N}^*}$ that converges weakly to a probability distribution, denoted as $m_\infty(\cdot)$:

$$\forall f \in \mathcal{C}_b(\Delta^n, \mathbb{R}), \quad \lim_{N' \uparrow \infty} \langle f, m_{N'} \rangle = \langle f, m_\infty \rangle$$

Using the fact that $\forall t \geq 0$, $|\mathcal{P}_t f|_\infty \leq |f|_\infty$, we have the following estimate:

$$\begin{aligned}
 & |\langle \mathcal{P}_u f, m_N \rangle - \langle f, m_N \rangle| \\
 &= \frac{1}{t_N} \left| \int_0^{t_N} \int_{\Delta^n} \left(\int_{\Delta^n} f(\boldsymbol{\mu}_{t+u}) p_u(\boldsymbol{\mu}_{t+u} | \boldsymbol{\mu}_t) d\boldsymbol{\mu}_{t+u} - f(\boldsymbol{\mu}_t) \right) p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) d\boldsymbol{\mu}_t dt \right| \\
 &\leq \frac{1}{t_N} \left| \int_u^{u+t_N} \int_{\Delta^n} f(\boldsymbol{\mu}_t) p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) d\boldsymbol{\mu}_t dt - \int_0^{t_N} \int_{\Delta^n} f(\boldsymbol{\mu}_t) p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) d\boldsymbol{\mu}_t dt \right| \\
 &= \frac{1}{t_N} \left(\int_{t_N}^{u+t_N} |\mathcal{P}_t f(\boldsymbol{\mu}_0)| dt + \int_0^u |\mathcal{P}_t f(\boldsymbol{\mu}_0)| dt \right) \leq \frac{2u|f|_\infty}{t_N} \rightarrow 0 \text{ as } N \uparrow \infty
 \end{aligned}$$

Since the semigroup associated to the process of market weights given by $MWE(\zeta)$ has Feller property, $\forall f \in \mathcal{C}_b(\Delta^n, \mathbb{R})$, $\forall u > 0$, $\mathcal{P}_u f \in \mathcal{C}_b(\Delta^n, \mathbb{R})$ then:

$$\langle \mathcal{P}_u f, m_\infty \rangle - \langle f, m_\infty \rangle = \lim_{N \rightarrow \infty} (\langle \mathcal{P}_u f, m_N \rangle - \langle f, m_N \rangle) = 0$$

Therefore $m_\infty(\cdot)$ is a stationary distribution associated to the process of market weights given by $MWE(\zeta)$.

- (ii) To show the uniqueness of stationary distribution under conditions stated above, we use the criterion from Theorem 4.1 in [Hai16], which works as follows. We first need to construct a Lyapunov function V defined on Δ^n with values in $[1, \infty]$ and such that there exist a constant $\kappa \in \mathbb{R}$ and an increasing and strictly concave function $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\varphi(0) = 0$ with $\varphi(x) \uparrow \infty$ as $x \uparrow \infty$, such that the condition $\mathcal{L}V \leq \kappa - \varphi \circ V$ is satisfied. Assume that for all $c > 0$, there exists $\alpha > 0$ and $t > 0$ such that

$$(\forall \boldsymbol{\mu}_0, \boldsymbol{\mu}'_0 \in \Delta^n, V(\boldsymbol{\mu}_0) + V(\boldsymbol{\mu}'_0) < c) \Rightarrow d_{TV}(m_t(\cdot | \boldsymbol{\mu}_0), m_t(\cdot | \boldsymbol{\mu}'_0)) \leq 2(1 - \alpha), \tag{3.2.14}$$

where d_{TV} is the total variation distance between two probability distributions, and $m_t(\cdot | \boldsymbol{\mu}_0)$ is the probability distribution at time t of market weights starting from $\boldsymbol{\mu}_0$ at time 0. Then the stationary distribution $m_\infty(\cdot)$ is unique and we also have the bound for the convergence speed to the stationary distribution:

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, d_{TV}(m_t(\cdot | \boldsymbol{\mu}_0), m_\infty(\cdot)) \leq \frac{C'V(\boldsymbol{\mu}_0)}{H_\varphi^{-1}(t)} + \frac{C'}{(\varphi \circ H_\varphi^{-1})(t)},$$

where $H_\varphi(u) = \int_1^u (\varphi(s))^{-1} ds$ and $C' > 0$ is a constant.

In view of the condition (3.2.14), we would like to have a function that explodes at the boundary of Δ^n , so here we take the opposite of sum of logarithms of all market weights:

$$V(\boldsymbol{\mu}) = - \sum_{i=1}^{n-1} \log(\mu^i) - \log \left(1 - \sum_{j=1}^{n-1} \mu^j \right) = - \sum_{i=1}^n \log(\mu^i)$$

We show in Appendix 3.5.3 that with this choice of function V , for any strictly concave function φ with $\varphi(0) = 0$ and $\varphi(x) \uparrow \infty$ as $x \uparrow \infty$, there exists $\kappa \in \mathbb{R}$

such that $\mathcal{L}V \leq \kappa - \varphi \circ V$. For example by choosing $\varphi(s) = u^\beta$, $\beta \in (0, 1)$, we have:

$$H_\varphi(u) = \frac{1}{1-\beta}(t^{1-\beta} - 1), \quad H_\varphi^{-1}(t) = (1 + (1-\beta)t)^{\frac{1}{1-\beta}},$$

and the convergence speed becomes explicit with $C(\beta) > 0$ a new constant depending on $\beta \in (0, 1)$ and independent of $\boldsymbol{\mu}_0$ and t :

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, \quad d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_\infty(\cdot)) \leq C(\beta) \cdot \left(\frac{V(\boldsymbol{\mu}_0)}{t^{1/(1-\beta)}} + \frac{1}{t^{\beta/(1-\beta)}} \right)$$

Taking $\gamma = \beta/(1-\beta)$, the function $x \mapsto x/(1-x)$, $(0, 1) \rightarrow (0, +\infty)$ is bijective, and for any $\beta \in (0, 1)$, we have $\gamma \in (0, +\infty)$ and denoting $C(\gamma) = C(\beta)$, the convergence rate estimate becomes,

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, \quad d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_\infty(\cdot)) \leq C(\gamma) \cdot \left(\frac{V(\boldsymbol{\mu}_0)}{t^{\gamma+1}} + \frac{1}{t^\gamma} \right)$$

To finish the proof, we need now to show that condition (3.2.14) holds true. It is easy to see that given $c > 0$, there exists $\epsilon_c \in (0, 1/n)$ such that if $\boldsymbol{\mu}_0, \boldsymbol{\mu}'_0$ satisfy $V(\boldsymbol{\mu}_0) + V(\boldsymbol{\mu}'_0) < c$, necessarily we have $\boldsymbol{\mu}_0, \boldsymbol{\mu}'_0 \in \Delta_{\epsilon_c}^n \subsetneq \Delta_+^n$. The condition on total variation: $\exists t > 0, \exists \alpha > 0, d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_t(\cdot|\boldsymbol{\mu}'_0)) \leq 2(1-\alpha)$ can be interpreted by saying that there exists a time $t > 0$ such that the probability distributions at t of two processes of market weights starting from $\boldsymbol{\mu}_0, \boldsymbol{\mu}'_0$ in $\Delta_{\epsilon_c}^n$ are not too far from each other. We can express the total variation distance in terms of transition density:

$$d_{TV}(m_t(\cdot|\boldsymbol{\mu}_0), m_t(\cdot|\boldsymbol{\mu}'_0)) = \int_{\Delta^n} |p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) - p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0)| d\boldsymbol{\mu}_t$$

We say that $(m_t(\cdot|\boldsymbol{\mu}_0))_{t \geq 0}$ satisfies localized Doeblin condition on $\Delta_{\epsilon_c}^n$ if $\exists t^* > 0, \exists p_c > 0$ such that $\forall \boldsymbol{\mu}_0 \in \Delta_{\epsilon_c}^n, m_{t^*}(\cdot|\boldsymbol{\mu}_0) \geq p_c \lambda(\cdot)$, where $\lambda(\cdot)$ is the Lebesgue measure on Δ^n . In terms of density we can equivalently write $\forall \boldsymbol{\mu} \in \Delta^n, \forall \boldsymbol{\mu}_0 \in \Delta_{\epsilon_c}^n, p_{t^*}(\boldsymbol{\mu}|\boldsymbol{\mu}_0) \geq p_c$. We define *restricted* localized Doeblin condition as $\forall \boldsymbol{\mu}_0 \in \Delta_{\epsilon_c}^n, m_{t^*}(\cdot|\boldsymbol{\mu}_0) \geq p_c \lambda(\cdot)$, but with $m_{t^*}(\cdot|\boldsymbol{\mu}_0)$ and $\lambda(\cdot)$ restricted to $\Delta_{\epsilon_c}^n$, i.e. $p_{t^*}(\boldsymbol{\mu}|\boldsymbol{\mu}_0) \geq p_c, \forall \boldsymbol{\mu}_0, \boldsymbol{\mu} \in \Delta_{\epsilon_c}^n$.

Let us show that if *restricted* localized Doeblin condition is satisfied on $\Delta_{\epsilon_c}^n$, then there exists a time $t > 0$ and $\alpha_c > 0$ such that the condition on total variation is satisfied: $\|m_t(\cdot|\boldsymbol{\mu}_0) - m_t(\cdot|\boldsymbol{\mu}'_0)\|_{TV} \leq 2(1-\alpha_c)$. We split the integral of the difference between densities as follows:

$$\begin{aligned} & \int_{\Delta^n} |p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) - p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0)| d\boldsymbol{\mu}_t \\ &= \int_{\Delta_{\epsilon_c}^n} |p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) - p_c - (p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0) - p_c)| d\boldsymbol{\mu}_t + \int_{\Delta^n \setminus \Delta_{\epsilon_c}^n} |p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) - p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0)| d\boldsymbol{\mu}_t \\ &\leq \int_{\Delta_{\epsilon_c}^n} (p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) - p_c + p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0) - p_c) d\boldsymbol{\mu}_t + \int_{\Delta^n \setminus \Delta_{\epsilon_c}^n} (p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) + p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0)) d\boldsymbol{\mu}_t \\ &= \int_{\Delta^n} (p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}_0) + p_t(\boldsymbol{\mu}_t|\boldsymbol{\mu}'_0)) d\boldsymbol{\mu}_t - 2p_c \lambda(\Delta_{\epsilon_c}^n) = 2(1 - p_c \lambda(\Delta_{\epsilon_c}^n)). \end{aligned}$$

The positive constant α_c is then defined as $\alpha_c = p_c \lambda(\Delta_{\epsilon_c}^n) > 0$.

As pointed out in [Lö13], local Doeblin condition is easily satisfied when the matrix of diffusion a is uniformly elliptic, which is not our case. However, in our the present locally elliptic setting, the restricted localized Doeblin condition follows from Theorem 21 in [Bal06], provided we can show that the assumptions of this theorem are satisfied.

Let us show that the assumption of the theorem are satisfied by the our process of market weights when $\zeta \in \mathbb{N}^*$: First using the same localization as for Hörmander's theorem, the boundedness, linear growth and differentiability of drift and volatility coefficients are obtained for polynomial drift and volatility. The set $\Delta_{\epsilon_c}^n$ being clearly convex, for any $t > 0$ and $\boldsymbol{\mu}_0, \boldsymbol{\mu}_t \in \Delta_{\epsilon_c}^n$ we take the segment $\{\boldsymbol{\mu}; \boldsymbol{\mu} = u \cdot \boldsymbol{\mu}_0 + (1 - u) \cdot \boldsymbol{\mu}_t, 0 \leq u \leq 1\} \subset \Delta_{\epsilon_c}^n$ as a differentiable path, then from Proposition 3.2.1 the matrix of diffusion a is nondegenerate for all $\boldsymbol{\mu}$ along this path and all its eigenvalues are positive and finite. □

Remark 3.2.4 (Equation of stationary density). Suppose that there exists a non-negative solution $p(\boldsymbol{\mu}) \in \mathcal{C}^2(\Delta^n, \mathbb{R}_+)$ to the equation $\mathcal{L}^*p = 0$ where \mathcal{L}^* is the adjoint of infinitesimal operator associated to market weights equation defined in (3.2.3), with boundary condition $p(\boldsymbol{\mu}) = 0$ for $\boldsymbol{\mu} \in \partial\Delta^n$ and normalization condition $\int_{\Delta^n} p(\boldsymbol{\mu}) d\boldsymbol{\mu} = 1$. Then from Proposition 8.2 in [Bas98] the probability distribution $m(\boldsymbol{\mu}) = p(\boldsymbol{\mu}) d\boldsymbol{\mu}$ of density $p(\boldsymbol{\mu})$ is stationary. The equation $\mathcal{L}^*p = 0$ could also be seen as the Fokker-Planck equation with time derivative equal to 0 and is written as follows:

$$\sum_{i=1}^{n-1} \partial_i [b^i(\boldsymbol{\mu}) p(\boldsymbol{\mu})] = \frac{1}{2} \sum_{i,j=1}^{n-1} \partial_{ij}^2 [a^{i,j}(\boldsymbol{\mu}) p(\boldsymbol{\mu})], \quad (3.2.15)$$

with the following nonnegativity, boundary and normalization conditions:

$$p(\boldsymbol{\mu}) \geq 0 \quad \forall \boldsymbol{\mu} \in \Delta^n; \quad p(\boldsymbol{\mu}) = 0 \quad \forall \boldsymbol{\mu} \in \partial\Delta^n; \quad \int_{\Delta^n} p(\boldsymbol{\mu}) d\boldsymbol{\mu} = 1 \quad (3.2.16)$$

In the example below, this equation can be solved explicitly for $n = 2$ under some additional model assumptions. However this equation is difficult to solve analytically in more general cases, and we cannot obtain uniqueness of the stationary distribution based on this equation.

Example 3.2.1 (Explicit solution for two assets). In the case of two assets $n = 2$, we have $\forall t \geq 0, \mu_t^1, \mu_t^2 > 0, \mu_t^1 + \mu_t^2 = 1$ and introduce the notation $x = \mu_t^1$. The explicit form of the stationary distribution is obtained by solving (3.2.15) in the following special cases:

1. For the example (3.2.13) in Remark 3.2.3, with exponent $\zeta = 1, \alpha, \beta \geq 0, \varsigma = [\varsigma^1 \ \varsigma^2]^\top$ and $\varsigma^1 \neq \varsigma^2$.

$$b(x) = \beta - (\alpha + \beta)x, \quad \sigma(x) = (\varsigma^1 - \varsigma^2)x(1 - x) > 0 \quad \forall x \in (0, 1)$$

The 1-dimensional case of (3.2.15) satisfied by $p(x)$ is written as follows, with corresponding conditions as in (3.2.16):

$$\frac{(\zeta^1 - \zeta^2)^2}{2} \frac{d^2}{dx^2} [x^2(1-x)^2 p(x)] = \frac{d}{dx} [(\beta - (\alpha + \beta)x)p(x)] \quad (3.2.17)$$

Then the density of stationary distribution is written as follows:

$$p(x) = C_p \cdot \exp \left(-\frac{2}{(\zeta^1 - \zeta^2)^2} \left[(\alpha - \beta) \log \left(\frac{x}{1-x} \right) + \frac{\alpha}{1-x} + \frac{\beta}{x} \right] \right)$$

where $C_p > 0$ is the normalizing constant to make $\int_0^1 p(x) dx = 1$

2. (2D Wright-Fisher/Jacobi diffusion or volatility-stabilized market weights) With $\zeta = 1/2$ and $d = n = 2$:

$$b(x) = \alpha(\xi - x), \quad \sigma(x) = \sqrt{cx}(1-x); \quad \alpha, c > 0, \quad \xi \in (0, 1)$$

The density of stationary distribution is written as:

$$p(x) = C_p \cdot x^{\alpha\xi/c-1} (1-x)^{\alpha(1-\xi)/c-1}$$

The constant C_p can be written with $B(\cdot, \cdot)$ and $\Gamma(\cdot)$, Beta and Gamma functions:

$$C_p = (B(\alpha\xi/c, \alpha(1-\xi)/c))^{-1} = \frac{\Gamma(\alpha\xi/c + \alpha(1-\xi)/c)}{\Gamma(\alpha\xi/c) \cdot \Gamma(\alpha(1-\xi)/c)}$$

This density is well defined if and only if $\alpha\xi/c > 1$ and $\alpha(1-\xi)/c > 1$. In fact, it has been shown in [EK81] that in the general case for any $n \geq 2$, the stationary distribution of Wright-Fisher/Jacobi diffusion is a Dirichlet distribution.

Proof. From equation (3.2.17), by integrating with respect to x we have:

$$\frac{d}{dx} [x^2(1-x)^2 p(x)] = \frac{2}{(\zeta^1 - \zeta^2)^2} (-\alpha x + \beta(1-x)) p(x) + C_1, \quad C_1 \in \mathbb{R}$$

It is a first-order ODE and its homogeneous equation can be written as:

$$d \log \tilde{p}(x) = \left[\frac{2}{(\zeta^1 - \zeta^2)^2} \left(\frac{\beta - \alpha}{x(1-x)} - \frac{\alpha}{(1-x)^2} + \frac{\beta}{x^2} \right) - 2 \left(\frac{1}{x} - \frac{1}{1-x} \right) \right] dx$$

By integrating the right hand side of the previous equality, we get its solution:

$$\tilde{p}(x) = \frac{C}{x^2(1-x)^2} \exp(-h(x)); \quad h(x) = \frac{2}{(\zeta^1 - \zeta^2)^2} \left[(\alpha - \beta) \log \left(\frac{x}{1-x} \right) + \frac{\alpha}{1-x} + \frac{\beta}{x} \right]$$

Then we apply the variation of constants to get the solution of the inhomogeneous equation:

$$p(x) = \frac{C(x)}{x^2(1-x)^2} \exp(-h(x)); \quad C(x) = C_p + C_1 \cdot \int_{x_0}^x \exp(h(u)) du$$

Let us show that in order to respect the boundary condition $p(0) = p(1) = 0$, the constant C_1 is necessarily equal to zero so $C(x) \equiv C_p$ is a constant. For example, considering the limit when $x \uparrow 1$, thanks to L'Hôpital's rule, we get:

$$\begin{aligned} h'(x) &= \frac{2}{(\zeta^1 - \zeta^2)^2} \left[\frac{(\alpha - \beta)}{x(1-x)} + \frac{\alpha}{(1-x)^2} - \frac{\beta}{x^2} \right] \\ \lim_{x \uparrow 1} p(x) &= \lim_{x \uparrow 1} [C_p + C_1 \cdot \int_{x_0}^x \exp(h(u)) du] \cdot [x^2(1-x)^2 \exp(h(x))]^{-1} \\ &= \lim_{x \uparrow 1} C_1 \cdot [2x(1-x)(2x-1) + x^2(1-x)^2 h'(x)]^{-1} = \frac{C_1(\zeta^1 - \zeta^2)^2}{2\alpha} \end{aligned}$$

With the assumption on the coefficients $\zeta^1 \neq \zeta^2$, $\alpha > 0$, the limit above is zero if and only if $C_1 = 0$. For the limit when $x \downarrow 0$, similar calculation yields the same necessary condition. From the normalization condition, the constant C_p is the unique positive constant such that the integral on $[0, 1]$ of $p(x)$ is equal to 1. Finally, the density function of the stationary distribution of the market weights in the case of two assets $n = 2$ is the unique solution of equation (3.2.17) satisfying the normalization condition:

$$p(x) = \tilde{p}(x) = C_p \cdot \exp \left(-\frac{2}{(\zeta^1 - \zeta^2)^2} \left[(\alpha - \beta) \log \left(\frac{x}{1-x} \right) + \frac{\alpha}{1-x} + \frac{\beta}{x} \right] \right)$$

The computations are similar for 2D Wright-Fisher/Jacobi diffusion and the normalizing constant C_p is given by the following integral which is nothing other than a Beta function and can be written as product and quotient of Gamma functions:

$$C_p^{-1} = \int_0^1 x^{\alpha\xi/c-1} (1-x)^{\alpha(1-\xi)/c-1} dx = B(\alpha\xi/c, \alpha(1-\xi)/c) = \frac{\Gamma(\alpha\xi/c + \alpha(1-\xi)/c)}{\Gamma(\alpha\xi/c) \cdot \Gamma(\alpha(1-\xi)/c)}$$

□

3.3 Structure of market and relative arbitrage

3.3.1 Market portfolio and growth rate

The market weights of assets $(\boldsymbol{\mu}_t)_{t \geq 0}$ given by the solution of market weights equation $MWE(\zeta)$ (3.2.9) studied above only represent the relative sizes of assets in the market, in order to have knowledge of the total market structure, we shall also consider the size of the market portfolio, or market total capitalization $(\Sigma_t)_{t \geq 0}$. The process given by the couple $(\boldsymbol{\mu}_t, \Sigma_t)_{t \geq 0}$ describes the complete structure of the market.

Definition 3.3.1 (Dynamics of total capitalization). The dynamics of the total market capitalization $(\Sigma_t)_{t \geq 0}$ is given by the following SDE:

$$d\Sigma_t = b^\Sigma(\Sigma_t) dt + \sum_{\nu=0}^d \sigma^{\Sigma, \nu}(\Sigma_t) dW_t^\nu, \quad \Sigma_0 > 0.$$

Here, $\sigma^{\Sigma, i}(\cdot)$, $i = 0, \dots, d$ are deterministic volatility functions, $b^\Sigma(\cdot)$ is a deterministic drift function and $W_t^0 = W_t^\Sigma$ is a Brownian motion independent of the d -dimensional Brownian motion $((W_t^\nu)_{\nu=1}^d)_{t \geq 0}$ in the dynamics of weights. We suppose that the total market capitalization does not vanish, *i.e.* $\Sigma_t > 0$, $\forall t \geq 0$ *a.s.* and $\sigma^\Sigma(\cdot) > 0$ on \mathbb{R}_+^* ,

Remark 3.3.1 (Dependency between market weights and total capitalization). For both volatility-stabilized markets and joint polynomial model introduced in [Cuc19], the total market capitalization process $(\Sigma_t)_{t \geq 0}$ is independent of the market weights process $(\boldsymbol{\mu}_t)_{t \geq 0}$. In our model, the dependency between the total market capitalization process and the market weights process is quantified by volatility coefficients $(\sigma^{\Sigma, \nu}(\cdot))_{\nu=1}^d$. However, we focus on the process of market weights and as we will see later, if we use the total capitalization as numéraire, the dynamics of total capitalization become rather symbolic in our discussions, and the assumption of positivity *i.e.* $\Sigma_t > 0, \forall t \geq 0$ *a.s.* is sufficient.

The individual asset capitalization processes are written as $S_t^i = \mu_t^i \cdot \Sigma_t$ with dynamics given by the following SDE:

$$\begin{aligned} dS_t^i &= \tilde{b}^i(\boldsymbol{\mu}_t, \Sigma_t)dt + \sum_{\nu=0}^d \tilde{\sigma}^{i, \nu}(\boldsymbol{\mu}_t) dW_t^\nu; \quad \tilde{\sigma}^{i, 0}(\cdot) = \sigma^\Sigma(\cdot), \quad W_t^0 = W_t^\Sigma \\ \tilde{b}^i(\boldsymbol{\mu}, \Sigma) &= b^\Sigma(\Sigma) + b^i(\boldsymbol{\mu}) + \sum_{\nu=1}^d \sigma^{i, \nu}(\boldsymbol{\mu}) \sigma^{\Sigma, \nu}(\Sigma), \quad (\boldsymbol{\mu}, \Sigma) \in \Delta^n \times \mathbb{R}_+^* \\ \tilde{\sigma}^{i, \nu}(\boldsymbol{\mu}, \Sigma) &= \sigma_t^{i, \nu}(\boldsymbol{\mu}) + \sigma_t^{\Sigma, \nu}(\Sigma), \quad \nu \in \{1, \dots, d\}, \end{aligned} \tag{3.3.1}$$

where $\tilde{\mathbf{b}}(\cdot) = (\tilde{b}^i(\cdot))_{i=1}^n$ is the total vector of drift and $\tilde{\sigma}(\cdot) = [\tilde{\sigma}^{i, \nu}(\cdot)] \in \mathbb{M}^{n, d+1}(\mathbb{R})$ is the total matrix of volatility corresponding to the process of capitalizations or prices.

Remark 3.3.2 (Decomposition of growth rate). The logarithm of the process of market capitalization of the i -th asset S_t^i is written as the sum of logarithm of the process of relative market weights and that of total market capitalization:

$$\begin{aligned} d \log S_t^i &= d \log(\mu_t^i \cdot \Sigma_t) = d \log \mu_t^i + d \log \Sigma_t, \quad a^{\Sigma, \Sigma}(\cdot) = \frac{d(\Sigma)_t}{dt} = \sum_{\nu=0}^d (\sigma^{\Sigma, \nu}(\cdot))^2 \\ &= \left((\mu_t^i)^{-1} b^i(\boldsymbol{\mu}_t) - \frac{1}{2} (\mu_t^i)^{-2} a^{i, i}(\boldsymbol{\mu}_t) + (\Sigma_t)^{-1} b^\Sigma(\Sigma_t) - \frac{1}{2} (\Sigma_t)^{-2} a^{\Sigma, \Sigma}(\Sigma_t) \right) dt \\ &\quad + \sum_{\nu=1}^d \left((\mu_t^i)^{-1} \sigma^{i, \nu}(\boldsymbol{\mu}_t) + (\Sigma_t)^{-1} \sigma^{\Sigma, \nu}(\Sigma_t) \right) dW_t^\nu + (\Sigma_t)^{-1} \sigma^\Sigma(\Sigma_t) dW_t^\Sigma \end{aligned}$$

The total growth rate of the i -th asset $\tilde{\gamma}^i(\cdot)$ is defined as the drift coefficient in the dynamics of $\log S_t^i$ and is decomposed into the growth rate $\gamma^i(\cdot)$ of relative market weights defined in (3.2.7) that is specific to the i -th asset, plus the growth rate $\gamma^\Sigma(\cdot)$ of the total market capitalization, or the size of market portfolio common to all assets:

$$\tilde{\gamma}^i(\boldsymbol{\mu}, \Sigma) = \gamma^i(\boldsymbol{\mu}) + \gamma^\Sigma(\Sigma), \quad \gamma^\Sigma(\Sigma) = \Sigma^{-1} b^\Sigma(\Sigma) - \frac{1}{2} \Sigma^{-2} a^{\Sigma, \Sigma}(\Sigma)$$

Under conditions (i) - (iii) in Theorem 3.2.1, as already discussed in step (3) of the proof, $\gamma^i(\boldsymbol{\mu}) > 0$ when $\mu^i \rightarrow 0$, which makes the weight μ^i get away from 0 and become positive. Similarly, when the i -th relative weight is close to dominate in the market, *i.e.* $\mu^i \rightarrow 1$, the first term $B^{i, i} < 0$ is predominant and we get negative growth rate that diminishes μ^i .

Remark 3.3.3 (*Effect of volatility stabilization with exponent ζ*). By setting drift and volatility of the total capitalization process to zero, *i.e.* $b^\Sigma(\cdot) = 0$, $\sigma^\Sigma(\cdot) = 0$, $\sigma^{\Sigma,\nu}(\cdot) = 0$, we recognize the logarithm representation of market weights:

$$d \log \mu_t^i = \gamma^i(\boldsymbol{\mu}) dt + \sum_{\nu=1}^d (\mu_t^i)^{-1} \sigma^{i,\nu}(\boldsymbol{\mu}_t) dW_t^\nu$$

The total volatility of the market weight process of i -th asset is defined as:

$$\begin{aligned} \sigma^i(\boldsymbol{\mu}_t)^2 &= \sum_{\nu=1}^d (\mu_t^i)^{-2} \sigma^{i,\nu}(\boldsymbol{\mu}_t)^2 = \sum_{\nu=1}^d (\mu_t^i)^{-2} \left((\mu_t^i)^\zeta \zeta^{i,\nu} - \mu_t^i \sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu} \right)^2 \\ &= (\mu_t^i)^{2(\zeta-1)} (1 - 2\mu_t^i) \sum_{\nu=1}^d (\zeta^{i,\nu})^2 + \sum_{\nu=1}^d \left(\sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu} \right)^2 - 2(\mu_t^i)^{(\zeta-1)} \sum_{\nu=1}^d \zeta^{i,\nu} \sum_{k=1, k \neq i}^n (\mu_t^k)^\zeta \zeta^{k,\nu} \end{aligned}$$

First recall that under Assumption 3.2.1, $c^i = \sum_{\nu=1}^d (\zeta^{i,\nu})^2 > 0$. For $\zeta \in (0, 1)$, when μ_t^i is small, the first term $(\mu_t^i)^{2(\zeta-1)} (1 - 2\mu_t^i) c^i$ in the expression above is predominant, so $\sigma^i(\boldsymbol{\mu}_t) \rightarrow \infty$ as $\mu_t^i \rightarrow 0$. As a result, the total volatility of i -th asset is larger when its weight is smaller so the overall market stability is maintained, see Proposition 4.1 in [FK05], and we call the phenomenon *asymptotic volatility stabilization*. Moreover, the smaller the exponent ζ , the higher the total volatility for fixed sufficiently small weight μ_t^i . On the contrary, for $\zeta \geq 1$, $\sigma^i(\boldsymbol{\mu}_t)$ is bounded for $\boldsymbol{\mu}_t \in \Delta_n$ and without additional analysis it is not clear whether $\sigma^i(\boldsymbol{\mu}_t)$ increases as μ_t^i decreases, thus we do not observe the effect of volatility stabilization.

3.3.2 Trading strategy and relative arbitrage

As often in SPT framework, a trading strategy $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t \geq 0}$ is defined by the quantities invested in risky assets $\boldsymbol{S} = (\boldsymbol{S}_t)_{t \geq 0}$, with no investment in the money market. The trading strategy $\boldsymbol{\pi}$ is a \mathbb{R}^n -valued predictable process, integrable with respect to the process \boldsymbol{S} , and we use the notation $\mathcal{L}(\boldsymbol{S})$ as in [KR17] for the collection of such strategies. Moreover we suppose that interest and dividend rates are equal to 0. The *wealth process*, or value of investment portfolio associated to $\boldsymbol{\pi}$ with initial value v is denoted as $\tilde{V}^{v,\boldsymbol{\pi}} = (\tilde{V}_t^{v,\boldsymbol{\pi}})_{t \geq 0}$:

$$\tilde{V}_t^{v,\boldsymbol{\pi}} = \boldsymbol{\pi}_t \cdot \boldsymbol{S}_t = \sum_{i=1}^n \pi_t^i S_t^i, \quad \tilde{V}_0^{v,\boldsymbol{\pi}} = v$$

Definition 3.3.2 (*Self-financing strategy and relative arbitrage*). A trading strategy $\boldsymbol{\pi} \in \mathcal{L}(\boldsymbol{S})$ with its associated wealth process $\tilde{V}^{v,\boldsymbol{\pi}}$ is self-financing with respect to \boldsymbol{S} if:

$$\tilde{V}_t^{v,\boldsymbol{\pi}} = \tilde{V}_0^{v,\boldsymbol{\pi}} + \int_0^t \sum_{i=1}^n \pi_u^i dS_u^i, \quad \tilde{V}_0^{v,\boldsymbol{\pi}} = v$$

In this chapter we only consider self-financing trading strategies, and the collection of such strategies will be denoted by $\mathcal{T}(\boldsymbol{S})$. Normalizing by v we may suppose that the initial investment v is equal to 1 and denote $\tilde{V}_t^\boldsymbol{\pi} = \tilde{V}_t^{1,\boldsymbol{\pi}}$.

By analogy with the weights of the market portfolio, we define the weights of assets $\tilde{\pi} = (\tilde{\pi}^i)_{i=1}^n$ associated to the trading strategy π . Clearly they are given by:

$$\tilde{\pi}_t^i = \frac{\pi_t^i S_t^i}{\tilde{V}_t^\pi} \quad (3.3.2)$$

We say that the trading strategy $\pi \in \mathcal{T}(\mathbf{S})$ realizes a *relative arbitrage* with respect to the investment strategy $\rho \in \mathcal{T}(\mathbf{S})$ over the horizon $[0, T]$ if:

$$\mathbb{P}[\tilde{V}_T^\pi \geq \tilde{V}_T^\rho] = 1 \text{ and } \mathbb{P}[\tilde{V}_T^\pi > \tilde{V}_T^\rho] > 0$$

It is a strong relative arbitrage if:

$$\mathbb{P}[\tilde{V}_T^\pi > \tilde{V}_T^\rho] = 1$$

Definition 3.3.3 (Change of numéraire). We introduce the value of the portfolio *in the numéraire* of the market portfolio $V_t^{v,\pi} = \tilde{V}_t^{v,\pi}/\Sigma_t$ with initial value $V_0^{v,\pi} = v/\Sigma_0$. Similarly, we define $V_t^\pi = V_t^{\Sigma_0,\pi}$ so that $V_0^\pi = 1$. The dynamics of this newly defined investment portfolio (or wealth process) V_t^π can be expressed in terms of the market weights $\boldsymbol{\mu}_t = (\mu_t^i)_{i=1}^n$:

$$dV_t^\pi = \boldsymbol{\pi}_t \cdot d\boldsymbol{\mu}_t = \sum_{i=1}^n \pi_t^i \cdot [b^i(\boldsymbol{\mu}_t)dt + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t)dW_t^\nu] \quad (3.3.3)$$

From Proposition 2.3 in [KR17], a \mathbb{R}^n -valued process trading strategy $\boldsymbol{\pi} = (\boldsymbol{\pi}_t)_{t \geq 0}$ is a trading strategy with respect to the process of market weights $\boldsymbol{\mu} = (\boldsymbol{\mu}_t)_{t \geq 0}$ if and only if it is a trading strategy with respect to the process of market prices $\mathbf{S} = (\mathbf{S}_t)_{t \geq 0}$, or equivalently in a simpler way, $\mathcal{T}(\boldsymbol{\mu}) = \mathcal{T}(\mathbf{S})$. The weights of assets $\tilde{\pi}$ defined in (3.3.2), in terms of market weights can be expressed by $\tilde{\pi}_t^i = \pi_t^i \mu_t^i / V_t^\pi$. From now on, we use the unique notation $\boldsymbol{\pi}$ for both quantities invested in the assets and percentage invested in the assets, depending on the context.

One particular trading strategy is that of the market portfolio, in this case we simply have $\forall t \geq 0$, $\boldsymbol{\pi}_t = \mathbf{1}$ or equivalently $\tilde{\pi}_t = \boldsymbol{\mu}_t$ and $V_t^\pi = 1$, the value market portfolio is always equal to 1 in its own numéraire. The relative arbitrage we are naturally interested in is that with respect to the market portfolio. Given a time horizon $[0, T]$, an auto-financing trading strategy $\boldsymbol{\pi} \in \mathcal{T}(\boldsymbol{\mu})$ with its wealth process $(V_t^\pi)_{t \geq 0}$ realizes a relative arbitrage with respect to the market, or simply “beats the market” if:

$$\begin{aligned} \mathbb{P}[V_T^\pi \geq 1] = 1 \text{ and } \mathbb{P}[V_T^\pi > 1] > 0 \\ \mathbb{P}[V_T^\pi > 1] = 1 \text{ for strong relative arbitrage} \end{aligned} \quad (3.3.4)$$

3.3.3 Conditions of arbitrage relative to the market

In this section we are interested in the conditions of existence of arbitrage relative to the market. There are two types of conditions: the first one is based on the the excess growth rate relative to the market, and for the second condition, we shall use the market risk premium and the corresponding market risk deflator.

Definition 3.3.4 (Growth and excess growth rates). Given a trading strategy $\pi \in \mathcal{T}(\boldsymbol{\mu})$, we write the logarithm of its wealth process:

$$d \log V_t^\pi = \gamma^\pi(\boldsymbol{\mu}_t) dt + \sum_{i=1}^n \pi_t^i (\mu_t^i)^{-1} \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_t) dW_t^\nu,$$

where $\gamma^\pi(\cdot)$ is the growth rate associated to π (recall that $\gamma^i(\cdot)$ is the growth rate of the i -th weight defined in (3.2.7)):

$$\gamma^\pi(\boldsymbol{\mu}_t) = \sum_{i=1}^n \pi_t^i \gamma^i(\boldsymbol{\mu}_t) + \frac{1}{2} \left(\sum_{i=1}^n \pi_t^i (\mu_t^i)^{-2} a^{i,i}(\boldsymbol{\mu}_t) - \sum_{i,j=1}^n \pi_t^i \pi_t^j (\mu_t^i \mu_t^j)^{-1} a^{i,j}(\boldsymbol{\mu}_t) \right)$$

On the right hand side, the first term corresponds to the weighted sum of growth rates of all assets, and the second term is what we call *excess growth rate* associated to π , denoted as $\gamma_*^\pi(\cdot)$:

$$\gamma_*^\pi(\boldsymbol{\mu}_t) = \gamma^\pi(\boldsymbol{\mu}_t) - \sum_{i=1}^n \pi_t^i \gamma^i(\boldsymbol{\mu}_t) = \frac{1}{2} \left(\sum_{i=1}^n \pi_t^i (\mu_t^i)^{-2} a^{i,i}(\boldsymbol{\mu}_t) - \sum_{i,j=1}^n \pi_t^i \pi_t^j (\mu_t^i \mu_t^j)^{-1} a^{i,j}(\boldsymbol{\mu}_t) \right)$$

The excess growth rate of the market is denoted as $\gamma_*(\cdot) := \gamma_*^\mu(\cdot)$ for $\pi = \boldsymbol{\mu}$. Using the fact that $\forall \boldsymbol{\mu}_t \in \Delta^n$, $\sum_{i=1}^n \sigma^{i,\nu}(\boldsymbol{\mu}_t) = 0 \forall \nu \in \{1, \dots, d\}$ so $\sum_{i,j=1}^n a^{i,j}(\boldsymbol{\mu}_t) = \sum_{\nu=1}^d \left(\sum_{i=1}^n \sigma^{i,\nu}(\boldsymbol{\mu}_t) \right)^2 = 0$, the excess growth rate of the market is written as:

$$\gamma_*(\boldsymbol{\mu}_t) = \frac{1}{2} \sum_{i=1}^n (\mu_t^i)^{-1} a^{i,i}(\boldsymbol{\mu}_t) = \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d \left((\mu_t^i)^{\zeta-1/2} \zeta^{i,\nu} - (\mu_t^i)^{1/2} \sum_{k=1}^n (\mu_t^k) \zeta^{k,\nu} \right)^2$$

The cumulative excess growth rate (of the market) $\Gamma_*(t)$ is defined as the integral of excess growth rate along the trajectory of market weights until time t :

$$\Gamma_*(t) = \Gamma_*((\boldsymbol{\mu}_u)_{u \in [0,t]}) = \int_0^t \gamma_*(\boldsymbol{\mu}_u) du = \frac{1}{2} \sum_{i=1}^n \int_0^t (\mu_u^i)^{-1} a^{i,i}(\boldsymbol{\mu}_u) du$$

Proposition 3.1 in [FK05] gives a *sufficient condition* for the existence of relative arbitrage: if the cumulative excess growth rate $\Gamma_*(t)$ is bounded from below almost surely by a continuous, strictly increasing function $\tilde{\Gamma}(t)$ with $\tilde{\Gamma}(0) = 0$, $\tilde{\Gamma}(\infty) = \infty$, then there is arbitrage opportunity relative to the market over the time horizon $[0, T]$, $\forall T > T_*$, where the critical horizon T_* is a function of $\Gamma_*(\cdot)$. A special case easier to check is when the excess growth is bounded from below by a positive constant, *i.e.* $\exists \underline{\gamma} > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma}, \forall t \in [0, T]$ a.s..

Proposition 3.3.1 (Excess growth rate estimates). For $\boldsymbol{\mu}_0 \in \Delta_+^n$, we may distinguish following cases according to the value of the exponent ζ :

1. When $\zeta \leq 1/2$ and Assumptions 3.2.1 and 3.2.3 hold true, there exists a constant $\underline{\gamma} > 0$ such that:

$$\forall t > 0 : \gamma_*(\boldsymbol{\mu}_t) \geq \underline{\gamma} \text{ a.s.}$$

In other words, the sufficient condition for the existence of relative arbitrage to the market on a large enough time horizon is satisfied.

2. When $\zeta \in \mathbb{N}^*$ we have:

$$\forall \epsilon > 0 : \forall t > 0 : \mathbb{P}[\gamma_*(\boldsymbol{\mu}_t) < \epsilon] > 0.$$

Proof. 1. For $\zeta \leq 1/2$ and $\zeta \boldsymbol{\zeta}^\top = \text{diag}(c^1, \dots, c^n)$ with $c^i > 0 \forall i \in \{1, \dots, n\}$, $\gamma_*(\boldsymbol{\mu}_t)$ can be written as follows:

$$\begin{aligned} \gamma_*(\boldsymbol{\mu}_t) &= \frac{1}{2} \sum_{i=1}^n ((\mu_t^i)^{2\zeta-1} c^i + \mu_t^i \sum_{k=1}^n (\mu_t^k)^{2\zeta} c^k - 2(\mu_t^i)^{2\zeta} c^i) \\ &= \frac{1}{2} \sum_{i=1}^n (\mu_t^i)^{2\zeta-1} (1 - \mu_t^i) c^i \geq \frac{1}{2} \sum_{i=1}^n (1 - \mu_t^i) \min_{k \in \{1, \dots, n\}} c^k = \frac{n-1}{2} \min_{k \in \{1, \dots, n\}} c^k > 0 \end{aligned}$$

2. In the case $\zeta \in \mathbb{N}^*$, the excess growth rate $\gamma_*(\boldsymbol{\mu}_t)$ may be written as follows. The following convergences take place uniformly on μ_t^j , $j \neq i$:

$$\begin{aligned} \gamma_*(\boldsymbol{\mu}_t) &= \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d ((\mu_t^i)^{\zeta-1/2} \zeta^{i,\nu} - (\mu_t^i)^{1/2} \sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu})^2 \\ &= \frac{1}{2} \sum_{i=1}^n (\mu_t^i)^{2\zeta-1} \sum_{\nu=1}^d ((1 - \mu_t^i) \zeta^{i,\nu} - (\mu_t^i)^{1-\zeta} \sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu})^2 \end{aligned}$$

At the boundary, we have following equivalences:

- When $\mu_t^i \rightarrow 0$:

$$\sum_{\nu=1}^d ((\mu_t^i)^{\zeta-1/2} \zeta^{i,\nu} - (\mu_t^i)^{1/2} \sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu})^2 \rightarrow 0$$

- When $\mu_t^i \rightarrow 1$:

$$(\mu_t^i)^{2\zeta-1} \sum_{\nu=1}^d ((1 - \mu_t^i) \zeta^{i,\nu} - (\mu_t^i)^{1-\zeta} \sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu})^2 \sim \mu_t^i \sum_{\nu=1}^d (\sum_{k=1}^n (\mu_t^k)^\zeta \zeta^{k,\nu})^2 \rightarrow 0$$

We deduce that for any $\epsilon > 0$, there exists $\delta > 0$ such that for $\boldsymbol{\mu}_t \in \Delta_+^n$ with $\mu_t^i \in (0, \delta] \cup [1 - \delta, 1)$, we have $\gamma_*(\boldsymbol{\mu}_t) \leq \epsilon$. Since $\boldsymbol{\mu}_0 \in \Delta_+^n$, there exists δ' such that $0 < \delta' < \delta$ and $\boldsymbol{\mu}_0 \in \Delta_{\delta'}^n$, i.e. $\mu_0^i \geq \delta'$, $\forall i \in \{1, \dots, n\}$. Let us estimate the probability of $\gamma_*(\boldsymbol{\mu}_t) \leq \epsilon$ given $\boldsymbol{\mu}_0$:

$$\begin{aligned} \mathbb{P}[\gamma_*(\boldsymbol{\mu}_t) \leq \epsilon | \boldsymbol{\mu}_0] &\geq \mathbb{P}[\mu_t^i \in (0, \delta] \cup [1 - \delta, 1) | \boldsymbol{\mu}_0] \\ &\geq \mathbb{P}[\mu_t^i \in [\delta', \delta] \cup [1 - \delta, 1 - \delta'] | \boldsymbol{\mu}_0] \\ &\geq \int_{D_{\delta, \delta'}^i} p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) d\boldsymbol{\mu}_t \geq p_{\delta, \delta', t} \lambda(D_{\delta, \delta'}^i) > 0 \end{aligned}$$

where the domain $D_{\delta, \delta'}^i$ with positive Lebesgue measure in \mathbb{R}^n defined as follows:

$$D_{\delta, \delta'}^i = \{\mu_t^i \in [\delta', \delta] \cup [1 - \delta, 1 - \delta']\} \cap \Delta_\epsilon^n, \lambda(D_{\delta, \delta'}^i) > 0$$

Using the same argument from Theorem 21 in [Bal06] as in the proof of Theorem 3.2.2, there exists a positive constant $p_{\delta, \delta', t}$ such that $p_t(\boldsymbol{\mu}_t | \boldsymbol{\mu}_0) \geq p_{\delta, \delta', t} > 0$ for all $\boldsymbol{\mu}_0 \in \Delta_\epsilon^n$ and $\boldsymbol{\mu}_t \in D_{\delta, \delta'}^i \subset \Delta_\epsilon^n$. \square

However, there is no necessary condition for the existence of relative arbitrage in terms of excess growth, alternatively we rather consider another approach by introducing the market risk premium and market risk deflator.

Definition 3.3.5 (Market risk premium). The total market risk premium (also called market price of risk) $\tilde{\boldsymbol{\theta}}(\cdot) = (\tilde{\theta}^\nu(\cdot))_{\nu=0}^d$ is a vector of \mathbb{R}^{d+1} that satisfies the following:

$$\begin{aligned} \forall (\boldsymbol{\mu}_0, \Sigma_0) \in \Delta_+^n \times \mathbb{R}_+^*, \forall t \geq 0, \tilde{\mathbf{b}}(\boldsymbol{\mu}_t, \Sigma_t) &= \tilde{\boldsymbol{\sigma}}(\boldsymbol{\mu}_t, \Sigma_t) \cdot \tilde{\boldsymbol{\theta}}(\boldsymbol{\mu}_t, \Sigma_t) \\ \forall T \geq 0, \int_0^T |\tilde{\boldsymbol{\theta}}(\boldsymbol{\mu}_t, \Sigma_t)|^2 dt &< \infty \end{aligned}$$

In terms of market risk premium, the i -th asset price S_t^i can be written as follows:

$$dS_t^i = \sum_{\nu=0}^d \tilde{\sigma}^{i, \nu}(\boldsymbol{\mu}_t, \Sigma_t) [\tilde{\theta}^\nu(\boldsymbol{\mu}_t, \Sigma_t) dt + dW_t^\nu].$$

Here we are rather define the market risk premium associated to weights $\boldsymbol{\theta}(\cdot) = (\theta^\nu(\cdot))_{\nu=1}^d$ depending only on the market weights, see for example [Cuc19] for similar definition:

$$\forall \boldsymbol{\mu}_0 \in \Delta_+^n, \forall t \geq 0, \mathbf{b}(\boldsymbol{\mu}_t) = \boldsymbol{\sigma}(\boldsymbol{\mu}_t) \cdot \boldsymbol{\theta}(\boldsymbol{\mu}_t) \text{ and } \forall T \geq 0, \int_0^T |\boldsymbol{\theta}(\boldsymbol{\mu}_t)|^2 dt < \infty$$

By ignoring the dynamics of the total capitalization Σ_t we have:

$$d\mu_t^i = \sum_{\nu=1}^d \sigma^{i, \nu}(\boldsymbol{\mu}_t) [\theta^\nu(\boldsymbol{\mu}_t) dt + dW_t^\nu].$$

Definition 3.3.6 (Market risk deflator, NUPBR and NFLVR). The market risk deflator $(Z_t)_{t \geq 0}$ associated to the market weights $(\boldsymbol{\mu}_t)_{t \geq 0}$ is an exponential local martingale defined from the market risk premium associated to market weights $(\boldsymbol{\theta}(\boldsymbol{\mu}_t))_{t \geq 0}$, in a similar way as the total market risk deflator defined in [FK09], also see Proposition 5.4 in [Cuc19]:

$$\begin{aligned} Z_t &= \mathcal{E} \left(-(\boldsymbol{\theta}(\boldsymbol{\mu}_u))_{u \in [0, t]} \right) = \exp \left(- \int_0^t \boldsymbol{\theta}(\boldsymbol{\mu}_u) \cdot d\mathbf{W}_u - \frac{1}{2} \int_0^t |\boldsymbol{\theta}(\boldsymbol{\mu}_u)|^2 du \right) \\ &= \exp \left(- \int_0^t \sum_{\nu=1}^d \theta^\nu(\boldsymbol{\mu}_u) dW_u^\nu - \frac{1}{2} \int_0^t \sum_{\nu=1}^d (\theta^\nu(\boldsymbol{\mu}_u))^2 du \right), \quad Z_0 = 1 \end{aligned} \tag{3.3.5}$$

Since $(Z_t)_{t \geq 0}$ is a positive local martingale, by Fatou's lemma it is a supermartingale so $\forall t \geq 0, \mathbb{E}[Z_t] \leq 1$. Moreover, $(Z_t)_{t \geq 0}$ is a true martingale on the horizon $[0, T]$ if and only if $\mathbb{E}[Z_T] = 1$.

The notion of *No unbounded profit with bounded risk* (NUPBR) condition was introduced in [DS94] and simply means that for $T > 0$, the set of wealth processes $\{V_T^\pi \mid \boldsymbol{\mu} \in \mathcal{T}(\boldsymbol{\mu})\}$ is bounded in probability. In fact, (NUPBR) on $[0, T]$ is equivalent to the existence of market risk deflator $(Z_t)_{t \geq 0}$. The condition of *No free lunch with vanishing risk* (NFLVR) is satisfied when (NUPBR) holds and there is no relative arbitrage.

In our market weights model given by market weights equation (3.2.9), for any exponent $\zeta > 0$ there is actually equivalence between the existence of market risk premium and the (NUPBR) condition as shown in the next theorem.

Theorem 3.3.1 (Existence of market risk premium and NUPBR). *Let $\boldsymbol{\mu}_0 \in \Delta_+^n$ and suppose that Assumption 3.2.1 holds true. The market risk premium $\boldsymbol{\theta}(\cdot)$, as well as the market risk deflator $(Z_t)_{t \geq 0}$ associated to weights exist almost surely, thus (NUPBR) is satisfied over the time horizon $[0, T]$, $T > 0$, if and only if one of the following conditions hold:*

- (1) *One of the conditions (i) - (iii) in Theorem 3.2.1 holds, recall that in this case $\forall t \geq 0$, $\boldsymbol{\mu}_t \in \Delta_+^n$ a.s..*
- (2) *One of the conditions (iv) - (vi) in Theorem 3.2.1 holds, and $b^i(\boldsymbol{\mu}) = 0$ on $\{\mu^i = 0\}$ for all indices i such that $\mathbb{P}[\exists t \in (0, T) \mid \mu_t^i = 0] > 0$.*

In this case, the market risk premium is written as follows:

$$\boldsymbol{\mu}_0 \in \Delta_+^n, \forall t \geq 0, \boldsymbol{\theta}(\boldsymbol{\mu}_t) = (\boldsymbol{\sigma}^\top \cdot \boldsymbol{\lambda})(\boldsymbol{\mu}_t), \boldsymbol{\lambda}(\boldsymbol{\mu}_t) = (a^+ \cdot \mathbf{b})(\boldsymbol{\mu}_t) \quad (3.3.6)$$

where $a^+(\cdot)$ denotes the Moore–Penrose pseudo-inverse of the covariance matrix $a(\cdot) \in \mathbb{M}^{n-1}(\mathbb{R})$. In addition, the total market risk premium $\boldsymbol{\theta}(\boldsymbol{\mu}_t, \Sigma_t)$ with σ^Σ can be defined from $\boldsymbol{\theta}(\boldsymbol{\mu}_t)$ through the following relation:

$$\begin{aligned} \tilde{\boldsymbol{\theta}}^0(\boldsymbol{\mu}_t, \Sigma_t) &= \left((b^\Sigma - \sum_{\nu=1}^d \sigma^{\Sigma, \nu} (\theta^\nu + \sigma^{\Sigma, \nu})) / \sigma^\Sigma \right) (\boldsymbol{\mu}_t, \Sigma_t) \\ \tilde{\theta}^\nu(\boldsymbol{\mu}_t, \Sigma_t) &= \theta^\nu(\boldsymbol{\mu}_t) + \sigma^{\Sigma, \nu}(\Sigma_t); \nu \in \{1, \dots, d\} \end{aligned} \quad (3.3.7)$$

Let one of the conditions (i)-(iii) in Theorem 3.2.1 be satisfied, and assume that $d = n - 1$ and the reduced covariance matrix $a(\cdot)$ is positive definite on Δ_+^n , or equivalently the reduced volatility matrix $\sigma(\cdot) \in \mathbb{M}^{n-1}(\mathbb{R})$ is invertible. In this case the market risk premium is unique and is written as follows:

$$\boldsymbol{\theta}(\boldsymbol{\mu}_t) = (\boldsymbol{\sigma}^\top a^{-1} \cdot \mathbf{b})(\boldsymbol{\mu}_t) = (\boldsymbol{\sigma}^{-1} \cdot \mathbf{b})(\boldsymbol{\mu}_t)$$

Finally, if none of the conditions (1) and (2) stated above holds, i.e. for some $i \in \{1, \dots, n\}$, $(\mu_t^i)_{t \geq 0}$ vanishes on $[0, T]$ with positive probability and there exists $\boldsymbol{\mu} \in \{\mu^i = 0\}$ such that $b^i(\boldsymbol{\mu}) > 0$, then the (NUPBR) condition is violated, in particular both the market risk premium $\boldsymbol{\theta}(\cdot)$ and deflator $(Z_t)_{t \geq 0}$ do not exist.

Proof. When one of the conditions (i)-(iii) of Theorem 3.2.1 is satisfied, for $\boldsymbol{\mu}_0 \in \Delta_+^n$, we have $\forall t \geq 0$, $\boldsymbol{\mu}_t \in \Delta_+^n$ and $a(\boldsymbol{\mu}_t)$ is invertible almost surely from Proposition 3.2.1.

The vector $\boldsymbol{\lambda}(\boldsymbol{\mu}_t) = (a^{-1} \cdot \mathbf{b})(\boldsymbol{\mu}_t)$ in (3.3.6) is then well defined on Δ_+^n and the market risk premium $\boldsymbol{\theta}(\boldsymbol{\mu}_t) = (\sigma^\top \cdot \boldsymbol{\lambda})(\boldsymbol{\mu}_t)$ satisfies $(\sigma \cdot \boldsymbol{\theta})(\boldsymbol{\mu}_t) = \mathbf{b}(\boldsymbol{\mu}_t)$, $\forall \boldsymbol{\mu}_t \in \Delta_+^n$.

Assume now that one of the conditions (iv)-(vi) of Theorem 3.2.1 is satisfied and $b^i(\boldsymbol{\mu}) = 0$ on $\{\mu^i = 0\}$ for all $i \in \{1, \dots, n\}$ such that $\forall T > 0$, $\mathbb{P}[\exists t \in (0, T] \mid \mu_t^i = 0] > 0$. Let us write $a(\cdot) = OdO^\top(\cdot)$ with $O(\cdot)$ an orthogonal matrix and $d(\cdot) = \text{diag}(d^1(\cdot), \dots, d^{n-1}(\cdot))$ a diagonal matrix with positive (resp. nonnegative) elements on Δ_+^n (resp. Δ^n). Its Moore-Penrose pseudo-inverse can be defined as $a_+(\cdot) = Od_+O^\top(\cdot)$ with $d_+(\cdot)$ the diagonal matrix defined by $d_+^j(\cdot) = 1/d^j(\cdot)$ if $d^j(\cdot) > 0$ and $d_+^j(\cdot) = 0$ otherwise. For $\boldsymbol{\mu} \in \{\mu^i = 0\}$, by changing $O(\cdot)$ we can suppose $d^i(\boldsymbol{\mu}) = 0$ then also $d_+^i(\boldsymbol{\mu}) = 0$ and similarly for j such that $\boldsymbol{\mu} \notin \{\mu^j = 0\}$ we suppose $d^j(\boldsymbol{\mu}) > 0$ and $d_+^j(\boldsymbol{\mu}) = 1/d^j(\boldsymbol{\mu})$. Then with $\boldsymbol{\theta}(\cdot)$ defined as in (3.3.6) we have:

$$(\sigma \cdot \boldsymbol{\theta})(\boldsymbol{\mu}) = (\sigma \sigma^\top a_+ \cdot \mathbf{b})(\boldsymbol{\mu}) = (aa_+ \cdot \mathbf{b})(\boldsymbol{\mu}) = (Odd_+O^\top \cdot \mathbf{b})(\boldsymbol{\mu})$$

A simple calculation shows that $Odd_+O^\top = \text{diag}(\epsilon^1, \dots, \epsilon^{n-1})$, with $\epsilon^i = 0$ when $\mu^i = 0$ and $\epsilon^j = 1$ when $\mu^j > 0$. Moreover since for such indices i we have $b^i(\boldsymbol{\mu}) = B^{i,i}\mu^i = 0$, we conclude that $(Odd_+O^\top \cdot \mathbf{b})(\boldsymbol{\mu}) = \mathbf{b}(\boldsymbol{\mu})$. It then remains to show the relation for the last component, *i.e.* $b^n(\boldsymbol{\mu}) = \sum_{\nu=1}^d \sigma^{n,\nu}(\boldsymbol{\mu})\theta^\nu(\boldsymbol{\mu})$, which can be easily deduced using the fact that $\sum_{i=1}^n \mu^i = 1$, $\sum_{i=1}^n b^i(\boldsymbol{\mu}) = 0$ and $\sum_{i=1}^n \sigma^{i,\nu}(\boldsymbol{\mu}) = 0 \forall \nu \in \{1, \dots, d\}$. Having the market risk premium constructed $\boldsymbol{\theta}(\cdot)$, the market risk deflator is well defined using the exponential martingale (3.3.5), the (NUPBR) condition is satisfied.

To show the non-uniqueness under Assumption 3.2.1, by applying the rank-nullity theorem to the operator $\sigma(\cdot) \in \mathbb{M}^{n-1,d}(\mathbb{R}) : \mathbb{R}^d \rightarrow \mathbb{R}^{n-1}$ with $d > n - 1$, we have $\forall t \geq 0$, $\exists \boldsymbol{\theta}'(\boldsymbol{\mu}_t) \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ such that $(\sigma \cdot \boldsymbol{\theta}')(\boldsymbol{\mu}_t) = \mathbf{0}$ and $(\boldsymbol{\theta} + \boldsymbol{\theta}')(\boldsymbol{\mu}_t)$ is a market risk premium different to $\boldsymbol{\theta}(\boldsymbol{\mu}_t)$. When $d = n - 1$ and under conditions (i)-(iii) of Theorem 3.2.1, $\forall t \geq 0$, the matrix $\sigma(\boldsymbol{\mu}_t)$ is invertible and $\boldsymbol{\theta}(\boldsymbol{\mu}_t)$ can be seen as the unique solution of the linear system $(\sigma \cdot \boldsymbol{\theta})(\boldsymbol{\mu}_t) = \mathbf{b}(\boldsymbol{\mu}_t)$. In addition when $d = n - 1$ and $\exists t > 0$ such that $a(\boldsymbol{\mu}_t)$ is not invertible, in particular under the conditions (iv)-(vi) of Theorem 3.2.1, the rank of $a(\boldsymbol{\mu}_t)$ is less than $n - 1$, the same argument as for $d > n - 1$ can be applied to show the non-uniqueness of the market risk premium. Finally with the relation between $\mathbf{b}(\boldsymbol{\mu}_t)$, $\sigma(\boldsymbol{\mu}_t)$ and $\tilde{\mathbf{b}}(\boldsymbol{\mu}_t, \Sigma_t)$, $\tilde{\sigma}(\boldsymbol{\mu}_t, \Sigma_t)$ in (3.3.1), a simple computation shows that $\tilde{\boldsymbol{\theta}}_t$ defined in (3.3.7) satisfies $\forall t \geq 0$, $\tilde{\mathbf{b}}(\boldsymbol{\mu}_t, \Sigma_t) = (\tilde{\sigma} \cdot \tilde{\boldsymbol{\theta}})(\boldsymbol{\mu}_t, \Sigma_t)$.

In the case where neither the condition (1) nor (2) holds, from Theorem 3.2.1 we know that for some $i \in \{1, \dots, n\}$, $(\mu_t^i)_{t \geq 0}$ vanishes with positive probability and there exists $\boldsymbol{\mu} \in \{\mu^i = 0\}$ such that $b^i(\boldsymbol{\mu}) > 0$. When $\mu_t^i = 0$, $t > 0$, consider the vector of weights $\boldsymbol{\mu}_t \in \{\mu_t^i = 0\}$ such that $b^i(\boldsymbol{\mu}_t) > 0$, also notice that $\sigma^{i,\nu}(\boldsymbol{\mu}_t) = 0 \forall \nu \in \{1, \dots, d\}$. Then there is positive probability that there exists $\tau > t$ such that $\mu_\tau^i > 0$. By buying i -th asset at time t and selling it at time $\tau > t$, an unbounded profit is then realized. The condition (NUPBR) is therefore violated and equivalently, the market risk deflator $(Z_t)_{t \geq 0}$ does not exist, and the market risk premium $\boldsymbol{\theta}(\cdot)$ could not exist either because otherwise we would be able to construct the market risk deflator using exponential martingale. \square

From Proposition 6.1 in [FK09], the necessary condition for the existence of arbitrage has been shown in terms of the total market risk deflator $(\tilde{Z}_t)_{t \geq 0}$ and it is actually a sufficient condition according to [FK10], where explicit strategies for relative arbitrage have been constructed. Here we have the similar result in terms of the market

risk deflator $(Z_t)_{t \geq 0}$ associated to the relative weights of assets, independently of the total capitalization $(\Sigma_t)_{t \geq 0}$ of the market.

Proposition 3.3.2 (A criterion for absence of arbitrage relative to the market portfolio). There is no arbitrage relative to the market in the sense of (3.3.4) over the time horizon $[0, T]$ if and only if the market risk deflator associated to the weights of assets $(Z_t)_{t \geq 0}$ (3.3.5) is a true martingale on $[0, T]$, i.e. $\mathbb{E}[Z_T] = 1$. Or equivalently, there exists a trading strategy that realizes relative arbitrage to the market on $[0, T]$ if and only if the local martingale $(Z_t)_{t \geq 0}$ is a strict local martingale on $[0, T]$ with $\mathbb{E}[Z_T] < 1$.

Proof. Suppose that $(Z_t)_{t \geq 0}$ is a true martingale on $[0, T]$ with $\mathbb{E}[Z_T] = 1$. We can then define the risk-neutral probability, or the equivalent martingale measure (EMM for short) \mathbb{Q}_T and the d -dimensional Brownian motion $(\widehat{W}_t)_{t \in [0, T]}$, $\widehat{W}_t = (\widehat{W}_t^\nu)_{\nu=1}^d$ under \mathbb{Q}_T :

$$\forall A \in \mathcal{F}_T, \mathbb{Q}_T[A] = \mathbb{E}[Z_T \cdot \mathbf{1}_A], \quad \widehat{W}_t^\nu = W_t^\nu + \int_0^t \theta^\nu(\boldsymbol{\mu}_s) ds, \quad t \in [0, T]$$

In the numéraire of the market portfolio, the dynamics of the portfolio value associated to the self-financing trading strategy $(\boldsymbol{\pi}_t)_{t \geq 0} \in \mathcal{T}(\boldsymbol{\mu})$ can be rewritten (3.3.3) as follows:

$$dV_t^\pi = \sum_{i=1}^n \pi_t^i \sum_{\nu=1}^d \sigma^{i, \nu}(\boldsymbol{\mu}_t) d\widehat{W}_t^\nu$$

From the nonnegativity of the portfolio value and Fatou's lemma we obtain that $\mathbb{E}[V_T^\pi] \leq 1$, which is a contradiction since the condition of relative arbitrage $\mathbb{P}[V_T^\pi \geq 1] = 1$, $\mathbb{P}[V_T^\pi > 1] > 0$ implies $\mathbb{E}[V_T^\pi] > 1$. As a consequence, there could not be relative arbitrage with respect to the market portfolio on $[0, T]$.

For the other direction, suppose that $(Z_t)_{t \geq 0}$ is a strict local martingale on $[0, T]$. Then there exists $\tau \in (0, T]$ such that $\mathbb{E}[Z_\tau] < 1$ and in particular $\mathbb{E}[Z_T] \leq \mathbb{E}[Z_\tau] < 1$. We will show in Remark 3.3.5 (also see [FK10]) that $\mathbb{E}[Z_T] < 1$ corresponds in fact to the minimum initial capital $v = V_0$ starting from which we are able to construct an explicit optimal arbitrage strategy using delta hedging. \square

It is worth investigating the absence of arbitrage opportunities relative to the market portfolio explicitly for some simple examples, for instance that of Remark 3.2.3 with $n = 2, d = 1$.

Theorem 3.3.2 (Arbitrage relative to the market for 2 assets under general MWE(ζ)). As in the example of Remark 3.2.3, we suppose that the matrix $\varsigma = [\varsigma^1 \ \varsigma^2]^\top$ satisfies $\varsigma^1 \geq 0 \geq \varsigma^2$, $(\varsigma^1, \varsigma^2) \neq (0, 0)$ and $\alpha, \beta \geq 0$, $(\alpha, \beta) \neq (0, 0)$. Then depending on the value of the exponent $\zeta > 0$ of MWE(ζ), we have the following results concerning the arbitrage relative to the market:

1. When $\zeta \in (0, 1)$, there exist arbitrage opportunities relative to the market, for long enough time horizons.
2. When $\zeta \geq 1$, we have absence of arbitrage relative to the market for any time horizon $[0, T]$, $T > 0$.

Proof. Using the same notations are in Remark 3.2.3, the market risk premium $\theta(\cdot)$ and the market risk deflator $(Z_t)_{t \geq 0}$ associated to market weights are written as follows:

$$\begin{aligned} \theta(x) &= \frac{b(x)}{\sigma(x)} = \frac{\beta - (\alpha + \beta)x}{\varsigma^1 x^\zeta (1-x) - \varsigma^2 x(1-x)^\zeta} \\ Z_t &= \exp\left(-\int_0^t \theta(\mu_s) dW_s - \frac{1}{2} \int_0^t \theta(\mu_s)^2 ds\right), \quad Z_0 = 1 \end{aligned} \quad (3.3.8)$$

From Theorem 3.2.1, we know that the process of market weights $(\mu_t)_{t \geq 0}$ does not exit its state space $\Delta^2 = [0, 1]$, in addition some simple computations show that the following condition is satisfied by the coefficients b, σ, θ of equation (3.3.8):

$$\forall x \in \Delta_+^2 = (0, 1), \quad \sigma(x) \neq 0 \text{ and } \frac{1}{\sigma^2}(x), \frac{b}{\sigma^2}(x), \frac{\theta}{\sigma^2}(x) \in L_{loc}^1(\Delta_+^2)$$

Then from Corollary 2.2 in [MU12], $(Z_t)_{t \geq 0}$ is a true martingale on $[0, \infty)$ if and only if the auxiliary process $(\tilde{\mu}_t)_{t \geq 0}$ does not exit Δ_+^2 , *i.e.* $\mathbb{P}[\tilde{\tau} = \infty] = 1$ or equivalently, $\forall t \geq 0, \mathbb{P}[\tilde{\tau} > t] = 1$, where $\tilde{\tau} := \inf\{t \geq 0, \tilde{\mu}_t \notin \Delta_+^2\}$. The auxiliary process $(\tilde{\mu}_t)_{t \geq 0}$ with initial condition $\tilde{\mu}_0 = \mu_0$ simply has zero drift and same volatility as $(\mu_t)_{t \geq 0}$

$$d\tilde{\mu}_t = (\varsigma^1 \tilde{\mu}_t^\zeta (1 - \tilde{\mu}_t) - \varsigma^2 \tilde{\mu}_t (1 - \tilde{\mu}_t)^\zeta) dW_t$$

We apply Feller's test for explosions from Theorem 5.29 in [KS88]: $(\tilde{\mu}_t)_{t \geq 0}$ does not exit $\Delta_+^2 = (0, 1)$ if and only if $\tilde{v}(0) = \infty$ and $\tilde{v}(1) = \infty$ with $\tilde{v}(\cdot)$ defined as follows, with a constant $c \in \Delta_+^2$:

$$\begin{aligned} \tilde{v}(x) &= \int_c^x \frac{\tilde{s}(x) - \tilde{s}(y)}{\tilde{\rho}(y) \sigma^2(y)} dy, \quad \tilde{s}(x) = \int_c^x \tilde{\rho}(y) dy \\ \tilde{\rho}(y) &= \exp\left(-2 \int_c^y \left(\frac{b}{\sigma^2} - \frac{\theta}{\sigma}\right)(z) dz\right) \end{aligned}$$

A simple calculation shows:

$$\tilde{\rho}(x) = 1, \quad \tilde{s}(x) = x - c, \quad \tilde{v}(x) = \int_c^x \frac{x - y}{(\varsigma^1 y^\zeta (1 - y) - \varsigma^2 y(1 - y)^\zeta)^2} dy \quad (3.3.9)$$

First notice that $x = 0$ and $x = 1$ are completely symmetric, up to interchanging ς^1 and ς^2 . It is therefore sufficient to treat the case $x = 0$.

Assume first that $\zeta \in (0, 1)$ and $\varsigma^1 \neq 0$. Then the integrand of $\tilde{v}(x)$ in (3.3.9) (we take the opposite sign to have nonnegative value) has the following equivalence as $y \rightarrow 0^+$:

$$\frac{y}{(\varsigma^1 y^\zeta (1 - y) - \varsigma^2 y(1 - y)^\zeta)^2} \underset{y \rightarrow 0^+}{\sim} \frac{1}{(\varsigma^1)^2 y^{2\zeta - 1}}$$

The integrand is thus integrable at 0^+ since $2\zeta - 1 < 1$ so $\tilde{v}(0) < \infty$. If $\varsigma^1 = 0$ but $\varsigma^2 \neq 0$ then $\tilde{v}(1) < \infty$, so that in both cases the auxiliary process exits the interval $(0, 1)$ and the market risk deflator is a strict local martingale.

Assume now that $\zeta = 1$. The equivalence above becomes

$$\frac{y}{(\varsigma^1 y^\zeta (1 - y) - \varsigma^2 y(1 - y)^\zeta)^2} \underset{y \rightarrow 0^+}{\sim} \frac{1}{(\varsigma^1 - \varsigma^2)^2 y}$$

The integral is thus not integrable at 0^+ so that $\tilde{v}(0) = \infty$, and by similar arguments, also $\tilde{v}(1) = \infty$. In this case, the auxiliary process does not exit the interval $(0, 1)$ and the market risk deflator is a true martingale.

Assume finally that $\zeta > 1$. If $\zeta^2 \neq 0$ then the equivalence becomes,

$$\frac{y}{(\zeta^1 y^\zeta (1-y) - \zeta^2 y(1-y)\zeta)^2} \underset{y \rightarrow 0^+}{\sim} \frac{1}{(\zeta^2)^2 y}.$$

If $\zeta^2 = 0$ (but $\zeta^1 \neq 0$) then,

$$\frac{y}{(\zeta^1 y^\zeta (1-y) - \zeta^2 y(1-y)\zeta)^2} \underset{y \rightarrow 0^+}{\sim} \frac{y^{1-2\zeta}}{(\zeta^1)^2}.$$

In both cases, the integrand is not integrable at 0^+ so $\tilde{v}(0) = \infty$, and by similar arguments, $\tilde{v}(1) = \infty$, so that the auxiliary process does not exit the interval $(0, 1)$ and the market risk deflator is a true martingale.

Finally we are able to conclude that in the case of two assets, there exist opportunities of relative arbitrage to the market when $\zeta \in (0, 1)$ and there is absence of relative arbitrage to the market when $\zeta \geq 1$ and the market portfolio cannot be outperformed in the sense of (3.3.4). \square

Unfortunately, for the general case $n \geq 3$, the analytical tools used in previous example are not applicable. In the next proposition, we formalize the result mentioned in [Ruf+15] to show the martingale property of the exponential local martingale which is defined as a change of measure between two diffusion processes, with the notions of a *generalized local martingale problem* and *candidate measure* introduced in Appendix 3.5.4.

Lemma 3.3.1 (Martingale defined by a change of measure). *Consider a closed set $\Delta \subset \mathbb{R}^n$ with its interior Δ_+ and two diffusion processes $X = (X_t)_{t \geq 0}$ and $\tilde{X} = (\tilde{X}_t)_{t \geq 0}$ of support Δ given by the unique weak solutions of following SDEs with same volatility $\sigma(t, \cdot)$, same initial condition $X_0 = \tilde{X}_0 = x_0 \in \Delta_+$ and different drift coefficients $b(t, \cdot)$ and $\tilde{b}(t, \cdot)$, with $(W_t)_{t \geq 0}, (\tilde{W}_t)_{t \geq 0}$ two d -dimensional Brownian motions:*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t; \quad d\tilde{X}_t = \tilde{b}(t, \tilde{X}_t)dt + \sigma(t, \tilde{X}_t)dW_t \quad (3.3.10)$$

Suppose that there exists a function θ defined on $\mathbb{R}_+ \times \Delta_+$ such that for all $t \geq 0$, $\theta(t, \cdot)$ is continuous so $|\theta(t, x)| < \infty \forall t \geq 0 \forall x \in \Delta_+$, and θ satisfies $\tilde{b}(t, X_t) = b(t, X_t) - \sigma(t, X_t)\theta(t, X_t)$ for $t \in [0, \tau(X))$ with $\tau(X)$ the stopping time defined as follows:

$$\tau(X) = \inf\{t \geq 0 \mid \int_0^t |\theta(u, X_u)|^2 du = \infty\}$$

We also suppose that $\forall x_0 \in \Delta_+$ and $X_0 = x_0$, we have $\forall t \geq 0$, $X_t^{x_0}, \tilde{X}_t^{x_0} \in \Delta_+$ a.s., that is for any starting point in the interior Δ_+ , both processes X and \tilde{X} will stay in Δ_+ almost surely for all time $t \geq 0$. Under these assumptions, the exponential local martingale $(Z_t)_{t \geq 0}$ defined by $Z_t = \mathcal{E}(-\int_0^t \theta(u, X_u)dW_u)_t$ is a true martingale on $[0, \infty)$.

Proof. With the definitions in Appendix 3.5.4, given the function θ , we write the associated candidate measure \mathbb{Q} and the \mathbb{Q} -Brownian motion $(\widehat{W}_t)_{t \geq 0}$, with $((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}), X)$ solution to the generalized local martingale problem corresponding to (E, x_0, a, \widehat{b}) on $[0, \tau(X))$, with $a(t, X_t) = \sigma \sigma^\top(t, X_t)$, $\widehat{b}(t, X_t) = b(t, X_t) - \sigma(t, X_t)\theta(t, X_t)\mathbf{1}_{\tau(X) > t}$. Then from Proposition 5.4.6 in [KS88], under \mathbb{Q} , $(X_t)_{t \geq 0}$ can be written as the solution to the following SDE on $[0, \tau(X))$ with $X_0 = x_0$:

$$dX_t = \widehat{b}(t, X_t)dt + \sigma(t, X_t)d\widehat{W}_t = (b(t, X_t) - \sigma(t, X_t)\theta(t, X_t))dt + \sigma(t, X_t)d\widehat{W}_t$$

Meanwhile, by comparing the equation above with (3.3.10), the process $(X_t)_{t \geq 0}$ on $[0, \tau(X))$ under the measure \mathbb{Q} has the same distribution as $(\widetilde{X}_t)_{t \geq 0}$ on $[0, \tau(\widetilde{X}))$ under the original measure \mathbb{P} . Under \mathbb{P} , with the assumptions $\forall x_0 \in \Delta_+$, $\widetilde{X}_t^{x_0} \in \Delta_+ \forall t \geq 0$ and $\forall t \geq 0$, $\theta(t, \cdot)$ is continuous on Δ^n and $|\theta(t, x)| < \infty \forall x \in \Delta_+$, we have $\int_0^t |\theta(u, \widetilde{X}_u)|^2 du < \infty \forall t \geq 0$ and:

$$\mathbb{P}[\tau(\widetilde{X}) = \infty] = 1; \forall t \geq 0, \widetilde{b}(t, \widetilde{X}_t) = b(t, \widetilde{X}_t) - \sigma(t, \widetilde{X}_t)\theta(t, \widetilde{X}_t) \mathbb{P}\text{-a.s.}$$

Then $\mathbb{Q}[\tau(X) = \infty] = 1$ and we can finally conclude with Theorem 1 in [Ruf+15], which states that the exponential local martingale $(Z_t)_{t \geq 0}$ is a true \mathbb{P} -martingale if and only if $\mathbb{Q}[\tau(X) = \infty] = 1$, or equivalently $\mathbb{Q}[\int_0^t |\theta(u, X_u)|^2 du < \infty] = 1 \forall t \geq 0$. \square

Corollary 3.3.1 (Absence of arbitrage for arbitrary number of assets under MWE(ζ), $\zeta \geq 1$). *For the MWE(ζ) with exponent $\zeta \geq 1$, with any number of assets $n \geq 2$, there is no opportunity of arbitrage relative to the market in the sense of (3.3.4) over any finite time horizon $[0, T]$.*

Proof. To apply Lemma 3.3.1, the process $(X_t)_{t \geq 0}$ should be replaced by the market weights process $(\mu_t)_{t \geq 0}$ given by (3.2.9), and the open and closed sets Δ_+ and Δ by Δ_+^n and Δ^n . Naturally, the function $\theta(t, \cdot)$ is the market risk premium associated to weights given for example by (3.3.6):

$$\theta(t, \mu_t) = \theta(\mu_t) = (\sigma^\top a^{-1} \cdot \mathbf{b})(\mu_t)$$

First, let us show that the market risk premium $\theta_t = \theta(t, \cdot)$ is continuous on Δ_+^n , thus satisfies the assumption in Lemma 3.3.1 and in particular $|\theta(t, \mu)| < \infty \forall \mu \in \Delta_+^n$. In fact, for $\mu \in \Delta_+^n$, the reduced covariance matrix a is positive definite, so its inverse a^{-1} is well-defined and continuous on Δ_+^n , with \mathbf{b} and σ being continuous, then $\theta(t, \cdot)$ is simply a product of continuous functions on the interior of unit simplex Δ_+^n which is also continuous.

The auxiliary process $(\widetilde{\mu}_t)_{t \geq 0}$ replacing $(\widetilde{X}_t)_{t \geq 0}$ is the martingale part of $(\mu_t)_{t \geq 0}$:

$$d\widetilde{\mu}_t = \sigma(\mu_t) \cdot d\mathbf{W}_t; \widetilde{\mu}_0 = \mu_0 \in \Delta_+^n$$

Feller's explosion test is not available in this multidimensional case but it is still possible to apply McKean's argument in step (3) of the proof of Theorem 3.2.1 to $(\widetilde{\mu}_t)_{t \geq 0}$, and obtain that $\forall t \geq 0$, $\widetilde{\mu}_t \in \Delta_+^n$ almost surely. By applying Lemma 3.3.1 above, the market risk deflator $(Z_t)_{t \geq 0}$ defined in (3.3.5) is a true martingale on $[0, \infty)$. Finally we conclude using Proposition 3.3.1, there is no arbitrage over the time horizon $[0, T]$, $\forall T > 0$ relative to the market. \square

Remark 3.3.4 (Diversity of the market). The market is said to be *diverse* over the horizon $[0, T]$ if no stock dominates the market at any time and *weakly diverse* if on average no stock dominates the market according to [FK09]. More precisely, the market is diverse or weakly diverse or asymptotically weakly diverse if there exists a constant $\delta \in (0, 1)$ such that almost surely:

$$\begin{aligned} \forall t \in [0, T], \mu_t^{\max} &= \max_{1 \leq i \leq n} \mu_t^i < 1 - \delta \text{ (diverse)} \\ \frac{1}{T} \int_0^T \mu_t^{\max} dt &< 1 - \delta \text{ (weakly diverse)} \\ \overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mu_t^{\max} dt &< 1 - \delta \text{ (asymptotically weakly diverse)} \end{aligned}$$

From Paragraph (7.3) in [FK09], weak diversity implies the existence of relative arbitrage over $[0, T]$ for large enough $T > 0$. From Corollary 3.3.1, we see that for $\zeta \geq 1$ and for any time horizon $[0, T]$, $T > 0$, there is no arbitrage opportunity relative to the market, thus the market is not weakly diverse and a fortiori it is not diverse.

Remark 3.3.5 (Super-hedging and optimal arbitrage). The definition of relative arbitrage (3.3.4) can be seen as the problem of finding a portfolio with trading strategy $\pi \in \mathcal{T}(\mu)$ that realizes an almost sure super-hedging relative to the portfolio with strategy $\rho \in \mathcal{T}(\mu)$ over the time horizon $[0, T]$, that is, the value V^ρ of the portfolio ρ is considered as the payoff to be (super-)replicated by another portfolio π at time T .

Recall that when the portfolio ρ corresponds to the market portfolio we have $V_T^\rho = 1$ almost surely. When relative arbitrage to the market exists for a time horizon $[0, T]$, let us consider the minimal initial capital U_T with which such a relative arbitrage can be found:

$$U_T = \inf\{v \geq 0 \mid \exists \pi \in \mathcal{T}(\mu), V_T^{v, \pi} \geq 1 \text{ a.s.}\}$$

U_T is also called the super-hedging price of payoff 1 for time horizon T . From [Ruf13] and [FK10] we know that U_T in fact corresponds to the expectation of market risk deflator at time T , *i.e.* $U_T = \mathbb{E}[Z_T]$. Furthermore, the optimal trading strategy associated to U_T is constructed by delta hedging. First, define the function u as follows:

$$u(t, \mu_t) = \frac{\mathbb{E}[Z_T | \mathcal{F}_t]}{Z_t}$$

Suppose that u is sufficiently regular, for instance \mathcal{C}^1 on its space variable μ_t , then the optimal trading strategy π is written as:

$$\tilde{\pi}_t^i = \mu_t^i \cdot (1 + \partial_i \log u(t, \mu_t) - \sum_{j=1}^n \mu_t^j \cdot \partial_j \log u(t, \mu_t))$$

where $\tilde{\pi}_t^i$ represents the percentage of value invested in the i -th asset at time t and ∂_i corresponds to the derivative with respect to the i -th weight. Equivalently, the corresponding number of units of asset π_t^i can be expressed as:

$$\pi_t^i = u(t, \mu_t) + \partial_i u(t, \mu_t) + \sum_{j=1}^n \mu_t^j \cdot \partial_j u(t, \mu_t)$$

3.4 Future works and conclusion

In this section we will discuss about some potential applications of market weights model that could be topics for future research.

3.4.1 Functionally generated portfolio

An important application of market weights models is the possibility to define functionally generated portfolios as introduced in [Fer99]. It is another advantage of modeling the market directly by the weights of assets, without the dynamic of total market capitalization involved. Given a function $G \in \mathcal{C}^2(\Delta^n, \mathbb{R}_+^*)$, a new portfolio of weights $(\pi_t)_{t \geq 0}$ is generated from the market weights $(\mu_t)_{t \geq 0}$ through the mapping $\Phi_G : \Delta^n \rightarrow \Delta^n$ defined as follows:

$$\begin{aligned} \pi_t &= \Phi_G(\mu_t), \quad \pi_t = (\pi_t^i)_{i=1}^n, \quad \mu_t = (\mu_t^i)_{i=1}^n \\ \pi_t^i &= \mu_t^i \cdot (1 + \partial_i \log G(\mu_t) - \sum_{j=1}^n \mu_t^j \cdot \partial_j \log G(\mu_t)) \end{aligned}$$

The function G is called portfolio generating function and $\pi = (\pi_t)_{t \geq 0}$ the associated self-financing trading strategy with π_t^i representing the percentage invested in the i -th asset at time t .

By using again the notation $(m_t^\mu(\cdot))_{t \geq 0}$ for the probability distribution of market weights $(\mu_t)_{t \geq 0}$ given by (3.2.9) under the conditions stated in Proposition 3.2.2, the distribution $(m_t^\pi(\cdot))_{t \geq 0}$ of the portfolio $(\pi_t)_{t \geq 0}$ generated from $(\mu_t)_{t \geq 0}$ by function G can be seen as the pushforward measure of distribution of $(\mu_t)_{t \geq 0}$ by the function Φ_G , that is:

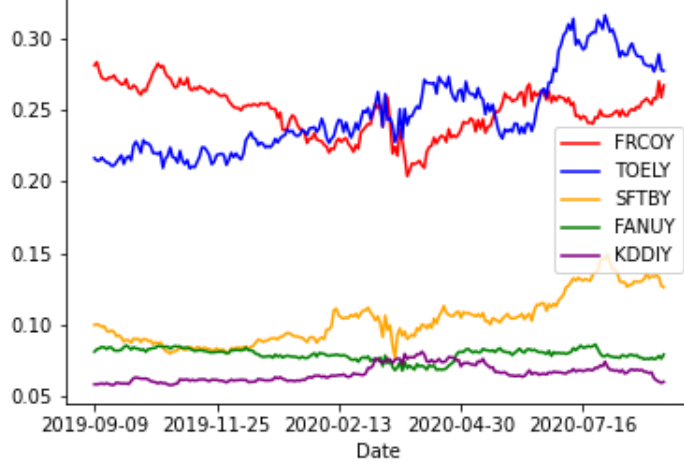
$$\forall t \geq 0, \quad \pi_t = \Phi_G(\mu_t) \text{ and } m_t^\pi(\cdot) = \Phi_G \# m_t^\mu(\cdot)$$

Moreover under the some conditions, the market portfolio has unique stationary distribution $m_\infty^\mu(\cdot)$, the generated portfolio $(\pi_t)_{t \geq 0}$ also has unique stationary distribution $m_\infty^\pi(\cdot)$ defined as the pushforward measure of $(\mu_t)_{t \geq 0}$ by the function Φ_G , denoted as $m_\infty^\pi(\cdot) = \Phi_G \# m_\infty^\mu(\cdot)$.

The functionally generated portfolios are widely used in constructing portfolios that realize arbitrage relative to the market, see [KR17] and [RX19] where specific trading strategies are constructed using Lyapunov functions. Also notice that the strategy of optimal arbitrage that minimizes the initial capital in Remark 3.3.5 is simply obtained with $G(\mu_t) = u(t, \mu_t) = \mathbb{E}[Z_T | \mathcal{F}_t] / Z_t$.

3.4.2 Estimation of model parameters with application to stock market indices

Since in practice, the market weights of assets in an index never fall to 0 (except when an asset quits the index), it is reasonable and actually simpler, for the purpose of parameters estimation, to suppose the exponent $\zeta = 1$. Using the notations for logarithm of market weights $z_t = (z_t^i)_{i=1}^n$, $z_t^i = \log(\mu_t^i)$, we define the process $\tilde{z}_t =$



Evolution of weights of 5 stocks in Nikkei 225 index

$(\tilde{z}_t^i)_{i=1}^n$ and we have following expressions of quadratic variations of z_t and \tilde{z}_t :

$$\begin{aligned} \frac{d\langle z^i, z^j \rangle_t}{dt} &= \sum_{\nu=1}^d (\zeta^{i,\nu} - \sum_{k=1}^n \mu_t^k \zeta^{k,\nu}) (\zeta^{j,\nu} - \sum_{k=1}^n \mu_t^k \zeta^{k,\nu}) \\ \frac{d\langle \tilde{z}^i, \tilde{z}^j \rangle_t}{dt} &= \sum_{\nu=1}^d \zeta^{i,\nu} \zeta^{j,\nu} = (\zeta \zeta^\top)^{i,j}; \quad \tilde{z}_t^i = z_t^i - \frac{1}{n} \sum_{k=1}^n z_t^k \end{aligned}$$

We remark that for $\zeta = 1$, thanks to the form of dependence $\sigma^{i,\nu}(\boldsymbol{\mu}_t) = \mu_t^i (\zeta^{i,\nu} - \sum_{k=1}^n \mu_t^k \zeta^{k,\nu})$ of volatility matrix σ on $\zeta = [\zeta^{i,\nu}] \in \mathbb{M}^{n,d}(\mathbb{R})$ and the fact that $\forall t \geq 0$, $\sum_{k=1}^n \mu_t^k = 1$ almost surely, subtracting a constant from all the coefficients of a column of the matrix, for instance C_ν for the ν -th column, will not change the equation. Thus we can suppose that the sum of each column of ζ is equal to 0, *i.e.* $\forall \nu \in \{1, \dots, d\}$, $\sum_{i=1}^n \zeta^{i,\nu} = 0$. Then it suffices to estimate the $(n-1)$ first lines of ζ using the notation $\zeta \in \mathbb{M}^{n-1,d}(\mathbb{R})$.

Over the time horizon $[0, T]$, the uniform grid \mathcal{T} corresponds to the $M_T + 1$ observation times of market weights. In practice, we have daily observations of market weights, Δt is equal to 1 day and $M_T + 1$ is the number of days of the observation period.

$$\mathcal{T} = \{t_m = m \cdot \Delta t, \Delta t = T/M_T; m = 0, \dots, M_T\}$$

From the expression of quadratic variation of \tilde{z}_t , the estimator of the symmetric matrix $\zeta \zeta^\top = [\zeta \zeta^\top]^{i,j} \in \mathbb{M}^{n-1}(\mathbb{R})$ can be written as follows:

$$(\zeta \zeta^\top)^{i,j} \approx \frac{1}{T} \sum_{m=1}^{M_T} (\tilde{z}_{t_m}^i - \tilde{z}_{t_{m-1}}^i) (\tilde{z}_{t_m}^j - \tilde{z}_{t_{m-1}}^j), \quad i \leq j$$

We fix the dimension of Brownian motion $d = n - 1$ so $\zeta \in \mathbb{M}^{n-1}(\mathbb{R})$ is a square matrix. If we suppose furthermore that ζ is symmetric we have $\zeta \zeta^\top = \zeta \cdot \zeta$, then ζ is the

square root of the symmetric positive semi-definite matrix $\varsigma\varsigma^\top$ and can be computed numerically.

For each $i \in \{1, \dots, n-1\}$ we write the Euler scheme for market weights equation (3.2.9):

$$\begin{aligned} \mu_{t_m}^i &\approx \mu_{t_{m-1}}^i + \left[\sum_{j=1}^{n-1} (B^{i,j} - B^{i,n}) \mu_{t_{m-1}}^j + B^{i,n} \right] \Delta t + \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu}_{t_{m-1}}) \Delta W_m^\nu \\ \Delta \mathbf{W}_m &= (\Delta W_m^\nu)_{\nu=1}^d, \quad \Delta \mathbf{W}_m = \mathbf{W}_{t_m} - \mathbf{W}_{t_{m-1}}, \quad m = 1, \dots, M_T \end{aligned}$$

Then for each $i \in \{1, \dots, n-1\}$, we write the analogue of a multiple linear regression:

$$\begin{aligned} y_m^{(i)} &= b^{(i),0} + \sum_{j=1}^{n-1} b^{(i),j} \cdot x_m^j + \epsilon_m^{(i)}, \quad m = 1, \dots, M_T \\ y_m^{(i)} &= \mu_{t_m}^i; \quad x_m^j = \mu_{t_{m-1}}^j, \quad j = 1, \dots, n-1 \end{aligned}$$

where $(y_m^{(i)})_{m=1}^{M_T}$ represents the dependent variable, $(x_m^j)_{m=1}^{M_T}$, $j = 1, \dots, n-1$ are explanatory variables and $(\epsilon_m^{(i)})_{m=1}^{M_T}$ the error term.

At last, we obtain an estimate of coefficients of the matrix $B \in \mathbb{M}^n(\mathbb{R})$. For each $i \in \{1, \dots, n-1\}$, the coefficients $\{B^{i,j}, j = 1, \dots, n\}$ are approximated from regression coefficients $(b^{(i),j})_{j=1}^{n-1}$ through the following relation:

$$\begin{aligned} b^{(i),j} &\approx (B^{i,j} - B^{i,n}) \Delta t \text{ for } j \in \{1, \dots, n-1\}, \quad j \neq i \\ b^{(i),0} &\approx B^{i,n} \Delta t, \quad b^{(i),i} \approx 1 + (B^{i,i} - B^{i,n}) \Delta t \end{aligned}$$

Stock market indices are portfolios where the assets are single stocks in the market, the weights of stocks are either given by relative market capitalizations of companies, called capitalization-weighted indices, such as CAC 40, DAX, HSI, FTSE-100 etc, or otherwise given by relative prices of companies, called price-weighted indices, with the most famous examples being Dow Jones Industrial Average and Nikkei 225.

3.4.3 Long-term portfolio optimization under market weights equation

Let us consider the market weights process $\boldsymbol{\mu} = (\boldsymbol{\mu}_t)_{t \geq 0}$ given by $MWE(\zeta)$ (3.2.9), a self-financing trading strategy $\boldsymbol{\pi} \in \mathcal{T}(\boldsymbol{\mu})$ and its wealth process $(V_t^\pi)_{t \geq 0}$, normalizing the initial value to $V_0^\pi = v = 1$. Let us define the risk-sensitive expected growth rate up to time horizon T , with $\theta > -2$, $\theta \neq 0$ a parameter of risk-sensitivity, also see [Nag02]:

$$J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T) = -\frac{2}{\theta} \log \mathbb{E}[e^{-\frac{\theta}{2} \log V_T^\pi}]$$

Define the value function $u(t, \boldsymbol{\mu}_0)$ for $t \in [0, T-t]$:

$$u(t, \boldsymbol{\mu}_0) = \sup_{\boldsymbol{\pi} \in \mathcal{T}(\boldsymbol{\mu}_0)} J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T-t)$$

The associated Bellman equation is obtained by applying the dynamic programming principle:

$$\frac{\partial u}{\partial t} + \sup_{\boldsymbol{\pi} \in \mathbb{R}^n} L^\pi = 0, \quad u(T, \boldsymbol{\mu}_0) = 0,$$

where L^π is a differential operator. This PDE has a unique solution under appropriate assumptions. Nevertheless, the most standard assumption of uniform ellipticity is not satisfied since for our market weights process only local ellipticity holds.

We can also define the long-term risk-sensitive growth rate as the limit superior of risk-sensitive growth rate per unit of time when the time horizon goes to infinity:

$$J(\boldsymbol{\mu}_0, \boldsymbol{\pi}) = \limsup_{T \rightarrow \infty} \frac{1}{T} J(\boldsymbol{\mu}_0, \boldsymbol{\pi}; T) = \limsup_{T \rightarrow \infty} -\frac{2}{\theta T} \log \mathbb{E}[e^{-\frac{\theta}{2} \log V_T^\pi}]$$

It is possible to write the corresponding Bellman equation of ergodic type, but again it is difficult to establish existence and uniqueness of solution in the absence of uniform ellipticity and this is an interesting question for future research.

3.4.4 Conclusion and perspectives

The model of market weights studied in the present chapter has nice properties from a probabilistic point of view: the unique solution to market weights equation takes values in the unit simplex and attains the boundary under well-stated conditions. The stationary distribution, which ensures the long-term behavior of market weights exists and is unique under adequate conditions. The market weights in classical volatility-stabilized market model can be recognized as a special case and our model has in some sense, richer covariance structure compared to polynomial model, with the exponent ζ being a measure of effect of volatility stabilization. Moreover, properties such as completeness of the market and existence of relative arbitrage to the market are studied for our model. However, more precise criterion for the existence of relative arbitrage in the general case still deserves further investigation.

The market weights model can be calibrated to real market data. New portfolios can be constructed from market weights using portfolio generating functions and the optimal arbitrage with minimum initial capital can be achieved. When the market weights process is ergodic, under conditions specified in the chapter, the optimization of long-term portfolio wealth becomes an ergodic control problem under degenerate process and could become an interesting problem for future research. Another possible extension to our model is the limit when the number of assets n in the market portfolio goes to infinity, which is reasonable when n is large, and the market weights equation becomes infinite dimensional and measure-valued. More links with rank-based model could also be considered, since we have interactions of mean-field type as the limit of a particle system where assets interact through their ranks.

3.5 Appendix

3.5.1 Attainment of boundary for Theorem 3.2.1

Under the conditions (iv)-(vi) in Theorem 3.2.1, we have showed that there exists $\boldsymbol{\mu}_* \in \{\mu^i = 0\}$ such that $0 \leq 2b^i(\boldsymbol{\mu}_*) < (\mu_*^i)^{-1} a^{i,i}(\boldsymbol{\mu}_*)$. Then there exist a constant $\delta > 0$ and an open ball in \mathbb{R}^n of radius $\rho > 0$ centered at $\boldsymbol{\mu}_*$, such that for $\boldsymbol{\mu} \in \Delta_+^n \cap U$ we have:

$$2b^i(\boldsymbol{\mu}) \leq (1 - \delta)(\mu^i)^{-1} a^{i,i}(\boldsymbol{\mu}) \quad \text{and} \quad (\mu^i)^{-1} a^{i,i}(\boldsymbol{\mu}) > 0$$

For $T > 0$ and $t < \tau(U) = \inf\{s \geq 0 \mid \boldsymbol{\mu}_s \notin U\} \wedge T$, using the dynamics of the i -th weight, with $\boldsymbol{\sigma}^i(\cdot)$ representing the i -th line of the volatility matrix we write:

$$\mu_t^i - \mu_0^i - \int_0^t b^i(\boldsymbol{\mu}_s) ds = \boldsymbol{\sigma}^i(\boldsymbol{\mu}_s) \cdot d\mathbf{W}_s = 2 \int_0^t \sqrt{\mu_s^i} \frac{1}{2} \sqrt{(\mu_s^i)^{-1} a^{i,i}(\boldsymbol{\mu}_s)} d\mathbf{B}_s$$

The process defined as $A_t = \frac{1}{4} \int_0^t (\mu_s^i)^{-1} a^{i,i}(\boldsymbol{\mu}_s) ds$ is strictly increasing on $[0, \tau(U))$ since $(\mu_s^i)^{-1} a^{i,i}(\boldsymbol{\mu}_s) > 0$, the time change $\lambda_u = \inf\{t \geq 0 \mid A_t > 0\}$ is continuous and strictly increasing on $[0, A_{\tau(U)})$. Define the time-changed process $Y_u = \mu_{\lambda_u}^i$, for $u < A_{\tau(U)}$:

$$Y_u = \mu_0^i + \int_0^u \frac{4b^i(\boldsymbol{\mu}_{\lambda_v})}{(\mu_{\lambda_v}^i)^{-1} a^{i,i}(\boldsymbol{\mu}_{\lambda_v})} dv + 2 \int_0^u \sqrt{Y_v} d\boldsymbol{\beta}_v$$

where $(\mathbf{B}_s)_{s \geq 0}$ and $(\boldsymbol{\beta}_v)_{v \geq 0}$ are d -dimensional Brownian motions.

Consider the $(2 - 2\delta)$ -dimensional squared Bessel process or BESQ($2 - 2\delta$) given by the unique strong solution of the equation:

$$Z_u = \mu_0^i + (2 - 2\delta)u + 2 \int_0^u \sqrt{Z_v} d\boldsymbol{\beta}_v$$

In fact it is well-known that BESQ(α) hits zero almost surely of and only if $\alpha < 2$, so it is the case for the process Z . Moreover, since

$$\frac{4b^i(\boldsymbol{\mu}_t)}{(\mu_t^i)^{-1} a^{i,i}(\boldsymbol{\mu}_t)} \leq 2 - 2\delta \text{ for } t < \tau(U)$$

By applying a standard comparison theorem we have $Y_u \leq Z_u$ for $u < A_{\tau(U)}$. Finally let us show that there exists $\epsilon > 0$ such that if $\boldsymbol{\mu}_0 \in \Delta_+^n$ and $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \epsilon$, then there is positive probability that $Z_u = 0$ for some $u < A_{\tau(U)}$. With $c = \inf_{\boldsymbol{\mu} \in U} (\mu^i)^{-1} a^{i,i}(\boldsymbol{\mu})/4 > 0$ so that $A_{\tau(U)} \geq c\tau(U)$ and $\eta > 0$ to be determined later, we write:

$$\begin{aligned} & \mathbb{P}[\eta < A_{\tau(U)} \text{ and } \inf_{u \leq \eta} Z_u = 0] \geq \mathbb{P}[\eta < A_{\tau(U)}] - \mathbb{P}[\inf_{u \leq \eta} Z_u > 0] \\ & \geq \mathbb{P}[\eta c^{-1} < \tau(U)] - \mathbb{P}[\inf_{u \leq \eta} Z_u > 0] \geq \mathbb{P}[\sup_{s \leq \eta c^{-1}} |\boldsymbol{\mu}_s - \boldsymbol{\mu}_0| < \rho/2] - \mathbb{P}[\inf_{u \leq \eta} Z_u > 0] \end{aligned} \quad (3.5.1)$$

The last inequality holds when $|\boldsymbol{\mu}_0 - \boldsymbol{\mu}_*| < \rho/2$, recall that U is the open ball centered at $\boldsymbol{\mu}_*$ of radius ρ . To show that for appropriate $\eta > 0$ and $\epsilon > 0$, the last quantity is positive, we use the Lemmas F.1 and F.2 in the proof of Theorem 5.7 in [FL16]. By applying Lemma F.1 to $(\boldsymbol{\mu}_t)_{t \geq 0}$, we have for any $r > 0$, there exist constants $c_1, c_2 > 0$ depending on r but not on $\boldsymbol{\mu}_0$, such that for $t \leq c_2$:

$$\mathbb{P}[\sup_{s \leq t} |\boldsymbol{\mu}_s - \boldsymbol{\mu}_0| < r] \geq 1 - tc_1(1 + |\boldsymbol{\mu}_0|^2)$$

It is then possible to choose $\eta > 0$ independently of $\boldsymbol{\mu}_0$ such that $\mathbb{P}[\sup_{s \leq \eta c^{-1}} |\boldsymbol{\mu}_s - \boldsymbol{\mu}_0| < \rho/2] \geq 1/2$.

Lemma F.2 states that for a BESQ(α) process Z with $\alpha \in (0, 2)$ with the time it first hits zero $\tau_0 = \inf\{t \geq 0 \mid Z_t = 0\}$, for any $\kappa > 0$ and $z \geq 0$ we have:

$$\lim_{z \rightarrow 0} \mathbb{P}[\tau_0 > \kappa \mid Z_0 = z] = \lim_{z \rightarrow 0^+} \mathbb{P}[\inf_{u \leq \kappa} Z_u > 0 \mid Z_0 = z] = 0$$

Then there exists $\rho', \rho'' > 0$ such that if $|\mu_0 - \mu_*| < \rho'$, we have $Z_0 = |\mu_0^i - \mu_*^i| < \rho''$ and $\mathbb{P}[\inf_{u \leq \eta} Z_u > 0] < 1/3$. With $\epsilon = \rho' \wedge \rho/2$ and the choice of $\eta > 0$ above, the last term in (3.5.1) is strictly positive, there is positive probability that Z hits zero before μ_{λ_u} leaves U . Finally we conclude that for any $T > 0$, there exists $\epsilon > 0$ such that if $|\mu_0 - \mu_*| < \epsilon$ almost surely, then $\mu_t^i = 0$ for some $t \leq T$ with positive probability.

3.5.2 Hörmander's condition and theorem for Proposition 3.2.2

We write the n -dimensional general form SDE with initial condition $X_0 = x_0$ in Stratonovich form:

$$dX_t = \tilde{b}(X_t)dt + \sigma(X_t) \circ dW_t, \quad \tilde{b}(\cdot) = (\tilde{b}^i(\cdot))_{i=1}^n$$

$$\tilde{b}^i(\cdot) = b^i(\cdot) - \frac{1}{2} \sum_{\nu=1}^d \sum_{j=1}^n \sigma^{j,\nu} \cdot \partial_j \sigma^{i,\nu}(\cdot)$$

Given two continuously differentiable functions $f, g \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$, the associated bracket (or Lie algebra) $[f, g]$ is defined by the following:

$$[f, g] = ([f, g]^i)_{i=1}^n = f \cdot \nabla g - g \cdot \nabla f, \quad [f, g]^i = \sum_{j=1}^n f^j \cdot \partial_j g^i - \partial_j f^i \cdot g^j$$

With $\sigma^\nu = [\sigma^{1,\nu}, \dots, \sigma^{n,\nu}]^\top$ representing the ν -th column of volatility matrix σ , the sequence of functions $\mathbb{R}^n \rightarrow \mathbb{R}^n$, $(L_k)_{k \in \mathbb{N}}$ is defined recursively and L_∞ as the union of all $(L_k)_{k \in \mathbb{N}}$:

$$L_0 = \{\sigma^1, \dots, \sigma^d\}, \quad L_{k+1} = \{[\tilde{b}, \phi], [\sigma^1, \phi], \dots, [\sigma^d, \phi] \mid \phi \in L_k\}, \quad L_\infty = \bigcup_{k=0}^{\infty} L_k$$

The k -th hypothesis at the fixed point $x_0 \in \mathbb{R}^n$ is denoted $H_k^{x_0}$ and we say that Hörmander's hypothesis H^{x_0} in x_0 is satisfied if there exist some $k_0 \in \mathbb{N}$ such that $H_{k_0}^{x_0}$ is true:

$$(H_k^{x_0}) : \text{Span}\{\phi(x_0); \phi \in \bigcup_{j=1}^k L_j\} = \mathbb{R}^n, \quad (H^{x_0}) : \text{Span}\{\phi(x_0); \phi \in L_\infty\} = \mathbb{R}^n \quad (3.5.2)$$

Finally from Theorem 2.3.2 in [Nua06], Hörmander's theorem states: if the drift and volatility are infinitely differential with bounded derivatives of all orders, *i.e.* $b^i, \sigma^{i,\nu} \in \mathcal{C}_b^\infty(\mathbb{R}^n) \forall i \in \{1, \dots, n\} \forall \nu \in \{1, \dots, d\}$ and Hörmander's hypothesis H^{x_0} at point x_0 is true, then for every $t > 0$, the law of X_t^{0,x_0} starting from x_0 at time 0 is absolutely continuous with respect to the Lebesgue measure, so the probability density function $p_t(\cdot|x_0)$ on \mathbb{R}^n exists. Furthermore, here b and σ do not depend on time t , Theorem 2.3.3 in [Nua06] ensures that the density function is infinitely differentiable with respect to the space variable, *i.e.* $x \mapsto p_t(x|x_0) \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}_+)$.

3.5.3 Choice of Lyapunov function V for Theorem 3.2.2

Let us show that in Theorem 3.2.2, $V : \Delta^n \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ chosen as follows is a Lyapunov function, *i.e.* the condition $\mathcal{L}V \leq \kappa - \varphi \circ V$ is satisfied on Δ^n , when

$\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is strictly concave function with $\varphi(0) = 0$ with $\varphi(x) \uparrow \infty$ as $x \uparrow \infty$:

$$V(\boldsymbol{\mu}) = -\sum_{i=1}^{n-1} \log(\mu^i) - \log\left(1 - \sum_{j=1}^{n-1} \mu^j\right) = \sum_{i=1}^n \log(\mu^i).$$

The partial derivatives $\partial_i V$, $\partial_{i,i}^2 V$ and $\partial_{i,j}^2 V$ with $i, j = 1, \dots, n-1$, $i \neq j$ are given by:

$$\partial_i V(\boldsymbol{\mu}) = (\mu^n)^{-1} - (\mu^i)^{-1}, \quad \partial_{i,i}^2 V(\boldsymbol{\mu}) = (\mu^n)^{-2} + (\mu^i)^{-2}, \quad \partial_{i,j}^2 V(\boldsymbol{\mu}) = (\mu^n)^{-2}$$

Let us calculate the quantity $(\mathcal{L}V + \varphi \circ V)(\boldsymbol{\mu})$ for $\boldsymbol{\mu} = (\mu^i)_{i=1}^{n-1} \in \Delta^n$:

$$\begin{aligned} (\varphi \circ V + \mathcal{L}V)(\boldsymbol{\mu}) &= (\varphi \circ V)(\boldsymbol{\mu}) + \sum_{i=1}^{n-1} b^i(\boldsymbol{\mu}) \partial_i V(\boldsymbol{\mu}) + \frac{1}{2} \sum_{i,j=1}^{n-1} a^{i,j}(\boldsymbol{\mu}) \partial_{i,j}^2 V(\boldsymbol{\mu}) \\ &= \varphi\left(-\sum_{i=1}^n \log(\mu^i)\right) + \sum_{i=1}^{n-1} b^i(\boldsymbol{\mu}) \left((\mu^n)^{-1} - (\mu^i)^{-1}\right) + \frac{1}{2} \sum_{i=1}^{n-1} (\mu^i)^{-2} a^{i,i}(\boldsymbol{\mu}) + \frac{1}{2} (\mu^n)^{-2} \sum_{i,j=1}^{n-1} a^{i,j}(\boldsymbol{\mu}) \\ &= \varphi\left(-\sum_{i=1}^n \log(\mu^i)\right) + \sum_{i=1}^n \left(\frac{1}{2} (\mu^i)^{-2} a^{i,i}(\boldsymbol{\mu}) - (\mu^i)^{-1} b^i(\boldsymbol{\mu})\right) = \varphi\left(-\sum_{i=1}^n \log(\mu^i)\right) - \sum_{i=1}^n \gamma^i(\boldsymbol{\mu}) \\ &= \varphi\left(-\sum_{i=1}^n \log(\mu^i)\right) + \sum_{i=1}^n \left(\frac{1}{2} \sum_{\nu=1}^d ((\mu^i)^{\zeta-1} \zeta^{i,\nu} - \sum_{k=1}^n (\mu^k)^{\zeta} \zeta^{k,\nu})^2 - (\mu^i)^{-1} b^i(\boldsymbol{\mu})\right) \end{aligned}$$

For the last equality, the following identities are used:

$$\sum_{i=1}^{n-1} b^i(\boldsymbol{\mu}) = -b^n(\boldsymbol{\mu}) \quad \text{and} \quad \sum_{i,j=1}^{n-1} a^{i,j}(\boldsymbol{\mu}) = \left(\sum_{i=1}^{n-1} \sum_{\nu=1}^d \sigma^{i,\nu}(\boldsymbol{\mu})\right)^2 = \left(-\sum_{\nu=1}^d \sigma^{n,\nu}(\boldsymbol{\mu})\right)^2 = a^{n,n}(\boldsymbol{\mu})$$

We choose the strictly concave function to be differentiable (or simply take $\varphi(x) = x^\beta$, $\beta \in (0, 1)$ as in the proof of Theorem 3.2.2), with the fact that $-\sum_{i=1}^n \log(\mu^i) \geq n \log(n) > 0$ for $\boldsymbol{\mu} \in \Delta^n$, we can find $\alpha_\varphi, \beta_\varphi \in \mathbb{R}^2$ such that $\varphi(x) \leq \alpha_\varphi - \beta_\varphi x$ for $x \geq n \log(n)$. Recall that $\zeta \in \mathbb{N}^*$, let us show that the function $(\varphi \circ V + \mathcal{L}V)(\boldsymbol{\mu})$ is bounded from above for $\boldsymbol{\mu} \in \Delta^n$:

$$\begin{aligned} (\varphi \circ V + \mathcal{L}V)(\boldsymbol{\mu}) &\leq \varphi\left(-\sum_{i=1}^n \log(\mu^i)\right) - \sum_{i=1}^n (\mu^i)^{-1} b^i(\boldsymbol{\mu}) + \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d (|\zeta^{i,\nu}| - \sum_{k=1}^n |\zeta^{k,\nu}|)^2 \\ &\leq \alpha_\varphi + \beta_\varphi \sum_{i=1}^n \log(\mu^i) - \sum_{i=1}^n (\mu^i)^{-1} b^i(\boldsymbol{\mu}) + \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d (|\zeta^{i,\nu}| - \sum_{k=1}^n |\zeta^{k,\nu}|)^2 \\ &\leq \sum_{i=1}^n (\beta_\varphi \log(\mu^i) - (\mu^i)^{-1} b^i(\boldsymbol{\mu})) + \alpha_\varphi + \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d (|\zeta^{i,\nu}| - \sum_{k=1}^n |\zeta^{k,\nu}|)^2 \\ &\leq \sum_{i=1}^n \kappa^i + \alpha_\varphi + \frac{1}{2} \sum_{i=1}^n \sum_{\nu=1}^d (|\zeta^{i,\nu}| - \sum_{k=1}^n |\zeta^{k,\nu}|)^2 =: \kappa \end{aligned}$$

where κ^i is defined as follows:

$$\begin{aligned} \beta_\varphi \log(\mu^i) - (\mu^i)^{-1} b^i(\boldsymbol{\mu}) &= \beta_\varphi \log(\mu^i) - B^{i,i} - (\mu^i)^{-1} \sum_{\substack{j=1 \\ j \neq i}}^n B^{i,j} \mu^j \\ &\leq -B^{i,i} + \beta_\varphi \log(\mu^i) - (\mu^i)^{-1} (1 - \mu^i) \min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} \leq \kappa^i \\ \kappa^i &:= \max_{x \in [0,1]} \left(-B^{i,i} + \beta_\varphi \log(x) - (x)^{-1} (1 - x) \min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} \right) \end{aligned}$$

The upper bound κ^i is chosen as the maximum on $[0, 1]$ of the function on the right hand side of the last equality, noted as $f(x)$. In fact, f is continuous on $(0, 1)$ with $\lim_{x \rightarrow 1^-} f(x) = -B^{i,i}$ and $\lim_{x \rightarrow 0^+} f(x) = -\infty$ given that $\min_{j \in \{1, \dots, n\}; j \neq i} B^{i,j} > 0$ so the maximum of f is actually attained on $[0, 1]$.

Consequently, we have $\varphi \circ V + \mathcal{L}V \leq \kappa$ so $\mathcal{L}V \leq \kappa - \varphi \circ V$ on Δ^n and V is a Lyapunov function as demanded.

3.5.4 Generalized local martingale problem and candidate measure for Lemma 3.3.1

Let us consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $X = (X_t)_{t \geq 0} \in \mathcal{X}^{x_0}$, with \mathcal{X}^{x_0} the set of all progressively measurable continuous processes taking values in a closed set Δ with interior Δ^+ and absorbing boundary $\partial\Delta = \Delta \setminus \Delta^+$ and $X_0 = x_0 \in \Delta^+$ \mathbb{P} -a.s. We say that $((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), X)$ is a solution to the *generalized local martingale problem* (GLMP for short) corresponding to (Δ, x_0, a, b) on $[0, \tau)$, with a, b two non-anticipative functions and τ a \mathbb{F} -stopping time, if there exists a non-decreasing sequence of \mathbb{F} -stopping times $(\tau_N)_{N \in \mathbb{N}^*}$ with $\lim_{N \uparrow \infty} \tau_N = \tau$ and $\forall f \in \mathcal{C}^2(\Delta, \mathbb{R})$, the following process $(M_t)_{t \geq 0}$ is a \mathbb{P} -local martingale:

$$M_{t \wedge \tau_N} = f(X_{t \wedge \tau_N}) - \int_0^{t \wedge \tau_N} \left(\sum_{i=1}^d b^i(u, X_u) \partial_i f(X_u) + \frac{1}{2} \sum_{i,j=1}^d a^{i,j} \partial_{i,j} f(X_u) \right) du$$

For a non-anticipative function θ , we define the exponential local martingale $(Z_t)_{t \geq 0}$, $Z_t = \mathcal{E}(\int_0^t \theta(u, X_u) dW_u)_t$ as in (3.3.5) and the stopping times $(\tau_N(X))_{N \in \mathbb{N}^*}$, $\tau(X)$ so that $\lim_{N \uparrow \infty} \tau_N(X) = \tau(X)$:

$$\begin{aligned} \tau_N(X) &= \inf \{ t \geq 0 \mid \int_0^t |\theta(u, X_u)|^2 du \geq N \} \\ \tau(X) &= \inf \{ t \geq 0 \mid \int_0^t |\theta(u, X_u)|^2 du = \infty \} \end{aligned}$$

If the function a can be written as $a = \sigma \sigma^\top$ with non-anticipative σ , then from Proposition 5.4.6 in [KS88] we can write $(X_{t \wedge \tau_N})$ as a stochastic integral, with $(W_t)_{t \geq 0}$ a \mathbb{P} -Brownian motion:

$$X_{t \wedge \tau_N} = x_0 + \int_0^{t \wedge \tau_N} b(u, X_u) du + \int_0^{t \wedge \tau_N} \sigma(u, X_u) dW_u$$

A probability measure \mathbb{Q} on (Ω, \mathcal{F}) is a *candidate measure* associated to the function θ if $((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}), X)$ is a solution to GLMP corresponding to (E, x_0, a, \hat{b}) on $[0, \tau(X))$

with $\widehat{b}(t, X) = b(t, X) - \sigma(t, X)\theta(t, X)\mathbf{1}_{\tau(X) > t}$. Proposition 1 in [Ruf+15] states that such a candidate measure exists and the stopped process $(X_{t \wedge \tau_N(X)})_{t \geq 0}$ has the same distribution under \mathbb{Q} and \mathbb{Q}_N with $d\mathbb{Q}_N = Z_{\cdot \wedge \tau_N(X)} \cdot d\mathbb{P}$.

From Theorem 1 in [Ruf+15], the exponential local martingale $(Z_t)_{t \geq 0}$ is a true \mathbb{P} -martingale if and only if:

$$\forall t \geq 0, \mathbb{Q}\left(\int_0^t |\theta(u, X_u)|^2 du < \infty\right) = 1; \text{ i.e. } \tau(X) = \infty \text{ } \mathbb{Q}\text{-a.s.}$$

Part II

A Probabilistic Scheme for McKean-Vlasov Equations

A Probabilistic Scheme for McKean-Vlasov Equations

The content of this chapter is a work in progress, in collaboration with my PhD supervisor Noufel Frikha.

4.1 Introduction

In this chapter, we are interested in the numerical approximation of the marginal law of a McKean-Vlasov stochastic differential equation (SDE for short) with dynamics

$$X_t^{s,\xi} = \xi + \int_s^t b(X_r^{s,\xi}, [X_r^{s,\xi}]) dr + \int_s^t \sigma(X_r^{s,\xi}, [X_r^{s,\xi}]) dW_r, [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d), \quad (4.1.1)$$

where W is a q -dimensional Brownian motion defined on some probability space $(\Omega, \mathcal{A}, \mathcal{P}(\mathbb{R}^d))$, $b : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times q}$ are smooth functions, $\mathcal{P}(\mathbb{R}^d)$ being the space of probability measures over \mathbb{R}^d . Here and throughout, for any random variable θ defined on $(\Omega, \mathcal{A}, \mathcal{P}(\mathbb{R}^d))$, we use $[\theta]$ to denote the law of θ .

For some results on the well-posedness in the weak and strong sense of such nonlinear SDEs, we refer the interested reader to Sznitman [Szn91], Gärtner [Gär88], Funaki [Fun84] and to the more recent contributions Jourdain [Jou97], Mishura and Veretenikov [MV18] and Chaudru de Raynal and Frikha [RF19] for a small samples.

The numerical approximation of McKean-Vlasov dynamics has attracted a lot of attention during the last decades. The classical method is the approximation using interacting particle systems and the law of the random variable is replaced by the empirical measure associated to the system of particles, see Antonelli and Kohatsu-Higa [AKH+02] and Bossy and Talay [BT97]. To improve the computational efficiency of the particle method, some variance reduction techniques for Monte-Carlo simulation have been considered in some more recent works, for example see Reiss, Smith and Tankov [RST18] for importance sampling and Szpruch, Tan and Tse [STT+19] for iterative multilevel particle approximation. In Crisan and McMurray [CM+19], the cubature on Wiener space is considered as an alternative to the particle system for McKean-Vlasov SDEs under the assumption of smooth scalar interaction.

The aim of this chapter is to provide a probabilistic representation for the approximation of the marginal law of the solution to McKean-Vlasov SDE over a given time horizon $[0, T]$, as an alternative to the aforementioned simulation methods. More specifically, we would like to approximate the probability density function $z \mapsto p(0, T, z)$ of the solution to the McKean-Vlasov SDE or equivalently, obtain an approximation of the expectation $\mathbb{E}[h(X_T)]$ for some measurable function h and starting at time 0 from point ξ with law $\mu \in \mathcal{P}(\mathbb{R}^d)$.

The probabilistic representation formula used and adapted in this chapter comes from the pioneering work of Bally and Kohatsu-Higa [BKH+15] for the marginal law of a standard multi-dimensional diffusion process. Furthermore, this formula provides an unbiased Monte-Carlo simulation method thus eliminates the error associated to the discretization of the diffusion process. However in McKean-Vlasov SDE, the drift and volatility coefficients depend on the distribution of the state variable, which make the method inapplicable. The idea is to regard the marginal laws of McKean-Vlasov SDE as a fixed point of some mapping and under suitable conditions, this fixed point can be approximated by applying Picard iteration to the mapping, with the convergence speed to the fixed point being geometric in number of Picard iterates L , see also [STT+19] and [Ang+19]. By introducing Picard iteration, we also manage to decouple the McKean-Vlasov SDE, which then becomes a time-inhomogeneous SDE and the probabilistic representation can be applied without difficulty.

To reduce the number of Picard iterates and thus reduce computational complexity, the method of continuation is applied by splitting the time intervals $[0, T]$ into uniform subintervals $\{0 = t_0 \leq t_1 \leq \dots \leq t_N = T\}$ and local Picard iteration is performed on each $[t_k, t_{k+1}]$, $k = 0, \dots, N - 1$ to get the flow of probability measure $(\mathbf{P}^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ at Picard iterate $\ell \geq 0$ and the measure $\mathbf{P}^{(L)}(t_{k+1})$ is used at the initial measure for the next interval $[t_k, t_{k+1}]$. From a simulation point of view, we are able to construct the sequence of finite signed measure $\left\{ \widehat{\mathbf{P}}^{(\ell)}(t)(dz) = \widehat{p}_\ell(0, t, z) dz, \ell \geq 0 \right\}$ as the approximation of the flow of probability measure $\left\{ \mathbf{P}^{(\ell)}(t)(dz) = p_\ell(0, t, z) dz, \ell \geq 0 \right\}$ for any $t \in [t_k, t_{k+1}]$. The probability density function $z \mapsto p(0, T, z)$ of the solution of McKean-Vlasov at time T is then approximated by $z \mapsto \widehat{p}_L(0, T, z)$.

The Monte-Carlo estimator for the density contains two sources of error, with the first one being the bias introduced by Picard iterations, more specifically the difference between the true fixed point and the output of Picard iterations as an approximation of the fixed point. The second one is a statistical error, due to the fact that the marginal laws are approximated by the Monte-Carlo estimator derived from the probabilistic representation.

The chapter is organized as follows: In Section 4.2, we first construct the Picard iteration scheme to approximate the solution of McKean-Vlasov SDE by applying the method of continuation. Then the probabilistic representation of marginal law is introduced for the Picard scheme and the Monte-Carlo estimator is naturally constructed. In Section 4.3.2, the errors of local and global Picard iteration schemes are studied. Finally in Section 4.4, some numerical examples are given to illustrate the efficiency of the algorithm, especially in term of density estimation.

Notations:

We introduce here some basic notations and definitions used throughout this chapter. For a sequence of linear operators $(S_i)_{1 \leq i \leq n}$, we define $\prod_{i=1}^n S_i = S_1 \cdots S_n$. We will often use the convention $\prod_{\emptyset} = 1$ which appears when we have for example $\prod_{i=0}^{-1}$. We denote by $\mathcal{C}_b^k(\mathbb{R}^d)$, the collection of bounded continuous functions defined on \mathbb{R}^d which are k -times continuously differentiable with bounded derivatives. The set $\mathcal{B}_b(\mathbb{R}^d)$ is the collection of real-valued bounded measurable maps defined on \mathbb{R}^d . Furthermore we will use the following notation for time and space variables $\mathbf{s}_p = (s_1, \dots, s_p)$, $\mathbf{z}_p = (z_1, \dots, z_p)$, the differentials $d\mathbf{s}_p = ds_1 \cdots ds_p$, $d\mathbf{z}_p = dz_1 \cdots dz_p$ and for a fixed times s, t such that $0 \leq s < t$, we denote by $\Delta_p([s, t]) = \{\mathbf{s}_p \in [s, t]^p : s_0 := s \leq s_1 \leq s_2 \leq \dots \leq s_p \leq t =: s_{p+1}\}$ and write $\Delta_p(t)$ for $\Delta_p([0, t])$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_\ell)$ of length ℓ , we sometimes write $\partial_\alpha^\ell f(x) = \partial_{x_{\alpha_1}} \cdots \partial_{x_{\alpha_\ell}} f(x)$. For a given positive symmetric matrix Σ , we denote by $y \mapsto g(\Sigma t, y)$ the density function taken at y of the Gaussian vector with mean 0 and covariance matrix Σ . With a slight abuse of notation, when $\Sigma = cI_d$ for some given positive constant c , we write $g(c, y) = g(cI_d, y)$. The first and second order Hermite polynomials are denoted by $H_1^i(\Sigma, y) = \partial_{y_i} g(\Sigma, y)(g(\Sigma, y))^{-1} = -(\Sigma^{-1}y)_i$ and $H_{i,j}^2(\Sigma, y) = \partial_{y_i, y_j}^2 g(\Sigma, y)(g(\Sigma, y))^{-1} = (\Sigma^{-1}y)_i(\Sigma^{-1}y)_j - (\Sigma^{-1})_{i,j}$. For a given $z \in \mathbb{R}^d$, the Dirac measure at point z is denoted by $\delta_z(dx)$.

We denote by $\mathcal{M}(\mathbb{R}^d)$ the space of finite signed measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, viewed as the dual of $\mathcal{C}_0(\mathbb{R}^d)$ and endowed with the associated weak- \star topology. The set of probability measures on \mathbb{R}^d is denoted by $\mathcal{P}(\mathbb{R}^d)$ and, for $p \geq 1$, $\mathcal{P}_p(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ is the subset of probability measures with finite p moment. Except otherwise stated we will equip $\mathcal{P}(\mathbb{R}^d)$ with the total variation metric while $\mathcal{P}_p(\mathbb{R}^d)$ will be equipped with the Wasserstein metric of order p .

One of the key inequality that will be used intensively in this work is the following: for any $p, q > 0$ and $x \in \mathbb{R}$, $|x|^p e^{-q|x|^2} \leq (p/(2qe))^{p/2}$. As a direct consequence of this inequality, we have the *space-time inequality*,

$$\forall x \in \mathbb{R}^d, \forall p, c > 0, \quad |x|^p g(ct, x) \leq Ct^{p/2} g(c't, x) \quad (4.1.2)$$

for any $c' > c$ and some positive constant $C := C(c)$ which in turn gives the *standard Gaussian estimates* for the derivatives of Gaussian density, namely

$$\forall c > 0, \quad |H_i^1(ct, x)|g(ct, x) \leq \frac{C}{t^{1/2}}g(c't, x), \quad |H_{i,j}^2(ct, x)|g(ct, x) \leq \frac{C}{t}g(c't, x)$$

for any $c' > c$ and some positive constant $C := C(c)$. As these two inequalities are applied on numerous occasions throughout the chapter, we will often omit to refer to them explicitly. We finally define for two positive parameters α, β , the Mittag-Leffler function $E_{\alpha, \beta}$ as $E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} z^k / \Gamma(\alpha k + \beta)$, $z \in \mathbb{R}$.

4.2 Description of the numerical probabilistic scheme

4.2.1 Assumptions and well-posedness of the McKean-Vlasov SDE

We will work under the following assumptions on the coefficients:

- (HR)** (i) For any $(i, j) \in \{1, \dots, d\}^2$, for any fixed $m \in \mathcal{P}(\mathbb{R}^d)$, the functions $\mathbb{R}^d \ni x \mapsto b_i(x, m)$ and $\mathbb{R}^d \ni x \mapsto a_{i,j}(x, m)$ are continuously differentiable and twice continuously differentiable respectively. Moreover, the functions

$$a_{i,j}, \quad \partial_{x_i} a_{i,j}, \quad \partial_{x_i, x_j}^2 a_{i,j}, \quad b_i, \quad \partial_{x_i} b_i$$

are bounded continuous functions in $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ and they are η -Hölder continuous with respect to the variable $x \in \mathbb{R}^d$, for some $\eta \in (0, 1]$, uniformly with respect to the measure variable m .

- (ii) For any $(i, j) \in \{1, \dots, d\}^2$ and any $x \in \mathbb{R}^d$, the maps $\mathcal{P}(\mathbb{R}^d) \ni m \mapsto b_i(x, m)$, $a_{i,j}(x, m)$ admit a bounded and continuous linear functional derivative denoted respectively by

$$\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \ni (y, m) \mapsto [\delta b_i(x, m)/\delta m](y), [\delta a_{i,j}(x, m)/\delta m](y)$$

thus satisfying the relation: for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and any $(m, m') \in \mathcal{P}(\mathbb{R}^d)$,

$$b_i(x, m) - b_i(x, m') = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b_i(x, \lambda m + (1 - \lambda)m')(y) (m - m')(dy) d\lambda \quad (4.2.1)$$

and

$$a_{i,j}(x, m) - a_{i,j}(x, m') = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} a_{i,j}(x, \lambda m + (1 - \lambda)m')(y) (m - m')(dy) d\lambda. \quad (4.2.2)$$

- (iii) For any $(i, j) \in \{1, \dots, d\}^2$ and any $m \in \mathcal{P}(\mathbb{R}^d)$, the maps

$$(x, y) \mapsto [\delta b_i(x, m)/\delta m](y) \in \mathcal{C}_b^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$$

and

$$(x, y) \mapsto [\delta a_{i,j}(x, m)/\delta m](y) \in \mathcal{C}_b^{2,2}(\mathbb{R}^d \times \mathbb{R}^d).$$

- (HE)** The diffusion coefficient $a(x, m) = (\sigma\sigma^T)(x, m)$ is uniformly elliptic, that is, there exists $\lambda > 1$ such that for all $(x, \xi, m) \in (\mathbb{R}^d)^2 \times \mathcal{P}(\mathbb{R}^d)$, $\lambda^{-1}|\xi|^2 \leq \langle a(x, m)\xi, \xi \rangle \leq \lambda|\xi|^2$.

We importantly point out that under **(HR)** and **(HE)**, the McKean-Vlasov SDE (4.1.1) admits a unique weak solution for any initial data $(s, \mu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$. We refer e.g. to Theorem 3.4 in [RF19] for a proof of the well-posedness of the non-linear martingale problem under weaker assumptions on the coefficients. It then turns out that the non-linear SDE (4.1.1) can be seen as a standard Itô SDE with time-inhomogeneous coefficients $\widehat{b}(t, x) = b(x, [X_t^{s,\xi}])$ and $\widehat{\sigma}(t, x) = \sigma(x, [X_t^{s,\xi}])$. Moreover, under **(HR)**(i), the maps $\widehat{\sigma}_{i,j}(t, \cdot)$ are Lipschitz-continuous on compact subsets of \mathbb{R}^d , see e.g. Theorem 1.2 [Fri10], so that, from Veretennikov [Ver81], pathwise uniqueness holds. It thus follows from the Yamada-Watanabe theorem that the non-linear SDE is well-posed in the strong sense.

Let us also emphasise that the coefficients satisfy some regularity properties with respect to the measure argument. Indeed, for any $x \in \mathbb{R}^d$ and any $m, m' \in \mathcal{P}(\mathbb{R}^d)$,

observing that both maps $y \mapsto [\delta b_i(x, m)/\delta m](y)$, $[\delta a_{i,j}(x, m)/\delta m](y)$ are bounded and Lipschitz-continuous, from (4.2.1) and (4.2.2), for any probability measure $\pi \in \Pi(\mu, \nu)$ ($\Pi(\mu, \nu)$ being the set of all couplings between μ and ν), it holds

$$\begin{aligned} & |b_i(x, m) - b_i(x, m')| + |a_{i,j}(x, m) - a_{i,j}(x, m')| \\ & \leq \sup_{(t,x,m'')} \left[\left| \frac{\delta}{\delta m} b_i(x, m'')(\cdot) \right|_{Lip} + \left| \frac{\delta}{\delta m} a_{i,j}(x, m'')(\cdot) \right|_{Lip} \right] \inf_{\pi \in \Pi(\mu, \nu)} \int_{(\mathbb{R}^d)^2} (|y - z| \wedge 1) \pi(dy, dz) \end{aligned}$$

where $|\frac{\delta}{\delta m} b_i(x, m'')(\cdot)|_{Lip}$ and $|\frac{\delta}{\delta m} a_{i,j}(x, m'')(\cdot)|_{Lip}$ stands for the Lipschitz modulus of the map $y \mapsto \frac{\delta}{\delta m} b_i(x, m)(y)$, $\frac{\delta}{\delta m} a_{i,j}(x, m)(y)$. It is also readily seen that the metric $\inf_{\pi \in \Pi(\mu, \nu)} \int_{(\mathbb{R}^d)^2} (|y - z| \wedge 1) \pi(dy, dz)$ is bounded by $W_1(m, m')$ as well as $d_{TV}(m, m')$. We thus deduce that both coefficients are Lipschitz-continuous in the measure variable with respect to the two aforementioned distances.

Before proceeding, let us give some commonly encountered examples of coefficients for which our assumptions are satisfied.

Example 4.2.1. 1. For some functions $\tilde{b} \in \mathcal{C}_b^{1+\eta, 2}(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$ and $\tilde{\sigma} \in \mathcal{C}_b^{2+\eta, 2}(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d \times \mathbb{R}^q)$, for some $\eta \in (0, 1]$, the coefficients b and σ are written as follows:

$$b(x, m) = \int_{\mathbb{R}^d} \tilde{b}(x, y) m(dy), \quad \sigma(x, m) = \int_{\mathbb{R}^d} \tilde{\sigma}(x, y) m(dy).$$

Then, it holds

$$\frac{\delta}{\delta m} b(x, m)(y) = \tilde{b}(x, y), \quad \frac{\delta}{\delta m} \sigma(x, m)(y) = \tilde{\sigma}(x, y).$$

2. For some functions $\tilde{b} \in \mathcal{C}_b^{1+\eta, 1}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$, $\tilde{\sigma} \in \mathcal{C}_b^{2+\eta, 1}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d \times \mathbb{R}^q)$, for some $\eta \in (0, 1]$, such that $\partial^2 \tilde{b}(x, y)/\partial x_i \partial y_k$, $\partial^3 \tilde{\sigma}(x, y)/\partial x_i \partial x_j \partial y_k$ exist, are continuous and bounded for any $(i, j, k) \in \{1, \dots, d\}^3$ and for some functions $\varphi^b, \varphi^\sigma \in \mathcal{C}_b^2(\mathbb{R}^d, \mathbb{R}^d)$, the coefficients b and σ are written as follows:

$$b(x, m) = \tilde{b}\left(x, \int_{\mathbb{R}^d} \varphi^b(y) m(dy)\right), \quad \sigma(x, m) = \tilde{\sigma}\left(x, \int_{\mathbb{R}^d} \varphi^\sigma(y) m(dy)\right).$$

Given $\mu \in \mathcal{P}(\mathbb{R}^d)$, by weak uniqueness, it makes sense to consider $([X_t^{s, \xi}])_{t \geq 0}$ as a function of the law μ of the r.v. ξ without specifying the choice of ξ . We then introduce, for any $x \in \mathbb{R}^d$, the following decoupled stochastic flow associated to the SDE (4.1.1)

$$X_t^{s, x, \mu} = x + \int_s^t b(X_r^{s, x, \mu}, [X_r^{s, \xi}]) dr + \int_s^t \sigma(X_r^{s, x, \mu}, [X_r^{s, \xi}]) dW_r. \quad (4.2.3)$$

Observe that the above dynamics is not a McKean-Vlasov SDE but simply a standard SDE with time-inhomogeneous coefficients involving the marginal law of the solution to the equation (4.1.1) which admits a unique weak solution under the aforementioned assumptions. Moreover, one may associate a time-inhomogeneous strong Markov semigroup defined for all bounded measurable function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ by $P_{s,t} h(x, \mu) = \mathbb{E}[h(X_t^{s, x, \mu})]$ so that, by weak uniqueness, $\mathbb{E}[h(X_t)] = \mathbb{E}[P_{0,t} h(X_0)] = \int_{\mathbb{R}^d} P_{0,t} h(x, \mu) \mu(dx)$.

4.2.2 Construction of the Picard iteration scheme

Our strategy to approximate the marginal law of the McKean-Vlasov SDE consists in a Picard iteration method based on the law of the process. Such type of approximation is known to converge at a geometric rate for any time horizon T , see e.g. [Szn91]. However, its computational complexity seems to be prohibitive when the number of iterate is large. To balance these two aspects, we employ the so-called method of continuation by regarding T as parameter that we increase step by step. At each step of our approximation scheme, we make use of a Picard iteration based on the approximation obtained at the previous steps.

We thus introduce an equidistant grid $\mathcal{S} = \{t_0 = 0 \leq t_1 \leq \dots \leq t_N = T\}$ of the time interval $[0, T]$ with $t_k = k\delta$ and $\delta = \frac{T}{N}$, for $N \geq 1$. In what follows, the parameter δ should be considered as small and we will state more precisely what we mean in our main results. For a prescribed integer k and starting probability distribution μ_k on \mathbb{R}^d , we perform local Picard iteration on the flow of probability measures on the considered time interval $[t_k, t_{k+1}]$. To be more specific, we introduce the following set

$$\mathcal{A}_{t_k, t_{k+1}, \mu_k} = \left\{ \mathbf{P} \in \mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d)) : \mathbf{P}(t_k) = \mu_k \right\}.$$

Recalling that $(\mathcal{P}(\mathbb{R}^d), d_{TV})$ is a complete metric space, the space $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$ is also complete when equipped with the uniform metric

$$d_{t_k, t_{k+1}}(\mathbf{P}, \mathbf{Q}) := \sup_{s \in [t_k, t_{k+1}]} |d_{TV}(\mathbf{P}(s), \mathbf{Q}(s))|.$$

Since $\mathcal{A}_{t_k, t_{k+1}, \mu_k}$ is a closed subspace of $\mathcal{C}([t_k, t_{k+1}], \mathcal{P}(\mathbb{R}^d))$, it is also complete when equipped with the above metric.

The approach that we follow consists in approximating the flow of probability measures on the time grid \mathcal{S} step by step by performing local Picard iterations on each time interval $[t_k, t_{k+1}]$, $k = 0, \dots, N-1$, as follows. For $k = 0, \dots, n-1$ and a given starting distribution μ_k , let us introduce the map $\mathcal{T}_k : \mathcal{A}_{t_k, t_{k+1}, \mu_k} \rightarrow \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ which to an element $\mathbf{Q} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$ associates the probability measure $\mathcal{T}_k(\mathbf{Q})(t) = [X_t^{\mathbf{Q}}]$ for $t \in [t_k, t_{k+1}]$ where $X^{\mathbf{Q}} = (X_t^{\mathbf{Q}})_{t \in [t_k, t_{k+1}]}$ is given by the unique weak solution to the SDE

$$X_t^{\mathbf{Q}} = \xi_k + \int_{t_k}^t b(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) ds + \int_{t_k}^t \sigma(X_s^{\mathbf{Q}}, \mathbf{Q}(s)) dW_s, \quad t \in [t_k, t_{k+1}]$$

where ξ_k is a random variable independent of the Brownian motion W with law μ_k . Observe that under **(HR)** and **(HE)** the above SDE admits a unique strong solution so that the map \mathcal{T}_k is well-defined. As a consequence, any fixed point of \mathcal{T}_k is the unique solution to the martingale problem associated to the SDE (4.1.1) on the time interval $[t_k, t_{k+1}]$ with initial data (t_k, μ_k) .

We select an initial condition $\mathbf{P}_k^{(0)} \in \mathcal{A}_{t_k, t_{k+1}, \mu_k}$, for the local Picard iteration schemes. For sake of simplicity, $\mathbf{P}_k^{(0)}$ is constructed in a similar way on each time interval $[t_k, t_{k+1}]$ ¹. From a practical point of view, the choice of the starting point $\mathbf{P}_k^{(0)}$

¹Of course, the first time marginal distribution $\mathbf{P}_k^{(0)}(0) = \mu_k$ will vary from one time interval to the next.

is important since it will control the bias of our scheme and one should select $\mathbf{P}_k^{(0)}$ as close as possible to the (unique) fixed point of \mathcal{T}_k , if any exists. A natural choice driven by the small time approximation of the SDE (4.1.1) on the time interval $[t_k, t_{k+1}]$ consists in taking the flow of marginal distributions induced by the limit dynamics with coefficients frozen at (ξ_k, μ_k) , that is, $\mathbf{P}_k^{(0)}(t) = [\bar{X}_t^k]$, $t \in [t_k, t_{k+1}]$, with

$$\bar{X}_t^k = \xi_k + b(\xi_k, \mu_k)(t - t_k) + \sigma(\xi_k, \mu_k)(W_t - W_{t_k})$$

so that

$$\begin{aligned} \mathbf{P}_k^{(0)}(t)(dz) &= g(a(\cdot, \mu_k)(t - t_k), z - b(\cdot, \mu_k)(t - t_k) - \cdot) \# \mu_k dz \\ &= \int_{\mathbb{R}^d} g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) \mu_k(dx) dz \end{aligned} \quad (4.2.4)$$

for any $t \in [t_k, t_{k+1}]$ with the convention $\mathbf{P}_k^{(0)}(t_k) = \lim_{t \downarrow t_k} \mathbf{P}_k^{(0)}(t) = \mu_k$.

Having this initial condition at hand, for a prescribed positive integer L , which represents the number of levels in our scheme, we iterate L times the map \mathcal{T}_k , $\mathcal{T}_k^{(L)} = \mathcal{T}_k(\mathcal{T}_k^{(L-1)})$ with the convention $\mathcal{T}_k^{(0)} = Id$, in order to approximate the flow of marginal distributions induced by the non-linear SDE (4.1.1). To be more specific, on the first time interval $[0, t_1]$, we construct the flow of probability measures

$$\left\{ \mathbf{P}_0^{(L)}(t) := \mathcal{T}_0^{(L)}(\mathbf{P}_0^{(0)}(t)), t \in [0, t_1] \right\}$$

where $\mathbf{P}_0^{(0)}$ is given by (4.2.4) with $\mu_0 = \mu = [\xi]$ and then iterate the above approach on the next time intervals. For any $k = 1, \dots, N-1$, on each time interval $[t_k, t_{k+1}]$, we approximate $([X_t])_{t \in [t_k, t_{k+1}]}$ by $\left\{ \mathbf{P}_k^{(L)}(t) = \mathcal{T}_k^{(L)}(\mathbf{P}_k^{(0)}(t)), t \in [t_k, t_{k+1}] \right\}$, $\mathbf{P}_k^{(0)}$ being the initial flow of probability measures given by (4.2.4) with $\mu_k = \mu_k^L := \mathbf{P}_k^{(L)}(t_k)$ is the marginal probability measure at time t_k after L Picard iteration that has been constructed at the previous step.

We importantly observe that for any time interval $[t_k, t_{k+1}]$ and any level $\ell = 1, \dots, L$, $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$ coincides with the flow of marginal distributions induced by the unique strong solution $X^{(\ell)}$ to the following SDE with dynamics

$$\begin{aligned} X_t^{(\ell)} &= \xi_k^{(\ell)} + \int_{t_k}^t b(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) ds + \int_{t_k}^t \sigma(X_s^{(\ell)}, \mathbf{P}_k^{(\ell-1)}(s)) dW_s, \\ t &\in [t_k, t_{k+1}], [X_{t_k}^{(\ell)}] = \mu_k^{(\ell)} = \mathbf{P}_k^{(\ell)}(t_k). \end{aligned} \quad (4.2.5)$$

Under **(HR)** and **(HE)**, for any $\ell = 1, \dots, L$, from Friedman [Fri08], it is known that the random vector $X_t^{(\ell)}$ given by the solution to (4.2.5) taken at time $t \in (t_k, t_{k+1}]$ admits a density function with respect to the Lebesgue measure, denoted by $z \mapsto p_\ell(t_k, t, z)$ and satisfying

$$\mathbf{P}_k^{(\ell)}(t)(dz) = p_\ell(t_k, t, z) dz = \int_{\mathbb{R}^d} p_\ell(t_k, t, x, z) \mu_k^{(\ell)}(dx) dz$$

where, $z \mapsto p_\ell(t_k, t, x, z)$ stands for the density function of the random variable $X_t^{(\ell), t_k, x}$, the process $(X_t^{(\ell), t_k, x})_{t \geq t_k}$ being the solution to the decoupled SDE associated to (4.2.5) starting from x at time t_k .

Moreover, the following Gaussian upper-bound holds: for any $t \in (t_k, t_{k+1}]$, for any $(x, z) \in (\mathbb{R}^d)^2$ and any positive integer ℓ

$$p_\ell(t_k, t, x, z) \leq E_{1/2,1}(C(|b|_\infty + 1))g(c(t - t_k), z - x) \quad (4.2.6)$$

where $C := C(T, \lambda)$ is a positive constant and $c := c(\lambda)$ is any positive constant satisfying $c > \lambda$, λ being the ellipticity constant appearing in **(HE)**, both being independent of ℓ . With a slight abuse of notation, we will write $z \mapsto p_0(t_k, t, x, z) = g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x)$ for the transition density associated to the level 0 so that, under our current assumptions, the previous Gaussian-upper bound is also satisfied for $\ell = 0$.

4.2.3 Probabilistic representation of the marginal law of the Picard iteration scheme

Our second step in order to compute the marginal law of the McKean-Vlasov SDE (4.1.1) consists in a numerical approximation of the sequence of probability measures $(\mathbf{P}^{(\ell)}(t_k))_{1 \leq \ell \leq L}$, $k = 1, \dots, N$. Our strategy here is based on a probabilistic representation of the marginal law of standard diffusion processes, in the spirit of the unbiased simulation method introduced by Bally and Kohatsu-Higa [BKH+15], see also Labor-dère and al. [HLTT+17], and investigated from a numerical perspective by Andersson and Kohatsu-Higa [AKH+17]. The first tool that we will employ is a renewal process N that we now introduce.

Definition 4.2.1. Let $\tau := (\tau_n)_{n \geq 1}$ be a sequence of random variables such that $(\tau_n - \tau_{n-1})_{n \geq 1}$, with the convention $\tau_0 = 0$, are i.i.d. with density function f and cumulant distribution function $t \mapsto F(t) = \int_{-\infty}^t f(s) ds$. Then, the renewal process $N := (N_t)_{t \geq 0}$ with jump times τ is defined by $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$.

It is readily seen that, for $0 \leq s \leq t$, $\{N_{t-s} = n\} = \{\tau_n \leq t - s < \tau_{n+1}\}$ and by an induction argument that we omit, one may prove that the joint distribution of (τ_1, \dots, τ_n) is given by

$$\mathbb{P}(\tau_1 \in ds_1, \dots, \tau_n \in ds_n) = \prod_{j=0}^{n-1} f(s_{j+1} - s_j) \mathbf{1}_{\{0 < s_1 < \dots < s_n\}}$$

which in turn implies

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}} \Phi(\tau_1, \dots, \tau_n)] &= \mathbb{E}[\mathbf{1}_{\{\tau_n \leq t-s < \tau_{n+1}\}} \Phi(\tau_1, \dots, \tau_n)] \\ &= \int_{t-s}^{\infty} \int_{\Delta_n(t-s)} \Phi(s_1, \dots, s_n) \prod_{j=0}^n f(s_{j+1} - s_j) ds_{n+1} \end{aligned}$$

with the convention $s_0 = 0$. Hence, by Fubini's theorem and the change of variables $s_j = r_j - s$, $j = 0, \dots, n$

$$\begin{aligned} &\mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}} \Phi(\tau_1, \dots, \tau_n)] \\ &= \int_{\Delta_n([s,t])} \Phi(r_1 - s, \dots, r_n - s) (1 - F(t - r_n)) \prod_{j=0}^{n-1} f(r_{j+1} - r_j) dx_n \quad (4.2.7) \end{aligned}$$

for any map $\Phi : \Delta_n(t-s) \rightarrow \mathbb{R}$ satisfying $\mathbb{E}[\mathbf{1}_{\{N_{t-s}=n\}}|\Phi(\tau_1, \dots, \tau_n)|] < \infty$.

Usual choices that we will consider are the followings.

- Example 4.2.2.** 1. If the density function f is given by $f(t) = \lambda e^{-\lambda t} \mathbf{1}_{[0, \infty)}(t)$ for some positive parameter λ , then N is a Poisson process with intensity λ .
2. If the density function f is given by $f(t) = \frac{1-\alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{t^\alpha} \mathbf{1}_{[0, \bar{\tau}]}(t)$ for some parameters $(\alpha, \bar{\tau}) \in (0, 1) \times (0, \infty)$, then N is a renewal process with $[0, \bar{\tau}]$ -valued Beta($1-\alpha, 1$) jump times.
3. More generally, if the density function f is given by $f(t) = \frac{\bar{\tau}^{1-\alpha-\beta}}{B(\alpha, \beta)} \frac{1}{t^{1-\alpha}(\bar{\tau}-t)^{1-\beta}} \mathbf{1}_{[0, \bar{\tau}]}(t)$ for some parameters $(\alpha, \beta, \bar{\tau}) \in (0, 1)^2 \times (0, \infty)$, then N is a renewal process with $[0, \bar{\tau}]$ -valued Beta(α, β) jump times.

Given a positive integer ℓ , an initial probability measure μ_k^ℓ , a Brownian motion W^ℓ , a renewal process N^ℓ with jump times $\tau^\ell = (\tau_n^\ell)_{n \geq 0}$ defined on $(\Omega, \mathcal{A}, \mathbb{P})^2$ and a fixed time $t \in [t_k, t_{k+1}]$, we introduce an Euler scheme $\bar{X}^\ell = (\bar{X}_{\zeta_n}^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$, evolving along the (random) time grid $\zeta^\ell := (\zeta_n^\ell)_{0 \leq n \leq N_{t-t_k}^\ell + 1}$, defined by $\zeta_n^\ell := t_k + \tau_n^\ell \wedge (t - t_k)$, $\zeta_0^\ell := t_k$, starting from $\bar{X}_{t_k}^\ell$ with distribution $[\bar{X}_{t_k}^\ell] = \mu_k^\ell$, and with dynamics

$$\bar{X}_{\zeta_{n+1}^\ell}^\ell = \bar{X}_{\zeta_n^\ell}^\ell + b_n^\ell(\zeta_{n+1}^\ell - \zeta_n^\ell) + \sigma_n^\ell(W_{\zeta_{n+1}^\ell}^\ell - W_{\zeta_n^\ell}^\ell), \quad 0 \leq n \leq N_{t-t_k}^\ell, \quad (4.2.8)$$

where we introduced the shorter notation $\sigma_n^\ell = \sigma(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell))$, $b_n^\ell = b(\bar{X}_{\zeta_n^\ell}^\ell, \mathbf{P}^{(\ell-1)}(\zeta_n^\ell))$ together with $a_n^\ell = \sigma_n^\ell(\sigma_n^\ell)^*$.

Here, for each level $\ell \geq 1$, W^ℓ , N^ℓ and $\bar{X}_{t_k}^\ell$ are independent random variables. Moreover, $(W^\ell, N^\ell, \bar{X}_{t_k}^\ell)_{\ell \geq 1}$ is an i.i.d. sequence of random variables with the same law as $(W, N, \bar{X}_{t_k}^\ell)$.

The main building block of our numerical scheme is the following probabilistic representation of the marginal laws $(\mathbf{P}^\ell(t))_{t \in [t_k, t_{k+1}]}$ whose proof is postponed to the Appendix, Section 4.5.1.

Theorem 4.2.1. Assume that **(HR)** and **(HE)** hold. Let $h : \mathbb{R}^d \rightarrow \mathbb{R}$ be a measurable function satisfying: $\forall z \in \mathbb{R}^d$, $|h(z)| \leq C \exp\left(\alpha \frac{|z|^2}{2T}\right)$ for some constants $C > 0$ and $\alpha < c^{-1}$, c being the constant prescribed in (4.2.6). Then, there exists a non-negative constant $\kappa := \kappa(c, \alpha)$ such that if $\int_{\mathbb{R}^d} \exp(\kappa|x|^2) \mu_k^\ell(dx) < \infty$, (in particular $\kappa = 0$ if $\alpha = 0$, i.e. if h is bounded), the expectation $\mathbb{E}[h(X_t^{(\ell)})]$, where $X^{(\ell)}$ is the solution of the SDE (4.2.5), satisfies the following probabilistic representation

$$\forall t \in [t_k, t_{k+1}], \quad \mathbb{E}[h(X_t^{(\ell)})] = \mathbb{E}\left[h(\bar{X}_{\zeta_{N_{t-t_k}^\ell}^\ell}^\ell) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^\ell\right], \quad \ell \geq 1 \quad (4.2.9)$$

where $\theta_r^\ell := \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}^\ell(\bar{X}_{\zeta_r^\ell}^\ell, \bar{X}_{\zeta_{r+1}^\ell}^\ell, \mathbf{P}_k^{(\ell-1)})$ for $r = 0, \dots, N_{t-t_k}^\ell - 1$ and $\theta_{N_{t-t_k}^\ell}^\ell = (1 - F(t - t_k))$.

²In what follows, we assume without loss of generality that the probability space is rich enough to support all random variables that we will employ.

$\zeta_{N_{t-t_k}^\ell}^\ell)^{-1}$ with

$$\theta_{s,t}(x, y, \mathbf{P}) := (f(t-s))^{-1} \left[\sum_{i,j=1}^d \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) - \sum_{i=1}^d \rho_{s,t}^i(x, y, \mathbf{P}) \right]$$

and where for $i, j = 1, \dots, d$, the maps

$$\Delta_2(t_k, t_{k+1}) \times (\mathbb{R}^d)^2 \times \mathcal{A}_{t_k, t_{k+1}, \mu_k} \ni (s, t, x, y, \mathbf{P}) \mapsto \kappa_{s,t}^{i,j}(x, y, \mathbf{P}), \rho_{s,t}^i(x, y, \mathbf{P})$$

are defined by

$$\begin{aligned} \kappa_{s,t}^{i,j}(x, y, \mathbf{P}) &:= \frac{1}{2} [a_{i,j}(y, \mathbf{P}(t)) - a_{i,j}(x, \mathbf{P}(s))] H_2^{i,j}(a(x, \mathbf{P}(s))(t-s), y-x) \\ &\quad + \partial_{y_i} a_{i,j}(y, \mathbf{P}(t)) H_1^i(a(x, \mathbf{P}(s))(t-s), y-x) + \frac{1}{2} \partial_{y_i y_j}^2 a_{i,j}(y, \mathbf{P}(t)), \\ \rho_{s,t}^i(x, y, \mathbf{P}) &:= [b_i(y, \mathbf{P}(t)) - b_i(x, \mathbf{P}(s))] H_1^i(a(x, \mathbf{P}(s))(t-s), y-x) + \partial_{y_i} b_i(y, \mathbf{P}(t)). \end{aligned}$$

Moreover, for any $t \in (t_k, t_{k+1}]$, the density functions $z \mapsto p_\ell(t_k, t, z)$, $p_\ell(t_k, t, x, z)$ satisfy the following probabilistic formulation

$$\forall t \in (t_k, t_{k+1}], \quad p_\ell(t_k, t, z) = \mathbb{E} \left[g(a_{N_{t-t_k}^\ell}^\ell(t - \zeta_{N_{t-t_k}^\ell}^\ell), z - \bar{X}_{N_{t-t_k}^\ell}^\ell) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^\ell \right], \quad \ell \geq 1, \quad (4.2.10)$$

and

$$\forall t \in (t_k, t_{k+1}], \quad p_\ell(t_k, t, x, z) = \mathbb{E} \left[g(a_{N_{t-t_k}^\ell}^\ell(t - \zeta_{N_{t-t_k}^\ell}^\ell), z - \bar{X}_{N_{t-t_k}^\ell}^{\ell,x}) \prod_{r=0}^{N_{t-t_k}^\ell} \theta_r^{\ell,x} \right], \quad \ell \geq 1,$$

with $\theta_r^{\ell,x} := \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}(\bar{X}_{\zeta_r^\ell}^{\ell,x}, \bar{X}_{\zeta_{r+1}^\ell}^{\ell,x}, \mathbf{P}_k^{(\ell-1)})$ for $r = 0, \dots, N_{t-t_k}^\ell - 1$ and $\theta_{N_{t-t_k}^\ell}^\ell = (1 - F(t - \zeta_{N_{t-t_k}^\ell}^\ell))^{-1}$ and where $(\bar{X}_{\zeta_r^\ell}^{\ell,x})_{0 \leq r \leq N_{t-t_k}^\ell}$ is the Euler scheme (5.2.4) starting from x at time t_k .

Corollary 4.2.1. Under the assumptions of the previous theorem, for any $k \in \{0, \dots, N-1\}$, it holds

$$\forall t \in [t_k, t_{k+1}], \quad \mathbb{E}[h(X_t^{(\ell)})] = \mathbb{E} \left[h(\bar{X}_{\zeta_{N_{t-t_k}^\ell}^\ell}^\ell) \prod_{r=0}^{N_{t-t_k}^{\ell,k+1}} \theta_r^{\ell,k+1} \times \Gamma_k^\ell \right], \quad \ell \geq 1 \quad (4.2.11)$$

where

$$\Gamma_k^\ell := \Gamma_{k-1}^\ell \prod_{r=0}^{N_\delta^{\ell,k}} \theta_r^{\ell,k} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell,j}} \theta_r^{\ell,j}, \quad \text{with} \quad \Gamma_0^\ell := 1$$

and for $j = 1, \dots, k$ and $r = 0, \dots, N_\delta^{\ell,j} - 1$

$$\begin{aligned}\theta_r^{\ell,j} &:= \theta_{\zeta_r^\ell, \zeta_{r+1}^\ell}^\ell(\bar{X}_r^\ell, \bar{X}_{r+1}^\ell, \mathbf{P}_{j-1}^{\ell-1}), \\ \theta_{N_\delta^{\ell,j}}^{\ell,j} &:= (1 - F(t_j - \zeta_{N_\delta^{\ell,j}}^\ell))^{-1}, \\ \theta_{N_{t-t_k}^{\ell,k+1}}^{\ell,k+1} &:= (1 - F(t - \zeta_{N_{t-t_k}^{\ell,k+1}}^\ell))^{-1}, \\ \bar{X}_{\zeta_{r+1}^\ell}^\ell &= \bar{X}_{\zeta_r^\ell}^\ell + b_r^\ell(\zeta_{r+1}^\ell - \zeta_r^\ell) + \sigma_r^\ell(W_{\zeta_{r+1}^\ell}^\ell - W_{\zeta_r^\ell}^\ell).\end{aligned}$$

Here $\{N^{\ell,j}, j = 1, \dots, k+1\}$ is a sequence of i.i.d. renewal processes with respective jump times $\tau^{\ell,j} = (\tau_r^{\ell,j})_{r \geq 0}$ so that $\zeta_r^\ell := t_{j-1} + \tau_r^{\ell,j} \wedge \delta$, for $r = 0, \dots, N_\delta^{\ell,j} + 1$, $j = 1, \dots, k$ and $\zeta_r^\ell := t_k + \tau_r^{\ell,k+1} \wedge (t - t_k)$, for $r = 0, \dots, N_{t-t_k}^{\ell,k+1} + 1$.

Remark 4.2.1. We here make an important remark concerning the sequence of weights $(\theta_r^\ell)_{r \geq 0}$ appearing in Theorem 4.2.1 and Corollary 4.2.1. From the proof of Theorem 4.2.1, it is crucial that $\theta_{s,t}(x, y, \mathbf{P}^{\ell-1})$ satisfies (4.5.3) in order to guarantee the validity of the probabilistic representation (4.2.9) and actually the $L^1(\mathbb{P})$ -boundedness (or $L^p(\mathbb{P})$ for $p > 1$ if the renewal process is suitably chosen) of the r.v. appearing inside the expectation.

In order to guarantee this important property from a numerical perspective for the Monte Carlo estimator, we will work under the modified version that we still denote by $\theta_{s,t}(x, y, \mathbf{P}^{\ell-1})$ and given by $-m_{s,t}(x, y) \vee \theta_{s,t}(x, y, \mathbf{P}^{\ell-1}) \wedge m_{s,t}(x, y)$ where $m_{s,t}(x, y)$ is the explicit upper-bound appearing in the right-hand side of (4.5.3).

4.2.4 Construction of the Monte Carlo estimator

It is then quite natural to approximate step by step the marginal laws $(\mathbf{P}^{(\ell)}(t_k))_{1 \leq k \leq n}$ using the previous probabilistic representation (4.2.11). Indeed, we importantly observe that given the previous flow of probability measures $(\mathbf{P}^{(\ell-1)}(t))_{t \in [0, t_{k+1}]}$, an *unbiased Monte Carlo simulation method* directly stems from (4.2.11) and allows to approximate the flow of probability measures at the next step, namely $(\mathbf{P}^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$. More specifically, for any given time $t \in [t_k, t_{k+1}]$, one may approximate the sequence of probability measures $\{\mathbf{P}^{(\ell)}(t)(dz) = p_\ell(0, t, z) dz, \ell \geq 0\}$ by a sequence of *finite signed measure* $\{\widehat{\mathbf{P}}^{(\ell)}(t)(dz) = \widehat{p}_\ell(0, t, z) dz, \ell \geq 0\}$ which are absolutely continuous with respect to the Lebesgue measure and constructed recursively as follows.

Given the initial flow of probability measures (constructed step by step from $k = 0$ up to $k = N - 1$) given by

$$\left\{ \widehat{\mu}^0(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \widehat{\mu}_k^L)(t - t_k), z - b(x, \widehat{\mu}_k^L)(t - t_k) - x) \widehat{\mu}_k^L(dx) dz, t \in [t_k, t_{k+1}], \right. \\ \left. k = 0, \dots, N - 1. \right\}$$

we first define recursively the sequence of signed measures $(\widehat{\mu}^\ell)_{t \in [0, T]}$, for $\ell \geq 1$, step by step from $k = 0$ up to $k = N - 1$, on the time interval $[t_k, t_{k+1}]$, $\forall t \in [t_k, t_{k+1}]$

$$\widehat{\mu}^\ell(t)(dz) := \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Gamma_k^{\ell, (m)} \prod_{r=0}^{N_{t-t_k}^{\ell, k+1, (m)}} \theta_r^{\ell, k+1, (m)} \delta_{\bar{X}_{N_{t-t_k}^{\ell, k+1, (m)}+1}^\ell}^{\ell, (m)}(dz) \quad (4.2.12)$$

with for $m \in \{1, \dots, M_\ell\}$

$$\Gamma_k^{\ell, (m)} := \Gamma_{k-1}^{\ell, (m)} \prod_{r=0}^{N_\delta^{\ell, k, (m)}} \theta_r^{\ell, k, (m)} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell, j, (m)}} \theta_r^{\ell, j, (m)}, \quad \text{with } \Gamma_0^{\ell, (m)} := 1,$$

and for any $j \in \{1, \dots, k+1\}$ and for any $r \in \left\{0, \dots, N_\delta^{\ell, j, (m)} - 1\right\}$ (if $1 \leq j \leq k$) or $r \in \left\{0, \dots, N_{t-t_k}^{\ell, j, (m)}\right\}$ (if $j = k+1$)

$$\begin{aligned} \theta_r^{\ell, j, (m)} &:= \theta_{\zeta_r^{\ell, (m)}, \zeta_{r+1}^{\ell, (m)}}^{\ell}(\bar{X}_r^{\ell, (m)}, \bar{X}_{r+1}^{\ell, (m)}, \hat{\mu}^{\ell-1}), \\ \theta_{N_\delta^{\ell, j}}^{\ell, j, (m)} &:= (1 - F(t_j - \zeta_{N_\delta^{\ell, j, (m)}}^{\ell, (m)}))^{-1}, \\ \theta_{N_{t-t_k}^{\ell, k+1}}^{\ell, k+1, (m)} &:= (1 - F(t - \zeta_{N_{t-t_k}^{\ell, k+1, (m)}}^{\ell, (m)}))^{-1}, \\ \bar{X}_{\zeta_{r+1}^{\ell, (m)}}^{\ell, (m)} &= \bar{X}_{\zeta_r^{\ell, (m)}}^{\ell, (m)} + b_r^{\ell}(\zeta_{r+1}^{\ell, (m)} - \zeta_r^{\ell, (m)}) + \sigma_r^{\ell, (m)}(W_{\zeta_{r+1}^{\ell, (m)}}^{\ell, (m)} - W_{\zeta_r^{\ell, (m)}}^{\ell, (m)}) \end{aligned} \quad (4.2.13)$$

with the notations $b_r^{\ell, (m)} := b(\bar{X}_r^{\ell, (m)}, \hat{\mu}^{\ell-1}(\zeta_r^{\ell, (m)}))$, $\sigma_r^{\ell, (m)} := \sigma(\bar{X}_r^{\ell, (m)}, \hat{\mu}^{\ell-1}(\zeta_r^{\ell, (m)}))$, $a_r^{\ell, (m)} = \sigma_r^{\ell, (m)}(\sigma_r^{\ell, (m)})^T$. Here, we require that $(W^{\ell, (m)}, N^{\ell, (m)}, \bar{X}_0^{\ell, (m)})$, $1 \leq m \leq M_\ell$, $\ell \geq 1$, is an i.i.d. sequence of random variables with the same law as (W, N, X_0) . Moreover, for any fixed $m \in \{1, \dots, M_\ell\}$, $\{N^{\ell, j, (m)}, j = 1, \dots, k+1\}$ is a sequence of i.i.d. renewal processes with respective jump times $\tau_r^{\ell, j, (m)} = (\tau_r^{\ell, j, (m)})_{r \geq 0}$ so that $\zeta_r^{\ell, (m)} := t_j + \tau_r^{\ell, j, (m)} \wedge \delta$, for $r = 0, \dots, N_\delta^{\ell, j, (m)} + 1$, $j = 1, \dots, k$ and $\zeta_r^{\ell, (m)} := t_k + \tau_r^{\ell, k+1, (m)} \wedge (t - t_k)$, for $r = 0, \dots, N_{t-t_k}^{\ell, k+1, (m)} + 1$. We then set $\hat{\mu}_k^L = \hat{\mu}^L(t_k)$ for $k = 1, \dots, N$ and $\hat{\mu}_0^L = \frac{1}{M_0} \sum_{i=1}^{M_0} \delta_{\xi^{(i)}}$, $(\xi^{(i)})_{1 \leq i \leq M_0}$ being an i.i.d sequence with common law μ .

We eventually approximate the quantity of interest $\mathbb{E}[h(X_T)]$ by the Monte Carlo estimator which consists in integrating the function h against the measure $\hat{\mu}^L(T)$, namely

$$\langle h, \hat{\mu}^L(T) \rangle = \frac{1}{M_L} \sum_{m=1}^{M_L} h(\bar{X}_{N_\delta^L, N, (m)+1}^L) \times \Gamma_N^{L, (m)}.$$

We importantly note that one may approximate the density function $z \mapsto p(0, T, z)$ of the random vector $X_T^{0, \xi}$ given by the solution of the SDE (4.1.1) taken at time T and starting at time 0 from ξ with law μ by the random signed measure $z \mapsto \hat{p}_L(0, T, z)$ defined by

$$\hat{p}_L(0, T, z) := \frac{1}{M_L} \sum_{m=1}^{M_L} g(a_{N_\delta^L, N, (m)}^{L, (m)}(T - \zeta_{N_\delta^L}^{L, (m)}), z - \bar{X}_{N_\delta^L}^{L, (m)}) \times \Gamma_N^{L, (m)}.$$

Observe that in order to construct $\hat{\mu}^L(t)$, for a given time $t \in [0, t_1]$ and a given positive level L , one needs to have access to the signed measure $\hat{\mu}^{L-1}$ on the time grid given by the jump times of the trajectories of the renewal processes $N^{L, (m)}$, $1 \leq m \leq M_L$. Having this measure at hand, one may then devise the Euler scheme with dynamics (4.2.13), compute the weights $(\theta_r^{L, (m)})_{0 \leq r \leq N_\delta^L}$ and then the quantity $\Gamma_1^{L, (m)}$.

However, the same procedure appears at level $L - 1$, since, in order to compute the measure $\widehat{\mu}^{L-1}(\zeta_r^{L,1,(m)})$, one needs to have access to the signed measure $\widehat{\mu}^{L-2}$ on the time grid of the renewal processes $N^{L-1,1,(m)}$, $1 \leq m \leq M_{L-1}$, so on and so forth. We thus observe that the (averaged) computational cost for constructing the measure $\widehat{\mu}^L(t_1)$ is of order $M_0 \times \cdots \times M_L$ ³. We then continue by constructing the measure $\widehat{\mu}^L(t)$, for a given time $t \in [t_1, t_2]$, the sequence of weights $\Gamma_1^{\ell,(m)}$, $1 \leq \ell \leq L$, $1 \leq m \leq M_L$, appearing in (4.2.12) for $k = 1$, being already computed at the preceding step. Clearly, the same methodology as for the previous time interval applies so that the computational cost for constructing the measure $\widehat{\mu}_{t_1}^L$ is of order $M_L \times M_1 \times \cdots \times M_L$. We proceed accordingly up to the last time interval $[t_{N-1}, T]$. Hence, the (averaged) total complexity of our algorithm is of order $M_0 \times \cdots \times M_L + (N - 1)M_L \times M_1 \times \cdots \times M_L$.

4.3 Convergence Analysis

In this section, we analyze the error associated to our numerical scheme.

4.3.1 Decomposition of the global error and complexity of the algorithm

Estimating the quantity $\mathbb{E}[h(X_T)]$ or the density function $z \mapsto p(0, T, z)$, X_T being the solution of the SDE (4.1.1) taken at time T and starting from the initial point ξ with distribution μ at time 0 and $z \mapsto p(0, T, z)$ being its density function, by $\langle h, \widehat{\mu}^L(T) \rangle = \int_{\mathbb{R}^d} h(z) \widehat{\mu}^L(T)(dz)$ or $\widehat{p}_L(0, T, z)$ introduces two different sources of error to be analyzed, namely

$$\mathbb{E}[h(X_T)] - \langle h, \widehat{\mu}^L(T) \rangle = \mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^L)] + \mathbb{E}[h(X_T^L)] - \langle h, \widehat{\mu}^L(T) \rangle$$

or

$$p(0, T, z) - \widehat{p}_L(0, T, z) = p(0, T, z) - p_L(0, T, z) + p_L(0, T, z) - \widehat{p}_L(0, T, z).$$

The first error in the above decomposition is the *bias* introduced by the approximation of the marginal law of the McKean-Vlasov equation by the Picard iteration scheme at level L . It is expected to converge to zero as $L \uparrow \infty$. We will prove that the convergence is geometric, namely of order $(C\delta)^L/\Gamma(1 + (L - 1)/2) \sim (C\delta)^L/([L/2])!$, see Proposition 4.3.1. This quantity is analyzed in details in Section 4.3.3 and the estimate of the global error (4.3.3) makes part of our current work.

The second error in the above decomposition corresponds to a statistical error and is due to the approximation at each step $\ell = 1, \dots, L$ of the flow of probability measures $(\mathbf{P}^\ell(s))_{s \in [0, T]}$ by the corresponding flow of signed measures $(\widehat{\mu}^\ell(s))_{s \in [0, T]}$ derived from the probabilistic representation of Theorem 4.2.1 and Corollary 4.2.1. In the best case, the $L^2(\mathbb{P})$ moment of this quantity, namely $\mathbb{E}[(\mathbb{E}[h(X_T^L)] - \langle h, \widehat{\mu}^L(T) \rangle)^2]^{1/2}$, is expected to be of order $M_L^{-\frac{1}{2}}$. We are currently investigating this statistical error.

Setting $M_\ell = M$, $\ell = 0, \dots, L$, for sake of simplicity, we thus observe that in order to obtain a global root mean square error of order ε , for any fixed positive integer L ,

³We assume that the computational cost of the numerical evaluation of a coefficient at a given empirical measure is of order the number of samples.

one has to choose N of order $\varepsilon^{-\frac{1}{L}}$ and M of order ε^{-2} so that our algorithm achieves a computation cost NM^{L+1} which is of order $\varepsilon^{-\frac{1}{L}}\varepsilon^{-2L-2}$. Optimizing the previous cost with respect to the variable L , we thus see that the optimal complexity is of order ε^{-5} and is actually achieved for $L = 1$.

4.3.2 Convergence analysis of the local Picard iteration schemes

Proposition 4.3.1. *Assume that (HR) and (HE) hold. Then, for any $k = 0, \dots, n-1$, the sequence of probability measures $\left\{ \mathbf{P}_k^{(\ell)} = (\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}, \ell \geq 0 \right\}$, $\mathbf{P}_k^{(\ell)}$ being the probability measure, with time marginals $(\mathbf{P}_k^{(\ell)}(t))_{t \in [t_k, t_{k+1}]}$, induced by the unique weak solution to the SDE (4.2.5) with starting distribution $\mu_k \in \mathcal{P}_2(\mathbb{R}^d)$, converges as $\ell \uparrow \infty$ to the probability measure \mathbf{P}_k^∞ , with time marginals $(\mathbf{P}_k^\infty(t))_{t \in [t_k, t_{k+1}]}$, given by the unique solution to the non-linear martingale problem associated to the SDE (4.1.1) on the time interval $[t_k, t_{k+1}]$ with starting distribution μ_k .*

Moreover, for all $h \in \mathcal{C}^2(\mathbb{R}^d)$, h and its derivatives being at most of quadratic growth, for any positive integer L , one has

$$\forall t \in [t_k, t_{k+1}], \quad |\mathbb{E}[h(X_t)] - \mathbb{E}[h(X_t^{(L)})]| \leq \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \quad (4.3.1)$$

where

$$\mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) := \sup_{t_k \leq r \leq t_{k+1}} \mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r))$$

with

$$\mathscr{W}(\mathbf{P}_k^\infty(r), \mathbf{P}_k^{(0)}(r)) := \sup_{|h|_{\mathcal{C}_b^2} \leq 1} \left(\int_{\mathbb{R}^d} h(z) \mathbf{P}_k^\infty(r)(dz) - \int_{\mathbb{R}^d} h(z) \mathbf{P}_k^{(0)}(r)(dz) \right).$$

Additionally, recalling that $z \mapsto p(t_k, t, x, z)$ (resp. $z \mapsto p_L(t_k, t, x, z)$) stands for the transition density function of the decoupled SDE associated to (4.1.1) (resp. associated to (4.2.5) at step L) starting from x at time t_k , for all $(t, x, z) \in (t_k, t_{k+1}] \times (\mathbb{R}^d)^2$, one has

$$|p(t_k, t, x, z) - p_L(t_k, t, x, z)| \leq \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) g(c(t-t_k), z-x) \quad (4.3.2)$$

for any constant $c > \lambda$ (independent of L and T), where $C := C(c, \lambda, a, b, T)$ is a positive constant, non-decreasing with respect to T .

The proof of the above proposition is postponed to the Appendix, Section 4.5.2.

Let us note that since $\mathbf{P}_k^{(0)}(t)(dz) = \int_{\mathbb{R}^d} g(a(x, \mu_k)(t-t_k), z - b(x, \mu_k)(t-t_k) - x) \mu_k(dx) dz$ and $\mathbf{P}_k^\infty(t)(dz) = \int_{\mathbb{R}^d} p(t_k, t, x, z) \mu_k(dx) dz$ for any $t \in [t_k, t_{k+1}]$, where $z \mapsto \cdot$ is the density of the r.v. $X_t^{t_k, x, \mu_k}$ given by the solution at time t of the decoupling SDE (4.2.3) starting from x at time t_k , it is possible to derive an error bound for the quantity $\mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)})$. Indeed, for $h \in \mathcal{C}_b^2(\mathbb{R}^d)$, introducing the notation $z^\lambda =$

$\lambda z + (1 - \lambda)x$, $\lambda \in [0, 1]$, and using a second order Taylor expansion, one obtains

$$\begin{aligned}
 & \int_{\mathbb{R}^d} h(z) [\mathbf{P}_k^{(0)}(t)(dz) - \mathbf{P}_k^\infty(t)(dz)] \\
 &= \int_{(\mathbb{R}^d)^2} h(z) [g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) - p(t_k, t, x, z)] dz \mu_k(dx) \\
 &= \int_{(\mathbb{R}^d)^2} [h(z) - h(x)] [g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) - p(t_k, t, x, z)] dz \mu_k(dx) \\
 &= \int_0^1 \int_{(\mathbb{R}^d)^2} (z - x)^* D^2 h(z^\lambda)(z - x) [g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) \\
 &\quad - p(t_k, t, x, z)] dz \mu_k(dx) d\lambda - \int_{(\mathbb{R}^d)^2} Dh(x) \cdot (z - x) p(t_k, t, x, z) dz \mu_k(dx) + \mathcal{O}(t - t_k)
 \end{aligned}$$

where we used the fact that $\int_{\mathbb{R}^d} (z - x) g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) dz = b(x, \mu_k)(t - t_k) = \mathcal{O}(t - t_k)$. From (4.2.6), the fact that $m_2(\phi) < \infty$ and $\mu_k \in \mathcal{P}^2(\mathbb{R}^d)$, one deduces

$$\left| \int_0^1 \int_{(\mathbb{R}^d)^2} (z - x)^* D^2 h(z^\lambda)(z - x) [g(a(x, \mu_k)(t - t_k), z - b(x, \mu_k)(t - t_k) - x) - p_1(t_k, t, x, z)] dz \mu_k(dx) d\lambda \right| \leq C |h|_{\mathcal{C}_b^2}(t - t_k).$$

Now, from Friedman [Fri08], one has $p(t_k, t, x, z) = g(\int_{t_k}^t a(x, \mathbf{P}_k^\infty(s)) ds, z - x) + \tilde{p}(t_k, t, x, z)$ with $|\tilde{p}(t_k, t, x, z)| \leq C(t - t_k)^{\frac{1}{2}} g(c(t - t_k), z - x)$ so that, using the fact that $\int_{\mathbb{R}^d} Dh(x) \cdot (z - x) g(\int_{t_k}^t a(x, \mathbf{P}_k^\infty(s)) ds, z - x) dz = 0$ and the space-time inequality we get (4.1.2)

$$\left| \int_{(\mathbb{R}^d)^2} Dh(x) \cdot (z - x) p(t_k, t, x, z) dz \mu_k(dx) \right| \leq C |Dh|_\infty (t - t_k).$$

Gathering the above estimates, we thus obtain for any $k = 0, \dots, n - 1$

$$\mathcal{W}_{k, k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \leq C\delta$$

for some positive constant C , independent of n and k .

4.3.3 Convergence analysis of the global Picard iteration scheme on $[0, T]$

We here analyse the global error induced by our local Picard iteration schemes. From the construction of our iterative scheme, we observe the following key relation

$$\begin{aligned}
 \mathbb{E}[h(X_T^L)] &= \int_{(\mathbb{R}^d)^2} h(z_N) p_L(t_{N-1}, T, z_{N-1}, z_N) dz_N \mu_{N-1}^L(dz_{N-1}) \\
 &= \int_{(\mathbb{R}^d)^{N+1}} h(z_N) \prod_{i=1}^N p_L(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}} \mu(dz_0) \\
 &= \int_{(\mathbb{R}^d)^2} h(z_N) p_L(0, T, z_0, z_N) dz_N \mu(dz_0)
 \end{aligned}$$

so that

$$p_L(0, T, z_N) = \int_{\mathbb{R}^d} p_L(0, T, z_0, z_N) \mu(dz_0) = \int_{(\mathbb{R}^d)^N} \prod_{i=1}^N p_L(t_{i-1}, t_i, z_{i-1}, z_i) dz_{\mathbf{N}-1} \mu(dz_0).$$

Assume that **(HR)** and **(HE)** hold. The goal of our current work in progress is to show that, for any positive integer L it holds

$$\begin{aligned} & |p(0, T, z) - p_L(0, T, z)| \\ & \leq CT^{-1} \frac{(C\delta)^{L-1}}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k, k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \int_{\mathbb{R}^d} g(cT, z - z_0) \mu(dz_0). \end{aligned} \quad (4.3.3)$$

The first idea is to use the following standard decomposition

$$\begin{aligned} & p(0, T, z) - p_L(0, T, z) \\ & = \int_{(\mathbb{R}^d)^N} \left[\prod_{i=1}^N p(t_{i-1}, t_i, z_{i-1}, z_i, \mu_{i-1}) - \prod_{i=1}^N p_L(t_{i-1}, t_i, z_{i-1}, z_i, \mu_{i-1}^L) \right] dz_{\mathbf{N}-1} \mu(dz_0) \\ & = \sum_{k=1}^N \int_{(\mathbb{R}^d)^3} p_L(0, t_{k-1}, z_0, z_{k-1}) [p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}) - p_L(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L)] \\ & \quad \times p(t_k, t_N, z_k, z) dz_{k-1} dz_k \mu(dz_0) \end{aligned}$$

where we specifically mention the dependence w.r.t the initial measure μ_{k-1} and μ_{k-1}^L in the two transition density functions $p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1})$ and $p_L(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L)$.

If $N = 1$, i.e. $T = \delta$, the result is a consequence of Proposition 4.3.1. We now assume that $N \geq 2$ for the rest of the proof. Recall that μ_{k-1} represents the marginal law of McKean-Vlasov SDE (4.1.1) and μ_{k-1}^L the marginal law of the L -th Picard iterate given by (4.2.5) at time t_{k-1} with initial measure μ_0 , the difference between densities can be decomposed as follows

$$\begin{aligned} & p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}) - p_L(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L) \\ & = \left(p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}) - p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L) \right) \\ & \quad + \left(p(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L) - p_L(t_{k-1}, t_k, z_{k-1}, z_k, \mu_{k-1}^L) \right) \end{aligned}$$

The first term appearing on the r.h.s of the above equality is the difference between probability densities at time t_k of solutions of McKean-Vlasov SDE (4.1.1) with different initial measures μ_{k-1} and μ_{k-1}^L at time t_{k-1} . We are currently studying this error.

The second term corresponds to the difference between the densities at time t_k of the marginal law of McKean-Vlasov SDE (4.1.1) and of the marginal law of the L -th Picard iterate given by (4.2.5) with initial measure μ_{k-1}^L at time t_{k-1} .

To analyze this error, we split the sum into two disjoint case: $\{k \in \{1, \dots, N\} : t_N - t_k \geq T/2\}$ and $\{k \in \{1, \dots, N\} : t_N - t_k < T/2\}$. Since there is no ambiguity, we remove below the dependence w.r.t the initial measure μ_{k-1}^L in the notation for both transition densities. In the first case, we use the fact that the map $x \mapsto p(t_k, t_N, x, z)$ is $\mathcal{C}_b^2(\mathbb{R}^d)$ with $|\partial_{x_i, x_j}^r p(t_k, t_N, x, z)| \leq C(t_N - t_k)^{-\frac{r}{2}} g(c(t_N - t_k), z - x) \leq CT^{-\frac{r}{2}} g(c(t_N - t_k), z - x)$,

$r = 1, 2$, so that

$$\begin{aligned}
 & \int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] p(t_k, t_N, z_k, z) dz_k \\
 &= \int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] [p(t_k, t_N, z_k, z) - p(t_k, t_N, z_{k-1}, z)] dz_k \\
 &= \left[\int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] (z_k - z_{k-1}) dz_k \right] \partial_x p(t_k, t_N, z_{k-1}, z) \\
 &+ \int_0^1 \int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] (z_k - z_{k-1})^* \\
 &\quad \times \partial_x^2 p(t_k, t_N, z_k^\lambda, z) (z_k - z_{k-1}) (1 - \lambda) dz_k d\lambda
 \end{aligned}$$

where we introduced the notation $z_k^\lambda = \lambda z_{k-1} + (1 - \lambda) z_k$. From (4.5.21) in the proof of Proposition 4.3.1 (note that $\kappa = 0$), one has

$$\begin{aligned}
 & \left| \int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] (z_k - z_{k-1}) dz_k \right| \\
 & \leq \frac{(C\delta)^L}{\Gamma(1 + (L - 1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}).
 \end{aligned}$$

We thus deduce

$$\begin{aligned}
 & \left| \left[\int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] (z_k - z_{k-1}) dz_k \right] \partial_x p(t_k, t_N, z_{k-1}, z) \right| \\
 & \leq CT^{-\frac{1}{2}} \frac{(C\delta)^L}{\Gamma(1 + (L - 1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) g(c(t_N - t_k), z - z_{k-1})
 \end{aligned}$$

By similar arguments, from (4.5.20) and the space-time inequality (4.1.2)

$$\begin{aligned}
 & \left| \int_0^1 \int_{\mathbb{R}^d} [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] \right. \\
 & \quad \times (z_k - z_{k-1})^* \partial_x^2 p(t_k, t_N, z_k^\lambda, z) (z_k - z_{k-1}) (1 - \lambda) dz_k d\lambda \left. \right| \\
 & \leq CT^{-1} \int_{\mathbb{R}^d} |p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)| |z_k - z_{k-1}|^2 \\
 & \quad \times [g(c(t_N - t_k), z - z_k) + g(c(t_N - t_k), z - z_{k-1})] dz_k \\
 & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L - 1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) g(c(t_N - t_k), z - z_{k-1}).
 \end{aligned}$$

From the semigroup property of the Gaussian transition density, we thus obtain

$$\begin{aligned}
 & \left| \int_{(\mathbb{R}^d)^3} p_L(0, t_{k-1}, z_0, z_{k-1}) [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] \right. \\
 & \quad \times p(t_k, t_N, z_k, z) dz_{k-1} dz_k \mu(dz_0) \left. \right| \\
 & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L - 1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^\infty, \mathbf{P}_k^{(0)}) \int_{\mathbb{R}^d} g(cT, z - z_0) \mu(dz) \tag{4.3.4}
 \end{aligned}$$

In the second case, that is, if $k \in \{k \in \{1, \dots, N\} : t_N - t_k < T/2\}$, we employ a duality argument that we now explain. For $h \in \mathcal{C}_b^2(\mathbb{R}^d)$, using the fact that $(s, z_{k-1}) \mapsto p(s, t_k, z_{k-1}, z_k)$ is the fundamental solution associated to the operator $\partial_s + \mathcal{L}_s = 0$ with the boundary condition $\lim_{s \uparrow t_k} p(s, t_k, \cdot, z_k) = \delta_{z_k}(\cdot)$ and then duality, we get

$$\begin{aligned} h(z_k) - \int_{\mathbb{R}^d} h(z_{k-1}) p(t_{k-1}, t_k, z_{k-1}, z_k) dz_{k-1} &= \int_{t_{k-1}}^{t_k} \partial_s \left[\int_{\mathbb{R}^d} h(z_{k-1}) p(s, t_k, z_{k-1}, z_k) dz_{k-1} \right] ds \\ &= \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} h(z_{k-1}) \mathcal{L}_s p(s, t_k, z_{k-1}, z_k) dz_{k-1} \\ &= \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} (\mathcal{L}_s^* h)(z_{k-1}) p(s, t_k, z_{k-1}, z_k) dz_{k-1} \end{aligned}$$

where \mathcal{L}_s^* stands for the dual operator of \mathcal{L}_s . In a completely similar way but for the map $(s, z_{k-1}) \mapsto p_L(s, t_k, z_{k-1}, z_k)$, one gets the identity

$$h(z_k) - \int_{\mathbb{R}^d} h(z_{k-1}) p_L(t_{k-1}, t_k, z_{k-1}, z_k) dz_{k-1} = \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} (\mathcal{L}_{s,L}^* h)(z_{k-1}) p_L(s, t_k, z_{k-1}, z_k) dz_{k-1}$$

where $\mathcal{L}_{s,L} h(x) = \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(s, x, \mathbf{P}_{k-1}^{(L-1)}(s)) \partial_{i,j}^2 h(x) + \sum_{i=1}^d b_i(s, x, \mathbf{P}_{k-1}^{(L-1)}(s)) \partial_i h(x)$ and $\mathcal{L}_{s,L}^*$ stands for its dual. If we take the difference of the two previous equalities, we get

$$\begin{aligned} &\int_{\mathbb{R}^d} h(z_{k-1}) (p - p_L)(t_{k-1}, t_k, z_{k-1}, z_k) dz_{k-1} \\ &= \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} (\mathcal{L}_{s,L}^* - \mathcal{L}_s^*) h(z_{k-1}) p_L(s, t_k, z_{k-1}, z_k) dz_{k-1} \\ &\quad + \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} \mathcal{L}_s^* h(z_{k-1}) (p_L - p)(s, t_k, z_{k-1}, z_k) dz_{k-1}. \end{aligned}$$

We then apply the previous identity to the map $h : z_{k-1} \mapsto p_L(0, t_{k-1}, z_0, z_{k-1})$. Combining **(HR)** with Proposition 4.3.1 and the inequalities $|\partial_{z_i, z_j}^r p_L(0, t_{k-1}, z_0, z)| \leq C t_{k-1}^{-\frac{r}{2}} g(ct_{k-1}, z - z_0) \leq C T^{-\frac{r}{2}} g(ct_{k-1}, z - z_0)$, $r = 0, 1, 2$, we get for $s \in [t_{k-1}, t_k]$

$$\begin{aligned} |(\mathcal{L}_{s,L}^* - \mathcal{L}_s^*) h(z_{k-1})| &\leq C t_{k-1}^{-1} \mathscr{W}(\mathbf{P}_{k-1}^{(L-1)}(s), \mathbf{P}(s)) g(ct_{k-1}, z_{k-1} - z_0) \\ &\leq C T^{-1} \mathscr{W}(\mathbf{P}_{k-1}^{(L-1)}(s), \mathbf{P}(s)) g(cs, z_{k-1} - z_0) \end{aligned}$$

where we used the fact that $k \geq 2$ for the last inequality. From the above computation, (4.3.1) and the semigroup property of Gaussian kernels, we thus deduce

$$\begin{aligned} &\left| \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} (\mathcal{L}_{s,L}^* - \mathcal{L}_s^*) h(z_{k-1}) p_L(s, t_k, z_{k-1}, z_k) dz_{k-1} \right| \\ &\leq C T^{-1} \int_{t_{k-1}}^{t_k} \mathscr{W}(\mathbf{P}_{k-1}^{(L-1)}(s), \mathbf{P}(s)) ds g(ct_k, z_k - z_0) \\ &\leq C T^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^{(1)}, \mathbf{P}_k^{(0)}) g(ct_k, z_k - z_0) \end{aligned}$$

Similarly, using the estimate $|\mathcal{L}_s^* h(z_{k-1})| \leq CT^{-1}g(cs, z_{k-1} - z_0)$ and the local estimate (4.3.2), one derives

$$\begin{aligned} & \left| \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} \mathcal{L}_s^* h(z_{k-1})(p_L - p)(s, t_k, z_{k-1}, z_k) dz_{k-1} ds \right| \\ & \leq CT^{-1} \int_{t_{k-1}}^{t_k} \int_{\mathbb{R}^d} g(cs, z_{k-1} - z_0) |(p_L - p)(s, t_k, z_{k-1}, z_k)| dz_{k-1} ds \\ & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N} \mathscr{W}_{k,k+1}(\mathbf{P}_k^{(1)}, \mathbf{P}_k^{(0)}) g(ct_k, z_k - z_0) \\ & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N} \mathscr{W}_{k,k+1}(\mathbf{P}_k^{(1)}, \mathbf{P}_k^{(0)}) g(ct_k, z_k - z_0). \end{aligned}$$

Gathering the two previous estimates, we conclude

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} p_L(0, t_{k-1}, z_0, z_{k-1})(p - p_L)(t_{k-1}, t_k, z_{k-1}, z_k) dz_{k-1} \right| \\ & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^{(1)}, \mathbf{P}_k^{(0)}) g(ct_k, z_k - z_0) \end{aligned}$$

which in turn implies

$$\begin{aligned} & \left| \int_{(\mathbb{R}^d)^3} p_L(0, t_{k-1}, z_0, z_{k-1}) [p(t_{k-1}, t_k, z_{k-1}, z_k) - p_L(t_{k-1}, t_k, z_{k-1}, z_k)] \right. \\ & \quad \left. \times p(t_k, t_N, z_k, z) dz_{k-1} dz_k \mu(dz_0) \right| \\ & \leq CT^{-1} \frac{(C\delta)^L}{\Gamma(1 + (L-1)/2)} \max_{0 \leq k \leq N-1} \mathscr{W}_{k,k+1}(\mathbf{P}_k^{(1)}, \mathbf{P}_k^{(0)}) \int_{\mathbb{R}^d} g(cT, z - z_0) \mu(dz) \end{aligned} \quad (4.3.5)$$

if $t_N - t_k < T/2$. Gathering (4.3.4) and (4.3.5), the estimate for the second difference is obtained.

4.4 Numerical results

In this section, we present some numerical simulations that illustrate the efficiency of our numerical scheme especially in terms of computational complexity comparing to classical particle system. Furthermore, numerical experiments indicate that the proposed probabilistic method works well even if the coefficients of the McKean-Vlasov SDE do not satisfy the previously stated regularity assumptions.

We are particularly interested in the mean squared error, or MSE for short

$$MSE^2 = \mathbb{E} [|\mathbb{E}[h(X_T)] - \langle h, \widehat{\mu}^L(T) \rangle|^2]$$

When the true density $p(0, t, \cdot)$, $t \in (0, T]$ is known, we consider the L^1 error of density $p_L(0, t, \cdot)$ obtained by probabilistic representation method

$$\mathbb{E} [|p_L(0, t, \cdot) - p(0, t, \cdot)|_{L^1(\mathbb{R}^d)}] = \mathbb{E} \left[\int_{\mathbb{R}^d} |p_L(0, t, z) - p(0, t, z)| dz \right]$$

For a measurable function h satisfying the condition stated in Theorem 4.2.1, i.e. $\forall z \in \mathbb{R}^d$, $|h(z)| \leq C \exp(\alpha z^2/2T)$ for some positive constants C, α , we have

$$\mathbb{E}[h(X_T^{(L)})] = \mathbb{E}\left[h(\bar{X}_{N_\delta^{L, N+1}}^L) \times \Gamma_N^L\right], \quad \langle h, \widehat{\mu}^L(T) \rangle = \frac{1}{M_L} \sum_{m=1}^{M_L} h(\bar{X}_{N_\delta^{L, N, (m)+1}}^L) \times \Gamma_N^{L, (m)}$$

Recall that $N_\delta^{L, k, (m)} = N_{t_k - t_{k-1}}^{L, (m)}$, $\Gamma_k^{\ell, (m)} = \prod_{j=1}^k \prod_{r=0}^{N_\delta^{\ell, j, (m)}} \theta_r^{\ell, j, (m)}$, $\Gamma_0^{\ell, (m)} = 1$ and

$$\begin{aligned} \theta_r^{\ell, j, (m)} &= \theta_{\zeta_r^{\ell, (m)}, \zeta_{r+1}^{\ell, (m)}}^{\ell}(\bar{X}_r^{\ell, (m)}, \bar{X}_{r+1}^{\ell, (m)}, \widehat{\mu}^{\ell-1}), \quad r = 0, \dots, N_\delta^{\ell, j, (m)} - 1, \\ \theta_{N_\delta^{\ell, j, (m)}}^{\ell, j, (m)} &= (1 - F(t - \zeta_{N_\delta^{\ell, j, (m)}}^{\ell, (m)}))^{-1}. \end{aligned}$$

For $\ell = 0$ and $t \in [t_k, t_{k+1}]$, the measure $\widehat{\mu}^0(t)(\cdot)$, together with its density $\widehat{p}_0(0, t, \cdot)$ are given by

$$\begin{aligned} \widehat{\mu}^0(t)(dz) &= \int_{\mathbb{R}^d} g(a(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k), z - b(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k) - \bar{X}_{N_\delta^{L, k, (m)}}^L) \widehat{\mu}_k^L(dx) dz \\ &= \frac{1}{M_L} \sum_{m=1}^{M_L} g(a(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k), z - b(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k) - \bar{X}_{N_\delta^{L, k, (m)}}^L) \times \Gamma_k^{L, (m)} dz \\ &= \widehat{p}_\ell(0, t, z) dz \end{aligned}$$

For $k = 1, \dots, N$, $t \in [t_k, t_{k+1}]$ and $\ell = 1, \dots, L$ we have

$$\begin{aligned} \widehat{\mu}^\ell(t)(dz) &= \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Gamma_k^{\ell, (m)} \prod_{r=0}^{N_{t-t_k}^{\ell, k+1, (m)}} \theta_r^{\ell, k+1, (m)} \delta_{\bar{X}_{N_{t-t_k+1}^{\ell, k+1, (m)}}^{\ell, k+1, (m)}}(dz) \\ \widehat{p}_\ell(0, t, z) &= \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} g(a(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k), z - b(\bar{X}_{N_\delta^{L, k, (m)}}^L, \widehat{\mu}_k^L)(t - t_k) - \bar{X}_{N_\delta^{L, k, (m)}}^L) \\ &\quad \times \Gamma_k^{\ell, (m)} \prod_{r=0}^{N_{t-t_k}^{\ell, k+1, (m)}} \theta_r^{\ell, k+1, (m)} \end{aligned}$$

Recall that from analysis in Section 3, it is optimal to take $L = 1$ and $M := M_1$ for our numerical tests.

4.4.1 Standard linear model

The drift is the expectation of the process at time t multiplied by a strictly negative factor $-\beta$ and the volatility is a constant σ

$$dX_t = -\beta \mathbb{E}[X_t] dt + \sigma dW_t, \quad X_0 = x_0 \in \mathbb{R}$$

By taking the expectation, the volatility term being bounded will vanish, the expectation $\mathbb{E}[X_t]$ is then known

$$\mathbb{E}[X_t] = x_0 \exp(-\beta t)$$

Table 4.1 – Linear model, beta sampling $\alpha = 0.5$, $\bar{\tau} = 2.0$, $M \propto N_t^{1.5}$, $h(x) = x$

M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
64000	1	0.312429	0.015596	0.265
181019	2	0.323830	0.011219	0.983
332553	3	0.326237	0.010517	2.043
512000	4	0.326753	0.010864	3.729
715541	5	0.326071	0.010748	6.036

$$T = 0.5, \sigma = 0.25, b = 0.4, X_0 = 0, 4, \mathbb{E}[h(X_T)] = 0.327492$$

Table 4.2 – Linear model, beta sampling $\alpha = 0.5$, $\bar{\tau} = 2.0$, $M \propto N_t^2$, $h(x) = x$

M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
64000	1	0.318420	0.0150587	0.243
256000	2	0.319011	0.00913713	1.253
576000	3	0.323656	0.00836158	3.567
1024000	4	0.323723	0.00744133	7.203
1600000	5	0.325007	0.00723796	13.25

$$T = 0.5, \sigma = 0.25, b = 0.4, X_0 = 0, 4, \mathbb{E}[h(X_T)] = 0.327492$$

The weight function is

$$\begin{aligned} \theta_{s,t}(x, y, \mathbf{P}) &= -(f(t-s))^{-1} b(y, \mathbf{P}(t)) H_1(a(x, \mathbf{P}(s))(t-s), y-x) \\ &= -(f(t-s))^{-1} \beta \int_{\mathbb{R}} z \mathbf{P}(t)(dz) \cdot \frac{(y-x)}{\sigma^2(t-s)} \end{aligned}$$

For $L = 1$, $M = M_1$, $t \in [t_k, t_{k+1}]$ and take $\mathbf{P} = \widehat{\mu}^0$ we have

$$\begin{aligned} \int_{\mathbb{R}} z \widehat{\mu}^0(t) dz &= \frac{1}{M} \sum_{m=1}^M (\bar{X}_{N_\delta^{1,k}+1}^{1,k,(m)} - \beta m_k^1(t-t_k)) \prod_{j=1}^k \left(\prod_{r=0}^{N_\delta^{1,j,(m)}} \theta_r^{1,j,(m)} \right) dz \\ m_k^1 &= \int_{\mathbb{R}} z \widehat{\mu}^1(t_k) = \frac{1}{M} \sum_{m=1}^M \bar{X}_{N_\delta^{1,k}+1}^{1,k,(m)} \prod_{j=1}^k \left(\prod_{r=0}^{N_\delta^{1,j,(m)}} \theta_r^{1,j,(m)} \right) \end{aligned}$$

The results are shown in Tables 4.1 to 4.4. From the discussions in Section 4.3.1 for the decomposition of errors, with the choice of $L = 1$, the bias corresponding to Picard iteration is of order $\delta = T/N_t$ and the statistical error of order $1/\sqrt{M}$ so M and N_t^2 should be of same order. To illustrate this fact, in Tables 4.1 and 4.2, the number of paths M in Monte-Carlo simulation is proportional to $N_t^{1.5}$ and N_t^2 respectively, with N_t the number of levels or time intervals. Tables 4.2 and 4.3 provide a comparison between beta sampling and exponential sampling and it is clear that beta sampling reduces the variance of the Monte-Carlo estimator.

The true density $p(0, t, \cdot)$ of the process $(X_t)_{t \geq 0}$ is also known, since it is a Gaussian density with mean $x_0 \exp(-\beta t)$ and variance $\sigma^2 t$. We compare the L^1 error of the density function $p_L(0, t, \cdot)$ obtained using probabilistic representation to the L^1 error

Table 4.3 – Linear model, exponential sampling $\lambda = 0.5$, $M \propto N_t^2$, $h(x) = x$

M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
64000	1	0.28518	0.0784653	0.196
256000	2	0.328188	0.0748352	1.53
576000	3	0.301085	0.055711	3.84
1024000	4	0.341605	0.0454096	8.43
1600000	5	0.360999	0.0469549	13.3

$$T = 0.5, \sigma = 0.25, b = 0.4, X_0 = 0, 4, \mathbb{E}[h(X_T)] = 0.327492$$

Table 4.4 – Linear model, beta sampling $\alpha = 0.5$, $\bar{\tau} = 2.0$, $M \propto N_t^2$, $h(x) = x$

M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
64000	1	0.233697	0.0343149	0.321
256000	2	0.241259	0.0236803	1.577
5760000	3	0.262055	0.022166	3.94
1024000	4	0.275469	0.0168063	8.026
1600000	5	0.260316	0.0160676	14.429

$$T = 1.0, \sigma = 0.25, b = 0.4, X_0 = 0, 4, \mathbb{E}[h(X_T)] = 0.327492$$

of the density obtained $\bar{p}(0, t, \cdot)$ using particle system $((\bar{X}_t^k)_{i=1}^n)_{t \geq 0}$ from [AKH+02]

$$\bar{p}(0, t, z) = \frac{1}{n} \sum_{k=1}^n g(\epsilon, \bar{X}_t^k - z)$$

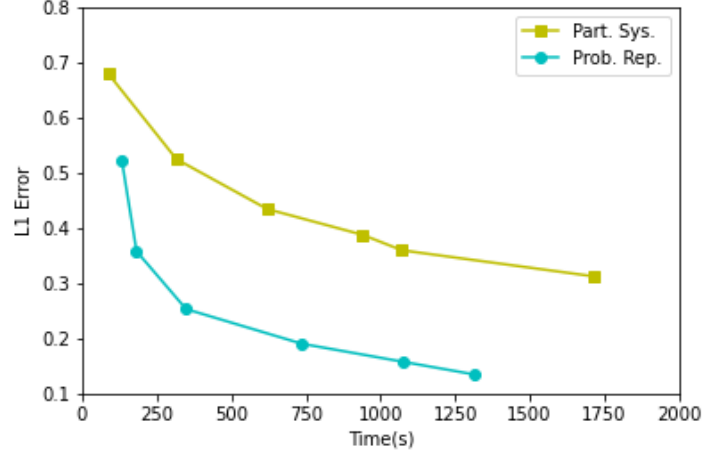
The L^1 error is estimated as

$$\mathbb{E}[|\bar{p}(0, t, \cdot) - p(0, t, \cdot)|_{L^1(\mathbb{R}^d)}] \leq C(\epsilon + \frac{1}{\sqrt{\epsilon}}(\frac{1}{\sqrt{n}} + \sqrt{h}))$$

With n the number of particles and h the size of time step, an easy optimization shows that in order to have an error of order ϵ , we need to have n of order ϵ^{-3} and h of order ϵ^3 . The complexity is n/h of order ϵ^{-6} , larger than the complexity ϵ^{-5} of probabilistic representation, which is confirmed by numerical results in Table 4.5.

Particle System				Probabilistic Representation			
Time(s)	L^1 error	$N_{part.}$	$N_{timestep}$	Time(s)	L^1 error	M	N_t
88.1	0.678167	22500	150	133.0	0.520676	200000	5
315.8	0.524177	40000	200	182.2	0.357589	400000	5
619.9	0.434062	62500	250	348.0	0.25326	800000	5
939.6	0.387467	78400	280	736.0	0.190606	1600000	5
1071.7	0.359487	90000	300	1077.8	0.157729	2400000	5
1714.5	0.312396	122500	350	1315.9	0.134896	3000000	5

Table 4.5 – Density comparison, $T = 0.5$, $\sigma = 0.15$, $b = 0.4$, $X_0 = 0, 4$

Figure 4.1 – Density comparison using L^1 error

4.4.2 Kuramoto model

Let us consider Kuramoto in which the drift has dependency on the measure through sine function:

$$dX_t = \beta \int_{\mathbb{R}} \sin(X_t - \xi) \mu_t^X(d\xi) dt + \sigma dW_t, \quad X_0 = x_0 \in \mathbb{R}.$$

We calculate the associated weight function $\theta_{s,t}(x, y, \mathbf{P})$

$$\begin{aligned} \theta_{s,t}(x, y, \mathbf{P}) &= -(f(t-s))^{-1} \left[(b(y, \mathbf{P}(t)) - b(x, \mathbf{P}(s))) H_1(a(x, \mathbf{P}(s))(t-s), y-x) + \partial_1 b(y, \mathbf{P}(t)) \right] \\ &= (f(t-s))^{-1} \beta \left[\left(\int_{\mathbb{R}} \sin(y-\xi) \mathbf{P}(t)(d\xi) - \int_{\mathbb{R}} \sin(y-\xi) \mathbf{P}(s)(d\xi) \right) \frac{y-x}{\sigma^2(t-s)} \right. \\ &\quad \left. - \int_{\mathbb{R}} \cos(y-\xi) \mathbf{P}(t)(d\xi) \right] \end{aligned}$$

Notice that we have the identity

$$\begin{aligned} \int_{\mathbb{R}} (\cos(y-z) + \mathbf{i} \sin(y-z)) g(c, z-m) dz &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi c}} \exp\left(\mathbf{i}(y-z) - \frac{(z-m)^2}{2c}\right) dz \\ &= \exp(\mathbf{i}(y-m)) \exp(-c/2) = (\cos(y-m) + \mathbf{i} \sin(y-m)) \exp(-c/2) \end{aligned}$$

Table 4.6 – Kuramoto model, beta sampling $\alpha = 0.7$, $\bar{\tau} = 1.0$

	M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
	6400	1	1.07534	0.0585993	0.017
$M \propto N_t$	12800	2	1.09519	0.0547775	37.96
	19200	3	1.09648	0.0575771	157.06
	25600	4	1.06784	0.0617268	342.52
$M \propto N_t^{1.5}$	18101	2	1.11311	0.0459845	85.72
	33255	3	1.11714	0.0439108	450.76
	51200	4	1.07849	0.0436546	1403.68
$M \propto N_t^2$	25600	2	1.0855	0.0384031	207.5
	57600	3	1.08744	0.0331515	1102
	102400	4	1.09767	0.0311312	5361.84

$$T = 0.5, \sigma = 0.25, b = 0.4, X_0 = 0, 4, h(x) = \sqrt{1+x^2}$$

Then with $L = 1$, $\mathbf{P} = \widehat{\mu}^0, \widehat{\mu}^1$ and $s, t \in [t_k, t_{k+1}]$, we calculate the following quantities needed for the expression of weight function $\theta_{s,t}(x, y, \mathbf{P})$:

$$\begin{aligned} \widehat{\mu}^0(t)(d\xi) &= \frac{1}{M} \sum_{m=1}^M g\left(\sigma^2(t-t_k), \xi - (\bar{X}_{N_k+1}^{1,k;m} + b(\bar{X}_{N_k+1}^{1,k;m}, \widehat{\mu}_k^1))(t-t_k)\right) \\ &\quad \times \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) d\xi \\ \int_{\mathbb{R}} \xi \widehat{\mu}^0(t)(d\xi) &= \frac{1}{M} \sum_{m=1}^M (\bar{X}_{N_k+1}^{1,k;m} + b(\bar{X}_{N_k+1}^{1,k;m}, \widehat{\mu}_k^1))(t-t_k) \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) \\ \int_{\mathbb{R}} \sin(y-\xi) \widehat{\mu}^0(t)(d\xi) &= \frac{1}{M} \sum_{m=1}^M \sin\left(y - (\bar{X}_{N_k+1}^{1,k;m} + b(\bar{X}_{N_k+1}^{1,k;m}, \widehat{\mu}_k^1))(t-t_k)\right) \\ &\quad \times \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) \times \exp(-\sigma^2(t-t_k)/2) \\ b(\bar{X}_{N_k+1}^{1,k;m}, \widehat{\mu}_k^1) &= \beta \int_{\mathbb{R}} \sin(\bar{X}_{N_k+1}^{1,k;m} - \xi) \widehat{\mu}_k^1(d\xi) \\ &= \beta \left(\sin(\bar{X}_{N_k+1}^{1,k;m}) \int_{\mathbb{R}} \cos(\xi) \widehat{\mu}_k^1(d\xi) - \cos(\bar{X}_{N_k+1}^{1,k;m}) \int_{\mathbb{R}} \sin(\xi) \widehat{\mu}_k^1(d\xi) \right) \end{aligned}$$

The results are shown in Table 4.6 with the number of Monte-Carlo simulations M proportional to N_t , $N_t^{1.5}$ and N_t^2 respectively.

4.4.3 Polynomial drift model

This model comes from [STT+19], the drift is a polynomial of X_t and its first and second order moments:

$$dX_t = \beta(2X_t + \mathbb{E}[X_t] - X_t \mathbb{E}[X_t^2])dt + \sigma X_t dW_t$$

Again by taking the expectation, we get a system of $\mathbb{E}[X_t]$ and $\mathbb{E}[X_t^2]$ that can be approximated for example using finite difference.

$$\begin{aligned} d\mathbb{E}[X_t] &= \beta\mathbb{E}[X_t](3 - \mathbb{E}[X_t^2])dt \\ d\mathbb{E}[X_t^2] &= ((4\beta + \sigma^2)\mathbb{E}[X_t^2] + 2\beta(\mathbb{E}[X_t]^2 - \mathbb{E}[X_t^2]^2))dt \end{aligned} \quad (4.4.1)$$

The associated weight function $\theta_{s,t}(x, y, \mathbf{P})$ is given as follows

$$\begin{aligned} a(x, \mathbf{P}) &= \sigma^2 x^2, \quad b(x, \mathbf{P}) = \beta(2x + \int_{\mathbb{R}} \xi \mathbf{P}(s)(d\xi) - x \int_{\mathbb{R}} \xi^2 \mathbf{P}(s)(d\xi)) \\ \partial_2 a(s, x, \mathbf{P}) &= 2\sigma^2 x, \quad \partial_{22}^2 a(s, x, \mathbf{P}) = 2\sigma^2, \quad \partial_2 b(s, x, \mathbf{P}) = \beta(2 - \int_{\mathbb{R}} \xi^2 P^{(\ell)}(s)(d\xi)) \\ \theta_{s,t}(x, y, \mathbf{P}) &= (f(t-s))^{-1} \left[\frac{1}{2} \partial_{22}^2 a(t, y, \mathbf{P}) + \partial_2 b(t, y, \mathbf{P}) \right. \\ &\quad \left. + (\partial_2 a(t, y, \mathbf{P}) - b(t, y, \mathbf{P})) \times H_1(a(s, x, \mathbf{P})(t-s), y-x) \right. \\ &\quad \left. + \frac{1}{2} (a(t, y, \mathbf{P}) - a(s, x, \mathbf{P})) \times H_2(a(s, x, \mathbf{P})(t-s), y-x) \right] \\ &= (f(t-s))^{-1} \left[\sigma^2 + \beta(2 - \int_{\mathbb{R}} \xi^2 \mathbf{P}(t)(d\xi)) + \frac{1}{2} (y^2 - x^2) \times \left(\frac{(y-x)^2}{\sigma^2 x^4 (t-s)^2} - \frac{1}{x^2 (t-s)} \right) \right. \\ &\quad \left. - (2\sigma^2 y - \beta(2y + \int_{\mathbb{R}} \xi \mathbf{P}(t)(d\xi) - y \int_{\mathbb{R}} \xi^2 \mathbf{P}(t)(d\xi))) \times \frac{(y-x)}{\sigma^2 x^2 (t-s)} \right] \end{aligned}$$

The measure $\widehat{\mu}^0(t)$ for $t \in [t_k, t_{k+1}]$, $k = 1, \dots, N-1$ is written as

$$\begin{aligned} \widehat{\mu}^0(t)(d\xi) &= \int_{\mathbb{R}} g(x^2(t-t_k), \xi - (2x + \int_{\mathbb{R}} z \widehat{\mu}_k^1(dz) - x \int_{\mathbb{R}} z^2 \widehat{\mu}_k^1(dz))(t-t_k) - x) \widehat{\mu}_k^1(dx) d\xi \\ &= \frac{1}{M} \sum_{m=1}^M g\left(\sigma^2 (\bar{X}_{N_{k+1}}^{1,k;m})^2 (t-t_k), \xi - (\bar{X}_{N_{k+1}}^{1,k;m} + \beta(2\bar{X}_{N_{k+1}}^{1,k;m} + m_k^2 - \bar{X}_{N_{k+1}}^{1,k;m} m_k^2)(t-t_k)) \right. \\ &\quad \left. \times \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) \right) d\xi \end{aligned}$$

Using the fact that $\widehat{\mu}^0(t_k) = \widehat{\mu}^1(t_k)$, the first and second order moments m_k^1 , m_k^2 at time t_k are given by

$$m_k^1 = \int_{\mathbb{R}} \xi \widehat{\mu}^0(t_k)(d\xi), \quad m_k^2 = \int_{\mathbb{R}} \xi^2 \widehat{\mu}^0(t_k)(d\xi)$$

Table 4.7 – Polynomial drift model, beta sampling $\alpha = 0.5$, $\bar{\tau} = 2.0$, $M \propto N_t^2$

M	N_t	$\langle h, \widehat{\mu}^L(T) \rangle$	CI	Time(s)
64000	1	0.663367	0.0509741	0.176
256000	2	0.675139	0.0443252	1.06
576000	3	0.668628	0.0238475	3.14
1024000	4	0.671675	0.0215759	6.63
1600000	5	0.671547	0.0199946	12.67

$$T = 0.5, \sigma = 0.25, b = 0.4, X_0 = 0, 4, h(x) = x$$

The first two moments corresponding to the measure $\widehat{\mu}^0(t)$, $t \in [t_k, t_{k+1}]$ are

$$\begin{aligned} \int_{\mathbb{R}} \xi \widehat{\mu}^0(t) d\xi &= \frac{1}{M} \sum_{m=1}^M (\bar{X}_{N_{k+1}}^{1,k;m} + \beta(2\bar{X}_{N_{k+1}}^{1,k;m} + m_k^1 - \bar{X}_{N_{k+1}}^{L,k;m} m_k^2)(t - t_k)) \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) \\ \int_{\mathbb{R}} \xi^2 \widehat{\mu}^0(t) d\xi &= \frac{1}{M} \sum_{m=1}^M (\bar{X}_{N_{k+1}}^{1,k;m} + \beta(2\bar{X}_{N_{k+1}}^{1,k;m} + m_k^1 - \bar{X}_{N_{k+1}}^{1,k;m} m_k^2)(t - t_k))^2 \\ &\quad + \sigma^2 (\bar{X}_{N_{k+1}}^{1,k;m})^2 (t - t_k) \prod_{k'=1}^k \left(\prod_{r=0}^{N_{k'}^{1;m}} \theta_r^{1,k';m} \right) \end{aligned}$$

The results can be found in Table 4.7 for M proportional to N_t^2 and we have the approximation of the expectation $\mathbb{E}[X_T] \approx 0.686183$ obtained by finite difference method applied to the system (4.4.1).

4.5 Appendix

4.5.1 Proof of Theorem 4.2.1

We here establish the probabilistic representation of Theorem 4.2.1 under **(HR)** and **(HE)**. We first establish the statement under the additional assumption that $h \in \mathcal{C}_b(\mathbb{R}^d)$ and then extend it to the current framework by an approximation argument. We introduce the time-inhomogeneous Markov semigroup with infinitesimal generator \mathcal{L}^ℓ associated to the decoupled SDE (4.2.5), defined for a measurable and bounded function h by $P_{s,t}^\ell h(x) = \int_{\mathbb{R}^d} h(z) p_\ell(s, t, x, z) dz = \mathbb{E} \left[h(X_t^{(\ell),s,x}) \right]$. Under the considered regularity assumptions on the maps $(t, x) \mapsto b(t, x, \mathbf{P}^{(\ell-1)}(t))$, $a(t, x, \mathbf{P}^{(\ell-1)}(t))$, we know that the mapping $[0, t] \times \mathbb{R}^d \ni (s, x) \mapsto p_\ell(s, t, x, z)$, for a fixed $(t, z) \in (0, T] \times \mathbb{R}^d$, belongs to $\mathcal{C}^2([0, t] \times \mathbb{R}^d)$ and satisfies the Backward Kolmogorov PDE: $(\partial_s + \mathcal{L}^\ell) p_\ell(s, t, x, z) = 0$ with the terminal condition $p_\ell(s, t, \cdot, z) \rightarrow \delta_z(\cdot)$ as $s \uparrow t$. Moreover, for any $\ell \geq 1$, the following estimates hold for $r = 0, 1, 2$

$$|\partial_x^r p_\ell(s, t, x, z)| \leq C t^{-\frac{r}{2}} g(c(t-s), z-x) \quad (4.5.1)$$

where $C := C(a, b, T, \lambda)$, $c := c(\lambda) > 0$ are two positive constants. We refer the reader to Friedman [Fri08] for more details.

The key idea to obtain the probabilistic representation consists in employing Itô's formula on an approximation process obtained from the dynamics (4.2.5) by removing the drift and freezing (w.r.t the time and space variables) the diffusion coefficient at the starting point. We denote by \bar{X}^ℓ this proxy process, that is, $\bar{X}_t^{\ell,s,x} = x + b(s, x, \mathbf{P}^{(\ell-1)}(s))(t-s) + \sigma(s, x, \mathbf{P}^{(\ell-1)}(s))(W_t - W_s)$. From Itô's formula, one obtains a one step expansion of the marginal law of $X^{\ell,s,x}$ around the marginal law of the Markov chain \bar{X}^ℓ . Then, one just needs to iterate the first step expansion in order to obtain a representation in infinite series for $P_{s,t}^\ell h(x)$. The final step consists in establishing a probabilistic representation for the resulting series. To simplify the notation in the computations below, for a function $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ and a $d \times d$ symmetric matrix A or a vector B of length d , we use the simplified notations $A\partial_x^2 f = \sum_{i,j=1}^d A_{i,j}\partial_{x_i x_j}^2 f$ and $B\partial_x f = \sum_{i=1}^d B_i \partial_{x_i} f$.

Let $\varepsilon > 0$. We thus starts by applying Itô's formula to $(P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}))_{r \in [s,t]}$. From the Backward Kolmogorov PDE satisfied by $(s, x) \mapsto p_\ell(s, t, x, z)$, we get

$$\begin{aligned}
& \mathbb{E}[P_{t,t+\varepsilon}^\ell h(\bar{X}_t^{\ell,s,x})] \\
&= P_{s,t+\varepsilon}^\ell h(x) + \mathbb{E}\left[\int_s^t \left(\partial_r P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) + b(s, x, \mathbf{P}^{(\ell-1)}(s))\partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) \right. \right. \\
&\quad \left. \left. + \frac{1}{2}a(s, x, \mathbf{P}^{(\ell-1)}(s))\partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) \right) dr\right] \\
&= P_{s,t}^\ell h(x) \\
&\quad + \mathbb{E}\left[\int_s^t \left(\left(b(x, \mathbf{P}^{(\ell-1)}(s)) - b(\bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(r)) \right) \partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) \right. \right. \\
&\quad \left. \left. + \frac{1}{2} \left(a(x, \mathbf{P}^{(\ell-1)}(s)) - a(\bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(r)) \right) \partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) \right) dr\right] \\
&= P_{s,t+\varepsilon}^\ell h(x) - \int_s^t \mathbb{E}\left[\left(b(\bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(r)) - b(x, \mathbf{P}^{(\ell-1)}(s)) \right) \partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right] dr \\
&\quad - \int_s^t \mathbb{E}\left[\frac{1}{2} \left(a(\bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(r)) - a(x, \mathbf{P}^{(\ell-1)}(s)) \right) \partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right] dr
\end{aligned}$$

where we used Fubini's theorem for the last but one equality, after observing that from (4.5.1), one has $|\partial_x^i P_{r,t+\varepsilon}^\ell h(\bar{X}_r^\ell)| \leq C|h|_\infty(t+\varepsilon-r)^{-\frac{i}{2}} \in L^1([s,t])$, for $i = 1, 2$. Remark that the regularization parameter ε allows to handle without any difficulty the time singularity induced by the spatial derivatives of the Markov semigroup $P_{s,t}^\ell$. We thus deduce

$$\begin{aligned}
P_{s,t+\varepsilon}^\ell h(x) &= \mathbb{E}[P_{t,t+\varepsilon}^\ell h(\bar{X}_t^{\ell,s,x})] \\
&\quad + \int_s^t \mathbb{E}\left[\left(b(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - b(x, \mathbf{P}^{(\ell-1)}(s)) \right) \partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right] dr \quad (4.5.2) \\
&\quad + \int_s^t \mathbb{E}\left[\frac{1}{2} \left(a(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - a(x, \mathbf{P}^{(\ell-1)}(s)) \right) \partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right] dr.
\end{aligned}$$

We then remove the spatial derivatives on the map $x \mapsto P_{s,t+\varepsilon}^\ell h(x)$ by applying the

integration by parts formula, that is

$$\begin{aligned}
 & \mathbb{E}\left[\frac{1}{2}\left(a(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - a(x, \mathbf{P}^{(\ell-1)}(s))\right)\partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^\ell)\right] \\
 &= \int_{\mathbb{R}^d} \frac{1}{2}\left(a(z, \mathbf{P}^{(\ell-1)}(r)) - a(x, \mathbf{P}^{(\ell-1)}(s))\right)\partial_x^2 P_{r,t+\varepsilon}^\ell h(z) g(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) dz \\
 &= \int_{\mathbb{R}^d} \frac{1}{2} \sum_{i,j=1}^d \partial_{z_i, z_j}^2 \left[\left(a_{i,j}(z, \mathbf{P}^{(\ell-1)}(r)) - a_{i,j}(x, \mathbf{P}^{(\ell-1)}(s)) \right) g(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) \right] P_{r,t+\varepsilon}^\ell h(z) dz \\
 &= \int_{\mathbb{R}^d} \frac{1}{2} \sum_{i,j=1}^d \left[\partial_{z_i, z_j}^2 a_{i,j}(z, \mathbf{P}^{(\ell-1)}(r)) + 2\partial_{z_i} a_{i,j}(z, \mathbf{P}^{(\ell-1)}(r)) H_1^i(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) \right. \\
 &\quad \left. + (a_{i,j}(z, \mathbf{P}^{(\ell-1)}(r)) - a_{i,j}(x, \mathbf{P}^{(\ell-1)}(s))) H_2^{i,j}(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) \right] \\
 &\quad \times g(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) P_{r,t+\varepsilon}^\ell h(z) dz
 \end{aligned}$$

and similarly,

$$\begin{aligned}
 & \mathbb{E}\left[\left(b(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - b(x, \mathbf{P}^{(\ell-1)}(s))\right)\partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right] \\
 &= \int_{\mathbb{R}^d} \sum_{i=1}^d \left[-\partial_{z_i} b_i(z, \mathbf{P}^{(\ell-1)}(r)) - (b_i(z, \mathbf{P}^{(\ell-1)}(r)) - b_i(x, \mathbf{P}^{(\ell-1)}(s))) H_1^i(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) \right] \\
 &\quad \times g(a(x, \mathbf{P}^{(\ell-1)}(s))(r-s), z-x) P_{r,t+\varepsilon}^\ell h(z) dz.
 \end{aligned}$$

Gathering the two previous expressions, we thus deduce

$$\begin{aligned}
 & \mathbb{E}\left[\left(\frac{1}{2}\left(a(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - a(x, \mathbf{P}^{(\ell-1)}(s))\right)\partial_x^2 P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x}) \right. \right. \\
 &\quad \left. \left. + \left(b(\bar{X}_r^\ell, \mathbf{P}^{(\ell-1)}(r)) - b(x, \mathbf{P}^{(\ell-1)}(s))\right)\partial_x P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})\right)\right] \\
 &= \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})]
 \end{aligned}$$

where we introduced the notation $\bar{\theta}_{s,r}(x, y, \mathbf{P}^{(\ell-1)}) := f(r-s)\theta_{s,r}(x, y, \mathbf{P}^{(\ell-1)})$.

We now quantify the time-degeneracy induced by the weight $\bar{\theta}_{s,r}$. Using **(HR)**, the chain rule formula on the space of probability measures $\mathcal{P}_2(\mathbb{R}^d)$ together with (4.2.5), we get

$$\begin{aligned}
 & b_i(x, \mathbf{P}^{(\ell-1)}(r)) - b_i(x, \mathbf{P}^{(\ell-1)}(s)) \\
 &= \int_s^r \int_{\mathbb{R}^d} \left\{ b(z, \mathbf{P}^{(\ell-2)}(u)) \partial_\mu b_i(x, \mathbf{P}^{(\ell-1)}(u))(z) + \frac{1}{2} a(z, \mathbf{P}^{(\ell-2)}(u)) \partial_z \partial_\mu b_i(x, \mathbf{P}^{(\ell-1)}(u))(z) \right\} \\
 &\quad du \mathbf{P}^{(\ell-1)}(dz)
 \end{aligned}$$

with the convention $\mathbf{P}^{(-1)}(u) = \mu$, which in turn, using the boundedness of the coefficients b , a and their derivatives, implies

$$|b_i(x, \mathbf{P}^{(\ell-1)}(r)) - b_i(x, \mathbf{P}^{(\ell-1)}(s))| \leq C_b(r-s), \quad C_b := |b|_\infty |\partial_\mu b|_\infty + \frac{1}{2} |a|_\infty |\partial_z \partial_\mu b|_\infty.$$

The previous upper-bound implies

$$|\rho_{s,r}^i(x, y, \mathbf{P}^{(\ell-1)})| \leq \left[|\partial_x b|_\infty + \lambda C_b |y - x| + \lambda \sup_\mu |b(\cdot, \mu)|_{Lip} \frac{|y - x|^2}{r - s} \right].$$

Similarly,

$$|a_{i,j}(x, \mathbf{P}^{(\ell-1)}(r) - a_{i,j}(x, \mathbf{P}^{(\ell-1)}(s))| \leq C_a(r - s), \quad C_a := |b|_\infty |\partial_\mu a|_\infty + \frac{1}{2} |a|_\infty |\partial_z \partial_\mu a|_\infty.$$

so that

$$\begin{aligned} |\kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{(\ell-1)})| &\leq \left[|\partial_x^2 a|_\infty + 2\lambda |\partial_x a|_\infty \frac{|y - x|}{r - s} + \lambda^2 \sup_\mu |a(\cdot, \mu)|_{Lip} \left\{ \frac{|y - x|^3}{(r - s)^2} + \frac{|y - x|}{r - s} \right\} \right. \\ &\quad \left. + \lambda^2 C_a(r - s) \left\{ \frac{|y - x|^2}{(r - s)^2} + \frac{1}{r - s} \right\} \right]. \end{aligned}$$

Gathering the previous estimates, we thus obtain

$$\begin{aligned} |\bar{\theta}_{s,r}(x, y, \mathbf{P}^{(\ell-1)})| &\leq d^2 \left\{ |\partial_x b|_\infty + |\partial_x^2 a|_\infty + \lambda^2 C_a + (\lambda C_b(r - s) \right. \\ &\quad \left. + 2\lambda |\partial_x a|_\infty + \lambda^2 \sup_\mu |a(\cdot, \mu)|_{Lip}) \frac{|y - x|}{r - s} \right. \\ &\quad \left. + (\lambda \sup_\mu |b(\cdot, \mu)|_{Lip} + \lambda^2 C_a) \frac{|y - x|^2}{r - s} + \lambda^2 \sup_\mu |a(\cdot, \mu)|_{Lip} \frac{|y - x|^3}{(r - s)^2} \right\}. \end{aligned} \quad (4.5.3)$$

Now, the space-time inequality (4.1.2) together with the previous bound yield

$$|\bar{\theta}_{s,r}(x, y, \mathbf{P}^{(\ell-1)})| g(a(x, \mathbf{P}^{(\ell-1)}(s))(r - s), y - x) \leq C(r - s)^{-\frac{1}{2}} g(c(r - s), y - x) \quad (4.5.4)$$

for some positive constant $C := C(T, a, b, \lambda)$, $c := c(\lambda)$, $T \mapsto C(T, a, b, \lambda)$ being non-decreasing. From the boundedness of h , we thus deduce that the map

$$r \mapsto \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})]$$

is in $L^1([s, t])$ so that coming back to (4.5.2) we conclude

$$P_{s,t+\varepsilon}^\ell h(x) = \mathbb{E}[P_{t,t+\varepsilon}^\ell h(\bar{X}_t^{\ell,s,x})] + \int_s^t \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})] dr.$$

We now pass to the limit as $\varepsilon \downarrow 0$ in the previous identity. From the continuity of $t \mapsto p_\ell(s, t, x, z)$ on $(s, T]$, one gets $\lim_{\varepsilon \downarrow 0} P_{s,t+\varepsilon}^\ell h(x) = P_{s,t}^\ell h(x)$ and $\lim_{\varepsilon \downarrow 0} P_{r,t+\varepsilon}^\ell(z) = P_{r,t}^\ell(z)$ so that, by dominated convergence, $\lim_{\varepsilon \downarrow 0} \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})] = \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t}^\ell h(\bar{X}_r^{\ell,s,x})]$. Therefore, from (4.5.4) and dominated convergence,

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \int_s^t \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t+\varepsilon}^\ell h(\bar{X}_r^{\ell,s,x})] dr \\ = \int_s^t \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t}^\ell h(\bar{X}_r^{\ell,s,x})] dr. \end{aligned}$$

Finally, from the continuity of h , one gets $\lim_{\varepsilon \downarrow 0} P_{t,t+\varepsilon}^\ell h(x) = h(x)$ so that again by the dominated convergence theorem, one has $\lim_{\varepsilon \downarrow 0} \mathbb{E}[P_{t,t+\varepsilon}^\ell h(\bar{X}_t^{\ell,s,x})] = \mathbb{E}[h(\bar{X}_t^{\ell,s,x})]$. We thus obtain the following first step expansion

$$P_{s,t}^\ell h(x) = \mathbb{E}[h(\bar{X}_t^{\ell,s,x})] + \int_s^t \mathbb{E}[\bar{\theta}_{s,r}(x, \bar{X}_r^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{r,t}^\ell h(\bar{X}_r^{\ell,s,x})] dr.$$

We now iterate the previous expansion n times by expanding the right-hand side of the previous expression. We derive

$$\begin{aligned} P_{s,t}^\ell h(x) &= \sum_{k=0}^n \int_{\Delta_k([s,t])} \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{k-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right] ds_k, \quad (4.5.5) \\ &+ \mathcal{R}_{s,t}^n h(x), \\ \mathcal{R}_{s,t}^n h(x) &:= \int_{\Delta_n([s,t])} \mathbb{E} \left[\prod_{j=0}^n \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{s_{n+1}, t}^\ell h(\bar{X}_{s_{n+1}}^{\ell,s,x}) \right] ds_{n+1} \end{aligned}$$

where we used the convention $s_0 = s$ and $\prod_{\emptyset} \dots = 1$. Observe that, in (4.5.5), we introduced the Euler scheme $(\bar{X}_{s_j}^{\ell,s,x})_{0 \leq j \leq k+1}$ on the time partition

$$\pi = \{s_0 = s \leq s_1 \leq s_2 \leq \dots \leq s_k \leq s_{k+1} = t\}$$

with transition $\bar{X}_{s_{j+1}}^{\ell,s,x} = \bar{X}_{s_j}^{\ell,s,x} + b(\bar{X}_{s_j}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(s_j))(s_{j+1} - s_j) + \sigma(\bar{X}_{s_j}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(s_j))(W_{s_{j+1}} - W_{s_j})$. We now pass to the limit as $n \uparrow \infty$ in (4.5.5). From similar arguments as those employed to derive (4.5.4), for any $h \in \mathcal{B}_b(\mathbb{R}^d)$, we get

$$|\mathbb{E}[h(\bar{X}_{s_{j+1}}^{\ell,s,x}) \bar{\theta}_{s_j, s_{j+1}}(x, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)})]| \leq C |h|_\infty (s_{j+1} - s_j)^{-\frac{1}{2}}$$

so that, from the Markov property satisfied by $(\bar{X}_{s_j}^{\ell,s,x})_{0 \leq j \leq k+1}$, by induction one gets

$$|\mathbb{E}[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{k-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)})]| \leq C^k |h|_\infty \prod_{j=0}^{k-1} (s_{j+1} - s_j)^{-\frac{1}{2}}$$

and similarly,

$$|\mathbb{E}[\prod_{j=0}^n \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) P_{s_{n+1}, t}^\ell h(\bar{X}_{s_{n+1}}^{\ell,s,x})]| \leq C^{n+1} |h|_\infty \prod_{j=0}^n (s_{j+1} - s_j)^{-\frac{1}{2}}.$$

Since $\int_{\Delta_k([s,t])} \prod_{j=0}^{k-1} (s_{j+1} - s_j)^{-\frac{1}{2}} ds_k = (t-s)^{\frac{k}{2}} \frac{\Gamma(\frac{1}{2})}{\Gamma(1+\frac{k}{2})}$, from the asymptotics of the Gamma function, we deduce that the series

$$\sum_{k=0}^n \int_{\Delta_k([s,t])} \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{k-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right] ds_k$$

converges absolutely and uniformly w.r.t $(s, t, x) \in \Delta_2(T) \times \mathbb{R}^d$ and that the remainder term $|\mathcal{R}_{s,t}^n h(x)|$ converges absolutely and uniformly w.r.t $(s, t, x) \in \Delta_2(T) \times \mathbb{R}^d$ to zero.

Hence, we obtain

$$\forall h \in \mathcal{C}_b(\mathbb{R}^d), \quad P_{s,t}^\ell h(x) = \sum_{n \geq 0} \int_{\Delta_n([s,t])} \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right] ds_n.$$

Finally, we provide a probabilistic representation of the time integrals in the above infinite series by employing the counting process N . To do that, it suffices to remark that from (5.2.3) with the convention $\zeta_0 = 0$

$$\begin{aligned} & \int_{\Delta_n([s,t])} \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right] ds_n \\ &= \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \frac{1}{1 - F(t - \zeta_n)} \prod_{j=0}^{n-1} \theta_{\zeta_j, \zeta_{j+1}}(\bar{X}_{\zeta_j}^{\ell,s,x}, \bar{X}_{\zeta_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \mathbf{1}_{\{N_{t-s}=n\}} \right] \end{aligned}$$

so that, summing over n from zero to infinity the previous identity yields

$$P_{s,t}^\ell h(x) = \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \frac{1}{1 - F(t - \zeta_{N_{t-s}})} \prod_{k=0}^{N_{t-s}-1} \theta_{\zeta_k, \zeta_{k+1}}(\bar{X}_{\zeta_k}^{\ell,s,x}, \bar{X}_{\zeta_{k+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right].$$

Taking $s = t_k$ and integrating both sides of the previous expression with respect to the initial distribution μ_k^ℓ allows to complete the proof of (4.2.9) for $h \in \mathcal{C}_b(\mathbb{R}^d)$. We are not so far from proving (4.2.10). Indeed, it suffices to remark that by Fubini's theorem

$$\begin{aligned} & \mathbb{E} \left[h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}) \right] = \int_{\mathbb{R}^d} h(z) \\ & \times \mathbb{E} \left[g(a(s_n, \bar{X}_{s_n}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(s_n))(t - s_n), z - \bar{X}_{\zeta_{N_t}}^{\ell,s,x}) \prod_{k=0}^{n-1} \bar{\theta}_{s_k, s_{k+1}}(\bar{X}_{s_k}^\ell, \bar{X}_{s_{k+1}}^\ell, \mathbf{P}^{(\ell-1)}(s_k)) \right] dz \end{aligned}$$

which can be justified using (4.5.4) to get

$$\begin{aligned} & \left| \mathbb{E} \left[g(a(t - s_n), z - \bar{X}_{s_n}) \prod_{k=0}^{n-1} \bar{\theta}_{s_k, s_{k+1}}(\bar{X}_{s_k}^\ell, \bar{X}_{s_{k+1}}^\ell, \mathbf{P}^{(\ell-1)}(s_k)) \right] \right| \\ & \leq C \int_{(\mathbb{R}^d)^n} g(c(t - s_n), z - z_n) \prod_{k=0}^{n-1} |\bar{\theta}_{s_k, s_{k+1}}(z_k, z_{k+1}, \mathbf{P}^{(\ell-1)}(s_k))| \\ & \quad \times g(a(s_k, z_k, \mathbf{P}^{(\ell-1)}(s_k))(s_{k+1} - s_k), z_{k+1} - z_k) dz_n \\ & \leq C^{n+1} \int_{(\mathbb{R}^d)^n} g(c(t - s_n), z - z_n) \prod_{k=0}^{n-1} (s_{k+1} - s_k)^{-\frac{1}{2}} g(c(s_{k+1} - s_k), z_{k+1} - z_k) dz_n \\ & \leq C^{n+1} \prod_{k=0}^{n-1} (s_{k+1} - s_k)^{-\frac{1}{2}} g(c(t - s), z - x) \end{aligned}$$

with the convention $s_0 = s$, $z_0 = x$ and where we used the semigroup property of Gaussian kernels for the last but one inequality. Hence, employing again the previous estimate and Fubini's theorem, we conclude

$$\begin{aligned}
& P_{s,t}^\ell h(x) \\
&= \sum_{n \geq 0} \int_{\Delta_n([s,t])} \mathbb{E} [h(\bar{X}_t^{\ell,s,x}) \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(\bar{X}_{s_j}^{\ell,s,x}, \bar{X}_{s_{j+1}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)})] d\mathbf{s}_n \\
&= \sum_{n \geq 0} \int_{\Delta_n([s,t])} \int_{\mathbb{R}^d} h(z) \mathbb{E} \left[g(a(\bar{X}_{s_n}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(s_n))(t - s_n), z - \bar{X}_{s_n}^{\ell,s,x}) \right. \\
&\quad \left. \times \prod_{k=0}^{n-1} \bar{\theta}_{s_k, s_{k+1}}(\bar{X}_{s_k}^\ell, \bar{X}_{s_{k+1}}^\ell, \mathbf{P}^{(\ell-1)}) \right] dz d\mathbf{s}_n \\
&= \int_{\mathbb{R}^d} h(z) \sum_{n \geq 0} \int_{\Delta_n([s,t])} \mathbb{E} \left[g(a(\bar{X}_{s_n}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(s_n))(t - s_n), z - \bar{X}_{s_n}^{\ell,s,x}) \right. \\
&\quad \left. \times \prod_{k=0}^{n-1} \bar{\theta}_{s_k, s_{k+1}}(\bar{X}_{s_k}^\ell, \bar{X}_{s_{k+1}}^\ell, \mathbf{P}^{(\ell-1)}) \right] d\mathbf{s}_n dz \\
&= \int_{\mathbb{R}^d} h(z) \mathbb{E} \left[g(a(\bar{X}_{\zeta_{N_t}}^{\ell,s,x}, \mathbf{P}^{(\ell-1)}(\zeta_{N_t}))(t - \zeta_{N_t}), z - \bar{X}_{\zeta_{N_t}}^{\ell,s,x}) \right. \\
&\quad \left. \times \frac{1}{1 - F(t - \zeta_{N_t})} \prod_{k=0}^{N_t-1} \theta_{\zeta_k, \zeta_{k+1}}(\bar{X}_{\zeta_k}^\ell, \bar{X}_{\zeta_{k+1}}^\ell, \mathbf{P}^{(\ell-1)}) \right] dz
\end{aligned}$$

for any $h \in \mathcal{C}_b(\mathbb{R}^d)$. Taking again $s = t_k$, integrating the previous equality w.r.t the initial distribution μ_k^ℓ and using the identity $\int_{\mathbb{R}^d} P_{s,t}^\ell h(x) \mu(dx) = \int_{\mathbb{R}^d} h(z) p_\ell(0, t, z) dz$ allow to complete the proof of (4.2.10). As a consequence, (4.2.9) holds for any $h \in \mathcal{B}_b(\mathbb{R}^d)$. It thus only remains to extend (4.2.9) for measurable h satisfying the growth assumption $|h(z)| \leq C \exp(\alpha \frac{|z|^2}{2T})$. This can be done by a standard approximation argument using a sequence of bounded measurable functions $(h_n)_{n \geq 1}$ satisfying: $h_n \rightarrow h$ a.e. and $|h_n(z)| \leq C \exp(\alpha \frac{|z|^2}{2T})$. This is where we use the condition $\int_{\mathbb{R}^d} \exp(\kappa|x|^2) \mu(dx) < \infty$ for some $\kappa := \kappa(\alpha, c)$ appears. The remaining technical details are omitted.

In the previous proof, one may replace without any additional difficulty the approximation process $\bar{X}_t^{\ell,s,x} = x + b(x, \mathbf{P}^{(\ell-1)}(s))(t - s) + \sigma(x, \mathbf{P}^{(\ell-1)}(s))(W_t - W_s)$ by the proxy $\bar{X}_t^{\ell,s,x} = x + \int_r^t b(x, \mathbf{P}^{(\ell-1)}(r)) dr + \int_s^t \sigma(x, \mathbf{P}^{(\ell-1)}(r)) dW_r$, as remarked in [Fri08] Chapter 1, Section 6. This will only change the definitions of the maps $\kappa_{s,t}^{i,j}(x, y, \mathbf{P})$ and $\rho_{s,t}^i(x, y, \mathbf{P})$ in the expression of $\theta_{s,t}$ as follows

$$\begin{aligned}
\kappa_{s,t}^{i,j}(x, y, \mathbf{P}) &:= \frac{1}{2} (a_{i,j}(y, \mathbf{P}(t)) - a_{i,j}(x, \mathbf{P}(t))) H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}(r)) dr, y - x \right) \\
&\quad + \partial_{x_i} a_{i,j}(y, \mathbf{P}(t)) H_1^i \left(\int_s^t a(x, \mathbf{P}(r)) dr, y - x \right) + \frac{1}{2} \partial_{x_i x_j}^2 a_{i,j}(y, \mathbf{P}(t)), \\
\rho_{s,t}^i(x, y, \mathbf{P}) &:= (b_i(y, \mathbf{P}(t)) - b_i(x, \mathbf{P}(t))) H_1^i \left(\int_s^t a(x, \mathbf{P}(r)) dr, y - x \right) + \partial_{x_i} b_i(y, \mathbf{P}(t)).
\end{aligned}$$

so that, from the Lipschitz regularity of the maps $x \mapsto a_{i,j}(x, \mathbf{P}(t)), b_i(x, \mathbf{P}(t))$, uniformly in the measure variable, we get

$$|\theta_{s,t}(x, y, \mathbf{P})| \leq C \left\{ 1 + \frac{|y-x|}{t-s} + \left\{ \frac{|y-x|^3}{(t-s)^2} + \frac{|y-x|}{t-s} \right\} \right\} \quad (4.5.6)$$

and (4.5.4) now becomes

$$|\theta_{s,t}(x, y, \mathbf{P})| g\left(\int_s^t a(x, \mathbf{P}(r)) dr, y-x\right) \leq C(r-s)^{-\frac{1}{2}} g(c(r-s), y-x).$$

The rest of the proof is similar and we omit the remaining technical details. We thus claim, see Theorem 15, Chapter 1 in [Fri08]:

Corollary 4.5.1. *Let $k \in \{0, \dots, n-1\}$. Under **(HR)** and **(HE)**, the random variable $X_t^{(\ell)}$, given by the unique weak solution to the SDE (4.2.5) taken at time $t \in (t_k, t_{k+1}]$ starting at time t_k from the initial distribution μ_k^ℓ , admits a density function $z \mapsto p_\ell(t_k, t, z) = \int_{\mathbb{R}^d} p_\ell(t_k, t, x, z) \mu_k^\ell(dx)$ with respect to the Lebesgue measure, $z \mapsto p_\ell(t_k, t, x, z)$ standing for the transition density function of the associated decoupled SDE which satisfies the following representation in infinite series*

$$\forall t \in (t_k, t_{k+1}], \quad p_\ell(t_k, t, x, z) := \sum_{n \geq 0} p_\ell^n(t_k, t, x, z)$$

where

$$\begin{aligned} p_\ell^n(s, t, x, z) &:= \int_{\Delta_n([s,t] \times (\mathbb{R}^d)^{n+1})} \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(z_j, z_{j+1}, \mathbf{P}^{(\ell-1)}) \\ &\quad \times g\left(\int_{s_j}^{s_{j+1}} a(z_j, \mathbf{P}^{(\ell-1)}(r)) dr, z_{j+1} - z_j\right) \\ &\quad \times g\left(\int_{s_n}^t a(z_n, \mathbf{P}^{(\ell-1)}(r)) dr, z - z_n\right) dz_n ds_n, \quad n \geq 1, \\ p_\ell^0(s, t, x, z) &:= g\left(\int_s^t a(x, \mathbf{P}^{(\ell-1)}(r)) dr, z - x\right) \end{aligned} \quad (4.5.7)$$

with the convention $z_0 = x, z_{n+1} = z, s_0 = s, s_{n+1} = t$ and $\prod_\emptyset = 1$.

The above corollary will play a key role in the proof of the convergence of the local Picard iteration schemes as presented in the next section.

4.5.2 Proof of Proposition 4.3.1

The first part of the proof relies on the Banach fixed point theorem applied to the map \mathcal{T}_k on the complete metric space $\mathcal{A}_{t_k, t_{k+1}, \mu}$ equipped with the total variation metric. We first introduce some notations. For two starting point given by two flows of probability measures $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{A}_{t_k, t_{k+1}, \mu}$, we consider the two corresponding sequences $\{\mathbf{P}^{\ell, i} \equiv \mathcal{F}^{(\ell)}(\mathbf{P}_i), \ell \geq 0\}$ and denote by $z \mapsto p_{\ell, i}(t_k, t, z) = \int_{\mathbb{R}^d} p_{\ell, i}(t_k, t, x, z) \mu(dx)$, for $i = 1, 2$ and $\ell \geq 1$, their respective density functions. Let us note that the maps

$z \mapsto p_{\ell,i}(t_k, t, x, z)$, for $i = 1, 2$ and $\ell \geq 1$, are the transition density functions of the corresponding decoupled SDEs.

From the very definition (4.5.7), writing $p_{\ell,i}^n(t_k, t, z) = \int_{\mathbb{R}^d} p_{\ell,i}^n(t_k, t, x, z) \mu(dx)$, we remark the following relation

$$\forall (s, t) \in \Delta_2(T), \quad p_{\ell,i}^{n+1}(s, t, x, z) = \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,i}) g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,i}(v)) dv, y - x\right) \times p_{\ell,i}^n(r, t, y, z) dy dr$$

which directly yields

$$\begin{aligned} & (p_{\ell,1}^{n+1} - p_{\ell,2}^{n+1})(s, t, x, z) \\ &= \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(v)) dr, y - x\right) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}(r)) \right. \\ & \quad \left. \times g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(v)) dv, y - x\right) \right] \times p_{\ell,1}^n(r, t, y, z) dy dr \\ & \quad + \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(v)) dv, y - x\right) [p_{\ell,1}^n(r, t, y, z) - p_{\ell,2}^n(r, t, y, z)] dy dr. \end{aligned}$$

Summing the previous equality from $n = 0$ to infinity, we thus get

$$\begin{aligned} & \sum_{n \geq 0} (p_{\ell,1}^{n+1} - p_{\ell,2}^{n+1})(s, t, x, z) \\ &= (p_{\ell,1} - p_{\ell,2})(s, t, x, z) - (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) \\ &= \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(v)) dv, y - x\right) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) \right. \\ & \quad \left. \times g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(v)) dv, y - x\right) \right] \times p_{\ell,1}(r, t, y, z) dy dr \\ & \quad + \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(v)) dr, y - x\right) [p_{\ell,1}(r, t, y, z) - p_{\ell,2}(r, t, y, z)] dy dr \\ &=: r_\ell^1(s, t, x, z) + r_\ell^2(s, t, x, z) \end{aligned}$$

or put in other words

$$(p_{\ell,1} - p_{\ell,2})(s, t, x, z) = (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) + r_\ell^1(s, t, x, z) + r_\ell^2(s, t, x, z).$$

We now give some estimates on the three terms appearing on the r.h.s. of the above identity. To this aim, we make use of the following decomposition

$$\begin{aligned} & r_\ell^1(s, t, x, z) \\ &= \int_s^t \int_{\mathbb{R}^d} \left\{ \sum_{i,j=1}^d [\kappa_{s,r}^{i,j}(x, y, \mathbf{P}^{\ell-1,1}) - \kappa_{s,r}^{i,j}(x, y, \mathbf{P}^{\ell-1,2})] - \sum_{i=1}^d [\rho_{s,r}^i(x, y, \mathbf{P}^{\ell-1,1}) - \rho_{s,r}^i(x, y, \mathbf{P}^{\ell-1,2})] \right\} \\ & \quad \times g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(v)) dv, y - x\right) p_{\ell,1}(r, t, y, z) dy dr \\ & \quad + \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) [g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(v)) dv, y - x\right) - g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(v)) dv, y - x\right)] \\ & \quad \times p_{\ell,1}(r, t, y, z) dy dr \\ &=: r_\ell^{1,1}(s, t, x, z) + r_\ell^{1,2}(s, t, x, z) \end{aligned}$$

with

$$\begin{aligned}
& \kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,1}) - \kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,2}) \\
&= \frac{1}{2} \left([a_{i,j}(y, \mathbf{P}^{\ell-1,1}(t)) - a_{i,j}(x, \mathbf{P}^{\ell-1,1}(t))] - [a_{i,j}(y, \mathbf{P}^{\ell-1,2}(t)) - a_{i,j}(x, \mathbf{P}^{\ell-1,2}(t))] \right) \\
& \quad \times H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) \\
& \quad + \frac{1}{2} \left[a_{i,j}(y, \mathbf{P}^{\ell-1,2}(t)) - a_{i,j}(x, \mathbf{P}^{\ell-1,2}(t)) \right] \\
& \quad \times \left[H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) - H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, y - x \right) \right] \\
& \quad + [\partial_{x_i} a_{i,j}(y, \mathbf{P}^{\ell-1,1}(t)) - \partial_{x_i} a_{i,j}(y, \mathbf{P}^{\ell-1,2}(t))] H_1^i \left(\int_s^t a(r, x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) \\
& \quad + \partial_{x_i} a_{i,j}(y, \mathbf{P}^{\ell-1,2}(t)) \left[H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) - H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, y - x \right) \right] \\
& \quad + \partial_{x_i x_j}^2 a_{i,j}(y, \mathbf{P}^{\ell-1,1}(t)) - \partial_{x_i x_j}^2 a_{i,j}(y, \mathbf{P}^{\ell-1,2}(t)) \\
&=: \text{A} + \text{B} + \text{C} + \text{D} + \text{E},
\end{aligned}$$

and

$$\begin{aligned}
& \rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,1}) - \rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,2}) \\
&= [b_i(y, \mathbf{P}^{\ell-1,1}(t)) H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) - b_i(y, \mathbf{P}^{\ell-1,2}(t)) \\
& \quad \times H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, y - x \right)] + [\partial_{x_i} b_i(y, \mathbf{P}^{\ell-1,1}(t)) - \partial_{x_i} b_i(y, \mathbf{P}^{\ell-1,2}(t))] \\
&=: \text{F} + \text{G}.
\end{aligned}$$

By the mean-value theorem, **(HR)**, i.e. taking the first time derivative w.r.t the variable x in both sides of (4.2.2), and using the definition of $\mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t))$, we obtain

$$\begin{aligned}
& \left| [a_{i,j}(t, y, \mathbf{P}^{\ell-1,1}(t)) - a_{i,j}(t, x, \mathbf{P}^{\ell-1,1}(t))] - [a_{i,j}(t, y, \mathbf{P}^{\ell-1,2}(t)) - a_{i,j}(t, x, \mathbf{P}^{\ell-1,2}(t))] \right| \\
&= \left| \int_0^1 [\partial_x a_{i,j}(\lambda x + (1-\lambda)y, \mathbf{P}^{\ell-1,1}(t)) - \partial_x a_{i,j}(\lambda x + (1-\lambda)y, \mathbf{P}^{\ell-1,2}(t))] \cdot (x - y) d\lambda \right| \\
&\leq C|x - y| \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t))
\end{aligned}$$

so that

$$|A| \leq C \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) \left\{ \frac{|y-x|^3}{(t-s)^2} + \frac{|y-x|}{t-s} \right\}.$$

By similar lines of reasoning, one gets

$$\begin{aligned}
& \left[H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y - x \right) - H_2^{i,j} \left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, y - x \right) \right] \\
&\leq C \left[\frac{|y-x|^2}{(t-s)^3} + \frac{1}{(t-s)^2} \right] \int_s^t \|a(x, \mathbf{P}^{\ell-1,1}(r)) - a(x, \mathbf{P}^{\ell-1,2}(r))\| dr \\
&\leq C \left[\frac{|y-x|^2}{(t-s)^3} + \frac{1}{(t-s)^2} \right] \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr
\end{aligned}$$

so that using the Lipschitz regularity of $x \mapsto a(x, m)$, uniformly in m , we obtain

$$|B| \leq C \left[\frac{|y-x|^3}{(t-s)^3} + \frac{|y-x|}{(t-s)^2} \right] \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr.$$

From **(HR)** and **(HE)**, i.e. taking again the first time derivative w.r.t the variable x in both sides of (4.2.2) and using the definition of $\mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t))$, we derive

$$|C| \leq C \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) \frac{|y-x|}{t-s}$$

and by the mean-value theorem,

$$\begin{aligned} |D| &\leq C \frac{|y-x|}{(t-s)^2} \int_s^t \|a(r, x, \mathbf{P}^{\ell-1,1}(r)) - a(r, x, \mathbf{P}^{\ell-1,2}(r))\| dr \\ &\leq C \frac{|y-x|}{(t-s)^2} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr. \end{aligned}$$

Finally, from **(HR)**, i.e. taking the second time derivative w.r.t the variable x in both sides of (4.2.2) and using the definition of $\mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t))$, we get

$$|E| \leq C \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)).$$

Gathering the two previous estimates, we get

$$\begin{aligned} &|\kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,1}) - \kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,2})| \\ &\leq C \left\{ \frac{|y-x|^3}{(t-s)^2} + \frac{|y-x|}{t-s} + 1 \right\} \\ &\quad \times \left\{ \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) + \frac{1}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \right\}. \end{aligned} \tag{4.5.8}$$

We treat F and G similarly to the previous terms. More precisely, from **(HR)**, i.e. using (4.2.1) we directly get

$$|b_i(y, \mathbf{P}^{\ell-1,1}(t)) - b_i(y, \mathbf{P}^{\ell-1,2}(t))| \leq C \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t))$$

and, by the mean value theorem, **(HR)** and **(HE)**,

$$\begin{aligned} &\left| H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, y-x \right) - H_1^i \left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, y-x \right) \right| \\ &\leq C \frac{|y-x|}{(t-s)^2} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \end{aligned}$$

so that

$$|F| \leq C \frac{|y-x|}{t-s} \left\{ \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) + \frac{1}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \right\}.$$

From **(HR)**, i.e. taking the first time derivative w.r.t. x in both sides of (4.2.1), we get

$$|G| \leq C \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)).$$

Gathering the previous estimates, we thus get

$$\begin{aligned} & |\rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,1}) - \rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,2})| \\ & \leq C \left\{ 1 + \frac{|y-x|}{t-s} \right\} \left\{ \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) + \frac{1}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \right\}. \end{aligned} \quad (4.5.9)$$

From (4.5.8) and (4.5.9), we finally obtain

$$\begin{aligned} & \left| \sum_{i,j=1}^d [\kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,1}) - \kappa_{s,t}^{i,j}(x, y, \mathbf{P}^{\ell-1,2})] - \sum_{i=1}^d [\rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,1}) - \rho_{s,t}^i(x, y, \mathbf{P}^{\ell-1,2})] \right| \\ & \leq C \left\{ \frac{|y-x|^3}{(t-s)^2} + \frac{|y-x|}{t-s} + 1 \right\} \left\{ \mathscr{W}(\mathbf{P}^{\ell-1,1}(t), \mathbf{P}^{\ell-1,2}(t)) + \frac{1}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \right\} \end{aligned} \quad (4.5.10)$$

which in turn, by the space-time inequality, the upper-estimate $p_{\ell,1}(r, t, y, z) \leq Cg(c(t-r), z-y)$, see [Fri08], the semigroup property of Gaussian kernels, yield

$$\begin{aligned} |r_\ell^{1,1}(s, t, x, y)| & \leq C \int_s^t \frac{1}{(r-s)^{\frac{1}{2}}} \left\{ \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) \right. \\ & \quad \left. + \frac{1}{r-s} \int_s^r \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du \right\} dr g(c(t-s), z-x) \\ & \leq C \int_s^t \frac{\mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r))}{(r-s)^{\frac{1}{2}}} dr g(c(t-s), z-x) \end{aligned}$$

where we used Fubini's theorem for the last inequality.

From the mean-value theorem and (HR), one gets

$$\begin{aligned} & \left| g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) - g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(u)) du, y-x\right) \right| \\ & \leq \frac{C}{r-s} \int_s^r \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du g(c(r-s), y-x) \end{aligned} \quad (4.5.11)$$

which combined with (4.5.6) and the upper-estimate $p_{\ell,1}(r, t, y, z) \leq Cg(c(t-r), z-y)$ yields

$$\begin{aligned} |r_\ell^{1,2}(s, t, x, y)| & \leq C \int_s^t \frac{1}{(r-s)^{\frac{3}{2}}} \int_s^r \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du dr g(c(t-s), z-x) \\ & \leq C \int_s^t \frac{\mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r))}{(r-s)^{\frac{1}{2}}} dr g(c(t-s), z-x) \end{aligned}$$

where we again used Fubini's theorem for the last inequality. Gathering the estimates, we get

$$|r_\ell^1(s, t, x, y)| \leq C \int_s^t \frac{\mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r))}{(r-s)^{\frac{1}{2}}} dr g(c(t-s), z-x).$$

Moreover, from the estimate (4.5.6) and the space-time inequality, one directly gets

$$|r_\ell^2(s, t, x, y)| \leq C \int_s^t \int_{\mathbb{R}^d} \frac{1}{(r-s)^{\frac{1}{2}}} g(c(r-s), y-x) |p_{\ell,1}(r, t, y, z) - p_{\ell,2}(r, t, y, z)| dy dr$$

and from the mean-value theorem, (HR) and (HE),

$$\begin{aligned} |p_{\ell,1}^0 - p_{\ell,2}^0|(s, t, x, z) &= \left| g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,1}(r)) dr, z-x\right) - g\left(\int_s^t a(x, \mathbf{P}^{\ell-1,2}(r)) dr, z-x\right) \right| \\ &\leq \frac{C}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr g(c(t-s), z-x). \end{aligned} \quad (4.5.12)$$

Gathering the above estimates, we finally get

$$\begin{aligned} |p_{\ell,1} - p_{\ell,2}|(s, t, x, z) &\leq \frac{C}{t-s} \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr g(c(t-s), z-x) \\ &\quad + C \int_s^t \frac{\mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r))}{(r-s)^{\frac{1}{2}}} dr g(c(t-s), z-x) \quad (4.5.13) \\ &\quad + C \int_s^t \int_{\mathbb{R}^d} \frac{1}{(r-s)^{\frac{1}{2}}} g(c(r-s), y-x) |p_{\ell,1} - p_{\ell,2}|(r, t, y, z) dy dr. \end{aligned}$$

Our aim now is to establish an estimate for the quantity $\mathscr{W}(\mathbf{P}^{\ell,1}(r), \mathbf{P}^{\ell,2}(r))$, for $r \in [s, t]$. For a function $h \in \mathcal{C}^2(\mathbb{R}^d)$ with derivatives of polynomial growth, namely $|\partial_x^\alpha h(x)| \leq C(1 + |x|^\kappa)$, for some $\kappa \geq 0$ and any $r \in \{0, 1, 2\}$, we thus investigate the quantity

$$\int_{\mathbb{R}^d} h(z) (p_{\ell,1} - p_{\ell,2})(s, t, x, z) dz = \int_{\mathbb{R}^d} [h(z) - h(x)] (p_{\ell,1} - p_{\ell,2})(s, t, x, z) dz$$

where we used the fact that $\int_{\mathbb{R}^d} h(x) (p_{\ell,1} - p_{\ell,2})(s, t, x, z) dz = h(x) - h(x) = 0$.

From a second order Taylor expansion, we get

$$\begin{aligned} &\int_{\mathbb{R}^d} [h(z) - h(x)] (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) dz \\ &= \int_0^1 (z-x)^* \partial_x^2 h(\lambda z + (1-\lambda)x) (z-x) (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) \lambda dz d\lambda \end{aligned}$$

where we used the fact that $\int_{\mathbb{R}^d} (z-x) (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) dz = 0$. Hence, from the previous computation, (4.5.12) and the space-time inequality (4.1.2), we thus get

$$\left| \int_{\mathbb{R}^d} [h(z) - h(x)] (p_{\ell,1}^0 - p_{\ell,2}^0)(s, t, x, z) dz \right| \leq C(1 + |x|^\kappa) \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr$$

We now deal with $\int_{\mathbb{R}^d} [h(z) - h(x)] r_\ell^1(s, t, x, z) dz$, namely

$$\begin{aligned}
& \int_{\mathbb{R}^d} [h(z) - h(x)] r_\ell^1(s, t, x, z) dz \\
&= \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}(r)) \right. \\
&\quad \left. \times g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(u)) du, y-x\right) \right] \times \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) dz [h(z) - h(x)] dy dr \\
&= \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}(r)) \right. \\
&\quad \left. \times g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(u)) du, y-x\right) \right] \times \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz dy dr
\end{aligned}$$

where we importantly observe that from the very definition of $\bar{\theta}_{s,r}$:

$$\int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,i}) g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,i}(r)) dr, y-x\right) dy = 0.$$

Let us also observe that the map $y \mapsto \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz$ has a first order derivative satisfying

$$\begin{aligned}
& \partial_y^2 \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz \\
&= \partial_y^2 \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) [h(z) - h(x_0)] dz + \partial_y^2 \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(x_0) dz \\
&= \int_{\mathbb{R}^d} \partial_y^2 p_{\ell,1}(r, t, y, z) [h(z) - h(x_0)] dz
\end{aligned}$$

for any $x_0 \in \mathbb{R}^d$. Thus, taking $x_0 = y$, using the Lipschitz regularity of h , the inequality $|\partial_y^2 p_{\ell,1}(r, t, y, z)| \leq C(t-r)^{-1} g(c(t-r), z-y)$ and finally the space-time inequality, we get

$$\left| \partial_y^2 \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz \right| \leq C(t-r)^{-1/2} (1 + |y|^\kappa).$$

Hence, using a second order Taylor expansion of the map $y \mapsto \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz$ around x , the fact that $\int_{\mathbb{R}^d} g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) \left[\int_{\mathbb{R}^d} \partial_y p_{\ell,1}(r, t, y, z) h(z) dz \right]_{|y=x} (y-x) dy = 0$, the inequality (4.5.10) and the space-time inequality, we obtain

$$\begin{aligned}
& \left| \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) \right] g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) \right. \\
&\quad \left. \times \left[\int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z) h(z) dz - \int_{\mathbb{R}^d} p_{\ell,1}(r, t, x, z) h(z) dz \right] dy dr \right| \\
&\leq C(1 + |x|^\kappa) \int_s^t (t-r)^{-\frac{1}{2}} \left[(r-s)^{\frac{1}{2}} \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) \right. \\
&\quad \left. + (r-s)^{-\frac{1}{2}} \int_s^r \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du \right] dr \\
&\leq C(1 + |x|^\kappa) \int_s^t \left[1 + (t-r)^{-\frac{1}{2}} (r-s)^{\frac{1}{2}} \right] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr.
\end{aligned}$$

Similarly, using a second order Taylor expansion of the map $y \mapsto \int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z)h(z) dz$ around x , the fact that $\int_{\mathbb{R}^d} g(\int_s^r a(x, \mathbf{P}^{\ell-1,i}(u)) du, y-x) [\int_{\mathbb{R}^d} \partial_y p_{\ell,1}(r, t, y, z)h(z) dz]_{y=x}(y-x)dy = 0$, $i \in \{1, 2\}$, the inequality (4.5.11) and finally the space-time inequality, we obtain

$$\begin{aligned} & \left| \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) \left[g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) \right. \right. \\ & \quad \left. \left. - g\left(\int_s^r a(x, \mathbf{P}^{\ell-1,2}(u)) du, y-x\right) \right] \right. \\ & \quad \left. \times \left[\int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z)h(z) dz - \int_{\mathbb{R}^d} p_{\ell,1}(r, t, x, z)h(z) dz \right] dy dr \right| \\ & \leq C(1 + |x|^\kappa) \int_s^t (r-s)^{-\frac{1}{2}}(t-r)^{-\frac{1}{2}} \int_s^r \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du \\ & \leq C(1 + |x|^\kappa) \int_s^t \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr. \end{aligned}$$

Gathering the above estimates, we deduce

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} [h(z) - h(x)] r_\ell^1(s, t, x, z) dz \right| \\ & = \left| \int_s^t \int_{\mathbb{R}^d} \left[\bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,1}) g\left(\int_s^r a(u, x, \mathbf{P}^{\ell-1,1}(u)) du, y-x\right) - \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}(r)) \right. \right. \\ & \quad \left. \left. \times g\left(\int_s^r a(u, x, \mathbf{P}^{\ell-1,2}(u)) du, y-x\right) \right] \left[\int_{\mathbb{R}^d} p_{\ell,1}(r, t, y, z)h(z) dz - \int_{\mathbb{R}^d} p_{\ell,1}(r, t, x, z)h(z) dz \right] dy dr \right| \\ & \leq C(1 + |x|^\kappa) \int_s^t [1 + (t-r)^{-\frac{1}{2}}(r-s)^{\frac{1}{2}}] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr. \end{aligned}$$

Finally, integrating the map $z \mapsto h(z) - h(x)$ against the kernel $z \mapsto r_\ell^2(s, t, x, z)$, by Fubini's theorem, we get

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} [h(z) - h(x)] r_\ell^2(s, t, x, z) dz \right| \\ & = \left| \int_s^t \int_{\mathbb{R}^d} \bar{\theta}_{s,r}(x, y, \mathbf{P}^{\ell-1,2}) g\left(\int_s^r a(r, x, \mathbf{P}^{\ell-1,2}(r)) dr, y-x\right) \right. \\ & \quad \left. \times \int_{\mathbb{R}^d} [p_{\ell,1}(r, t, y, z) - p_{\ell,2}(r, t, y, z)] h(z) dz dy dr \right| \\ & \leq C \int_s^t \int_{\mathbb{R}^d} \frac{1}{(r-s)^{\frac{1}{2}}} g(c(r-s), y-x) \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(r, t, y, z) - p_{\ell,2}(r, t, y, z)] dz \right| dy dr \end{aligned}$$

so that, combining the previous estimates, we get

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell,2}(s, t, x, z)] dz \right| \\ & \leq C(1 + |x|^\kappa) \int_s^t [1 + (t-r)^{-\frac{1}{2}}(r-s)^{\frac{1}{2}}] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \quad (4.5.14) \\ & + C \int_s^t \int_{\mathbb{R}^d} \frac{1}{(r-s)^{\frac{1}{2}}} g(c(r-s), y-x) \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(r, t, y, z) - p_{\ell,2}(r, t, y, z)] dz \right| dy dr. \end{aligned}$$

For $0 \leq s_0 < s$ and $x_0 \in \mathbb{R}^d$ fixed, we introduce the non-negative, bounded and measurable map $f(\cdot, t) : (s_0, t] \ni s \mapsto \int_{\mathbb{R}^d} g(c(s - s_0), x - x_0) \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell,2}(s, t, x, z)] dz \right| dx$. Integrating both sides of the previous bound w.r.t $g(c(s - s_0), x - x_0) dx$ yields

$$f(s, t) \leq C(1 + |x_0|^\kappa) \int_s^t (t - r)^{-\frac{1}{2}} \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr + C \int_s^t (r - s)^{-\frac{1}{2}} f(r, t) dr.$$

Iterating the previous inequality, we obtain

$$f(s, t) \leq C(1 + |x_0|^\kappa) \int_s^t (t - r)^{-\frac{1}{2}} \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr$$

so that, coming back to (4.5.14), yields

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell,2}(s, t, x, z)] dz \right| \\ & \leq C(1 + |x|^\kappa) \left[\int_s^t [1 + (t - r)^{-\frac{1}{2}} (r - s)^{\frac{1}{2}}] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \right. \\ & \quad \left. + \int_s^t (r - s)^{-\frac{1}{2}} \int_r^t (t - u)^{-\frac{1}{2}} \mathscr{W}(\mathbf{P}^{\ell-1,1}(u), \mathbf{P}^{\ell-1,2}(u)) du dr. \right] \\ & \leq C(1 + |x|^\kappa) \int_s^t [1 + (t - r)^{-\frac{1}{2}} (r - s)^{\frac{1}{2}}] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \quad (4.5.15) \end{aligned}$$

for any $h \in \mathcal{C}^2(\mathbb{R}^d)$ with derivatives of polynomial growth of order $\kappa \geq 0$.

From the previous inequality, it directly follows

$$\begin{aligned} \mathscr{W}(\mathbf{P}^{\ell,1}(t), \mathbf{P}^{\ell,2}(t)) &= \sup_{h \in \mathcal{C}_b^2(\mathbb{R}^d)} \left| \int_{(\mathbb{R}^d)^2} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell,2}(s, t, x, z)] dz \mu(dx) \right| \\ &\leq C \int_s^t [1 + (t - r)^{-\frac{1}{2}} (r - s)^{\frac{1}{2}}] \mathscr{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) dr \end{aligned}$$

so that, by an induction that we omit, we get

$$\mathscr{W}(\mathbf{P}^{\ell,1}(t), \mathbf{P}^{\ell,2}(t)) \leq \frac{(C(t - s))^\ell}{\Gamma(1 + \ell/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}_1(r), \mathbf{P}_2(r)).$$

Plugging the previous estimate into (4.5.13) and then iterating, for any positive integer ℓ we obtain

$$\begin{aligned} & |p_{\ell,1} - p_{\ell,2}|(s, t, x, z) \\ & \leq \frac{(C(t - s))^{\ell-1}}{\Gamma(1 + (\ell - 1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}_1(r), \mathbf{P}_2(r)) g(c(t - s), z - x). \quad (4.5.16) \end{aligned}$$

and (4.5.15) now becomes

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell,2}(s, t, x, z)] dz \right| \\ & \leq C(1 + |x|^\kappa) \frac{(C(t - s))^\ell}{\Gamma(1 + (\ell - 1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}_1(r), \mathbf{P}_2(r)). \quad (4.5.17) \end{aligned}$$

We finally deduce

$$\begin{aligned} d_{s,t}(\mathbf{P}^{\ell,1}, \mathbf{P}^{\ell,2}) &\leq \sup_{s \leq r \leq t} \int_{(\mathbb{R}^d)^2} |p_{\ell,1} - p_{\ell,2}|(s, r, x, z) dz \mu(dx) \\ &\leq \frac{(C(t-s))^{\ell-1}}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}_1(r), \mathbf{P}_2(r)) \end{aligned} \quad (4.5.18)$$

and, since $\sum_{\ell \geq 1} \frac{(C(t-s))^{\ell-1}}{\Gamma(1 + (\ell-1)/2)} < \infty$, by the Banach fixed point theorem, one deduces that \mathcal{T} admits a unique fixed point \mathbf{P}_∞ in $\mathcal{A}_{s,t,\mu}$ for any fixed $(s, t) \in \Delta_2(T)$ for all $T > 0$, which is the unique solution to the martingale problem associated to the McKean-Vlasov equation (4.1.1) on the interval $[s, t]$. This concludes the proof of the first part of our proposition. Observe now that by taking $\mathbf{P}_2 = \mathcal{T}(\mathbf{P}_1)$, so that $\mathbf{P}^{\ell,2} = \mathbf{P}^{\ell+1,1}$, by (4.5.18) one gets

$$d_{s,t}(\mathbf{P}^{\ell,1}, \mathbf{P}^{\ell+1,1}) \leq \frac{(C(t-s))^{\ell-1}}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}^{1,1}(r), \mathbf{P}_1(r))$$

so that, for any positive integer m

$$d_{s,t}(\mathbf{P}^{\ell,1}, \mathbf{P}^{\ell+m,1}) \leq \sum_{n=\ell}^{m-1} d_{s,t}(\mathbf{P}^{n,1}, \mathbf{P}^{n+1,1}) \leq \sum_{n \geq \ell} \frac{(C(t-s))^{n-1}}{\Gamma(1 + (n-1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}^{1,1}(r), \mathbf{P}_1(r))$$

and, since the sequence $(\mathbf{P}^{\ell+m,1})_{m \geq 0}$ converges w.r.t the distance d to the unique fixed point \mathbf{P}_∞ , passing to the limit as $m \uparrow \infty$

$$d_{s,t}(\mathbf{P}^{\ell,1}, \mathbf{P}_\infty) \leq C \frac{(C(t-s))^{\ell-1}}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}^{1,1}(r), \mathbf{P}^1(r)).$$

Similarly, one gets

$$\mathscr{W}(\mathbf{P}^{\ell,1}, \mathbf{P}_\infty) \leq C \frac{(C(t-s))^\ell}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathscr{W}(\mathbf{P}^{1,1}(r), \mathbf{P}^1(r)). \quad (4.5.19)$$

From the continuity of the maps $\mathcal{A}_{s,t,\mu} \ni \mathbf{P} \mapsto \bar{\theta}_{s,t}(x, z, \mathbf{P})$, $g(\int_s^t a(r, x, \mathbf{P}(r)) dr, z - x)$, the estimate (4.5.1), by the dominated convergence theorem, one may pass to the limit as $\ell \uparrow \infty$ in (4.5.7) of Corollary 4.5.1 to deduce that the sequence $(p_\ell^n(s, t, x, z))_{\ell \geq 1}$ converges to $p^n(s, t, x, z)$, for each integer n , given by

$$\begin{aligned} p^n(s, t, x, z) &:= \int_{\Delta_n([s,t] \times (\mathbb{R}^d)^{n+1})} \prod_{j=0}^{n-1} \bar{\theta}_{s_j, s_{j+1}}(z_j, z_{j+1}, \mathbf{P}_\infty) \\ &\quad \times g\left(\int_{s_j}^{s_{j+1}} a(r, z_j, \mathbf{P}_\infty(r)) dr, z_{j+1} - z_j\right) \\ &\quad \times g\left(\int_{s_n}^t a(r, z_n, \mathbf{P}_\infty(r)) dr, z - z_n\right) dz_n ds_n, \quad n \geq 1 \\ p^0(s, t, x, z) &:= g\left(\int_s^t a(r, x, \mathbf{P}_\infty(r)) dr, z - x\right). \end{aligned}$$

Therefore, from the representation in infinite series of Corollary 4.5.1 and again the dominated convergence theorem, the sequence $(p_\ell(s, t, x, z))_{\ell \geq 1}$ converges to $p(s, t, x, z)$ given by

$$p(s, t, x, z) = \sum_{n \geq 0} p^n(s, t, x, z).$$

Observe now that taking $\mathbf{P}_2 = \mathcal{F}^{(m)}(\mathbf{P}_1)$ for a positive integer m so that $\mathbf{P}^{\ell,2} = \mathcal{F}^{(\ell+m)}(\mathbf{P}_1) = \mathbf{P}^{\ell+m,1}$ and $p_{\ell,2}(s, t, x, z) = p_{\ell+m,1}(s, t, x, z)$, one gets $\mathcal{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell-1,2}(r)) = \mathcal{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}^{\ell+m-1,1}(r)) \rightarrow \mathcal{W}(\mathbf{P}^{\ell-1,1}(r), \mathbf{P}_\infty(r))$ as $m \uparrow \infty$. Combining the previous observation with (4.5.16) and (4.5.19), we get

$$\begin{aligned} & |p_{\ell,1}(s, t, x, z) - p(s, t, x, z)| \\ &= \lim_{m \uparrow \infty} |p_{\ell,1}(s, t, x, z) - p_{\ell+m,1}(s, t, x, z)| \\ &\leq \frac{(C(t-s))^{\ell-1}}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathcal{W}(\mathbf{P}_1(r), \mathbf{P}_\infty(r)) g(c(t-s), z-x). \end{aligned} \quad (4.5.20)$$

and from (4.5.17)

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p(s, t, x, z)] dz \right| \\ &= \lim_{m \uparrow \infty} \left| \int_{\mathbb{R}^d} h(z) [p_{\ell,1}(s, t, x, z) - p_{\ell+m,1}(s, t, x, z)] dz \right| \\ &\leq C(1 + |x|^\kappa) \frac{(C(t-s))^\ell}{\Gamma(1 + (\ell-1)/2)} \sup_{s \leq r \leq t} \mathcal{W}(\mathbf{P}_1(r), \mathbf{P}_\infty(r)) \end{aligned} \quad (4.5.21)$$

for any $h \in \mathcal{C}^2(\mathbb{R}^d)$ with derivatives of polynomial growth of order $\kappa \geq 0$. This completes the proof.

Part III

Probabilistic Representation and Integration by Parts Formulae for some Stochastic Volatility Models

Probabilistic Representation and Integration by Parts Formulae for some Stochastic Volatility Models

5.1 Introduction

In this work, we consider a two dimensional stochastic volatility model given by the solution of the following stochastic differential equation (SDE for short) with dynamics

$$\begin{cases} S_t &= s_0 + \int_0^t r S_s ds + \int_0^t \sigma_S(Y_s) S_s dW_s, \\ Y_t &= y_0 + \int_0^t b_Y(Y_s) ds + \int_0^t \sigma_Y(Y_s) dB_s, \\ d\langle B, W \rangle_s &= \rho ds \end{cases} \quad (5.1.1)$$

where the coefficients $b_Y, \sigma_S, \sigma_Y : \mathbb{R} \rightarrow \mathbb{R}$ are smooth functions, W and B are one-dimensional standard Brownian motions with correlation factor $\rho \in (-1, 1)$ both being defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The aim of this chapter is to prove a probabilistic representation formula for two integration by parts (IBP) formulae for the marginal law of the process (S, Y) at a given time maturity T . To be more specific, for a given starting point $(s_0, y_0) \in (0, \infty) \times \mathbb{R}$ and a given finite time horizon $T > 0$, we establish two Bismut-Elworthy-Li (BEL) type formulae for the two following quantities

$$\partial_{s_0} \mathbb{E}[h(S_T, Y_T)] \quad \text{and} \quad \partial_{y_0} \mathbb{E}[h(S_T, Y_T)] \quad (5.1.2)$$

where h is a real-valued possibly non-smooth payoff function defined on $[0, \infty) \times \mathbb{R}$.

Such IBP formulae have attracted a lot of interest during the last decades both from a theoretical and a practical point of views as they can be further analyzed to derive properties related to the transition density of the underlying process or to develop Monte Carlo simulation algorithm among other practical applications, see e.g. Nualart [Nua06], Malliavin and Thalmaier [MT06] and the references therein. They are also of major interest for computing sensitivities, also referred as to *Greeks* in finance, of arbitrage price of financial derivatives which is the keystone for hedging purpose, i.e. for protecting the value of a portfolio against some possible changes in sources of risk.

The two quantities appearing in (5.1.2) corresponds respectively to the Delta and Vega of the European option with payoff $h(S_T, Y_T)$. For a more detailed discussion on this topic, we refer the interested reader to Fournié and al. [Fou+99],[Fou+01] for IBP formulae related to European, Asian options and conditional expectations, Gobet and al.[GKH+03], [BGKH03] for IBP formulae related to some barrier or lookback options. Let us importantly point out that, from a numerical point of view, the aforementioned IBP formulae will inevitably involve a time discretization procedure of the underlying process and Malliavin weights, thus introducing two sources of error given by a bias and a statistical error, as it is already the case for the computation of the price $\mathbb{E}[h(S_T, Y_T)]$.

Relying on a perturbation argument for the Markov semigroup generated by the couple (X, Y) , we first establish a probabilistic representation formula for the marginal law (S_T, Y_T) for a fixed prescribed maturity $T > 0$ based on a simple Markov chain evolving along a random time grid given by the jump times of an independent renewal process. Such probabilistic representation formula was first derived in Bally and Kohatsu-Higa [BKH+15] for the marginal law of a multi-dimensional diffusion process and of some Lévy driven SDEs with bounded drift, diffusion and jump coefficients. Still in the case of bounded coefficients, it was then further investigated in Labordère and al. [HLTT+17], Agarwal and Gobet [AG17] for multi-dimensional diffusion processes and in Frikha and al. [FKHL+19] for one-dimensional killed processes. The major advantage of the aforementioned probabilistic formulae lies in the fact that an unbiased Monte Carlo simulation method directly stems from it. Thus, it may be used to numerically compute an option price with optimal complexity since its computation will be only affected by the statistical error. However, let us emphasize that in general the variance of the Monte Carlo estimator tends to be large or even infinite. In order to circumvent this issue, an importance sampling scheme based on the law of the jump times of the underlying renewal process has been proposed in Anderson and Kohatsu-Higa [AKH+17] in the multi-dimensional diffusion framework and in [FKHL+19] for one-dimensional killed processes.

The main novelty of our approach in comparison with the aforementioned works is that we allow the drift coefficient b_Y to be possibly unbounded as it is the case in most stochastic volatility models (Stein-Stein, Heston, ...). Such boundedness condition on the drift coefficient has appeared persistently in the previous contributions and is actually essential since basically it allows to remove the drift in the choice of the approximation process in order to derive the probabilistic representation formula. The key ingredient that we here develop in order to remove this restriction consists in choosing adequately the approximation process around which the original perturbation argument of the Markov semigroup (X, Y) is done by taking into account the transport of the initial condition by the deterministic ordinary differential equation (ODE) having unbounded coefficient¹. The approximation process, or equivalently the underlying Markov chain on which the probabilistic representation is based, is then obtained from the original dynamics (5.1.1) by *freezing the coefficients* b_Y , σ_S and σ_Y along the flow of this ODE. We stress that the previous choice is here crucial since it provides the adequate approximation process on which some good controls can be established. To the best of our knowledge, this feature appears to be new in this context.

¹This dynamical system is obtained by removing the noise, that is, by setting $\sigma_Y \equiv 0$, from the dynamics of Y in (5.1.1).

Having this probabilistic representation formula at hand together with the tailor-made Malliavin calculus machinery for this well-chosen underlying Markov chain, in the spirit of the BEL formula established in [FKHL+19] for killed diffusion processes with bounded drift coefficient, we rely on a propagation of the spatial derivatives forward in time then perform local IBP formulas on each time interval of the random time grid and finally merge them in a suitable manner in order to establish the two BEL formulae for the two quantities (5.1.2). Following the ideas developed in [AKH+17], we achieve finite variance for the Monte Carlo estimators obtained from the probabilistic representation formulas of the couple (S_T, Y_T) and of both IBP formulae by selecting adequately the law of the jump times of the renewal process. We finally provide some numerical tests illustrating our previous analysis.

The chapter is organized as follows. In Section 5.2, we introduce our assumptions on the coefficients, present the approximation process that will be the main building block for our perturbation argument as well as the Markov chain that will play a central role in our probabilistic representation for the marginal law of the process (X, Y) and for our IBP formulae. In addition, we construct the tailor-made Malliavin calculus machinery related to the underlying Markov chain upon which both IBP formulae are made. In Section 5.3, relying on the Markov chain introduced in Section 5.2, we establish in Theorem 5.3.1 the probabilistic representation formula for the coupled (S_T, Y_T) . In Section 5.4, we establish the BEL formulae for the two quantities appearing in (5.1.2). The main result of this section is Theorem 5.4.1. Some numerical results are presented in Section 5.5. The proofs of Theorem 5.3.1 and of some other technical but important results are postponed to the appendix of Section 5.6.

Notations:

For a fixed time T and positive integer n , we introduce the simplex $\Delta_n(T) := \{s_n \in [0, T]^n : 0 \leq s_1 < \dots < s_n \leq T\}$.

In order to deal with time-degeneracy estimates, we will often use the following space-time inequality:

$$\forall p, q > 0, \forall x \in \mathbb{R}, \quad |x|^p e^{-q|x|^2} \leq (p/(2qe))^{p/2}. \quad (5.1.3)$$

For two positive real numbers α and β , we define the Mittag-Leffler function $z \mapsto E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} z^k / \Gamma(\alpha k + \beta)$. For a positive integer d , we denote by $\mathcal{C}_p^\infty(\mathbb{R}^d)$ the space of real-valued functions which are infinitely differentiable on \mathbb{R}^d with derivatives of any order having polynomial growth.

5.2 Preliminaries: assumptions, definition of the underlying Markov chain and related Malliavin calculus

5.2.1 Assumptions

Throughout the chapter, we work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is assumed to be rich enough to support all random variables that we will consider in what follows. We will work under the following assumptions on the coefficients:

AR The coefficients σ_S and σ_Y are bounded and smooth, in particular a_S and a_Y belong to $\mathcal{C}_b^\infty(\mathbb{R})$. The drift coefficient b_Y belongs to $\mathcal{C}^\infty(\mathbb{R})$ and admits bounded derivatives of any order greater than or equal to one. In particular, the drift coefficient b_Y may be unbounded.

ND There exists $\kappa \geq 1$ such that for all $x \in \mathbb{R}$,

$$\kappa^{-1} \leq a_S(x) \leq \kappa, \quad \kappa^{-1} \leq a_Y(x) \leq \kappa$$

where $a_S = \sigma_S^2$ and $a_Y = \sigma_Y^2$. Therefore, without loss of generality, we will assume that both σ_S and σ_Y are positive function.

Apply Itô's Lemma to $X_t = \ln(S_t)$. We get

$$\begin{cases} X_t &= x_0 + \int_0^t \left(r - \frac{1}{2} a_S(Y_s) \right) ds + \int_0^t \sigma_S(Y_s) dW_s, \\ Y_t &= y_0 + \int_0^t b_Y(Y_s) ds + \int_0^t \sigma_Y(Y_s) dB_s, \\ d\langle B, W \rangle_s &= \rho ds, \end{cases} \quad (5.2.1)$$

with $x_0 = \ln(s_0)$. Without loss of generality, we will thus work with the Markov semigroup associated to the process (X, Y) , namely $P_t h(x_0, y_0) = \mathbb{E}[h(X_t, Y_t)]$.

5.2.2 Choice of the approximation process

As already mentioned in the introduction, our strategy here is based on a probabilistic representation of the marginal law of standard diffusion processes, in the spirit of the unbiased simulation method introduced for diffusion processes by Bally and Kohatsu-Higa [BKH+15], see also Labordère and al. [HLTT+17], and investigated from a numerical perspective by Andersson and Kohatsu-Higa [AKH+17]. We also mention the recent contribution of one the author with Kohatsu-Higa and Li [FKHL+19] for IBP formulae for the marginal law of one-dimensional killed diffusion processes.

However, at this stage, it is important to point out that our choice of approximation process significantly differs from the four aforementioned references. Indeed, in the previous contributions, the drift is assumed to be bounded and basically plays no role so that one usually removes it in the dynamics of the approximation process. In order to handle the unbounded drift term b_Y appearing in the dynamics of the volatility process, one has to take into account the transport of the initial condition by the ODE obtained by removing the noise in the dynamics of Y . To be more specific, we denote by $(m_t(s, y))_{t \in [s, T]}$, $0 \leq s \leq T$, the unique solution to the ODE $\dot{m}_t = b_Y(m_t)$ with initial condition $m_s = y$. Observe that by time-homogeneity of the coefficient b_Y , one has $m_t(s, y) = m_{t-s}(0, y)$. We will simplify the notation when $s = 0$ and write $m_t(y_0)$ for $m_t(0, y_0)$. When there is no ambiguity, we will often omit the dependence with respect to the initial point y_0 and we only write m_t for $m_t(y_0)$. We now introduce the approximation process (\bar{X}, \bar{Y}) defined by

$$\begin{cases} \bar{X}_t^{x_0} &= x_0 + \int_0^t \left(r - \frac{1}{2} a_S(m_s) \right) ds + \int_0^t \sigma_S(m_s) dW_s, \\ \bar{Y}_t^{y_0} &= y_0 + \int_0^t b_Y(m_s) ds + \int_0^t \sigma_Y(m_s) dB_s, \\ d\langle B, W \rangle_s &= \rho ds. \end{cases}$$

We will make intensive use of the explicit form of the Markov semigroup $(\bar{P}_t)_{t \in [0, T]}$ defined for any bounded measurable map $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $\bar{P}_t h(x_0, y_0) = \mathbb{E}[h(\bar{X}_t^{x_0}, \bar{Y}_t^{y_0})]$.

Lemma 5.2.1. *Let $(x_0, y_0) \in \mathbb{R}^2$, $\rho \in (-1, 1)$ and $t \in (0, \infty)$. Then, for any bounded and measurable map $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, it holds*

$$\bar{P}_t h(x_0, y_0) = \int_{\mathbb{R}^2} h(x, y) \bar{p}(t, x_0, y_0, x, y) dx dy$$

with

$$\begin{aligned} \bar{p}(t, x_0, y_0, x, y) &= \frac{1}{2\pi\sigma_{S,t}\sigma_{Y,t}\sqrt{1-\rho_t^2}} \exp\left(-\frac{1}{2} \frac{(x-x_0-(rt-\frac{1}{2}a_{S,t}))^2}{a_{S,t}(1-\rho_t^2)} - \frac{1}{2} \frac{(y-m_t)^2}{a_{Y,t}(1-\rho_t^2)}\right) \\ &\times \exp\left(\rho_t \frac{(x-x_0-(rt-\frac{1}{2}a_{S,t}))(y-m_t)}{\sigma_{S,t}\sigma_{Y,t}(1-\rho_t^2)}\right) \end{aligned}$$

where we introduced the notations

$$\begin{aligned} a_{S,t} &= a_{S,t}(y_0) := \sigma_{S,t}^2 := \int_0^t a_S(m_s(y_0)) ds, \\ a_{Y,t} &= a_{Y,t}(y_0) := \sigma_{Y,t}^2 := \int_0^t a_Y(m_s(y_0)) ds, \\ \sigma_{S,Y,t} &= \sigma_{S,Y,t}(y_0) := \int_0^t (\sigma_S \sigma_Y)(m_s(y_0)) ds, \\ \rho_t &:= \rho \sigma_{S,Y,t} / (\sigma_{S,t} \sigma_{Y,t}). \end{aligned}$$

Moreover, for any $t \in (0, T]$, there exists some positive constant $C := C(T, \rho, a, r, \kappa)$ such that

$$\forall t \in (0, T], \quad \bar{p}(t, x_0, y_0, x, y) \leq C \bar{q}_{4\kappa}(t, x_0, y_0, x, y) \quad (5.2.2)$$

where, for a positive parameter c , we introduced the density function

$$(x, y) \mapsto \bar{q}_c(t, x_0, y_0, x, y) := \frac{1}{2\pi ct} \exp\left(-\frac{(x-x_0)^2}{2ct} - \frac{(y-m_t)^2}{2ct}\right).$$

Proof. We write

$$(\bar{X}_t^{x_0}, \bar{Y}_t^{y_0}) = \left(x_0 + rt - \frac{1}{2}a_{S,t} + \int_0^t \sigma_S(m_s) dW_s, m_t + \int_0^t \sigma_Y(m_s) (\rho dW_s + \sqrt{1-\rho^2} d\widetilde{W}_s)\right)$$

where \widetilde{W} is a one-dimensional standard Brownian motion independent of W . We thus deduce that $(\bar{X}_t^{x_0}, \bar{Y}_t^{y_0}) \sim \mathcal{N}(\mu(t, x_0, y_0), \Sigma_t)$ with

$$\mu(t, x_0, y_0) = \left(x_0 + \left(rt - \frac{1}{2}a_{S,t}\right), m_t\right) \quad \text{and} \quad \Sigma_t = \begin{pmatrix} a_{S,t} & \rho\sigma_{S,Y,t} \\ \rho\sigma_{S,Y,t} & a_{Y,t} \end{pmatrix}.$$

The expression of the transition density then readily follows. Now, from **ND**, it is readily seen that $a_{S,t}, a_{Y,t} \leq \kappa t$ so that using the inequalities $(a-b)^2 \geq \frac{1}{2}a^2 - b^2$ and $\rho_t^2 \leq \rho^2$, it follows

$$\begin{aligned} \bar{p}(t, x_0, y_0, x, y) &= \frac{1}{2\pi\sigma_{S,t}\sigma_{Y,t}\sqrt{1-\rho_t^2}} \exp\left(-\frac{1}{2} \frac{(x-x_0-(rt-\frac{1}{2}a_{S,t}))^2}{a_{S,t}(1-\rho_t^2)} - \frac{1}{2} \frac{(y-m_t)^2}{a_{Y,t}(1-\rho_t^2)}\right) \\ &\times \exp\left(\frac{\rho_t}{1-\rho_t^2} \frac{(x-x_0-(rt-\frac{1}{2}a_{S,t}))(y-m_t)}{\sigma_{S,t}\sigma_{Y,t}}\right) \\ &\leq C \frac{1}{2\pi(2\kappa)t} \exp\left(- (4\kappa)^{-1} \frac{(x-x_0)^2}{2t} - (4\kappa)^{-1} \frac{(y-m_t)^2}{2t}\right) \\ &=: C \bar{q}_{4\kappa}(t, x_0, y_0, x, y) \end{aligned}$$

for some positive constants $C := C(T, \lambda, \rho, a, r, \kappa)$. □

We will also use the notation $(\bar{X}_t^{s,x}, \bar{Y}_t^{s,y})_{t \geq s}$ for the approximation process starting from (x, y) at time s and with coefficients frozen along the deterministic flow $\{m_t(s, y) = m_{t-s}(y), t \geq s\}$. Note that the corresponding Markov semigroup satisfies $\bar{P}_{s,t}h(x, y) := \mathbb{E}[h(\bar{X}_t^{s,y}, \bar{Y}_t^{s,y})] = \mathbb{E}[h(\bar{X}_{t-s}^y, \bar{Y}_{t-s}^y)] = \bar{P}_{t-s}h(x, y)$.

5.2.3 Markov chain on random time grid

The first tool that we will employ is a renewal process N that we now introduce.

Definition 5.2.1. Let $\tau := (\tau_n)_{n \geq 1}$ be a sequence of random variables such that $(\tau_n - \tau_{n-1})_{n \geq 1}$, with the convention $\tau_0 = 0$, are i.i.d. with positive density function f and cumulant distribution function $t \mapsto F(t) = \int_{-\infty}^t f(s) ds$. Then, the renewal process $N := (N_t)_{t \geq 0}$ with jump times τ is defined by $N_t := \sum_{n \geq 1} \mathbf{1}_{\{\tau_n \leq t\}}$.

It is readily seen that, for any $t > 0$, $\{N_t = n\} = \{\tau_n \leq t < \tau_{n+1}\}$ and by an induction argument that we omit, one may prove that the joint distribution of (τ_1, \dots, τ_n) is given by

$$\mathbb{P}(\tau_1 \in ds_1, \dots, \tau_n \in ds_n) = \prod_{j=0}^{n-1} f(s_{j+1} - s_j) \mathbf{1}_{\{0 < s_1 < \dots < s_n\}}$$

which in turn implies

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\{N_t=n\}} \Phi(\tau_1, \dots, \tau_n)] &= \mathbb{E}[\mathbf{1}_{\{\tau_n \leq t < \tau_{n+1}\}} \Phi(\tau_1, \dots, \tau_n)] \\ &= \int_t^\infty \int_{\Delta_n(t)} \Phi(s_1, \dots, s_n) \prod_{j=0}^n f(s_{j+1} - s_j) ds_{n+1} \end{aligned}$$

with the convention $s_0 = 0$. Hence, by Fubini's theorem, it holds

$$\mathbb{E}[\mathbf{1}_{\{N_t=n\}} \Phi(\tau_1, \dots, \tau_n)] = \int_{\Delta_n(t)} \Phi(s_1, \dots, s_n) (1 - F(t - s_n)) \prod_{j=0}^{n-1} f(s_{j+1} - s_j) ds_n$$

for any measurable map $\Phi : \Delta_n(t) \rightarrow \mathbb{R}$ satisfying $\mathbb{E}[\mathbf{1}_{\{N_t=n\}} |\Phi(\tau_1, \dots, \tau_n)|] < \infty$.

Usual choices that we will consider are the followings.

Example 5.2.1. 1. If the density function f is given by $f(t) = \lambda e^{-\lambda t} \mathbf{1}_{[0, \infty)}(t)$ for some positive parameter λ , then N is a Poisson process with intensity λ .

2. If the density function f is given by $f(t) = \frac{1-\alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{t^\alpha} \mathbf{1}_{[0, \bar{\tau}]}(t)$ for some parameters $(\alpha, \bar{\tau}) \in (0, 1) \times (0, \infty)$, then N is a renewal process with $[0, \bar{\tau}]$ -valued Beta $(1-\alpha, 1)$ jump times.

3. More generally, if the density function f is given by $f(t) = \frac{\bar{\tau}^{1-\alpha-\beta}}{B(\alpha, \beta)} \frac{1}{t^{1-\alpha}(\bar{\tau}-t)^{1-\beta}} \mathbf{1}_{[0, \bar{\tau}]}(t)$ for some parameters $(\alpha, \beta, \bar{\tau}) \in (0, 1)^2 \times (0, \infty)$, then N is a renewal process with $[0, \bar{\tau}]$ -valued Beta (α, β) jump times.

Given a sequence $Z = (Z_n^1, Z_n^2)_{n \geq 1}$ of i.i.d. random vector with law $\mathcal{N}(0, I_2)$ and a renewal process N independent of Z with jump times $(\tau_i)_{i \geq 0}$, we set $\zeta_i = \tau_i \wedge T$, with the convention $\zeta_0 = 0$, and we consider the two-dimensional Markov chain (\bar{X}, \bar{Y}) starting from the initial point (x_0, y_0) at time 0 (evolving on the random time grid $(\zeta_i)_{i \geq 0}$) and with dynamics for any $0 \leq i \leq N_T$

$$\begin{cases} \bar{X}_{i+1} &= \bar{X}_i + \left(r(\zeta_{i+1} - \zeta_i) - \frac{1}{2} a_{S,i} \right) + \sigma_{S,i} Z_{i+1}^1, \\ \bar{Y}_{i+1} &= m_i + \sigma_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right), \end{cases} \quad (5.2.4)$$

where we introduced the notations

$$\begin{aligned} a_{S,i} &:= \sigma_{S,i}^2 := a_{S, \zeta_{i+1} - \zeta_i}(\bar{Y}_i) = \int_0^{\zeta_{i+1} - \zeta_i} a_S(m_s(\bar{Y}_i)) ds, \\ a_{Y,i} &:= \sigma_{Y,i}^2 := a_{Y, \zeta_{i+1} - \zeta_i}(\bar{Y}_i) = \int_0^{\zeta_{i+1} - \zeta_i} a_Y(m_s(\bar{Y}_i)) ds, \\ \sigma_{S,Y,i} &:= \int_0^{\zeta_{i+1} - \zeta_i} (\sigma_S \sigma_Y)(m_s(\bar{Y}_i)) ds, \\ \rho_i &:= \rho_{\zeta_{i+1} - \zeta_i}(\bar{Y}_i) = \rho \frac{\sigma_{S,Y,i}}{\sigma_{S,i} \sigma_{Y,i}}, \\ m_i &:= m_{\zeta_{i+1} - \zeta_i}(\bar{Y}_i). \end{aligned}$$

We will denote by $\sigma'_{S,i}$ the first derivative of $y \mapsto \sigma_{S,i}(y)$ taken at \bar{Y}_i and proceed similarly for the quantities $\sigma'_{Y,i}$, $\sigma'_{S,Y,i}$, ρ'_i and m'_i . We define the filtration $\mathcal{G} = (\mathcal{G}_i)_{i \geq 0}$ where $\mathcal{G}_i = \sigma(Z_j, 0 \leq j \leq i)$, for $i \geq 1$ and \mathcal{G}_0 stands for the trivial σ -field. We assume that the filtration \mathcal{G} satisfies the usual conditions. For an integer n , we will use the notations $\zeta^n = (\zeta_0, \dots, \zeta_n)$ and $\tau^n = (\tau_0, \dots, \tau_n)$.

5.2.4 Tailor-made Malliavin calculus for the Markov chain.

In this section we introduce a tailor-made Malliavin calculus for the underlying Markov chain (\bar{X}, \bar{Y}) defined by (5.2.4) which will be employed in order to establish our IBP formulae. Instead of using an infinite dimensional calculus as it is usually done in the literature, see e.g. Nualart [Nua06], the approach developed below is based on a finite dimensional calculus for which the dimension is given by the number of jumps of the underlying renewal process involved in the Markov chain (\bar{X}, \bar{Y}) .

Definition 5.2.2. *Let $n \in \mathbb{N}$. For any $i \in \{0, \dots, n\}$, we define the set $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, as the space of random variables H such that*

- $H = h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1})$, on the set $\{N_T = n\}$, where we recall $\zeta^{n+1} := (\zeta_0, \dots, \zeta_{n+1}) = (0, \zeta_1, \dots, \zeta_n, T)$.
- For any $\mathbf{s}_{n+1} \in \Delta_{n+1}(T)$, the map $h(\cdot, \cdot, \cdot, \cdot, \mathbf{s}_{n+1}) \in \mathcal{C}_p^\infty(\mathbb{R}^4)$.

For a r.v. $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, we will often abuse the notations and write

$$H \equiv H(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1})$$

that is the same symbol H may denote the r.v. or the function in the set $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$. One can easily define the flow derivatives for $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$ as follows

$$\begin{aligned}\partial_{\bar{X}_{i+1}} H &= \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}), \\ \partial_{\bar{Y}_{i+1}} H &= \partial_4 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}), \\ \partial_{\bar{X}_i} H &= \partial_1 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) + \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) \partial_{\bar{X}_i} \bar{X}_{i+1}, \\ \partial_{\bar{Y}_i} H &= \partial_2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) + \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) \partial_{\bar{Y}_i} \bar{X}_{i+1} + \\ &\quad \partial_4 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1},\end{aligned}$$

and

$$\begin{aligned}\partial_{\bar{X}_i} \bar{X}_{i+1} &= 1, \\ \partial_{\bar{Y}_i} \bar{Y}_{i+1} &= m'_i + \sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) + \sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right),\end{aligned}\tag{5.2.5}$$

$$\partial_{\bar{Y}_i} \bar{X}_{i+1} = -\frac{1}{2} a'_{S,i} + \sigma'_{S,i} Z_{i+1}^1 = -\frac{1}{2} a'_{S,i} + \frac{\sigma'_{S,i}}{\sigma_{S,i}} \left(\bar{X}_{i+1} - \bar{X}_i - \left(r(\zeta_{i+1} - \zeta_i) - \frac{1}{2} a_{S,i} \right) \right).\tag{5.2.6}$$

We now define the integral and derivative operators for $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, as

$$\mathcal{I}_{i+1}^{(1)}(H) = H \left[\frac{Z_{i+1}^1}{\sigma_{S,i}(1 - \rho_i^2)} - \frac{\rho_i}{1 - \rho_i^2} \frac{\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2}{\sigma_{S,i}} \right] - \mathcal{D}_{i+1}^{(1)} H,\tag{5.2.7}$$

$$\mathcal{I}_{i+1}^{(2)}(H) = H \left[\frac{\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2}{\sigma_{Y,i}(1 - \rho_i^2)} - \frac{\rho_i}{1 - \rho_i^2} \frac{Z_{i+1}^1}{\sigma_{Y,i}} \right] - \mathcal{D}_{i+1}^{(2)} H,\tag{5.2.8}$$

$$\mathcal{D}_{i+1}^{(1)} H = \partial_{\bar{X}_{i+1}} H,$$

$$\mathcal{D}_{i+1}^{(2)} H = \partial_{\bar{Y}_{i+1}} H.$$

Note that due to the above definitions and assumption **H**, it is readily checked that $\mathcal{I}_{i+1}^{(1)}(H), \mathcal{I}_{i+1}^{(2)}(H), \mathcal{D}_{i+1}^{(1)} H$ and $\mathcal{D}_{i+1}^{(2)} H$ are elements of $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$ so that we can define iterations of the above operators. Namely, by induction, for a multi-index $\alpha = (\alpha_1, \dots, \alpha_p)$ of length p with $\alpha_i \in \{1, 2\}$ and $\alpha_{p+1} \in \{1, 2\}$, we define

$$\mathcal{I}_{i+1}^{(\alpha, \alpha_{p+1})}(H) = \mathcal{I}_{i+1}^{(\alpha_{p+1})}(\mathcal{I}_{i+1}^{(\alpha)}(H)), \quad \mathcal{D}_{i+1}^{(\alpha, \alpha_{p+1})} H = \mathcal{D}_{i+1}^{(\alpha_{p+1})}(\mathcal{D}_{i+1}^{(\alpha)} H)$$

with the intuitive notation $(\alpha, \alpha_{p+1}) = (\alpha_1, \dots, \alpha_{p+1})$.

Throughout the chapter, we will use the following notation for a certain type of conditional expectation that will be frequently employed. For any $X \in L^1(\mathbb{P})$ and any $i \in \{0, \dots, n\}$,

$$\mathbb{E}_{i,n}[X] = \mathbb{E}[X | \mathcal{G}_i, \tau^{n+1}, N_T = n]$$

where we recall that we employ the notation $\tau^{n+1} = (\tau_0, \dots, \tau_{n+1})$. Having the above definitions and notations at hand, the following duality formula is satisfied: for any

non-empty multi-index α of length p , with $\alpha_i \in \{1, 2\}$, for any $i \in \{1, \dots, p\}$, p being a positive integer, it holds

$$\mathbb{E}_{i,n} \left[\mathcal{D}_{i+1}^{(\alpha)} f(\bar{X}_{i+1}, \bar{Y}_{i+1}) H \right] = \mathbb{E}_{i,n} \left[f(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(\alpha)}(H) \right]. \quad (5.2.9)$$

In order to obtain explicit norm estimates for random variables in $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, it is useful to define for $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, $i \in \{0, \dots, n\}$ and $p \geq 1$

$$\|H\|_{p,i,n}^p = \mathbb{E}_{i,n}[|H|^p].$$

We will also employ a chain rule formula for the integral operators defined above.

Lemma 5.2.2. *Let $H \equiv H(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, for some $i \in \{0, \dots, n\}$. The following chain rule formulae hold for any $(\alpha_1, \alpha_2) \in \{1, 2\}^2$*

$$\partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1)}(H) = \mathcal{I}_{i+1}^{(\alpha_1)}(\partial_{\bar{X}_i} H), \quad \partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1, \alpha_2)}(H) = \mathcal{I}_{i+1}^{(\alpha_1, \alpha_2)}(\partial_{\bar{X}_i} H). \quad (5.2.10)$$

Moreover, one has

$$\partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1)}(H) = \mathcal{I}_{i+1}^{(1)}(\partial_{\bar{Y}_i} H) - \frac{\sigma'_{S,i} \mathcal{I}_{i+1}^{(1)}(H) - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(2)}(H)}, \quad (5.2.11)$$

$$\partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(2)}(H) = \mathcal{I}_{i+1}^{(2)}(\partial_{\bar{Y}_i} H) - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(2)}(H), \quad (5.2.12)$$

$$\begin{aligned} \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1,1)}(H) &= \mathcal{I}_{i+1}^{(1,1)}(\partial_{\bar{Y}_i} H) - 2 \frac{\sigma'_{S,i} \mathcal{I}_{i+1}^{(1,1)}(H)}{\sigma_{S,i}} \\ &\quad - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \left(\mathcal{I}_{i+1}^{(1,2)}(H) + \mathcal{I}_{i+1}^{(2,1)}(H) \right), \end{aligned} \quad (5.2.13)$$

$$\partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(2,2)}(H) = \mathcal{I}_{i+1}^{(2,2)}(\partial_{\bar{Y}_i} H) - 2 \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(2,2)}(H), \quad (5.2.14)$$

$$\begin{aligned} \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1,2)}(H) &= \mathcal{I}_{i+1}^{(1,2)}(\partial_{\bar{Y}_i} H) - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(2,2)}(H) \\ &\quad - \left(\frac{\sigma'_{S,i}}{\sigma_{S,i}} + \frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(1,2)}(H). \end{aligned} \quad (5.2.15)$$

Proof. Observe that from the very definitions (5.2.7) and (5.2.8), one directly gets

$$\partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(1)}(1) = \partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(2)}(1) = 0$$

while, also by direct computation, we obtain

$$\begin{aligned} \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1)}(1) &= - \frac{\sigma'_{S,i} \mathcal{I}_{i+1}^{(1)}(1) - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(2)}(1)}, \\ \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(2)}(1) &= - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(2)}(1). \end{aligned}$$

We thus deduce

$$\begin{aligned}\partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1)}(H) &= \partial_{\bar{X}_i} H \mathcal{I}_{i+1}^{(\alpha_1)}(1) + H \partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1)}(1) - \partial_{\bar{X}_i} \mathcal{D}_{i+1}^{(\alpha_1)} H \\ &= \partial_{\bar{X}_i} H \mathcal{I}_{i+1}^{(\alpha_1)}(1) - \mathcal{D}_{i+1}^{(\alpha_1)}(\partial_{\bar{X}_i} H) \\ &= \mathcal{I}_{i+1}^{(\alpha_1)}(\partial_{\bar{X}_i} H)\end{aligned}$$

where we used the fact $\mathcal{D}_{i+1}^{(\alpha_1)} \partial_{\bar{X}_i} H = \partial_{\bar{X}_i} \mathcal{D}_{i+1}^{(\alpha_1)} H$ which easily follows by direct computation. As a consequence, it is readily seen

$$\begin{aligned}\partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1, \alpha_2)}(H) &= \partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_2)}(\mathcal{I}_{i+1}^{(\alpha_1)}(H)) = \mathcal{I}_{i+1}^{(\alpha_2)}(\partial_{\bar{X}_i} \mathcal{I}_{i+1}^{(\alpha_1)}(H)) \\ &= \mathcal{I}_{i+1}^{(\alpha_2)}(\mathcal{I}_{i+1}^{(\alpha_1)}(\partial_{\bar{X}_i} H)) = \mathcal{I}_{i+1}^{(\alpha_1, \alpha_2)}(\partial_{\bar{X}_i} H).\end{aligned}$$

This concludes the proof of (5.2.10). The chain rule formulae (5.2.11), (5.2.12), (5.2.13), (5.2.14) and (5.2.15) follow from similar arguments. Let us prove (5.2.11) and (5.2.12). The proofs of (5.2.13), (5.2.14) and (5.2.15) are omitted. Observe first that in general $\mathcal{D}_{i+1}^{(\alpha_1)} \partial_{\bar{Y}_i} H \neq \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(\alpha_1)} H$. Indeed, by standard computations, it holds

$$\begin{aligned}\partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(1)} H &= \partial_{\bar{Y}_i} \partial_{\bar{X}_{i+1}} h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \\ &= \partial_{2,3}^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) + \partial_3^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{X}_{i+1} + \partial_{4,3}^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{Y}_{i+1}, \\ \mathcal{D}_{i+1}^{(1)} \partial_{\bar{Y}_i} H &= \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \\ &= \partial_{\bar{X}_{i+1}} (\partial_2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) + \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{X}_{i+1} + \partial_4 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{Y}_{i+1}) \\ &= \partial_{3,2}^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) + \partial_3^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{X}_{i+1} + \partial_{3,4}^2 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{Y}_i} \bar{Y}_{i+1} \\ &\quad + \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{X}_{i+1} + \partial_4 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} \\ &= \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(1)} H + \partial_3 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{X}_{i+1} + \partial_4 h(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^n) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} \\ &= \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(1)} H + \mathcal{D}_{i+1}^{(1)} H \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{X}_{i+1} + \mathcal{D}_{i+1}^{(2)} H \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} \\ &= \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(1)} H + \frac{\sigma'_{S,i}}{\sigma_{S,i}} \mathcal{D}_{i+1}^{(1)} H + \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{D}_{i+1}^{(2)} H\end{aligned}$$

where we used the two identities $\partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{X}_{i+1} = \frac{\sigma'_{S,i}}{\sigma_{S,i}}$ and $\partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} = \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}}$ which readily stems from (5.2.5), (5.2.6) and the dynamics (5.2.4).

From (5.2.7) and the previous identity, we thus obtain

$$\begin{aligned}\partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1)}(H) &= \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(1)}(1) H + \mathcal{I}_{i+1}^{(1)}(1) \partial_{\bar{Y}_i} H - \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(1)} H \\ &= -\frac{\sigma'_{S,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(1)}(1) H - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(2)}(1) H + \mathcal{I}_{i+1}^{(1)}(1) \partial_{\bar{Y}_i} H - \mathcal{D}_{i+1}^{(1)} \partial_{\bar{Y}_i} H + \frac{\sigma'_{S,i}}{\sigma_{S,i}} \mathcal{D}_{i+1}^{(1)} H \\ &\quad + \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{D}_{i+1}^{(2)} H \\ &= -\frac{\sigma'_{S,i}}{\sigma_{S,i}} \left(\mathcal{I}_{i+1}^{(1)}(1) H - \mathcal{D}_{i+1}^{(1)} H \right) + \mathcal{I}_{i+1}^{(1)}(1) \partial_{\bar{Y}_i} H - \mathcal{D}_{i+1}^{(1)} \partial_{\bar{Y}_i} H - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \left(\mathcal{I}_{i+1}^{(2)}(1) H - \mathcal{D}_{i+1}^{(2)} H \right) \\ &= \mathcal{I}_{i+1}^{(1)}(\partial_{\bar{Y}_i} H) - \frac{\sigma'_{S,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(1)}(H) - \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \mathcal{I}_{i+1}^{(2)}(H).\end{aligned}$$

Similarly, after some algebraic manipulations using (5.2.4) and (5.2.5), we get $\partial_{\bar{Y}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} = \frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2}$ so that

$$\mathcal{D}_{i+1}^{(2)} \partial_{\bar{Y}_i} H = \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(2)} H + \mathcal{D}_{i+1}^{(2)} H \partial_{\bar{Y}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} = \partial_{\bar{Y}_i} \mathcal{D}_{i+1}^{(2)} H + \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{D}_{i+1}^{(2)} H$$

so that, omitting some technical details, we get

$$\begin{aligned} \partial_{\bar{Y}_i} \mathcal{I}_{i+1}^{(2)}(H) &= \mathcal{I}_{i+1}^{(2)}(\partial_{\bar{Y}_i} H) - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(2)}(1)H + \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{D}_{i+1}^{(2)} H \\ &= \mathcal{I}_{i+1}^{(2)}(\partial_{\bar{Y}_i} H) - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \mathcal{I}_{i+1}^{(2)}(H). \end{aligned}$$

The identities (5.2.13), (5.2.14) and (5.2.15) eventually follows from (5.2.11) and (5.2.12) using some simple algebraic computations. \square

We conclude this section by introducing the following space of random variables which satisfy some time regularity estimates.

Definition 5.2.3. *Let $\ell \in \mathbb{Z}$ and $n \in \mathbb{N}$. For any $i \in \{0, \dots, n\}$, we define the space $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$ as the set of finite random variables $H \in \mathbb{S}_{i,n}(\bar{X}, \bar{Y})$ satisfying the following time regularity estimate: for any $p \geq 1$, for any $c > 0$, there exists some positive constants $C := C(T)$, c' , $T \mapsto C(T)$ being non-decreasing and c' being independent of T , such that for any $(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1}) \in \mathbb{R}^4 \times \Delta_{n+1}(T)$,*

$$\begin{aligned} |H(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1})|^p \bar{q}_c(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) & \quad (5.2.16) \\ & \leq C(s_{i+1} - s_i)^{p\ell/2} \bar{q}_{c'}(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) \end{aligned}$$

where the density function $\mathbb{R}^2 \ni (x_{i+1}, y_{i+1}) \mapsto \bar{q}_c(\zeta_{i+1} - \zeta_i, x_i, y_i, x_{i+1}, y_{i+1})$ is defined in Lemma 5.2.1.

We again remark that since the space $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$ is a subset of $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$, when we say that a random variable $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$ this statement is always understood on the set $\{N_T = n\}$.

Before proceeding, let us provide a simple example of some random variables that belong to the aforementioned space. From (5.2.7) and the dynamics (5.2.4) of the Markov chain (\bar{X}, \bar{Y}) , it holds

$$\begin{aligned} \mathcal{I}_{i+1}^{(1)}(1) &= \left[\frac{\bar{X}_{i+1} - \bar{X}_i - (r(\zeta_{i+1} - \zeta_i) - \frac{1}{2}a_{S,i})}{a_{S,i}(1 - \rho_i^2)} - \frac{\rho_i}{1 - \rho_i^2} \frac{\bar{Y}_{i+1} - m_i}{\sigma_{S,i}\sigma_{Y,i}} \right], \\ \mathcal{I}_{i+1}^{(1,1)}(1) &= (\mathcal{I}_{i+1}^{(1)}(1))^2 - \mathcal{D}_{i+1}^1(\mathcal{I}_{i+1}^{(1)}(1)) = (\mathcal{I}_{i+1}^{(1)}(1))^2 - \frac{1}{a_{S,i}(1 - \rho_i^2)}, \end{aligned}$$

so that, $\mathcal{I}_{i+1}^{(1)}(1)$ and $\mathcal{I}_{i+1}^{(1,1)}(1)$ belong to $\mathbb{S}_{i,n}(\bar{X}, \bar{Y})$. Moreover, under **ND**, for any $p \geq 1$, it holds

$$\left| \mathcal{I}_{i+1}^{(1)}(1)(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1}) \right|^p \leq C \left(1 + \frac{|x_{i+1} - x_i|^p}{(s_{i+1} - s_i)^p} + \frac{|y_{i+1} - m_i(y_i)|^p}{(s_{i+1} - s_i)^p} \right)$$

and similarly,

$$\left| \mathcal{I}_{i+1}^{(1,1)}(1)(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1}) \right|^p \leq C \left(1 + \frac{1}{(s_{i+1} - s_i)^p} + \frac{|x_{i+1} - x_i|^{2p}}{(s_{i+1} - s_i)^{2p}} + \frac{|y_{i+1} - m_i(y_i)|^{2p}}{(s_{i+1} - s_i)^{2p}} \right).$$

Hence, from the space-time inequality (5.1.3), for any $c > 0$ and any $c' > c$, it holds

$$\begin{aligned} & \left| \mathcal{I}_{i+1}^{(1)}(1)(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1}) \right|^p \bar{q}_c(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) \\ & \leq C(s_{i+1} - s_i)^{-p/2} \bar{q}_{c'}(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) \end{aligned}$$

and

$$\begin{aligned} & \left| \mathcal{I}_{i+1}^{(1,1)}(1)(x_i, y_i, x_{i+1}, y_{i+1}, \mathbf{s}_{n+1}) \right|^p \bar{q}_c(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) \\ & \leq C(s_{i+1} - s_i)^{-p} \bar{q}_{c'}(s_{i+1} - s_i, x_i, y_i, x_{i+1}, y_{i+1}) \end{aligned}$$

for some positive constant $C := C(T)$, $T \mapsto C(T)$ being non-decreasing. We thus conclude that $\mathcal{I}_{i+1}^{(1)}(1) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$ and $\mathcal{I}_{i+1}^{(1,1)}(1) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1)$ for any $i \in \{0, \dots, n\}$.

A straightforward generalization of the above example is the following property that will be frequently used in the sequel. We omit its proof.

Lemma 5.2.3. *Let $n \in \mathbb{N}$ and $i \in \{0, \dots, n\}$. Assume that $H \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$ and $\mathcal{D}_i^{(\alpha_1)} H \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell'/2)$, for some $(\ell, \ell') \in \mathbb{Z}^2$ and some $\alpha_1 \in \{1, 2\}$. Then, it holds $\mathcal{I}_i^{(\alpha_1)}(H) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, ((\ell - 1) \wedge \ell')/2)$. Additionally, if $\mathcal{D}_i^{(\alpha_2)} H \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell'/2)$ and $\mathcal{D}_i^{(\alpha_1, \alpha_2)} H \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell''/2)$, for some $\ell'' \in \mathbb{Z}$ and $\alpha_2 \in \{1, 2\}$, then it holds $\mathcal{I}_i^{(\alpha_1, \alpha_2)}(H) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, ((\ell - 2) \wedge (\ell' - 1) \wedge \ell'')/2)$. Finally, if $H_1 \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell_1/2)$ and $H_2 \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell_2/2)$ for some $(\ell_1, \ell_2) \in \mathbb{Z}^2$ then $H_1 H_2 \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, (\ell_1 + \ell_2)/2)$ and $(\zeta_{i+1} - \zeta_i) \mathcal{I}_i^{(\alpha_1)}(H_1) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, (\ell_1 + 1)/2)$.*

Finally, we importantly emphasize that if $H \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$ for some $n \in \mathbb{N}$, $i \in \{0, \dots, n\}$ and $\ell \in \mathbb{Z}$, then, its conditional $L^p(\mathbb{P})$ -moment is finite and also satisfies a time regularity estimate. More precisely, for any $p \geq 1$, it holds

$$\|H\|_{p,i,n} \leq C(\zeta_{i+1} - \zeta_i)^{\ell/2} \quad (5.2.17)$$

for some positive constant $C := C(T)$, $T \mapsto C(T)$ being non-decreasing. Indeed, using the fact that the sequence Z is independent of N as well as the upper-estimate (5.2.2) of Lemma 5.2.1 and finally (5.2.16), one directly gets

$$\begin{aligned} \|H\|_{p,i,n}^p &= \mathbb{E}_{i,n} \left[\left| H(\bar{X}_i, \bar{Y}_i, \bar{X}_{i+1}, \bar{Y}_{i+1}, \zeta^{n+1}) \right|^p \middle| \bar{X}_i, \bar{Y}_i, \tau^{n+1}, N_T = n \right] \\ &= \int_{\mathbb{R}^2} |H(\bar{X}_i, \bar{Y}_i, x_{i+1}, y_{i+1}, \zeta^{n+1})|^p \bar{p}(\zeta_{i+1} - \zeta_i, \bar{X}_i, \bar{Y}_i, x_{i+1}, y_{i+1}) dx_{i+1} dy_{i+1} \\ &\leq C \int_{\mathbb{R}^2} |H(\bar{X}_i, \bar{Y}_i, x_{i+1}, y_{i+1}, \zeta^{n+1})|^p \bar{q}_{4\kappa}(\zeta_{i+1} - \zeta_i, \bar{X}_i, \bar{Y}_i, x_{i+1}, y_{i+1}) dx_{i+1} dy_{i+1} \\ &\leq C(\zeta_{i+1} - \zeta_i)^{p\ell/2} \end{aligned}$$

so that (5.2.17) directly follows. The previous conditional $L^p(\mathbb{P})$ -moment estimate will be used at several places in the sequel.

5.3 Probabilistic representation for the couple (S_T, Y_T) .

In this section, we establish a probabilistic representation for the marginal law (S_T, Y_T) , or equivalently, for the law of (X_T, Y_T) which is based on the Markov chain (\bar{X}, \bar{Y}) introduced in the previous section. For $\gamma > 0$, we denote by $\mathcal{B}_\gamma(\mathbb{R}^2)$ the set of Borel measurable map $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying the following exponential growth assumption at infinity, namely, for some positive constant C , for any $(x, y) \in \mathbb{R}^2$,

$$|h(x, y)| \leq C \exp(\gamma(|x|^2 + |y|^2)). \quad (5.3.1)$$

Theorem 5.3.1. *Let $T > 0$. Under assumptions **AR** and **ND**, the law of the couple (X_T, Y_T) given by the unique solution to the SDE (5.2.1) at time T starting from $(x_0 = \ln(s_0), y_0)$ at time 0 satisfies the following probabilistic representation: there exists a positive constant $c := c(T, b_Y, \kappa)$ such that for any $0 \leq \gamma < c^{-1}$ and any $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, it holds*

$$\mathbb{E}[h(X_T, Y_T)] = \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i\right] \quad (5.3.2)$$

where the random variables $\theta_i \in \mathbb{S}_{i-1, n}(\bar{X}, \bar{Y})$ are defined by

$$1 \leq i \leq N_T, \quad \theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{I}_i^{(1)}(c_S^i) + \mathcal{I}_i^{(2,2)}(c_Y^i) + \mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{I}_i^{(1,2)}(c_{Y,S}^i) \right], \quad (5.3.3)$$

$$\theta_{N_T+1} = (1 - F(T - \zeta_{N_T}))^{-1}, \quad (5.3.4)$$

with

$$\begin{aligned} c_S^i &:= \frac{1}{2} \left(a_S(\bar{Y}_i) - a_S(m_{i-1}) \right), \\ c_Y^i &:= \frac{1}{2} \left(a_Y(\bar{Y}_i) - a_Y(m_{i-1}) \right), \\ b_Y^i &:= b_Y(\bar{Y}_i) - b_Y(m_{i-1}), \\ c_{Y,S}^i &:= \rho((\sigma_S \sigma_Y)(\bar{Y}_i) - (\sigma_S \sigma_Y)(m_{i-1})). \end{aligned}$$

Assume furthermore that N is a renewal process with Beta($\alpha, 1$) jump times. For any $p \geq 1$ satisfying $p(\frac{1}{2} - \alpha) \leq 1 - \alpha$, for any γ such that $p\gamma < c^{-1}$ and any $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, the random variable appearing inside the expectation in the right-hand side of (5.3.2) admits a finite $L^p(\mathbb{P})$ -moment. In particular, if $\alpha = 1/2$ then for any $p \geq 1$, the $L^p(\mathbb{P})$ -moment is finite.

The proof of Theorem 5.3.1 is postponed to Appendix 5.6.1.

5.4 Integration by parts formulae

In this section, we establish two IBP formulae for the law of the couple (S_T, Y_T) . More precisely, we are interested in providing a Bismut-Elworthy-Li formula for the two quantities

$$\partial_{s_0} \mathbb{E}[h(S_T, Y_T)], \quad \partial_{y_0} \mathbb{E}[h(S_T, Y_T)].$$

Our strategy is divided into two steps:

Step 1: The first step was performed with the probabilistic representation established in Theorem 5.3.1 for the couple (X_T, Y_T) involving the two-dimensional Markov chain (\bar{X}, \bar{Y}) evolving on a time grid governed by the jump times of the renewal process N . Introducing $h(x, y) = f(e^x, y)$ and assuming that f is of polynomial growth at infinity, it is sufficient to consider the two quantities

$$\partial_{s_0} \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \right], \quad \partial_{y_0} \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \right]$$

for $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ for some $\gamma > 0$.

Step 2: At this stage, one might be tempted to perform a standard IBP formula as presented in Nualart [Nua06] on the whole time interval $[0, T]$. However, we were not able to do this. The main reason is that the Skorokhod integral of the product of weights $\prod_{i=1}^{N_T} \theta_i$ will inevitably involve the Malliavin derivative of θ_i which will in turn raise some integrability issues of the resulting Malliavin weight. The key idea that we use consists in performing local IBP formulae on each of the random intervals $[\zeta_i, \zeta_{i+1}]$, $i = 0, \dots, N_T$, that is, by using the noise of the Markov chain on this specific time interval and then by combining all these local IBP formulae in a suitable way.

To implement successfully our strategy, two main ingredients are needed. Our first ingredient consists in transferring the partial derivatives ∂_{s_0} and ∂_{y_0} on the expectation forward in time from the first time interval $[0, \zeta_1]$ to the interval on which we perform the local IBP formula, say $[\zeta_i, \zeta_{i+1}]$. Our second ingredient consists in combining these various local IBP formulae in an adequate manner. Roughly speaking, we will consider a weighted sum of each IBP formula, the weight being precisely the length of the corresponding time interval.

5.4.1 The transfer of derivative formula

Lemma 5.4.1. *Let $h \in \mathcal{C}_p^1(\mathbb{R}^2)$ and n an integer. The maps*

$$\mathbb{R}^2 \ni (x, y) \mapsto \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} | (\bar{X}_i, \bar{Y}_i) = (x, y) \right], \quad i \in \{0, \dots, n\}$$

belong to $\mathcal{C}_p^1(\mathbb{R}^2)$ a.s. Moreover, the following transfer of derivative formulae hold

$$\partial_{s_0} \mathbb{E}_{0,n} \left[h(\bar{X}_1, \bar{Y}_1) \theta_1 \right] = \mathbb{E}_{0,n} \left[\partial_{\bar{X}_1} h(\bar{X}_1, \bar{Y}_1) \frac{\theta_1}{s_0} \right] \quad (5.4.1)$$

while for $1 \leq i \leq n$,

$$\partial_{\bar{X}_i} \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} \right] = \mathbb{E}_{i,n} \left[\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} \right]. \quad (5.4.2)$$

Similarly, the following transfer of derivative formulae hold: for any $0 \leq i \leq n-1$

$$\begin{aligned} \partial_{\bar{Y}_i} \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} \right] &= \mathbb{E}_{i,n} \left[\partial_{\bar{Y}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,Y} \right] + \mathbb{E}_{i,n} \left[\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,X} \right] \\ &\quad + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^c \right] \end{aligned} \quad (5.4.3)$$

with

$$\begin{aligned}
 \vec{\theta}_{i+1}^{e,Y} &= (f(\zeta_{i+1} - \zeta_i))^{-1} \left[\mathcal{I}_{i+1}^{(1,1)}(d_S^{i+1}) + \mathcal{I}_{i+1}^{(2,2)}(d_Y^{i+1}) + \mathcal{I}_{i+1}^{(1)}(e_S^{Y,i+1}) + \mathcal{I}_{i+1}^{(2)}(e_Y^{Y,i+1}) + \mathcal{I}_{i+1}^{(1,2)}(d_{Y,S}^{i+1}) \right], \\
 \vec{\theta}_{i+1}^{e,X} &= (f(\zeta_{i+1} - \zeta_i))^{-1} \mathcal{I}_{i+1}^{(1)}(e_S^{X,i+1}), \\
 \vec{\theta}_{i+1}^c &= \mathcal{I}_{i+1}^{(1)} \left(\partial_{\bar{Y}_i} \bar{X}_{i+1} \theta_{i+1} - \vec{\theta}_{i+1}^{e,X} \right) + \partial_{\bar{Y}_i} \theta_{i+1} \\
 &\quad + \mathcal{I}_{i+1}^{(2)} \left(m'_i \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} + \left(\sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) \right. \right. \\
 &\quad \left. \left. + \sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right) \right) \theta_{i+1} \right), \\
 d_S^{i+1} &= m'_i c_S^{i+1}, \\
 d_Y^{i+1} &= m'_i c_Y^{i+1}, \\
 d_{Y,S}^{i+1} &= m'_i c_{Y,S}^{i+1}, \\
 e_S^{Y,i+1} &= -m'_i c_S^{i+1} + \partial_{\bar{Y}_i} c_{Y,S}^{i+1}, \\
 e_Y^{Y,i+1} &= m'_i b_Y^{i+1} + \partial_{\bar{Y}_i} c_Y^{i+1}, \\
 e_S^{X,i+1} &= \partial_{\bar{Y}_i} c_S^{i+1}.
 \end{aligned}$$

For $i = n$, one also has

$$\begin{aligned}
 \partial_{\bar{Y}_n} \mathbb{E}_{n,n} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \theta_{n+1} \right] &= \mathbb{E}_{n,n} \left[\partial_{\bar{Y}_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \vec{\theta}_{n+1}^{e,Y} \right] + \mathbb{E}_{n,n} \left[\partial_{\bar{X}_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \vec{\theta}_{n+1}^{e,X} \right] \\
 &\quad + \mathbb{E}_{n,n} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \vec{\theta}_{n+1}^c \right] \tag{5.4.4}
 \end{aligned}$$

with

$$\begin{aligned}
 \vec{\theta}_{n+1}^{e,Y} &= (1 - F(T - \zeta_n))^{-1} \left(m'_n + \sigma'_{Y,n} \left(\rho_n Z_{n+1}^1 + \sqrt{1 - \rho_n^2} Z_{n+1}^2 \right) \right. \\
 &\quad \left. + \sigma_{Y,n} \frac{\rho'_n}{\sqrt{1 - \rho_n^2}} \left(\sqrt{1 - \rho_n^2} Z_{n+1}^1 - \rho_n Z_{n+1}^2 \right) \right), \\
 \vec{\theta}_{n+1}^{e,X} &= (1 - F(T - \zeta_n))^{-1} \left(-\frac{1}{2} a'_{S,n} + \sigma'_{S,n} Z_{n+1}^1 \right),
 \end{aligned}$$

and we set $\vec{\theta}_{n+1}^c = 0$ for notational convenience.

Finally, the weight sequences $(\vec{\theta}_i^{e,Y})_{1 \leq i \leq n+1}$, $(\vec{\theta}_i^{e,X})_{1 \leq i \leq n+1}$ and $(\vec{\theta}_i^c)_{1 \leq i \leq n+1}$ defined above satisfy

$$f(\zeta_i - \zeta_{i-1}) \vec{\theta}_i^{e,Y}, f(\zeta_i - \zeta_{i-1}) \vec{\theta}_i^{e,X}, f(\zeta_i - \zeta_{i-1}) \vec{\theta}_i^c \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1/2), \quad i \in \{1, \dots, n\}$$

$$\text{and } (1 - F(T - \zeta_n)) \vec{\theta}_{n+1}^{e,Y} \in \mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 0), (1 - F(T - \zeta_n)) \vec{\theta}_{n+1}^{e,X} \in \mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 1/2).$$

The proof of Lemma 5.4.1 is postponed to Appendix 5.6.2. The transfer of derivatives procedure starts on the first time interval $[0, \zeta_1]$ according to formulae (5.4.1) and (5.4.3) (for $i = 0$). It expresses the fact that the flow derivatives ∂_{s_0} and ∂_{y_0} of the conditional expectations on the left-hand side of the equations are transferred to derivative operators $\partial_{\bar{X}_1}$ and $\partial_{\bar{Y}_1}$ on the test function h appearing on the right-hand side. Remark

that the first derivatives of h have been written ubiquitously as $\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1})$ and $\partial_{\bar{Y}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1})$.

Then, by the Markov property satisfied by the process (\bar{X}, \bar{Y}) , the function h appearing inside the (conditional) expectations on the right-hand side of (5.4.1) and (5.4.3) (for $i = 0$) will be given by the conditional expectation appearing on the left-hand side of the same equations but for $i = 1$. The transfer of derivative formulae for the following time intervals are obtained by induction using (5.4.2) and (5.4.3) up to the last time interval. Doing so, we obtain various transfer of derivatives formulae by transferring successively the derivative operators through all intervals forward in time.

5.4.2 The integration by parts formulae

We first define the weights that will be used in our IBP formulae. For an integer n , on the set $\{N_T = n\}$, for any $k \in \{1, \dots, n+1\}$ and any $j \in \{1, \dots, k\}$, we define

$$\begin{aligned} \vec{\theta}^{\mathcal{I}_k^{(1),n+1}} &:= \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\theta_k) \times \prod_{i=1}^{k-1} \theta_i, \\ \vec{\theta}^{C_j^{n+1}} &:= \prod_{i=j+1}^{n+1} \theta_i \times \vec{\theta}_j^c \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y}, \\ \vec{\theta}^{\mathcal{I}_k^{(2),n+1}} &:= \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y}, \\ \vec{\theta}_j^{\mathcal{I}_k^{(1),n+1}} &:= \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\theta_k) \times \prod_{i=j+1}^{k-1} \theta_i \times \vec{\theta}_j^{e,X} \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y}, \quad j = 1, \dots, k-1, \\ \vec{\theta}_k^{\mathcal{I}_k^{(1),n+1}} &:= \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y}. \end{aligned}$$

With the above definitions at hand, we are now able to state our IBP formulae.

Theorem 5.4.1. *Let $T > 0$. Under assumptions **AR** and **ND**, the law of the couple (X_T, Y_T) , given by the unique solution to the SDE (5.2.1) at time T starting from $(x_0 = \ln(s_0), y_0)$ at time 0, satisfies the following Bismut-Elworthy-Li IBP formulae: there exists some positive constant $c := c(T, \text{by}, \kappa)$ such that for any $0 \leq \gamma < c^{-1}$ and any $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, for any $(s_0, y_0) \in (0, \infty) \times \mathbb{R}$, it holds*

$$s_0 T \partial_{s_0} \mathbb{E} \left[h(X_T, Y_T) \right] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \vec{\theta}^{\mathcal{I}_k^{(1),N_T+1}} \right] \quad (5.4.5)$$

and

$$T \partial_{y_0} \mathbb{E} \left[h(X_T, Y_T) \right] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\vec{\theta}^{\mathcal{I}_k^{(2),N_T+1}} + \sum_{j=1}^k \vec{\theta}_j^{C_j^{N_T+1}} + \vec{\theta}_j^{\mathcal{I}_k^{(1),N_T+1}} \right) \right]. \quad (5.4.6)$$

Moreover, if N is a renewal process with Beta($\alpha, 1$) jump times, then, for any $p \geq 1$ satisfying $p(\frac{1}{2} - \alpha) \leq 1 - \alpha$, for any γ such that $0 \leq p\gamma < c^{-1}$ and any $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, the random variables appearing inside the expectation in the right-hand side of (5.4.5) and (5.4.6) admit a finite $L^p(\mathbb{P})$ -moment. In particular, if $\alpha = 1/2$ then for any $p \geq 1$, the $L^p(\mathbb{P})$ -moment is finite.

Proof. We only prove the IBP formula (5.4.6). The proof of (5.4.5) follows by completely analogous (and actually more simple) arguments and is thus omitted.

Step 1: proof of the IBP formula (5.4.6) for $h \in \mathcal{C}_b^1(\mathbb{R}^2)$.
Let $h \in \mathcal{C}_b^1(\mathbb{R}^2)$. From Theorem 5.3.1, we write

$$\mathbb{E}[h(X_T, Y_T)] = \sum_{n \geq 0} \mathbb{E} \left[\mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1} \right] \mathbf{1}_{\{N_T=n\}} \right] \quad (5.4.7)$$

using the fact that $\{N_T = n\} = \{\tau_{n+1} > T\} \cap \{\tau_n \leq T\}$. In most of the arguments below, we will work on the set $\{N_T = n\}$. In order to perform our induction argument forward in time through the Markov chain structure, we define for $k \in \{0, \dots, n+1\}$ the functions

$$H_k(\bar{X}_k, \bar{Y}_k) := \mathbb{E}_{k,n} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \right] = \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \middle| \bar{X}_k, \bar{Y}_k, \tau^{n+1}, N_T = n \right].$$

We also let $H_{n+1}(\bar{X}_{n+1}, \bar{Y}_{n+1}) := h(\bar{X}_{n+1}, \bar{Y}_{n+1})$. Note that we omit the dependence w.r.t the sequence τ^{n+1} in the definition of the (random) maps $(H_k)_{0 \leq k \leq n+1}$. From the above definition and using **ND**, **AR**, it follows that the map H_k belongs to $\mathcal{C}_p^1(\mathbb{R}^2)$ a.s. for any $0 \leq k \leq n+1$. Moreover, from the tower property of conditional expectation the following relation is satisfied for any $k \in \{0, \dots, n\}$

$$H_k(\bar{X}_k, \bar{Y}_k) = \mathbb{E}_{k,n} [H_{k+1}(\bar{X}_{k+1}, \bar{Y}_{k+1}) \theta_{k+1}].$$

Now, using first the Lebesgue differentiation theorem and then iterating the transfer of derivative formula (5.4.3) in Lemma 5.4.1, we obtain² for any $k \in \{1, \dots, n\}$,

$$\begin{aligned} & \partial_{y_0} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1} \right] \quad (5.4.8) \\ &= \partial_{y_0} \mathbb{E} \left[\mathbb{E}_{0,n} [H_1(\bar{X}_1, \bar{Y}_1) \theta_1] \middle| \tau^{n+1} \right] = \mathbb{E} \left[\partial_{y_0} \mathbb{E}_{0,n} [H_1(\bar{X}_1, \bar{Y}_1) \theta_1] \middle| \tau^{n+1} \right] \\ &= \mathbb{E} \left[\mathcal{D}_k^{(2)} H_k(\bar{X}_k, \bar{Y}_k) \prod_{i=1}^k \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] + \sum_{j=1}^k \mathbb{E} \left[H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^c \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\ & \quad + \sum_{j=1}^k \mathbb{E} \left[\mathcal{D}_j^{(1)} H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right]. \end{aligned}$$

²As before, we use the convention $\sum_{\emptyset} \dots = 0$, $\prod_{\emptyset} \dots = 1$.

To further simplify the first term appearing on the right-hand side of (5.4.8), we use the tower property of conditional expectation (w.r.t $\mathbb{E}_{k-1,n}[\cdot]$) and the integration by parts formula (5.2.9). For any $k \in \{1, \dots, n\}$, we obtain

$$\mathbb{E}\left[\mathcal{D}_k^{(2)} H_k(\bar{X}_k, \bar{Y}_k) \vec{\theta}_k^{e,Y} \middle| \mathcal{G}_{k-1}, \tau^{n+1}\right] = \mathbb{E}\left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \middle| \mathcal{G}_{k-1}, \tau^{n+1}\right].$$

We also simplify the third term appearing on the right-hand side of (5.4.8), by using the transfer of derivatives formula (5.4.2) up to the time interval $[\zeta_{k-1}, \zeta_k]$. For any $j \in \{1, \dots, k\}$, it holds

$$\mathbb{E}\left[\mathcal{D}_j^{(1)} H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right] = \mathbb{E}\left[\mathcal{D}_k^{(1)} H_k(\bar{X}_k, \bar{Y}_k) \prod_{i=j+1}^k \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right]$$

so that, if $j \in \{1, \dots, k-1\}$, taking conditional expectation (using again $\mathbb{E}_{k-1,n}[\cdot]$) and then performing an IBP formula on the last time interval $[\zeta_{k-1}, \zeta_k]$ yield

$$\mathbb{E}\left[\mathcal{D}_k^{(1)} H_k(\bar{X}_k, \bar{Y}_k) \prod_{i=j+1}^k \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right] = \mathbb{E}\left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(1)}(\theta_k) \prod_{i=j+1}^{k-1} \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right].$$

while if $j = k$, we obtain

$$\mathbb{E}\left[\mathcal{D}_k^{(1)} H_k(\bar{X}_k, \bar{Y}_k) \prod_{i=j+1}^k \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right] = \mathbb{E}\left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X}) \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1}\right].$$

Coming back to (5.4.8) and using the definition of the maps $(H_k)_{0 \leq k \leq n+1}$, we thus

deduce

$$\begin{aligned}
 & \partial_{y_0} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1} \right] \tag{5.4.9} \\
 &= \mathbb{E} \left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{j=1}^k \mathbb{E} \left[H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^c \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{j=1}^{k-1} \mathbb{E} \left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(1)}(\theta_k) \times \prod_{i=j+1}^{k-1} \theta_i \times \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \mathbb{E} \left[H_k(\bar{X}_k, \bar{Y}_k) \mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X}) \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &= \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{j=1}^k \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=j+1}^{n+1} \theta_i \times \vec{\theta}_j^c \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{j=1}^{k-1} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\theta_k) \times \prod_{i=j+1}^{k-1} \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X}) \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right].
 \end{aligned}$$

In the case $k = n + 1$, using the transfer of derivative formula (5.4.4) of Lemma 5.4.1 on the last time interval and then performing the IBP formula (5.2.9), we obtain

the representation

$$\begin{aligned}
& \partial_{y_0} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1} \right] \\
&= \mathbb{E} \left[\mathcal{D}_{n+1}^{(2)} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] + \sum_{j=1}^{n+1} \mathbb{E} \left[H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^c \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&\quad + \sum_{j=1}^{n+1} \mathbb{E} \left[\mathcal{D}_j^{(1)} H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&= \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathcal{I}_{n+1}^{(2)}(\vec{\theta}_{n+1}^{e,Y}) \prod_{i=1}^n \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] + \sum_{j=1}^{n+1} \mathbb{E} \left[H_j(\bar{X}_j, \bar{Y}_j) \vec{\theta}_j^c \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&\quad + \sum_{j=1}^{n+1} \mathbb{E} \left[\mathcal{D}_{n+1}^{(1)} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=j+1}^{n+1} \theta_i \vec{\theta}_j^{e,X} \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&= \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathcal{I}_{n+1}^{(2)}(\vec{\theta}_{n+1}^{e,Y}) \times \prod_{i=1}^n \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&\quad + \sum_{j=1}^{n+1} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=j+1}^{n+1} \theta_i \times \vec{\theta}_j^c \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&\quad + \sum_{j=1}^n \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathcal{I}_{n+1}^{(1)}(\theta_{n+1}) \times \prod_{i=j+1}^n \theta_i \times \vec{\theta}_j^{e,X} \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
&\quad + \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathcal{I}_{n+1}^{(1)}(\vec{\theta}_{n+1}^{e,X}) \times \prod_{i=1}^n \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \tag{5.4.10}
\end{aligned}$$

where, for the last term appearing in the right-hand side of the above identities, we employed the transfer of derivative formula (5.4.2) up to the last time interval and then performed an IBP formula.

Now, the key point in order to establish the IBP formula is to combine in a suitable way the identities (5.4.9) and (5.4.10). For each $k \in \{0, \dots, n\}$, we multiply the above formulae by the length of the interval on which the local IBP formula is performed, namely we multiply by $\zeta_k - \zeta_{k-1}$ both sides of (5.4.9), $k = 1, \dots, n-1$, and we multiply by $T - \zeta_n$ both side of (5.4.10). We then sum them over all k . Recalling that

$\sum_{k=1}^{n+1} \zeta_k - \zeta_{k-1} = T - \zeta_0 = T$, we deduce

$$\begin{aligned}
 & T \partial_{y_0} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1} \right] \\
 &= \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=j+1}^{n+1} \theta_i \times \vec{\theta}_j^c \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \\
 &+ \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \left(\sum_{j=1}^{k-1} \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\theta_k) \times \prod_{i=j+1}^{k-1} \theta_i \times \vec{\theta}_j^{e,X} \times \prod_{i=1}^{j-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \right. \\
 &\left. + \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X}) \times \prod_{i=1}^n \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \right) \\
 &= \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \left(\vec{\theta}^{\mathcal{I}_k^{(2),n+1}} + \sum_{j=1}^k \vec{\theta}^{C_j^{n+1}} + \vec{\theta}^{\mathcal{I}_j^{(1),n+1}} \right) \middle| \tau^{n+1} \right].
 \end{aligned}$$

We now provide a sharp upper-estimate for the above quantity. From Lemma 5.7.2 and Lemma 5.4.1, it follows that $f(\zeta_i - \zeta_{i-1})\theta_i$, $f(\zeta_i - \zeta_{i-1})\vec{\theta}_i^{e,Y}$, $f(\zeta_i - \zeta_{i-1})\vec{\theta}_i^c \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1/2)$ and $f(\zeta_i - \zeta_{i-1})\vec{\theta}_i^{e,X} \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 0)$ for any $i \in \{1, \dots, n\}$. Moreover, from the very definition of the weights θ_i , $\vec{\theta}_i^{e,X}$ and $\vec{\theta}_i^{e,Y}$, after some simple but cumbersome computations that we omit, one has $f(\zeta_i - \zeta_{i-1})\mathcal{D}_i^{(1)}(\theta_i)$, $f(\zeta_i - \zeta_{i-1})\mathcal{D}_i^{(2)}(\vec{\theta}_i^{e,Y}) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1)$ and $f(\zeta_i - \zeta_{i-1})\mathcal{D}_i^{(1)}(\vec{\theta}_i^{e,X}) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1/2)$ so that from Lemma 5.2.3 we conclude $f(\zeta_i - \zeta_{i-1})(\zeta_i - \zeta_{i-1})\mathcal{I}_i^{(1)}(\theta_i) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 0)$, $f(\zeta_i - \zeta_{i-1})(\zeta_i - \zeta_{i-1})\mathcal{I}_i^{(2)}(\vec{\theta}_i^{e,Y}) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 0)$ and $f(\zeta_i - \zeta_{i-1})(\zeta_i - \zeta_{i-1})\mathcal{I}_i^{(1)}(\vec{\theta}_i^{e,X}) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 1/2)$. Hence, from the boundedness of h , the tower property of conditional expectation and (5.2.17), it holds

$$\begin{aligned}
 & \left| (\zeta_k - \zeta_{k-1}) \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \right| \\
 & \leq C^{n+1} (1 - F(T - \zeta_n))^{-1} \prod_{i=1}^{k-1} (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \\
 & \quad \times (f(\zeta_k - \zeta_{k-1}))^{-1} \prod_{i=k+1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}}
 \end{aligned}$$

so that using (5.2.3)

$$\begin{aligned}
& \sum_{n \geq 0} \mathbb{E} \left[\sum_{k=1}^{n+1} \left| (\zeta_k - \zeta_{k-1}) \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=k+1}^{n+1} \theta_i \times \mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y}) \times \prod_{i=1}^{k-1} \vec{\theta}_i^{e,Y} \middle| \tau^{n+1} \right] \right| \mathbf{1}_{\{N_T=n\}} \right] \\
& \leq \sum_{n \geq 0} C^{n+1} \sum_{k=1}^{n+1} \mathbb{E} \left[(1 - F(T - \zeta_n))^{-1} (f(\zeta_k - \zeta_{k-1}))^{-1} \prod_{i=1, i \neq k}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \mathbf{1}_{\{N_T=n\}} \right] \\
& \leq \sum_{n \geq 0} C^{n+1} \sum_{k=1}^{n+1} \int_{\Delta_n(T)} \prod_{i=1, i \neq k}^n (s_i - s_{i-1})^{-1/2} d\sigma_n \\
& \leq \sum_{n \geq 0} (n+1) C^{n+1} T^{(n+1)/2} \frac{\Gamma^n(1/2)}{\Gamma(1+n/2)} < \infty.
\end{aligned}$$

From similar arguments that we omit, it follows

$$\begin{aligned}
& \left| (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \left(\vec{\theta}^{C_j^{n+1}} + \vec{\theta}_j^{T_k^{(1),n+1}} \right) \middle| \tau^{n+1} \right] \right| \\
& \leq (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k (1 - F(T - \zeta_n))^{-1} \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} [1 + \mathbf{1}_{\{i=k\}} (\zeta_i - \zeta_{i-1})^{-1/2}].
\end{aligned}$$

so that using again (5.2.3)

$$\begin{aligned}
& \sum_{n \geq 0} \mathbb{E} \left[\sum_{k=1}^{n+1} \left| (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \left(\vec{\theta}^{C_j^{n+1}} + \vec{\theta}_j^{T_k^{(1),n+1}} \right) \middle| \tau^{n+1} \right] \right| \mathbf{1}_{\{N_T=n\}} \right] \\
& \leq \sum_{n \geq 0} C^{n+1} \sum_{k=1}^{n+1} \mathbb{E} \left[(\zeta_k - \zeta_{k-1}) \sum_{j=1}^k (1 - F(T - \zeta_n))^{-1} \right. \\
& \quad \times \left. \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} [1 + \mathbf{1}_{\{i=k\}} (\zeta_i - \zeta_{i-1})^{-1/2}] \mathbf{1}_{\{N_T=n\}} \right] \\
& \leq \sum_{n \geq 0} C^{n+1} (n+1)(n+2) T^{(1+n)/2} \frac{\Gamma^n(1/2)}{\Gamma(1+n/2)} < \infty.
\end{aligned}$$

The preceding estimates combined with (5.4.7) and the Lebesgue dominated convergence theorem allows to conclude that $y_0 \mapsto \mathbb{E}[h(X_T, Y_T)]$ is continuously differentiable

with

$$\begin{aligned}
 T\partial_{y_0}\mathbb{E}[h(X_T, Y_T)] &= T\partial_{y_0}\mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i\right] \\
 &= \sum_{n \geq 0} \mathbb{E}\left[T\partial_{y_0}\mathbb{E}\left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| \tau^{n+1}\right] \mathbf{1}_{\{N_T=n\}}\right] \\
 &= \sum_{n \geq 0} \mathbb{E}\left[\mathbb{E}\left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \left(\overrightarrow{\theta}^{\mathcal{I}_k^{(2), n+1}} + \sum_{j=1}^k \overrightarrow{\theta}^{C_j^{n+1}} + \overrightarrow{\theta}^{\mathcal{I}_j^{(1), n+1}}\right) \middle| \tau^{n+1}\right] \mathbf{1}_{\{N_T=n\}}\right] \\
 &= \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\overrightarrow{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \overrightarrow{\theta}^{C_j^{n+1}} + \overrightarrow{\theta}^{\mathcal{I}_j^{(1), N_T+1}}\right)\right]
 \end{aligned}$$

where we used Fubini's theorem for the last equality. This completes the proof of the IBP formula (5.4.6) for $h \in \mathcal{C}_b^1(\mathbb{R}^2)$.

Step 2: Extension to $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ for some positive γ .

We now extend the two IBP formulae that we have established in the previous step to the case of a test function $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ for some sufficiently small $\gamma > 0$. Let us note that under assumption **H**, from Kusuoka and Stroock [KS85], Corollary (3.25) and the upper-estimate (3.27) therein, the process $(X_t, Y_t)_{t \geq 0}$ admits a smooth transition density $(t, x_0, y_0, x, y) \mapsto p(t, x_0, y_0, x, y) \in \mathcal{C}^\infty((0, \infty) \times \mathbb{R}^2 \times \mathbb{R}^2)$ and for any $h \in \mathcal{C}_b^1(\mathbb{R}^2)$, it holds

$$\partial_{s_0}^\alpha \partial_{y_0}^\beta \mathbb{E}[h(X_T, Y_T)] = \int_{\mathbb{R}^2} h(x, y) \partial_{s_0}^\alpha \partial_{y_0}^\beta p(T, x_0, y_0, x, y) dx dy$$

for any $T > 0$ and any integers α and β .

We then proceed as in step 2 of the proof of Theorem 5.3.1. Namely, we prove that

$$\begin{aligned}
 &T\partial_{y_0}\mathbb{E}[h(X_T, Y_T)] \\
 &= \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\overrightarrow{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \overrightarrow{\theta}^{C_j^{n+1}} + \overrightarrow{\theta}^{\mathcal{I}_j^{(1), N_T+1}}\right)\right] \\
 &= \int_{\mathbb{R}^2} h(x, y) \mathbb{E}\left[\bar{p}(T - \zeta_{N_T}, \bar{X}_{N_T}, \bar{Y}_{N_T}, x, y) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \right. \\
 &\quad \left. \times \left(\overrightarrow{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \overrightarrow{\theta}^{C_j^{n+1}} + \overrightarrow{\theta}^{\mathcal{I}_j^{(1), N_T+1}}\right)\right] dx dy \tag{5.4.11}
 \end{aligned}$$

for any $h \in \mathcal{C}_b^1(\mathbb{R}^2)$.

Indeed, since $f(\zeta_i - \zeta_{i-1})\theta_i \in \mathbb{M}_{i-1, n}(\bar{X}, \bar{Y}, -1/2)$ and $f(\zeta_k - \zeta_{k-1})\mathcal{I}_k^{(2)}(\overrightarrow{\theta}_k^{e, Y}) \in$

$\mathbb{M}_{k-1,n}(\bar{X}, \bar{Y}, -1)$, for some $c := c(T, b_Y, \kappa) > 4\kappa$, it holds

$$\begin{aligned}
& \mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \left| \vec{\theta}^{\mathcal{I}_k^{(2),n+1}} \right| \middle| \tau^{n+1} \right] \\
& \leq C^{n+1} \int_{(\mathbb{R}^2)^n} \bar{q}_{4\kappa}(T - \zeta_n, x_n, y_n, x, y) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) (1 - F(T - \zeta_n))^{-1} \\
& \quad \times (f(\zeta_k - \zeta_{k-1}))^{-1} (\zeta_k - \zeta_{k-1})^{-1} \prod_{i=1}^{k-1} (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} \\
& \quad \times \prod_{i=k+1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} \prod_{i=1}^n \bar{q}_{4\kappa}(\zeta_i - \zeta_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) dx_n dy_n \\
& \leq C^{n+1} \bar{q}_c(T, x_0, y_0, x, y) \sum_{k=1}^{n+1} (1 - F(T - \zeta_n))^{-1} \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} \prod_{i=1, i \neq k}^n (\zeta_i - \zeta_{i-1})^{-1/2}
\end{aligned} \tag{5.4.12}$$

where, for the first inequality we used the upper-estimate (5.2.2) and for the last inequality we used Lemma 5.7.3. From similar arguments, one gets

$$\begin{aligned}
& \mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k \left| \vec{\theta}^{C_j^{n+1}} \right| + \left| \vec{\theta}_j^{\mathcal{I}_k^{(1),N_T+1}} \right| \middle| \tau^{n+1} \right] \\
& \leq C^{n+1} \bar{q}_c(T, x_0, y_0, x, y) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k (1 - F(T - \zeta_n))^{-1} \\
& \quad \times \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} [1 + \mathbf{1}_{\{i=k\}} (\zeta_i - \zeta_{i-1})^{-1/2}].
\end{aligned} \tag{5.4.13}$$

Now, from the upper-bounds (5.4.12) and (5.4.13) as well as the identity (5.2.3), we conclude

$$\begin{aligned}
& \sum_{n \geq 0} \mathbb{E} \left[\mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \left(\left| \vec{\theta}^{\mathcal{I}_k^{(2),n+1}} \right| + \sum_{j=1}^k \left| \vec{\theta}_j^{C_j^{n+1}} \right| + \left| \vec{\theta}_j^{\mathcal{I}_k^{(1),n+1}} \right| \right) \middle| \tau^{n+1} \right] \mathbf{1}_{\{N_T=n\}} \right] \\
& \leq \bar{q}_c(T, x_0, y_0, x, y) \sum_{n \geq 0} C^{n+1} \mathbb{E} \left[\sum_{k=1}^{n+1} (1 - F(T - \zeta_n))^{-1} \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} \prod_{i=1, i \neq k}^n (\zeta_i - \zeta_{i-1})^{-1/2} \right. \\
& \quad \left. + \sum_{k=1}^{n+1} (\zeta_k - \zeta_{k-1}) \sum_{j=1}^k (1 - F(T - \zeta_n))^{-1} \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-1/2} [1 + \mathbf{1}_{\{i=k\}} (\zeta_i - \zeta_{i-1})^{-1/2}] \right] \\
& \leq \bar{q}_c(T, x_0, y_0, x, y) \sum_{n \geq 0} C^{n+1} [(n+1) + (n+1)(n+2)/2] T^{(1+n)/2} \frac{\Gamma^n(1/2)}{\Gamma(1+n/2)} \\
& = CT^{1/2} \bar{q}_c(T, x_0, y_0, x, y).
\end{aligned} \tag{5.4.14}$$

From the preceding inequality and Fubini's theorem, we thus get

$$\begin{aligned} & \left| \mathbb{E} \left[\bar{p}(T - \zeta_{N_T}, \bar{X}_{N_T}, \bar{Y}_{N_T}, x, y) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\vec{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \vec{\theta}^{C_j^{N_T+1}} + \vec{\theta}_j^{\mathcal{I}_k^{(1), N_T+1}} \right) \right] \right| \\ & \leq CT^{1/2} \bar{q}_c(T, x_0, y_0, x, y) \end{aligned} \quad (5.4.15)$$

for some positive constant $C := C(T)$ such that $T \mapsto C(T)$ is non-decreasing. Applying again Fubini's theorem allows to complete the proof of (5.4.11). Hence,

$$\begin{aligned} T \partial_{y_0} \mathbb{E}[h(X_T, Y_T)] &= \int_{\mathbb{R}^2} h(x, y) \partial_{y_0} p(T, x_0, y_0, x, y) dx dy \\ &= \int_{\mathbb{R}^2} h(x, y) \mathbb{E} \left[\bar{p}(T - \zeta_{N_T}, \bar{X}_{N_T}, \bar{Y}_{N_T}, x, y) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\vec{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \vec{\theta}^{C_j^{N_T+1}} + \vec{\theta}_j^{\mathcal{I}_k^{(1), N_T+1}} \right) \right] \end{aligned}$$

for any $h \in \mathcal{C}_b^1(\mathbb{R}^2)$. A monotone class argument allows to conclude that the preceding identity is still valid for any bounded and measurable map h defined over \mathbb{R}^2 and a standard approximation argument allows to extend it to $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$ for $\gamma < c^{-1}$, c being the positive constant appearing in (5.4.15). We eventually conclude from the preceding identity, (5.4.14) combined with Fubini's theorem that

$$T \partial_{y_0} \mathbb{E}[h(X_T, Y_T)] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \left(\vec{\theta}^{\mathcal{I}_k^{(2), N_T+1}} + \sum_{j=1}^k \vec{\theta}^{C_j^{N_T+1}} + \vec{\theta}_j^{\mathcal{I}_k^{(1), N_T+1}} \right) \right]$$

for any $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, with $\gamma < c^{-1}$.

Step 3: $L^p(\mathbb{P})$ -moments for a renewal process with Beta jump times.

From the above formula, the proof of the $L^p(\mathbb{P})$ -moment estimate when N is a renewal process with Beta jump times follows by similar arguments as those employed at step 3 of the proof of Theorem 5.3.1. We omit the remaining technical details. \square

5.5 Numerical Results

In this section, we provide some numerical results for the unbiased Monte Carlo algorithm that stems from the probabilistic representation formula established in Theorem 5.3.1 and the Bismut-Elworthy-Li formulae of Theorem 5.4.1 for the couple (S_T, Y_T) that allows to compute the Delta and the Vega related to the option price of the vanilla option with payoff $h(S_T)$. We will consider the unique strong solution associated to the SDE (5.1.1) for three different models corresponding to three different diffusion coefficient function σ_S and two different payoff functions: Call and Digital call options with maturity T and strike K , namely $h(x, y) = (\exp(x) - K)_+$ and $h(x, y) = \mathbf{1}_{\{\exp(x) \geq K\}}$. For these three models, the drift function of the volatility process is defined by $b_Y(x) = \lambda(\mu - x)$ and the parameters are selected as follows: $T = 0.5$, $r = 0.03$, $K = 1.5$, $x_0 = \ln(s_0) = 0.4$, $Y_0 = 0.2$, $\sigma_Y(\cdot) \equiv \sigma_Y = 0.2$, $\lambda = 0.5$, $\mu = 0.3$ and $\rho = 0.6$. We also consider two type of renewal process N : a Poisson process with intensity parameter $\lambda = 0.5$ and a renewal process with $Beta(1 - \alpha, 1)$ jump times with parameters $\alpha = 0.1$ and $\bar{\tau} = 2$.

5.5.1 Black-Scholes Model

We first consider the simple (toy) example corresponding to Black-Scholes dynamics

$$dS_t = rS_t dt + \sigma_S S_t dW_t, \quad dY_t = b_Y(Y_t) dt + \sigma_Y(Y_t) dB_t, \quad d\langle B, W \rangle_t = \rho dt, \quad \rho \in (-1, 1).$$

with constant diffusion coefficient function $\sigma_S(\cdot) \equiv \sigma_S > 0$. The law of (S_T, Y_T) can be computed explicitly so that analytical formulas are available for the price, Delta and Vega. Note that the discount factor e^{-rT} has been added in our probabilistic representation formula for comparison purposes. Then the probabilistic representation in (5.3.2) of Theorem 5.3.1 and (5.4.5), (5.4.6) of Theorem 5.4.1 then becomes

$$\begin{aligned} \mathbb{E}[(S_T - K)_+] &= \mathbb{E}[(e^{X_T} - K)_+] = \mathbb{E}[h(X_T, Y_T)] = \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i\right] \\ \partial_{s_0} \mathbb{E}[(S_T - K)_+] &= \frac{1}{s_0 T} \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \left(\prod_{i=1}^{N_T+1} \theta_i\right) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \frac{\mathcal{I}_k^{(1)} \theta_k}{\theta_k}\right] \\ \partial_{y_0} \mathbb{E}[(S_T - K)_+] &= \frac{1}{T} \mathbb{E}\left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \left(\prod_{i=1}^{N_T+1} \theta_i\right) \sum_{k=1}^{N_T+1} (\zeta_k - \zeta_{k-1}) \right. \\ &\quad + \left(\frac{\mathcal{I}_k^{(2)}(\vec{\theta}_k^{e,Y})}{\theta_k} \times \prod_{i=1}^{k-1} \frac{\vec{\theta}_i^{e,Y}}{\theta_i} + \sum_{j=1}^k \frac{\vec{\theta}_j^c}{\theta_j} \times \prod_{i=1}^{j-1} \frac{\vec{\theta}_i^{e,Y}}{\theta_i} \right. \\ &\quad \left. \left. + \frac{\mathcal{I}_k^{(1)}(\theta_k)}{\theta_k} \times \sum_{j=1}^{k-1} \left(\frac{\vec{\theta}_j^{e,X}}{\theta_j} \times \prod_{i=1}^{j-1} \frac{\vec{\theta}_i^{e,Y}}{\theta_i}\right) + \frac{\mathcal{I}_k^{(1)}(\vec{\theta}_k^{e,X})}{\theta_k} \times \prod_{i=1}^{k-1} \frac{\vec{\theta}_i^{e,Y}}{\theta_i}\right)\right] \end{aligned}$$

with

$$\begin{aligned} \bar{X}_{i+1} &= \bar{X}_i + \left(r(\zeta_{i+1} - \zeta_i) - \frac{1}{2} a_{S,i}\right) + \sigma_{S,i} Z_{i+1}^1, \\ \bar{Y}_{i+1} &= m_i + \sigma_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2\right), \\ m_i &= m_{\zeta_{i+1} - \zeta_i}(\bar{Y}_i) = \mu + (\bar{Y}_i - \mu) e^{-\lambda(\zeta_{i+1} - \zeta_i)}, \\ \theta_i &= (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{I}_i^{(1)}(c_S^i) + \mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{I}_i^{(1,2)}(c_{Y,S}^i)\right], \quad 1 \leq i \leq N_T, \\ \theta_{N_T+1} &= (1 - F(T - \zeta_{N_T}))^{-1}. \end{aligned}$$

We perform $M_1 = 10^7$ Monte Carlo path simulations to approximate the price as well as the two Greeks and compare them with the corresponding values obtained using the standard Monte Carlo method combined with an Euler-Maruyama approximation scheme for the dynamics (5.1.1) with $M_2 = 160000$ Monte Carlo simulations paths and mesh size $\delta = T/n$ where $n = 200$. The Delta and Vega are obtained using the Monte Carlo finite difference approach combined with the Euler-Maruyama discretization scheme, that is, denoting by $E_{M_2}^n(s_0, y_0)$ the Monte Carlo algorithm with Euler-Maruyama scheme, we compute $(E_{M_2}^n(s_0 + \varepsilon, y_0) - E_{M_2}^n(s_0, y_0))/\varepsilon$ and

$(E_{M_2}^n(s_0, y_0 + \varepsilon) - E_{M_2}^n(s_0, y_0))/\varepsilon$ respectively with $\varepsilon = 10^{-2}$. The numerical results for the three different quantities are summarized in Table 5.1, Table 5.2, Table 5.3 respectively. The first column provides the value of the parameter σ_S . The second column stands for the value of the price, Delta or Vega obtained by the corresponding Black-Scholes formula. The third and fourth columns correspond to the value obtained by the Euler-Maruyama discretization scheme together with its 95% confidence interval. The fifth and sixth (resp. seventh and eighth) columns provide the estimated value with its 95% confidence interval by our method in the case of Exponential sampling (resp. Beta sampling). We observe a good behaviour of the unbiased estimators for all three quantities and for all the values of the parameter σ_S .

σ_S	B-S formula	Euler Scheme		Exponential sampling		Beta sampling	
		Price	95% CI	Price	95% CI	Price	95% CI
0.25	0.111804	0.111467	[0.110699, 0.112235]	0.112285	[0.111781, 0.112789]	0.112159	[0.111734, 0.112584]
0.3	0.132621	0.13293	[0.132337, 0.133524]	0.133054	[0.132394, 0.133713]	0.132954	[0.132482, 0.133425]
0.4	0.174152	0.17392	[0.173103, 0.174737]	0.175346	[0.174557, 0.176135]	0.174584	[0.173912, 0.175255]
0.6	0.256572	0.258063	[0.256727, 0.259399]	0.255934	[0.254592, 0.257277]	0.256514	[0.255419, 0.257608]

Table 5.1 – Comparison between the unbiased Monte Carlo estimation and the Monte Carlo Euler-Maruyama scheme for the price of a Call option in the Black-Scholes model for different values of σ_S .

σ_S	B-S formula	Euler Scheme		Exponential sampling		Beta sampling	
		Delta	95% CI	Delta	95% CI	Delta	95% CI
0.25	0.556589	0.555686	[0.553178, 0.558194]	0.554613	[0.551336, 0.557891]	0.556314	[0.553141, 0.559488]
0.3	0.560018	0.561099	[0.559455, 0.562742]	0.557398	[0.553517, 0.56128]	0.557561	[0.554848, 0.560274]
0.4	0.569512	0.570293	[0.568533, 0.572053]	0.569098	[0.565706, 0.572489]	0.56731	[0.564279, 0.570341]
0.6	0.592743	0.594988	[0.592957, 0.59702]	0.586245	[0.582428, 0.590062]	0.588015	[0.584572, 0.591457]

Table 5.2 – Comparison between the unbiased Monte Carlo estimation and the Monte Carlo Euler-Maruyama scheme for the Delta of a Call option in the Black-Scholes model for different values of σ_S .

Before continuing with the other two examples, let us compute the following quantities that are useful for both examples, always under the assumption that the volatility of volatility $\sigma_Y(y) = \sigma_Y > 0$ is constant, and the remaining useful but cumbersome

σ_S	B-S formula	Exponential sampling		Beta sampling	
		Vega	95% CI	Vega	95% CI
0.25	0	0.000745386	[-0.00102979, 0.00252057]	-0.000438032	[-0.00211468, 0.00123862]
0.3	0	-0.0013932	[-0.0036299, 0.000843502]	-0.000491083	[-0.00249688, 0.00151471]
0.4	0	0.00331309	[0.000258292, 0.00636788]	-0.00117019	[-0.00393975, 0.00159938]
0.6	0	-0.00286877	[-0.00777679, 0.00203925]	-0.0027807	[-0.00718374, 0.00162235]

Table 5.3 – Comparison between the unbiased Monte Carlo estimation for the Vega of a Call option in the Black-Scholes model for different values of σ_S .

computations are postponed in Appendix 5.7.2.

$$\begin{aligned}
 m_s(\bar{Y}_i) &= \mu + (\bar{Y}_i - \mu)e^{-\lambda s}, \quad m'_i = e^{-\lambda(\zeta_{i+1} - \zeta_i)}, \\
 \partial_{\bar{Y}_i} \bar{X}_{i+1} &= -\sigma_{S,i} \sigma'_{S,i} + \sigma'_{S,i} Z_{i+1}^1, \quad \partial_{\bar{Y}_i} \bar{Y}_{i+1} = m'_i + \sigma_{Y,i} \rho'_i \left(Z_{i+1}^1 - \frac{\rho_i Z_{i+1}^2}{\sqrt{1 - \rho_i^2}} \right), \\
 \partial_{\bar{X}_{i+1}} (\partial_{\bar{Y}_i} \bar{X}_{i+1}) &= \frac{\sigma'_{S,i}}{\sigma_{S,i}}, \quad \partial_{\bar{Y}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} = -\frac{\rho_i \rho'_i}{1 - \rho_i^2}, \quad \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1} = \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}}, \\
 c_S^i &= \frac{1}{2} (\sigma_S(\bar{Y}_i)^2 - \sigma_S(m_{i-1})^2), \quad b_Y^i = \lambda(m_{i-1} - \bar{Y}_i), \\
 c_{Y,S}^i &= \rho \sigma_Y (\sigma_S(\bar{Y}_i) - \sigma_S(m_{i-1})), \quad d_S^{i+1} = m'_i c_S^{i+1}, \\
 d_{Y,S}^{i+1} &= m'_i c_{Y,S}^{i+1}, \quad e_S^{Y,i+1} = -m'_i c_S^{i+1} + \mathcal{D}_i^{(2)}(c_{Y,S}^{i+1}), \\
 e_Y^{Y,i+1} &= m'_i b_Y^{i+1}, \quad e_S^{X,i+1} = \mathcal{D}_i^{(2)}(c_S^{i+1}).
 \end{aligned}$$

And the following quantities are associated to the derivative operators

$$\begin{aligned}
 \mathcal{D}_i^{(2)}(c_S^i) &= \sigma_S(\bar{Y}_i) \sigma'_S(\bar{Y}_i), \\
 \mathcal{D}_i^{(2)}(c_S^{i+1}) &= \sigma_S(\bar{Y}_{i+1}) \sigma'_S(\bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1} - \sigma_S(m_i) \sigma'_S(m_i) m'_i, \\
 \mathcal{D}_{i+1}^{(1)} \mathcal{D}_i^{(2)}(c_S^{i+1}) &= \sigma_S(\bar{Y}_{i+1}) \sigma'_S(\bar{Y}_{i+1}) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1}, \\
 \mathcal{D}_i^{(2)}(b_Y^{i+1}) &= -\lambda(\partial_{\bar{Y}_i} \bar{Y}_{i+1} - m'_i), \\
 \mathcal{D}_i^{(2)}(c_{Y,S}^i) &= \rho \sigma_Y \sigma'_S(\bar{Y}_i), \\
 \mathcal{D}_i^{(2,2)}(c_{Y,S}^i) &= \rho \sigma_Y \sigma''_S(\bar{Y}_i), \\
 \mathcal{D}_i^{(2)}(c_{Y,S}^{i+1}) &= \rho \sigma_Y (\sigma'_S(\bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1} - m'_i \sigma'_S(m_i)), \\
 \mathcal{D}_i^{(2)}(\mathcal{D}_{i+1}^{(2)}(c_{Y,S}^{i+1})) &= \rho \sigma_Y \sigma''_S(\bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1}, \\
 \mathcal{D}_{i+1}^{(2)}(\mathcal{D}_i^{(2)}(c_{Y,S}^{i+1})) &= \rho \sigma_Y \sigma''_S(\bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1} + \rho \sigma_Y \sigma'_S(\bar{Y}_{i+1}) \partial_{\bar{Y}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1}, \\
 \mathcal{D}_{i+1}^{(1)}(\mathcal{D}_i^{(2)}(c_{Y,S}^{i+1})) &= \rho \sigma_Y \sigma'_S(\bar{Y}_{i+1}) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1}, \\
 \mathcal{D}_{i+1}^{(2)} \mathcal{D}_{i+1}^{(1)}(\mathcal{D}_i^{(2)}(c_{Y,S}^{i+1})) &= \rho \sigma_Y \sigma''_S(\bar{Y}_{i+1}) \partial_{\bar{X}_{i+1}} \partial_{\bar{Y}_i} \bar{Y}_{i+1},
 \end{aligned}$$

When the integral operators are added we have

$$\begin{aligned}\mathcal{I}_{i+1}^{(1)}(1) &= \frac{Z_{i+1}^1}{\sigma_{S,i}(1-\rho_i^2)} - \frac{\rho_i}{1-\rho_i^2} \frac{\rho_i Z_{i+1}^1 + \sqrt{1-\rho_i^2} Z_{i+1}^2}{\sigma_{S,i}}, \\ \mathcal{I}_{i+1}^{(2)}(1) &= \frac{\rho_i Z_{i+1}^1 + \sqrt{1-\rho_i^2} Z_{i+1}^2}{\sigma_{Y,i}(1-\rho_i^2)} - \frac{\rho_i}{1-\rho_i^2} \frac{Z_{i+1}^1}{\sigma_{Y,i}}, \\ \mathcal{D}_{i+1}^{(1)}\mathcal{I}_{i+1}^{(1)}(1) &= \frac{1}{(1-\rho_i^2)\sigma_{S,i}^2}, \\ \mathcal{D}_{i+1}^{(2)}\mathcal{I}_{i+1}^{(1)}(1) &= -\frac{\rho_i}{1-\rho_i^2} \frac{1}{\sigma_{S,i}\sigma_{Y,i}}, \\ \mathcal{D}_{i+1}^{(1)}\mathcal{I}_{i+1}^{(2)}(1) &= -\frac{\rho_i}{1-\rho_i^2} \frac{1}{\sigma_{S,i}\sigma_{Y,i}}, \\ \mathcal{D}_{i+1}^{(2)}\mathcal{I}_{i+1}^{(2)}(1) &= \frac{1}{(1-\rho_i^2)\sigma_{Y,i}^2}, \\ \mathcal{D}_i^{(2)}\mathcal{I}_{i+1}^{(1)}(1) &= -\frac{\sigma'_{S,i}}{\sigma_{S,i}}\mathcal{I}_{i+1}^{(1)}(1) - \frac{\rho'_i}{1-\rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}}\mathcal{I}_{i+1}^{(2)}(1), \\ \mathcal{D}_i^{(2)}\mathcal{I}_{i+1}^{(2)}(1) &= \frac{\rho'_i\rho_i}{1-\rho_i^2}\mathcal{I}_{i+1}^{(2)}(1) \\ \mathcal{D}_i^{(2)}\mathcal{D}_{i+1}^{(1)}\mathcal{I}_{i+1}^{(1)}(1) &= 2\mathcal{D}_{i+1}^{(1)}\mathcal{I}_{i+1}^{(1)}(1) \left(\frac{\rho_i\rho'_i}{1-\rho_i^2} - \frac{\sigma'_{S,i}}{\sigma_{S,i}} \right), \\ \mathcal{D}_i^{(2)}\mathcal{D}_{i+1}^{(2)}\mathcal{I}_{i+1}^{(1)}(1) &= \mathcal{D}_{i+1}^{(2)}\mathcal{I}_{i+1}^{(1)}(1) \left(\frac{\rho'_i(1+\rho_i^2)}{\rho_i(1-\rho_i^2)} - \frac{\sigma'_{S,i}}{\sigma_{S,i}} \right).\end{aligned}$$

Finally, let us calculate the quantities and their derivatives associated to our approximation process:

$$\begin{aligned}a_{S,i} &= \int_0^{\zeta_{i+1}-\zeta_i} a_S(m_s(\bar{Y}_i)) ds = \int_0^{\zeta_{i+1}-\zeta_i} \sigma_S^2(m_s(\bar{Y}_i)) ds, \\ a'_{S,i} &= \int_0^{\zeta_{i+1}-\zeta_i} a'_S(m_s(\bar{Y}_i)) ds = 2 \int_0^{\zeta_{i+1}-\zeta_i} m'_s(\bar{Y}_i)\sigma'_S(m_s(\bar{Y}_i))\sigma_S(m_s(\bar{Y}_i)) ds, \\ a_{Y,i} &= \sigma_Y^2(\zeta_{i+1}-\zeta_i), \quad \sigma_{Y,i} = \sigma_Y\sqrt{\zeta_{i+1}-\zeta_i}, \quad a'_{Y,i} = \sigma'_{Y,i} = 0 \\ \sigma_{S,i} &= \sqrt{a_{S,i}}, \quad \sigma'_{S,i} = \frac{a'_{S,i}}{2\sqrt{a_{S,i}}} = \frac{a'_{S,i}}{2\sigma_{S,i}}, \\ \rho_i &= \rho \frac{\int_0^{\zeta_{i+1}-\zeta_i} \sigma_S(m_s(\bar{Y}_i)) ds}{\sigma_{S,i}\sqrt{\zeta_{i+1}-\zeta_i}}, \\ \rho'_i &= \rho \frac{\sigma_{S,i} \int_0^{\zeta_{i+1}-\zeta_i} \sigma'_S(m_s(\bar{Y}_i))m'_s(\bar{Y}_i) ds - \sigma'_{S,i} \int_0^{\zeta_{i+1}-\zeta_i} \sigma_S(m_s(\bar{Y}_i)) ds}{a_{S,i}\sqrt{\zeta_{i+1}-\zeta_i}}.\end{aligned}$$

5.5.2 A Stein-Stein type model

In this second example, we consider a Stein-Stein type model where the diffusion coefficient function for the spot price is an affine function, namely $\sigma_S(x) = \sigma_1 x + \sigma_2$ where σ_1 and σ_2 are two positive constants. Note carefully that σ_S is not uniformly elliptic and bounded so that **AR** and **ND** are not satisfied. However, we heuristically choose σ_1 and σ_2 so that $\sigma_S(\bar{Y}_i)$ is bounded and strictly positive with high probability. Also, analytical expressions for the coefficients are available

$$\begin{aligned}
 a_{S,i} &= \int_0^{\zeta_{i+1}-\zeta_i} \left[\alpha(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + C \right]^2 ds, \\
 &= (\alpha\mu + C)^2(\zeta_{i+1} - \zeta_i) + \alpha^2(\bar{Y}_i - \mu)^2 \frac{1 - e^{-2\lambda(\zeta_{i+1}-\zeta_i)}}{2\lambda} + 2\alpha(\alpha\mu + C)(\bar{Y}_i - \mu) \frac{1 - e^{-\lambda(\zeta_{i+1}-\zeta_i)}}{\lambda}, \\
 a'_{S,i} &= \alpha^2(\bar{Y}_i - \mu) \frac{1 - e^{-2\lambda(\zeta_{i+1}-\zeta_i)}}{\lambda} + 2\alpha(\alpha\mu + C) \frac{1 - e^{-\lambda(\zeta_{i+1}-\zeta_i)}}{\lambda}, \\
 \rho_i &= \rho \frac{\int_0^{\zeta_{i+1}-\zeta_i} \left[\alpha(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + C \right] ds}{\sigma_{S,i} \sqrt{\zeta_{i+1} - \zeta_i}} = \rho \frac{\alpha(\bar{Y}_i - \mu)(1 - e^{-\lambda(\zeta_{i+1}-\zeta_i)})/\lambda + (\alpha\mu + C)(\zeta_{i+1} - \zeta_i)}{\sigma_{S,i} \sqrt{\zeta_{i+1} - \zeta_i}}, \\
 \rho'_i &= \rho \frac{\sigma_{S,i}(\alpha(1 - e^{-\lambda(\zeta_{i+1}-\zeta_i)})/\lambda) - \sigma'_{S,i}(\alpha(\bar{Y}_i - \mu)(1 - e^{-\lambda(\zeta_{i+1}-\zeta_i)})/\lambda + (\alpha\mu + C)(\zeta_{i+1} - \zeta_i))}{a_{S,i} \sqrt{\zeta_{i+1} - \zeta_i}}.
 \end{aligned}$$

The parameters for the unbiased Monte Carlo method and the Monte Carlo method combined with an Euler-Maruyama approximation scheme are chosen as in the first example. The numerical results related to the price, Delta and Vega are provided in Table 5.4, Table 5.5, Table 5.6 respectively for the Call option and in Table 5.7, Table 5.8, Table 5.9 for the digital Call option. In spite of the fact that the main assumptions are not satisfied, we again observe a good performance of the unbiased estimators for all three quantities and for all the values of the parameters σ_1, σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Price	95% CI	Price	95% CI	Price	95% CI
0.1	0.15	0.0790885	[0.0784919, 0.0796851]	0.0794717	[0.0785927, 0.0803508]	0.0791559	[0.0786344, 0.0796774]
0.2	0.25	0.129665	[0.128602, 0.130728]	0.129818	[0.128292, 0.131345]	0.129055	[0.126215, 0.131895]
0.3	0.4	0.202324	[0.200507, 0.20414]	0.200155	[0.199442, 0.200868]	0.200371	[0.199675, 0.201066]
0.4	0.5	0.249249	[0.246866, 0.251632]	0.249114	[0.248217, 0.250012]	0.249279	[0.248237, 0.250322]

Table 5.4 – Comparison between the unbiased Monte Carlo estimation for the price of a Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

5.5.3 A model with a periodic diffusion coefficient function

In our last example, the volatility of spot price takes the following form $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ where σ_1 and σ_2 are two positive constants such that $\sigma_2 - \sigma_1 > 0$ in order to ensure that **ND** is satisfied. Here, the coefficients appearing in the dynamics

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Delta	95% CI	Delta	95% CI	Delta	95% CI
0.1	0.15	0.545257	[0.542601, 0.547914]	0.542838	[0.534227, 0.551448]	0.540304	[0.535519, 0.545088]
0.2	0.25	0.548642	[0.54574, 0.551544]	0.541611	[0.533812, 0.549409]	0.535165	[0.51798, 0.552351]
0.3	0.4	0.566919	[0.563629, 0.570208]	0.555688	[0.548827, 0.56255]	0.558808	[0.553545, 0.56407]
0.4	0.5	0.579445	[0.575861, 0.583028]	0.569003	[0.559972, 0.578034]	0.568666	[0.561328, 0.576004]

Table 5.5 – Comparison between the unbiased Monte Carlo estimation for the Delta of a Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Vega	95% CI	Vega	95% CI	Vega	95% CI
0.1	0.15	0.0370801	[0.0367679, 0.0373923]	0.0340152	[0.0317984, 0.036232]	0.0350342	[0.0333134, 0.036755]
0.2	0.25	0.0738723	[0.0731769, 0.0745676]	0.0704958	[0.0662385, 0.074753]	0.0652897	[0.0612527, 0.0693267]
0.3	0.4	0.11114	[0.109907, 0.112373]	0.0899367	[0.0830599, 0.0968136]	0.10303	[0.0912086, 0.114851]
0.4	0.5	0.14385	[0.142055, 0.145645]	0.122496	[0.109106, 0.135885]	0.133235	[0.125356, 0.141114]

Table 5.6 – Comparison between the unbiased Monte Carlo estimation for the Vega of a Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Price	95% CI	Price	95% CI	Price	95% CI
0	0.3	0.469652	[0.467241, 0.472063]	0.469251	[0.468668, 0.469834]	0.468994	[0.468439, 0.469548]
0.1	0.15	0.491121	[0.488708, 0.493535]	0.489251	[0.488158, 0.490344]	0.489883	[0.489044, 0.490723]
0.2	0.25	0.459518	[0.45711, 0.461926]	0.458577	[0.457324, 0.459829]	0.458555	[0.457574, 0.459535]
0.3	0.4	0.430451	[0.428057, 0.432845]	0.428559	[0.42734, 0.429779]	0.428941	[0.428074, 0.429807]
0.4	0.5	0.408908	[0.406529, 0.411286]	0.40788	[0.404907, 0.410852]	0.409511	[0.408526, 0.410496]

Table 5.7 – Comparison between the unbiased Monte Carlo estimation for the price of a digital Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Delta	95% CI	Delta	95% CI	Delta	95% CI
0	0.3	1.22307	[1.19252, 1.25363]	1.24579	[1.24326, 1.24833]	1.24408	[1.24165, 1.24651]
0.1	0.15	2.17706	[2.13691, 2.21721]	2.17577	[2.16695, 2.18459]	2.18049	[2.17398, 2.18701]
0.2	0.25	1.29839	[1.26695, 1.32984]	1.26832	[1.26267, 1.27397]	1.26854	[1.26428, 1.27279]
0.3	0.4	0.776519	[0.752002, 0.801036]	0.792788	[0.789598, 0.795978]	0.793139	[0.790702, 0.795577]
0.4	0.5	0.606688	[0.58496, 0.628416]	0.618031	[0.610424, 0.625638]	0.621753	[0.619061, 0.624446]

Table 5.8 – Comparison between the unbiased Monte Carlo estimation for the Delta of a digital Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Vega	95% CI	Vega	95% CI	Vega	95% CI
0	0.3	0	[0, 0]	0.000481062	[-0.00499082, 0.00595295]	-0.000755278	[-0.00588091, 0.00437035]
0.1	0.15	-0.0200101	[-0.0308873, -0.00913292]	-0.0249364	[-0.0346885, -0.0151842]	-0.0286496	[-0.0358769, -0.0214223]
0.2	0.25	-0.0246278	[-0.0366948, -0.0125608]	-0.032211	[-0.0436935, -0.0207285]	-0.0311689	[-0.0392428, -0.023095]
0.3	0.4	-0.0354025	[-0.04987, -0.0209349]	-0.0422004	[-0.0518535, -0.0325472]	-0.0413018	[-0.0489346, -0.0336691]
0.4	0.5	-0.0492556	[-0.0663201, -0.0321911]	-0.0512594	[-0.0638074, -0.0387114]	-0.0517876	[-0.0597881, -0.0437871]

Table 5.9 – Comparison between the unbiased Monte Carlo estimation for the Vega of a digital Call option in the Stein-Stein type model for different values of the parameters σ_1 and σ_2 .

(5.2.4) write

$$\begin{aligned}
 a_{S,i} &= \int_0^{\zeta_{i+1}-\zeta_i} \left[\sigma_1 \cos(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + \sigma_2 \right]^2 ds, \\
 a'_{S,i} &= -2\alpha \int_0^{\zeta_{i+1}-\zeta_i} e^{-\lambda s} \sin(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) \left[\sigma_1 \cos(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + \sigma_2 \right] ds, \\
 \rho_i &= \rho \frac{\int_0^{\zeta_{i+1}-\zeta_i} \left[\sigma_1 \cos(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + \sigma_2 \right] ds}{\sigma_{S,i} \sqrt{\zeta_{i+1} - \zeta_i}}, \\
 \rho'_i &= -\rho \frac{\sigma_1 \sigma_{S,i} \int_0^{\zeta_{i+1}-\zeta_i} e^{-\lambda s} \sin(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) ds + \sigma'_{S,i} \int_0^{\zeta_{i+1}-\zeta_i} \left[\sigma_1 \cos(\mu + (\bar{Y}_i - \mu)e^{-\lambda s}) + \sigma_2 \right] ds}{a_{S,i} \sqrt{\zeta_{i+1} - \zeta_i}}
 \end{aligned}$$

and no analytical expressions are available. However, a simple numerical integration method can be employed for the computation of the above integrals. We here use Simpson's 3/8 rule which for a real-valued $\mathcal{C}^4([0, T])$ function g writes as follows

$$\forall t \in [0, T], \quad \int_0^t g(s) ds \approx \frac{t}{8} \left(g(0) + 3g\left(\frac{t}{3}\right) + 3g\left(\frac{2t}{3}\right) + g(t) \right)$$

with an error given by $g^{(4)}(t')T^5/6480$ for some $t' \in [0, T]$.

The number of samples for the unbiased Monte Carlo method and the Monte Carlo Euler-Maruyama scheme remains unchanged. The numerical results related to the price, Delta and Vega are provided in Table 5.10, Table 5.11, Table 5.12 respectively for the Call option and in Table 5.13, Table 5.14, Table 5.15 for the digital Call option. Here again, the unbiased estimators perform very well for all range of values of the parameters.

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Price	95% CI	Price	95% CI	Price	95% CI
0.1	0.15	0.110649	[0.109801, 0.111497]	0.111245	[0.110746, 0.111745]	0.111163	[0.11071, 0.111617]
0.2	0.25	0.193525	[0.191897, 0.195154]	0.19476	[0.19378, 0.19574]	0.193705	[0.192939, 0.19447]
0.3	0.4	0.294275	[0.291444, 0.297106]	0.294418	[0.292502, 0.296333]	0.294724	[0.293178, 0.296269]
0.4	0.5	0.371509	[0.367579, 0.375439]	0.3739	[0.371509, 0.376292]	0.373974	[0.372141, 0.375806]

Table 5.10 – Comparison between the unbiased Monte Carlo estimation for the price of a Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

5.6 Proof of Theorem 5.3.1 and Lemma 5.4.1

5.6.1 Proof of Theorem 5.3.1

The proof is divided into three steps. In the first part, we establish the probabilistic representation for a bounded and continuous function h . We then provide the extension to measurable maps satisfying the growth assumption 5.3.1. We eventually conclude

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Delta	95% CI	Delta	95% CI	Delta	95% CI
0.1	0.15	0.556917	[0.554118, 0.559717]	0.560077	[0.556733, 0.563422]	0.555364	[0.552636, 0.558092]
0.2	0.25	0.577937	[0.574727, 0.581148]	0.579704	[0.575622, 0.583785]	0.577287	[0.574331, 0.580243]
0.3	0.4	0.605788	[0.601947, 0.60963]	0.604575	[0.602771, 0.606379]	0.601188	[0.599354, 0.603021]
0.4	0.5	0.62865	[0.624204, 0.633096]	0.623698	[0.618519, 0.628878]	0.626259	[0.622246, 0.630271]

Table 5.11 – Comparison between the unbiased Monte Carlo estimation for the Delta of a Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Vega	95% CI	Vega	95% CI	Vega	95% CI
0.1	0.15	-0.00773549	[-0.00782648, -0.00764449]	-0.00805159	[-0.00985368, -0.0062495]	-0.00846248	[-0.0101504, -0.00677453]
0.2	0.25	-0.0156691	[-0.0158849, -0.0154532]	-0.0161045	[-0.0194751, -0.0127339]	-0.0137565	[-0.0169305, -0.0105825]
0.3	0.4	-0.0235822	[-0.0240098, -0.0231547]	-0.0177797	[-0.0236385, -0.0119209]	-0.0232616	[-0.0288379, -0.0176852]
0.4	0.5	-0.030774	[-0.0314484, -0.0300996]	-0.031729	[-0.0405267, -0.0229313]	-0.0327252	[-0.0402293, -0.0252211]

Table 5.12 – Comparison between the unbiased Monte Carlo estimation for the Vega of a Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Price	95% CI	Price	95% CI	Price	95% CI
0	0.3	0.470206	[0.467795, 0.472617]	0.468756	[0.468174, 0.469338]	0.46883	[0.468273, 0.469387]
0.1	0.15	0.481972	[0.479559, 0.484385]	0.481373	[0.480778, 0.481968]	0.481499	[0.480937, 0.482061]
0.2	0.25	0.446163	[0.443761, 0.448566]	0.445241	[0.444661, 0.445821]	0.445228	[0.444679, 0.445778]
0.3	0.4	0.407307	[0.40493, 0.409684]	0.407842	[0.407286, 0.408398]	0.407422	[0.40689, 0.407954]
0.4	0.5	0.379459	[0.37711, 0.381808]	0.380161	[0.379604, 0.380717]	0.379669	[0.379146, 0.380193]

Table 5.13 – Comparison between the unbiased Monte Carlo estimation for the price of a digital Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Delta	95% CI	Delta	95% CI	Delta	95% CI
0	0.3	1.24454	[1.21372, 1.27535]	1.24282	[1.24031, 1.24533]	1.24579	[1.24335, 1.24822]
0.1	0.15	1.5064	[1.47264, 1.54016]	1.51062	[1.50751, 1.51373]	1.51253	[1.50958, 1.51549]
0.2	0.25	0.846474	[0.820904, 0.872044]	0.835956	[0.834251, 0.83766]	0.834907	[0.833284, 0.836531]
0.3	0.4	0.512589	[0.492588, 0.53259]	0.528498	[0.527433, 0.529563]	0.527844	[0.526825, 0.528863]
0.4	0.5	0.399506	[0.381818, 0.417194]	0.40319	[0.402367, 0.404012]	0.402428	[0.401642, 0.403214]

Table 5.14 – Comparison between the unbiased Monte Carlo estimation for the Delta of a digital Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

σ_1	σ_2	Euler Scheme		Exponential sampling		Beta sampling	
		Vega	95% CI	Vega	95% CI	Vega	95% CI
0	0.3	0	[0, 0]	0.00122669	[-0.00426466, 0.00671804]	-0.00228675	[-0.00742253, 0.00284903]
0.1	0.15	0.00769619	[-0.000285705, 0.0156781]	0.00900717	[0.00339831, 0.014616]	0.00730275	[0.0021288, 0.0124767]
0.2	0.25	0.0138531	[0.00480268, 0.0229036]	0.0146584	[0.00906018, 0.0202566]	0.0131819	[0.00819735, 0.0181665]
0.3	0.4	0.0107747	[0.00279286, 0.0187565]	0.00807909	[0.00292412, 0.0132341]	0.0116897	[0.00689717, 0.0164823]
0.4	0.5	0.0153924	[0.00585238, 0.0249324]	0.0152859	[0.0102097, 0.0203621]	0.0164134	[0.0117414, 0.0210853]

Table 5.15 – Comparison between the unbiased Monte Carlo estimation for the Vega of a digital Call option in the model with $\sigma_S(x) = \sigma_1 \cos(x) + \sigma_2$ for different values of the parameters σ_1 and σ_2 .

by establishing the L^p -moments when the jump times are distributed according to the Beta law.

Denote by \mathcal{L} and $(\bar{\mathcal{L}}_t)_{t \geq 0}$ the infinitesimal generators of $(P_t)_{t \geq 0}$ and $(\bar{P}_t)_{t \geq 0}$ respectively given by

$$\begin{aligned}\mathcal{L}f(x, y) &= \left(r - \frac{1}{2}a_S(y)\right)\partial_x f(x, y) + \frac{1}{2}a_S(y)\partial_x^2 f(x, y) + b_Y(y)\partial_y f(x, y) \\ &\quad + \frac{1}{2}a_Y(y)\partial_y^2 f(x, y) + \rho(\sigma_S\sigma_Y)(y)\partial_{x,y}^2 f(x, y), \\ \bar{\mathcal{L}}_t f(x, y) &= \left(r - \frac{1}{2}a_S(m_t(y_0))\right)\partial_x f(x, y) + \frac{1}{2}a_S(m_t(y_0))\partial_x^2 f(x, y) + b_Y(m_t(y_0))\partial_y f(x, y) \\ &\quad + \frac{1}{2}a_Y(m_t(y_0))\partial_y^2 f(x, y) + \rho(\sigma_S\sigma_Y)(m_t(y_0))\partial_{x,y}^2 f(x, y)\end{aligned}$$

for any $f \in \mathcal{C}_b^2(\mathbb{R}^2)$.

Step 1: Probabilistic representation for a bounded and continuous map h

We establish a first order expansion of the Markov semigroup $(P_t)_{t \geq 0}$ around $(\bar{P}_t)_{t \geq 0}$. We apply Itô's rule to the map $[0, t] \times \mathbb{R}^2 \ni (s, x, y) \mapsto P_{t-s}h(x, y) \in \mathcal{C}_b^{1,2}([0, t] \times \mathbb{R}^2)$ for $h \in \mathcal{C}_b^\infty(\mathbb{R}^2)$, observing that $\partial_s P_{t-s}h(x, y) = -\mathcal{L}P_{t-s}h(x, y)$. We obtain

$$\begin{aligned}h(\bar{X}_t, \bar{Y}_t) &= P_t h(x_0, y_0) + \int_0^t \left(\partial_s P_{t-s}h(\bar{X}_s, \bar{Y}_s) + \bar{\mathcal{L}}_s P_{t-s}h(\bar{X}_s, \bar{Y}_s) \right) ds + M_t \\ &= P_t h(x_0, y_0) + \int_0^t (\bar{\mathcal{L}}_s - \mathcal{L})P_{t-s}h(\bar{X}_s, \bar{Y}_s) ds + M_t\end{aligned}$$

where $M := (M_t)_{t \geq 0}$ is a square integrable martingale. We then take expectation in the previous expression, make use of Fubini's theorem and finally let $t \uparrow T$ by dominated convergence theorem so that

$$\begin{aligned}P_T h(x_0, y_0) &= \bar{P}_T h(x_0, y_0) + \int_0^T \mathbb{E}[(\mathcal{L} - \bar{\mathcal{L}}_s)P_{T-s}h(\bar{X}_s, \bar{Y}_s)] ds \\ &= \mathbb{E}[h(\bar{X}_T^{x_0}, \bar{Y}_T^{y_0})] + \int_0^T \mathbb{E} \left[\frac{1}{2}(a_S(\bar{Y}_s^{y_0}) - a_S(m_s(y_0)))[\partial_x^2 P_{T-s}h(\bar{X}_s^{x_0}, \bar{Y}_s^{y_0}) - \partial_x P_{T-s}h(\bar{X}_s^{x_0}, \bar{Y}_s^{y_0})] \right. \\ &\quad \left. + \int_0^T \mathbb{E} \left[\frac{1}{2}(a_Y(\bar{Y}_s^{y_0}) - a_Y(m_s(y_0)))\partial_y^2 P_{T-s}h(\bar{X}_s^{x_0}, \bar{Y}_s^{y_0}) ds \right. \right. \\ &\quad \left. \left. + (b_Y(\bar{Y}_s^{y_0}) - b_Y(m_s(y_0)))\partial_y P_{T-s}h(\bar{X}_s^{x_0}, \bar{Y}_s^{y_0}) \right] \right. \\ &\quad \left. + \int_0^T \mathbb{E} \left[\rho((\sigma_S\sigma_Y)(\bar{Y}_s^{x_0}) - (\sigma_S\sigma_Y)(m_s(y_0)))\partial_{x,y}^2 P_{T-s}h(\bar{X}_s^{x_0}, \bar{Y}_s^{y_0}) \right] ds. \quad (5.6.1)\end{aligned}$$

We now rewrite the previous first order expansion using the Markov chain $(\bar{X}_i, \bar{Y}_i)_{0 \leq i \leq N_T+1}$ and the renewal process N . From the previous identity, the definition of θ_{N_T+1} in (5.3.4)

and the identity (5.2.3), we directly obtain

$$\begin{aligned}
 P_T h(x_0, y_0) &= \mathbb{E}[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1})\theta_{N_T+1}\mathbf{1}_{\{N_T=0\}}] \\
 &\quad + \mathbb{E}\left[\left((1 - F(T - \zeta_1))f(\zeta_1)\right)^{-1}\mathbf{1}_{\{N_T=1\}}\left[\frac{1}{2}(a_S(\bar{Y}_1) - a_S(m_0))\mathcal{D}_1^{(1,1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right.\right. \\
 &\quad \left. - \frac{1}{2}(a_S(\bar{Y}_1) - a_S(m_0))\mathcal{D}_1^{(1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + \frac{1}{2}(a_Y(\bar{Y}_1) - a_Y(m_0))\mathcal{D}_1^{(2,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right. \\
 &\quad \left. + (b_Y(\bar{Y}_1) - b_Y(m_0))\mathcal{D}_1^{(2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + \rho((\sigma_S\sigma_Y)(\bar{Y}_1) - (\sigma_S\sigma_Y)(m_0))\mathcal{D}_1^{(1,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right]\Big] \\
 &= \mathbb{E}[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1})\theta_{N_T+1}\mathbf{1}_{\{N_T=0\}}] + \mathbb{E}\left[\left((1 - F(T - \zeta_1))f(\zeta_1)\right)^{-1}\mathbf{1}_{\{N_T=1\}}\right. \\
 &\quad \times \left[c_S^1\mathcal{D}_1^{(1,1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) - c_S^1\mathcal{D}_1^{(1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + c_Y^1\mathcal{D}_1^{(2,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right. \\
 &\quad \left. + b_Y^1\mathcal{D}_1^{(2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + c_{S,Y}^1\mathcal{D}_1^{(1,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right]\Big]. \tag{5.6.2}
 \end{aligned}$$

Next, we apply the IBP formula (5.2.9) with respect to the random vector (\bar{X}_1, \bar{Y}_1) in the above expression. In order to do that rigorously, one first has to take the conditional expectation $\mathbb{E}_{0,1}[\cdot]$ in the second term of the above equality. We thus obtain

$$\begin{aligned}
 &\mathbb{E}_{0,1}\left[c_S^1\mathcal{D}_1^{(1,1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) - c_S^1\mathcal{D}_1^{(1)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + c_Y^1\mathcal{D}_1^{(2,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right. \\
 &\quad \left. + b_Y^1\mathcal{D}_1^{(2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1) + c_{S,Y}^1\mathcal{D}_1^{(1,2)}P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right] \\
 &= \mathbb{E}_{0,1}\left[\left[\mathcal{I}_1^{(1,1)}(c_S^1) - \mathcal{I}_1^{(1)}(c_S^1) + \mathcal{I}_1^{(2,2)}(c_Y^1) + \mathcal{I}_1^{(2)}(b_Y^1) + \mathcal{I}_1^{(1,2)}(c_{S,Y}^1)\right]P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right].
 \end{aligned}$$

From Lemma 5.7.2 and the estimate (5.2.17), we get

$$\begin{aligned}
 &\mathbb{E}_{0,1}\left[\left|\left[\mathcal{I}_1^{(1,1)}(c_S^1) - \mathcal{I}_1^{(1)}(c_S^1) + \mathcal{I}_1^{(2,2)}(c_Y^1) + \mathcal{I}_1^{(2)}(b_Y^1) + \mathcal{I}_1^{(1,2)}(c_{S,Y}^1)\right]P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right|\right] \\
 &\leq C_T|h|_\infty\zeta_1^{-1/2}
 \end{aligned}$$

for some positive constant C_T such that $T \mapsto C_T$ is non-decreasing. The previous estimate yields an integrable time singularity. Indeed, from the previous estimate and (5.2.3), one directly gets

$$\begin{aligned}
 &\mathbb{E}\left[\left((1 - F(T - \zeta_1))f(\zeta_1)\right)^{-1}\mathbf{1}_{\{N_T=1\}}\left|\mathbb{E}_{0,1}\left[\left[\mathcal{I}_1^{(1,1)}(c_S^1) - \mathcal{I}_1^{(1)}(c_S^1) + \mathcal{I}_1^{(2)}(b_Y^1)\right.\right.\right.\right. \\
 &\quad \left. \left. \left. + \mathcal{I}_1^{(1,2)}(c_{S,Y}^1) + \mathcal{I}_1^{(2,2)}(c_Y^1)\right]P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right|\right]\Big] \\
 &\leq C\mathbb{E}\left[\left((1 - F(T - \zeta_1))f(\zeta_1)\right)^{-1}\zeta_1^{-1/2}\mathbf{1}_{\{N_T=1\}}\right] \\
 &\leq C\int_0^T s_1^{-1/2} ds_1 < \infty.
 \end{aligned}$$

Coming back to (5.6.2), we thus derive

$$\begin{aligned}
 &P_T h(x_0, y_0) \\
 &= \mathbb{E}[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1})\theta_{N_T+1}\mathbf{1}_{\{N_T=0\}}] + \mathbb{E}\left[\left((1 - F(T - \zeta_1))f(\zeta_1)\right)^{-1}\mathbf{1}_{\{N_T=1\}}\right. \\
 &\quad \left.\times \left[\mathcal{I}_1^{(1,1)}(c_S^1) - \mathcal{I}_1^{(1)}(c_S^1) + \mathcal{I}_1^{(2,2)}(c_Y^1) + \mathcal{I}_1^{(2)}(b_Y^1) + \mathcal{I}_1^{(1,2)}(c_{S,Y}^1)\right]P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\right] \\
 &= \mathbb{E}[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1})\theta_{N_T+1}\mathbf{1}_{\{N_T=0\}}] + \mathbb{E}\left[P_{T-\zeta_1}h(\bar{X}_1, \bar{Y}_1)\theta_2\theta_1\mathbf{1}_{\{N_T=1\}}\right]. \tag{5.6.3}
 \end{aligned}$$

Our aim now is to iterate the above first order expansion. We prove by induction the following formula: for any positive integer n , one has

$$P_T h(x_0, y_0) = \sum_{j=0}^{n-1} \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i 1_{\{N_T=j\}} \right] + \mathbb{E} \left[P_{T-\zeta_n} h(\bar{X}_n, \bar{Y}_n) \prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \right]. \quad (5.6.4)$$

The case $n = 1$ corresponds to (5.6.3). We thus assume that (5.6.4) holds at step n . We expand the last term appearing in the right-hand side of the previous equality using again (5.6.1), by then applying Lemma 5.7.1 and by finally performing IBPs as before.

To be more specific, using the notations introduced in Subsection 5.2.2, from (5.6.1) and a change of variable, for any $\zeta \in [0, T]$, one has

$$\begin{aligned} P_{T-\zeta} h(x, y) &= \mathbb{E}[h(\bar{X}_T^{\zeta, x}, \bar{Y}_T^{\zeta, x})] \\ &+ \int_{\zeta}^T \mathbb{E} \left[\frac{1}{2} (a_S(\bar{Y}_s^{\zeta, y}) - a_S(m_{s-\zeta}(y))) [\partial_x^2 P_{T-s} h(\bar{X}_s^{\zeta, x}, \bar{Y}_s^{\zeta, y}) - \partial_x P_{T-s} h(\bar{X}_s^{\zeta, x}, \bar{Y}_s^{\zeta, y})] \right] ds \\ &+ \int_{\zeta}^T \mathbb{E} \left[\frac{1}{2} (a_Y(\bar{Y}_s^{\zeta, y}) - a_Y(m_{s-\zeta}(y))) \partial_y^2 P_{T-s} h(\bar{X}_s^{\zeta, x}, \bar{Y}_s^{\zeta, y}) \right. \\ &\quad \left. + (b_Y(\bar{Y}_s^{\zeta, y}) - b_Y(m_{s-\zeta}(y))) \partial_y P_{T-s} h(\bar{X}_s^{\zeta, x}, \bar{Y}_s^{\zeta, y}) \right] ds \\ &+ \int_{\zeta}^T \mathbb{E} \left[\rho((\sigma_S \sigma_Y)(\bar{Y}_s^{\zeta, x}) - (\sigma_S \sigma_Y)(m_{s-\zeta}(y))) \partial_{x,y}^2 P_{T-s} h(\bar{X}_s^{\zeta, x}, \bar{Y}_s^{\zeta, y}) \right] ds. \end{aligned}$$

We take $\zeta = \zeta_n$, $(x, y) = (\bar{X}_{N_T}, \bar{Y}_{N_T})$ in the previous equality, then multiply it by $\prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}}$ and finally take expectation. We obtain

$$\begin{aligned} &\mathbb{E} \left[P_{T-\zeta_n} h(\bar{X}_n, \bar{Y}_n) \prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \right] \\ &= \mathbb{E} \left[h(\bar{X}_T^{\zeta_n, \bar{X}_n}, \bar{Y}_T^{\zeta_n, \bar{Y}_n}) \prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \right] \\ &+ \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \int_{\zeta_n}^T \frac{1}{2} (a_S(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - a_S(m_{s-\zeta_n}(\bar{Y}_n))) [\partial_x^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) \right. \\ &\quad \left. - \partial_x P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n})] ds \right] \\ &+ \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \int_{\zeta_n}^T \frac{1}{2} (a_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - a_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) ds \right] \\ &+ \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \int_{\zeta_n}^T (b_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - b_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) ds \right] \\ &+ \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i 1_{\{N_T=n\}} \int_{\zeta_n}^T \rho((\sigma_S \sigma_Y)(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - (\sigma_S \sigma_Y)(m_{s-\zeta_n}(\bar{Y}_n))) \partial_{x,y}^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) ds \right]. \end{aligned} \quad (5.6.5)$$

Now, from the very definition of the Markov chain $(\bar{X}_i, \bar{Y}_i)_{0 \leq i \leq N_T+1}$ and of the weight sequence $(\theta_i)_{1 \leq i \leq N_T+1}$ of Theorem 5.3.1, the first term of the above equality can be written as

$$\mathbb{E} \left[h(\bar{X}_T^{\zeta_n, \bar{X}_n}, \bar{Y}_T^{\zeta_n, \bar{Y}_n}) \prod_{i=1}^{n+1} \theta_i \mathbf{1}_{\{N_T=n\}} \right] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \mathbf{1}_{\{N_T=n\}} \right].$$

We now look at the second, third, fourth and fifth terms. Let us deal with the third and fourth terms. The others are treated in a similar manner and we will omit some technical details. We first take its conditional conditional expectation w.r.t $\{\zeta_1 = t_1, \dots, \zeta_n = t_n, N_T = n\}$ and introduce the measurable function

$$\begin{aligned} G(t_1, \dots, t_n, s, T) := & \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i \int_{\zeta_n}^T \left[\frac{1}{2} (a_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - a_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) \right. \right. \\ & \left. \left. + (b_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - b_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) \right] ds \middle| \zeta_1 = t_1, \dots, \zeta_n = t_n, N_T = n \right] \end{aligned}$$

which satisfies

$$\begin{aligned} |G(t_1, \dots, t_n, s, T)| & \leq C \mathbb{E} \left[\prod_{i=1}^{n+1} |\theta_i| \left(1 + |W_s - W_{\zeta_n}| + |\widetilde{W}_s - \widetilde{W}_{\zeta_n}| \right) \middle| \zeta_1 = t_1, \dots, \zeta_n = t_n, N_T = n \right] \\ & \leq C \mathbb{E} \left[\prod_{i=1}^{n+1} |\theta_i| \middle| \zeta_1 = t_1, \dots, \zeta_n = t_n, N_T = n \right] \end{aligned}$$

where we used the boundedness of a_Y , the Lipschitz regularity of b_Y , the inequalities $\sup_{0 \leq t \leq T} |\partial_y^\ell P_t h|_\infty \leq C$ for $\ell = 1, 2$ and, for the last inequality the fact that, conditionally on $\{\zeta_1 = t_1, \dots, \zeta_n = t_n, N_T = n\}$, the random variables $(W_s - W_{\zeta_n}, \widetilde{W}_s - \widetilde{W}_{\zeta_n})$ are independent of the sigma-field \mathcal{G}_n . Recall now that $\mathbb{P}(N_T = n, \zeta_1 \in dt_1, \dots, \zeta_n \in dt_n) = (1 - F(T - t_n)) \prod_{j=0}^{n-1} f(t_{j+1} - t_j) dt_1, \dots, dt_n$ on the set $\Delta_n(T)$ so that from Lemma 5.7.2 and the estimate (5.2.17), we obtain

$$\mathbb{E} \left[\int_{\zeta_n}^T |G(\zeta_1, \dots, \zeta_n, s, T)| \mathbf{1}_{\{N_T=n\}} ds \right] \leq C \int_{t_n}^T \int_{\Delta_n(T)} \prod_{i=1}^n (t_i - t_{i-1})^{-1/2} dt_1 \dots dt_n dt_{n+1} < \infty.$$

Hence, by Lemma 5.7.1, it holds

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_{\{N_T=n\}} \prod_{i=1}^{n+1} \theta_i \int_{\zeta_n}^T \left[\frac{1}{2} (a_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - a_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) \right. \right. \\ & \quad \left. \left. + (b_Y(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - b_Y(m_{s-\zeta_n}(\bar{Y}_n))) \partial_y P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) \right] ds \right] \\ & = \mathbb{E} \left[\mathbf{1}_{\{N_T=n\}} \int_{\zeta_n}^T G(\zeta_1, \dots, \zeta_n, s, T) ds \right] \\ & = \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} \left[\frac{1}{2} (a_Y(\bar{Y}_{n+1}) - a_Y(m_n)) \mathcal{D}_{n+1}^{(2,2)} P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \right. \right. \\ & \quad \left. \left. + (b_Y(\bar{Y}_{n+1}) - b_Y(m_n)) \mathcal{D}_{n+1}^{(2)} P_{T-s} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \right] \mathbf{1}_{\{N_T=n+1\}} \right]. \end{aligned}$$

Finally, we take the conditional expectation $\mathbb{E}_{n,n+1}[\cdot]$ inside the above expectation and then employ the IBP formula (5.2.9), two times w.r.t. the diffusion coefficient and one time w.r.t the drift coefficient as done before. We obtain

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} \left[\frac{1}{2} (a_Y(\bar{Y}_{n+1}) - a_Y(m_n)) \mathcal{D}_{n+1}^{(2,2)} P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \right. \right. \\ & \quad \left. \left. + (b_Y(\bar{Y}_{n+1}) - b_Y(m_n)) \mathcal{D}_{n+1}^{(2)} P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \right] \mathbf{1}_{\{N_T=n+1\}} \right] \\ &= \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} [\mathcal{I}_{n+1}^{(2,2)}(c_Y^{n+1}) + \mathcal{I}_{n+1}^{(2)}(b_Y^{n+1})] \right. \\ & \quad \left. \times P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathbf{1}_{\{N_T=n+1\}} \right]. \end{aligned}$$

In a completely analogous manner, we derive

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i \mathbf{1}_{\{N_T=n\}} \int_{\zeta_n}^T \frac{1}{2} (a_S(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - a_S(m_{s-\zeta_n}(\bar{Y}_n))) \right. \\ & \quad \left. \times [\partial_x^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) - \partial_x P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n})] ds \right] \\ &= \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} [\mathcal{I}_{n+1}^{(1,1)}(c_S^{n+1}) - \mathcal{I}_{n+1}^{(1)}(c_S^{n+1})] \right. \\ & \quad \left. \times P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathbf{1}_{\{N_T=n+1\}} \right] \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^{n+1} \theta_i \mathbf{1}_{\{N_T=n\}} \int_{\zeta_n}^T \rho((\sigma_S \sigma_Y)(\bar{Y}_s^{\zeta_n, \bar{Y}_n}) - (\sigma_S \sigma_Y)(m_{s-\zeta_n}(\bar{Y}_n))) \partial_{x,y}^2 P_{T-s} h(\bar{X}_s^{\zeta_n, \bar{X}_n}, \bar{Y}_s^{\zeta_n, \bar{Y}_n}) ds \right] \\ &= \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} \mathcal{I}_{n+1}^{(1,2)}(c_{Y,S}^{n+1}) P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathbf{1}_{\{N_T=n+1\}} \right]. \end{aligned}$$

Summing the three previous identities, we obtain that the sum of the second, third, fourth and fifth term in the right-hand side of (5.6.5) is equal to

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^n \theta_i (1 - F(T - \zeta_{n+1}))^{-1} (f(\zeta_{n+1} - \zeta_n))^{-1} \right. \\ & \quad \times \left[\mathcal{I}_{i+1}^{(1,1)}(a_S^{n+1}) - \mathcal{I}_{n+1}^{(1)}(a_S^{n+1}) + \mathcal{I}_{n+1}^{(2,2)}(a_Y^{n+1}) + \mathcal{I}_{n+1}^{(2)}(b_Y^{n+1}) + \mathcal{I}_{n+1}^{(1,2)}(a_{Y,S}^{n+1}) \right] \\ & \quad \left. \times P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathbf{1}_{\{N_T=n+1\}} \right] \\ &= \mathbb{E} \left[\prod_{i=1}^{n+2} \theta_i P_{T-\zeta_{n+1}} h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \mathbf{1}_{\{N_T=n+1\}} \right] \end{aligned}$$

where we used the very definitions (5.3.3) and (5.3.4) of the weights $(\theta_i)_{1 \leq i \leq N_T+1}$ on the set $\{N_T = n+1\}$. This concludes the proof of (5.6.4) at step $n+1$.

To conclude it remains to prove the absolute convergence of the first sum and the convergence to zero of the last term in (5.6.4). These two results follow directly from

the boundedness of h and the general estimates on the product of weights established in Lemma 5.7.2.

Indeed, from Lemma 5.7.2, the estimate (5.2.17), the tower property of conditional expectation and (5.2.3), we obtain

$$\begin{aligned}
 & \left| \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \mathbf{1}_{\{N_T=j\}} \right] \right| \\
 & \leq C^j |h|_\infty \mathbb{E} \left[(1 - F(T - \zeta_j))^{-1} \prod_{i=1}^j (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \mathbf{1}_{\{N_T=j\}} \right] \\
 & = C^j |h|_\infty \int_{\Delta_j(T)} \prod_{i=1}^j (s_i - s_{i-1})^{-\frac{1}{2}} ds_j \\
 & = C^j |h|_\infty T^{\frac{j}{2}} \frac{\Gamma^j(1/2)}{\Gamma(1 + j/2)}
 \end{aligned}$$

which in turn yields

$$\sum_{j=0}^{n-1} \left| \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \mathbf{1}_{\{N_T=j\}} \right] \right| \leq |h|_\infty \sum_{n \geq 0} \frac{(CT^{1/2})^j}{\Gamma(1 + j/2)} = |h|_\infty E_{1/2,1}(CT^{1/2})$$

so that the series converge absolutely. Similarly,

$$\left| \mathbb{E} \left[P_{T-\zeta_n} h(\bar{X}_n, \bar{Y}_n) \prod_{i=1}^{n+1} \theta_i \mathbf{1}_{\{N_T=n\}} \right] \right| \leq C^n |h|_\infty T^{\frac{n}{2}} \frac{\Gamma^n(1/2)}{\Gamma(1 + n/2)}$$

so that the remainder indeed vanishes as n goes to infinity. We thus conclude

$$P_T h(x_0, y_0) = \sum_{n \geq 0} \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \mathbf{1}_{\{N_T=n\}} \right] = \mathbb{E} \left[h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1}) \prod_{i=1}^{N_T+1} \theta_i \right] \tag{5.6.6}$$

for any $h \in \mathcal{C}_b^2(\mathbb{R}^2)$. We eventually extend the above representation formula to any bounded and continuous function h using a standard approximation argument. The remaining technical details are omitted.

Step 2: Extension to measurable map h satisfying the growth assumption (5.3.1)

We first extend the previous result to any bounded and measurable h . This follows from a monotone class argument that we now detail.

Let us first consider $h \in \mathcal{C}_b(\mathbb{R}^2)$. From Fubini's theorem, it holds

$$\begin{aligned}
 & \mathbb{E} \left[h(\bar{X}_{n+1}, \bar{Y}_{n+1}) \prod_{i=1}^{n+1} \theta_i \middle| N_T = n, \zeta^n \right] \\
 & = \int_{\mathbb{R}^2} h(x, y) \mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \prod_{i=1}^{n+1} \theta_i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^n) \middle| N_T = n, \zeta^n \right] dx dy
 \end{aligned}$$

which can be justified as follows. From Lemma 5.7.2, Lemma 5.7.3 and the upper-bound estimate (5.2.2), it holds

$$\begin{aligned}
& \left| \mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \prod_{i=1}^{n+1} \theta_i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^n) \middle| N_T = n, \zeta^n \right] \right| \\
& \leq \int_{(\mathbb{R}^2)^n} \bar{p}(T - \zeta_n, x_n, y_n, x, y) \prod_{i=1}^n |\theta_i(x_{i-1}, y_{i-1}, x_i, y_i, \zeta^n)| \bar{p}(\zeta_i - \zeta_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) dx_n dy_n \\
& \leq (C_T)^n \int_{(\mathbb{R}^2)^n} \bar{q}_{4\kappa}(T - \zeta_n, x_n, y_n, x, y) \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \bar{q}_{4\kappa}(\zeta_i - \zeta_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) dx_n dy_n \\
& \leq (C_T)^n \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \bar{q}_c(T, x_0, y_0, x, y)
\end{aligned}$$

for some $c := c(T, b_Y, \kappa) \geq 4\kappa$. Hence, from (5.6.6) and again Fubini's theorem, justified by the previous estimate and the fact that $\mathbb{E} \left[(C_T)^{N_T} \prod_{i=1}^{N_T} (f(\zeta_i - \zeta_{i-1}))^{-1} (\zeta_i - \zeta_{i-1})^{-\frac{1}{2}} \right] < \infty$, one has

$$P_T h(x_0, y_0) = \int_{\mathbb{R}^2} h(x, y) \mathbb{E} \left[\bar{p}(T - \zeta_{N_T}, \bar{X}_{N_T}, \bar{Y}_{N_T}, x, y) \prod_{i=1}^{N_T+1} \theta_i \right] dx dy \quad (5.6.7)$$

for any $h \in \mathcal{C}_b(\mathbb{R}^2)$. Moreover, from the previous computations, the following upper-bound holds

$$\begin{aligned}
& \left| \mathbb{E} \left[\bar{p}(T - \zeta_{N_T}, \bar{X}_{N_T}, \bar{Y}_{N_T}, x, y) \prod_{i=1}^{N_T+1} \theta_i \right] \right| \\
& = \left| \sum_{n \geq 0} \int_{\Delta_n(T)} \mathbb{E} \left[\bar{p}(T - s_n, \bar{X}_n, \bar{Y}_n, x, y) \prod_{i=1}^{n+1} \theta_i \middle| N_T = n, \zeta^n = (s_1, \dots, s_n) \right] \right. \\
& \quad \left. \times (1 - F(T - s_n)) \prod_{i=1}^n f(s_i - s_{i-1}) ds_n \right| \\
& \leq \left(\sum_{n \geq 0} \int_{\Delta_n(T)} (C_T)^n \prod_{i=1}^n (s_i - s_{i-1})^{-\frac{1}{2}} ds_n \right) \bar{q}_c(T, x_0, y_0, x, y) \\
& = E_{1/2,1}(CT^{1/2}) \bar{q}_c(T, x_0, y_0, x, y) \quad (5.6.8)
\end{aligned}$$

It now follows from (5.6.7) and a monotone class argument that (5.3.2) is valid for any real-valued bounded and measurable map h defined over \mathbb{R}^2 . The extension to any measurable map h satisfying the growth assumption: $|h(x, y)| \leq C \exp(\gamma(|x|^2 + |y|^2))$ for any $\gamma < (2\kappa)^{-1}$, follows from the integral representation (5.6.7), the upper-bound (5.6.8) combined with a standard approximation argument. Remaining technical details are omitted.

Step 3: Finite $L^p(\mathbb{P})$ -moment for the probabilistic representation

If N is a renewal process with $Beta(\alpha, 1)$ jump times then

$$\begin{aligned} f(s_i - s_{i-1}) &= \frac{1 - \alpha}{\bar{\tau}^{1-\alpha}} \frac{1}{(s_i - s_{i-1})^\alpha} \mathbf{1}_{[0, \bar{\tau}]}(s_i - s_{i-1}) \geq 1 - \left(\frac{T}{\bar{\tau}}\right)^{1-\alpha} \\ 1 - F(T - s_n) &= 1 - \left(\frac{T - s_n}{\bar{\tau}}\right)^{1-\alpha} \geq 1 - \left(\frac{T}{\bar{\tau}}\right)^{1-\alpha} \end{aligned}$$

similarly to step 2, by Fubini's theorem, we get

$$\begin{aligned} &\mathbb{E} \left[|h(\bar{X}_{n+1}, \bar{Y}_{n+1})|^p \prod_{i=1}^{n+1} |\theta_i|^p \middle| N_T = n, \zeta^n \right] \\ &= \int_{\mathbb{R}^2} |h(x, y)|^p \mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \prod_{i=1}^{n+1} |\theta_i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^n)|^p \middle| N_T = n, \zeta^n \right] dx dy. \end{aligned}$$

The above formula is justified by Lemma 5.7.2 and Lemma 5.7.3 which yield

$$\begin{aligned} &\mathbb{E} \left[\bar{p}(T - \zeta_n, \bar{X}_n, \bar{Y}_n, x, y) \prod_{i=1}^{n+1} |\theta_i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^n)|^p \middle| N_T = n, \zeta^n \right] \\ &\leq \int_{(\mathbb{R}^2)^n} \bar{p}(T - \zeta_n, x_n, y_n, x, y) \prod_{i=1}^n |\theta_i(x_{i-1}, y_{i-1}, x_i, y_i, \zeta^n)|^p \bar{p}(\zeta_i - \zeta_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) d\mathbf{x}_n d\mathbf{y}_n \\ &\leq C^n \int_{(\mathbb{R}^2)^n} \bar{q}_{4\kappa}(T - \zeta_n, x_n, y_n, x, y) \prod_{i=1}^n (f(\zeta_i - \zeta_{i-1}))^{-p} (\zeta_i - \zeta_{i-1})^{-\frac{p}{2}} \\ &\quad \times \bar{q}_{4\kappa}(\zeta_i - \zeta_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) d\mathbf{x}_n d\mathbf{y}_n \\ &\leq C^n \prod_{i=1}^n (\zeta_i - \zeta_{i-1})^{\alpha p} (\zeta_i - \zeta_{i-1})^{-\frac{p}{2}} \bar{q}_c(T, x_0, y_0, x, y) \end{aligned}$$

for some $c := c(T, b_Y, \kappa) \geq 4\kappa$. Now, using the fact that $\mathbb{E} \left[C^n \prod_{i=1}^{N_T} (\zeta_i - \zeta_{i-1})^{\alpha p} (\zeta_i - \zeta_{i-1})^{-\frac{p}{2}} \right] < \infty$ as soon as $p(\frac{1}{2} - \alpha) < 1 - \alpha$ and that $h \in \mathcal{B}_\gamma(\mathbb{R}^2)$, from the previous computation, we obtain

$$\begin{aligned} &\mathbb{E} \left[|h(\bar{X}_{N_T+1}, \bar{Y}_{N_T+1})|^p \prod_{i=1}^{N_T+1} |\theta_i|^p \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[|h(\bar{X}_{n+1}, \bar{Y}_{n+1})|^p \prod_{i=1}^{n+1} |\theta_i|^p \middle| N_T, \zeta^n \right] \right] \\ &\leq \mathbb{E} \left[C^n \prod_{i=1}^{N_T} (\zeta_i - \zeta_{i-1})^{\alpha p} (\zeta_i - \zeta_{i-1})^{-\frac{p}{2}} \right] \int_{\mathbb{R}^2} e^{\gamma p(|x|^2 + |y|^2)} \bar{q}_c(T, x_0, y_0, x, y) dx dy. \end{aligned}$$

To conclude the proof, it suffices to note that the above space integral is finite as soon as $\gamma p < c^{-1}$.

5.6.2 Proof of Lemma 5.4.1

Since $h \in \mathcal{C}_p^1(\mathbb{R}^2)$ and $\mathbb{E}_{i,n} \left[|\bar{X}_{i+1}|^q + |\bar{Y}_{i+1}|^q \right] < \infty$ a.s., for any $q \geq 1$, under **AR**, one may differentiate under the (conditional) expectation and deduce that $(x, y) \mapsto \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} | (\bar{X}_i, \bar{Y}_i) = (x, y) \right] \in \mathcal{C}_p^1(\mathbb{R}^2)$ for any $i \in \{0, \dots, n\}$. The rest of the proof is divided into three parts.

Step 1: proofs of (5.4.1) and (5.4.2)

The transfer of derivatives formulae (5.4.1) and (5.4.2) are easily obtained by differentiating under expectation (which is allowed by the polynomial growth at infinity of h) noting from the definition of the Markov chain \bar{X} that $\partial_{s_0} \bar{X}_1 = \partial_{s_0} \ln(s_0) = \frac{1}{s_0}$ and $\partial_{\bar{X}_i} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) = \partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{X}_i} \bar{X}_{i+1} = \partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1})$. Observe as well that from (5.2.10), the fact that $\partial_{\bar{X}_i} c_S^{i+1} = \partial_{\bar{X}_i} c_Y^{i+1} = \partial_{\bar{X}_i} b_Y^{i+1} = \partial_{\bar{X}_i} c_{Y,S}^{i+1} = 0$ and the very definition of the random variables $(\theta_i)_{1 \leq i \leq n+1}$, one has $\partial_{\bar{X}_i} \theta_{i+1} = 0$. This gives the identities (5.4.1) and (5.4.2).

Step 2: proofs of (5.4.3) and (5.4.4)

The proofs of (5.4.3) and (5.4.4) are more involved. Let us prove (5.4.3). We proceed by considering the difference between the term appearing on the left-hand side and the first two terms appearing on the right-hand side of (5.4.3). On the one hand, using the IBP formula (5.2.9) and (5.2.5), we get

$$\begin{aligned} \partial_{\bar{Y}_i} \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} \right] &= \mathbb{E}_{i,n} \left[\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{X}_{i+1} \theta_{i+1} \right] \\ &\quad + \mathbb{E}_{i,n} \left[\partial_{\bar{Y}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{Y}_i} \bar{Y}_{i+1} \theta_{i+1} \right] + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{Y}_i} \theta_{i+1} \right] \\ &= \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \left[\mathcal{I}_{i+1}^{(1)} (\partial_{\bar{Y}_i} \bar{X}_{i+1} \theta_{i+1}) + \mathcal{I}_{i+1}^{(2)} (m'_i \theta_{i+1}) \right] \right] \\ &\quad + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(2)} \left(\sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) \theta_{i+1} \right) \right] \\ &\quad + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(2)} \left(\sigma_{Y,i} \frac{\rho'_i}{1 - \rho_i^2} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right) \theta_{i+1} \right) \right] \\ &\quad + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{Y}_i} \theta_{i+1} \right]. \end{aligned}$$

On the other hand, again from the IBP formula (5.2.9), we obtain

$$\begin{aligned} \mathbb{E}_{i,n} \left[\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,X} \right] &+ \mathbb{E}_{i,n} \left[\partial_{\bar{Y}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,Y} \right] + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^c \right] \\ &= \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \left[\mathcal{I}_{i+1}^{(1)} (\vec{\theta}_{i+1}^{e,X}) + \mathcal{I}_{i+1}^{(2)} (\vec{\theta}_{i+1}^{e,Y}) + \vec{\theta}_{i+1}^c \right] \right]. \end{aligned}$$

Combining the two previous identities, we see that the difference

$$\begin{aligned} \partial_{\bar{Y}_i} \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \theta_{i+1} \right] &- \left(\mathbb{E}_{i,n} \left[\partial_{\bar{X}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,X} \right] \right. \\ &\quad \left. + \mathbb{E}_{i,n} \left[\partial_{\bar{Y}_{i+1}} h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^{e,Y} \right] + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^c \right] \right) \end{aligned}$$

can be written as

$$\begin{aligned}
 & \mathbb{E} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(2)} \left(m'_i \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) \right] + \mathbb{E} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \partial_{\bar{Y}_i} \theta_{i+1} \right] - \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) \right] \\
 & + \mathbb{E}_{i,n} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(1)} \left(\partial_{\bar{Y}_i} \bar{X}_{i+1} \theta_{i+1} \right) \right] + \mathbb{E} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(2)} \left(\sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) \theta_{i+1} \right) \right] \\
 & + \mathbb{E} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \mathcal{I}_{i+1}^{(2)} \left(\sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right) \theta_{i+1} \right) \right] \quad (5.6.9) \\
 & - \mathbb{E} \left[h(\bar{X}_{i+1}, \bar{Y}_{i+1}) \vec{\theta}_{i+1}^c \right].
 \end{aligned}$$

Before proceeding, we provide the explicit expression for the quantity $\partial_{\bar{Y}_i} \theta_{i+1}$. Using the chain rule formula of Lemma 5.2.2, after some standard but cumbersome computations, we obtain

$$\begin{aligned}
 \partial_{\bar{Y}_i} \theta_{i+1} &= (f(\zeta_{i+1} - \zeta_i))^{-1} \left[\mathcal{I}_{i+1}^{(1,1)} (\partial_{\bar{Y}_i} c_S^{i+1}) - \mathcal{I}_{i+1}^{(1)} (\partial_{\bar{Y}_i} c_S^{i+1}) + \mathcal{I}_{i+1}^{(2)} (\partial_{\bar{Y}_i} b_Y^{i+1}) + \mathcal{I}_{i+1}^{(1,2)} (\partial_{\bar{Y}_i} c_{Y,S}^{i+1}) \right. \\
 & + \mathcal{I}_{i+1}^{(2,2)} (\partial_{\bar{Y}_i} c_Y^{i+1}) - \frac{\sigma'_{S,i}}{\sigma_{S,i}} \left(2\mathcal{I}_{i+1}^{(1,1)} (c_S^{i+1}) - \mathcal{I}_{i+1}^{(1)} (c_S^{i+1}) + \mathcal{I}_{i+1}^{(1,2)} (c_{Y,S}^{i+1}) \right) \\
 & - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \left(2\mathcal{I}_{i+1}^{(2,2)} (c_Y^{i+1}) + \mathcal{I}_{i+1}^{(2)} (b_Y^{i+1}) + \mathcal{I}_{i+1}^{(1,2)} (c_{Y,S}^{i+1}) \right) \\
 & \left. + \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \left(\mathcal{I}_{i+1}^{(1,2)} (c_S^{i+1}) + \mathcal{I}_{i+1}^{(2,1)} (c_S^{i+1}) - \mathcal{I}_{i+1}^{(2)} (c_S^{i+1}) \right) \right].
 \end{aligned}$$

Also, after some simple algebraic simplifications using the definitions of $\vec{\theta}_{i+1}^{e,Y}$ and $\vec{\theta}_{i+1}^{e,X}$ in (5.4.3), one obtains

$$\mathcal{I}_{i+1}^{(2)} \left(m'_i \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) = -(f(\zeta_{i+1} - \zeta_i))^{-1} \left[\mathcal{I}_{i+1}^{(2,2)} \left(\partial_{\bar{Y}_i} c_Y^{i+1} \right) + \mathcal{I}_{i+1}^{(1,2)} \left(\partial_{\bar{Y}_i} c_{Y,S}^{i+1} \right) \right]$$

and

$$\mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) = (f(\zeta_{i+1} - \zeta_i))^{-1} \mathcal{I}_{i+1}^{(1,1)} \left(\partial_{\bar{Y}_i} c_S^{i+1} \right).$$

Combining the three previous identities and gathering similar terms, we obtain

$$\begin{aligned}
 & \mathcal{I}_{i+1}^{(2)} \left(m'_i \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) + \partial_{\bar{Y}_i} \theta_{i+1} - \mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) \\
 &= (f(\zeta_{i+1} - \zeta_i))^{-1} \left[-\mathcal{I}_{i+1}^{(1)} (\partial_{\bar{Y}_i} c_S^{i+1}) + \mathcal{I}_{i+1}^{(2)} (\partial_{\bar{Y}_i} b_Y^{i+1}) \right. \quad (5.6.10) \\
 & - \frac{\sigma'_{S,i}}{\sigma_{S,i}} \left(2\mathcal{I}_{i+1}^{(1,1)} (c_S^{i+1}) - \mathcal{I}_{i+1}^{(1)} (c_S^{i+1}) + \mathcal{I}_{i+1}^{(1,2)} (c_{Y,S}^{i+1}) \right) \\
 & - \left(\frac{\sigma'_{Y,i}}{\sigma_{Y,i}} - \frac{\rho'_i \rho_i}{1 - \rho_i^2} \right) \left(2\mathcal{I}_{i+1}^{(2,2)} (c_Y^{i+1}) + \mathcal{I}_{i+1}^{(2)} (b_Y^{i+1}) + \mathcal{I}_{i+1}^{(1,2)} (c_{Y,S}^{i+1}) \right) \\
 & \left. + \frac{\rho'_i}{1 - \rho_i^2} \frac{\sigma_{Y,i}}{\sigma_{S,i}} \left(\mathcal{I}_{i+1}^{(1,2)} (c_S^{i+1}) + \mathcal{I}_{i+1}^{(2,1)} (c_S^{i+1}) - \mathcal{I}_{i+1}^{(2)} (c_S^{i+1}) \right) \right].
 \end{aligned}$$

The previous identity will be used in the next step of the proof. Coming back to (5.6.9) and using the definition of the weight $\vec{\theta}_{i+1}^c$ allows to conclude the proof of the identity (5.4.3).

Step 3: The weight sequences $(\vec{\theta}_i^{e,Y})_{1 \leq i \leq n+1}$, $(\vec{\theta}_i^{e,X})_{1 \leq i \leq n+1}$ and $(\vec{\theta}_i^c)_{1 \leq i \leq n+1}$ and the related spaces $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, \ell/2)$, $\ell \in \mathbb{Z}$.

In this last step, we prove the last statement of Lemma 5.4.1 concerning the weight sequences $(\vec{\theta}_i^{e,Y})_{1 \leq i \leq n+1}$, $(\vec{\theta}_i^{e,X})_{1 \leq i \leq n+1}$ and $(\vec{\theta}_i^c)_{1 \leq i \leq n+1}$.

Following similar lines of reasonings as those used in the proof of Lemma 5.7.2, namely using the fact that

$$d_S^{i+1}, d_Y^{i+1}, d_{Y,S}^{i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 1/2)$$

and

$$\begin{aligned} & \mathcal{D}_{i+1}^{(1)} d_S^{i+1}, \mathcal{D}_{i+1}^{(1,1)} d_S^{i+1}, \mathcal{D}_{i+1}^{(2)} d_Y^{i+1}, \mathcal{D}_{i+1}^{(2,2)} d_Y^{i+1}, \mathcal{D}_{i+1}^{(1)} d_Y^{i+1}, \mathcal{D}_{i+1}^{(2)} d_Y^{i+1}, \mathcal{D}_{i+1}^{(1,2)} d_Y^{i+1}, \\ & e_S^{Y,i+1}, \mathcal{D}_{i+1}^{(1)} e_S^{Y,i+1}, e_Y^{Y,i+1}, \mathcal{D}_{i+1}^{(2)} e_Y^{Y,i+1}, e_S^{X,i+1}, \mathcal{D}_{i+1}^{(1)} e_S^{Y,i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0) \end{aligned}$$

as well as Lemma 5.2.3, we conclude

$$f(\zeta_{i+1} - \zeta_i) \vec{\theta}_{i+1}^{e,Y}, f(\zeta_{i+1} - \zeta_i) \vec{\theta}_{i+1}^{e,X} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2), \quad i \in \{0, \dots, n-1\}.$$

We now prove that $f(\zeta_{i+1} - \zeta_i) \vec{\theta}_{i+1}^c \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$ for any $i \in \{0, \dots, n-1\}$. We use the decomposition

$$\begin{aligned} f(\zeta_{i+1} - \zeta_i) \vec{\theta}_{i+1}^c &= f(\zeta_{i+1} - \zeta_i) \left(\mathcal{I}_{i+1}^{(2)} \left(m_i' \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) + \partial_{\bar{Y}_i} \theta_{i+1} - \mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) \right) \\ &+ \mathcal{I}_{i+1}^{(1)} \left(\partial_{\bar{Y}_i} \bar{X}_{i+1} f(\zeta_{i+1} - \zeta_i) \theta_{i+1} \right) + \mathcal{I}_{i+1}^{(2)} \left(\left(\sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) \right. \right. \\ &\left. \left. + \sigma_{Y,i} \frac{\rho_i'}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right) \right) f(\zeta_{i+1} - \zeta_i) \theta_{i+1} \right). \end{aligned}$$

We first prove that $f(\zeta_{i+1} - \zeta_i) \left(\mathcal{I}_{i+1}^{(2)} \left(m_i' \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) + \partial_{\bar{Y}_i} \theta_{i+1} - \mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) \right) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$. We investigate each term appearing on the right-hand side of (5.6.10).

In particular, we first use the fact that $c_S^{i+1}, c_Y^{i+1}, b_Y^{i+1}, c_{Y,S}^{i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 1/2)$, $\partial_{\bar{Y}_i} c_S^{i+1}, \partial_{\bar{Y}_i} b_Y^{i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)$ and that when one applies the differential operators $\mathcal{D}_{i+1}^{(\alpha_1)}, \mathcal{D}_{i+1}^{(\alpha_1, \alpha_2)}$ to these elements the resulting random variables belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)$ for any $(\alpha_1, \alpha_2) \in \{1, 2\}^2$. From Lemma 5.2.3, we thus conclude that the elements $\mathcal{I}_{i+1}^{(1)}(\partial_{\bar{Y}_i} c_S^{i+1}), \mathcal{I}_{i+1}^{(1,1)}(c_S^{i+1}), \mathcal{I}_{i+1}^{(1,2)}(c_S^{i+1}), \mathcal{I}_{i+1}^{(2,1)}(c_S^{i+1}), \mathcal{I}_{i+1}^{(2)}(\partial_{\bar{Y}_i} b_Y^{i+1}), \mathcal{I}_{i+1}^{(2,2)}(c_Y^{i+1}), \mathcal{I}_{i+1}^{(1,2)}(c_{Y,S}^{i+1})$ belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$ and that $\mathcal{I}_{i+1}^{(1)}(c_S^{i+1}), \mathcal{I}_{i+1}^{(2)}(b_Y^{i+1})$ belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)$. Moreover, using **ND**, one gets that there exists $C > 0$ such that for any $i \in \{0, \dots, n-1\}$, $|\sigma'_{S,i}/\sigma_{S,i}| + |\sigma'_{Y,i}/\sigma_{Y,i}| + |\sigma_{Y,i}/\sigma_{S,i}| + |\rho_i'|/(1 - \rho_i^2)| + |\rho_i' \rho_i|/(1 - \rho_i^2)| \leq C$. We thus conclude that $f(\zeta_{i+1} - \zeta_i) \left(\mathcal{I}_{i+1}^{(2)} \left(m_i' \theta_{i+1} - \vec{\theta}_{i+1}^{e,Y} \right) + \partial_{\bar{Y}_i} \theta_{i+1} - \mathcal{I}_{i+1}^{(1)} \left(\vec{\theta}_{i+1}^{e,X} \right) \right) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$.

It thus suffices to prove $\mathcal{I}_{i+1}^{(1)} \left(\partial_{\bar{Y}_i} \bar{X}_{i+1} f(\zeta_{i+1} - \zeta_i) \theta_{i+1} \right), \mathcal{I}_{i+1}^{(2)} \left(\sigma'_{Y,i} \left(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2 \right) f(\zeta_{i+1} - \zeta_i) \theta_{i+1} \right)$ and $\mathcal{I}_{i+1}^{(2)} \left(\sigma_{Y,i} \frac{\rho_i'}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2 \right) f(\zeta_{i+1} - \zeta_i) \theta_{i+1} \right)$ belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$.

In order to do this, we remark that

$$\begin{aligned}
 \partial_{\bar{Y}_i} \bar{X}_{i+1} &= -\frac{1}{2} a'_{S,i} + \sigma'_{S,i} Z_{i+1}^1 \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 1/2), \\
 \mathcal{D}_{i+1}^{(1)}(\partial_{\bar{Y}_i} \bar{X}_{i+1}) &= \frac{\sigma'_{S,i}}{\sigma_{S,i}} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0), \\
 \sigma'_{Y,i}(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2) &\in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 1/2), \\
 \mathcal{D}_{i+1}^{(2)}\left(\sigma'_{Y,i}(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2)\right) &= \frac{\sigma'_{Y,i}}{\sigma_{Y,i}} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0), \\
 \sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2\right) &\in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 1/2), \\
 \mathcal{D}_{i+1}^{(2)}\left(\sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} \left(\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2\right)\right) &= \rho'_i \frac{\rho_i}{1 - \rho_i^2} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)
 \end{aligned}$$

and from Lemma 5.7.2, $f(\zeta_{i+1} - \zeta_i)\theta_{i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$. From Lemma 5.2.3, it follows that $f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\partial_{\bar{Y}_i} \bar{X}_{i+1} \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)$ and $f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\sigma'_Y(m_i)m'_i(\rho(W_{\zeta_{i+1}} - W_{\zeta_i}) + \sqrt{1 - \rho^2}(\tilde{W}_{\zeta_{i+1}} - \tilde{W}_{\zeta_i})) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0)$. Now following similar computations as those employed in the proof of Lemma 5.7.2 and omitting some technical details we obtain $\mathcal{D}_{i+1}^{(\alpha)}(f(\zeta_{i+1} - \zeta_i)\theta_{i+1}) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1)$ so that from the chain rule formula and Lemma 5.2.3, the random variables

$$\begin{aligned}
 &\mathcal{D}_{i+1}^{(\alpha)}(f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\partial_{\bar{Y}_i} \bar{X}_{i+1}); \\
 &\mathcal{D}_{i+1}^{(\alpha)}(\sigma'_{Y,i}(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2)f(\zeta_{i+1} - \zeta_i)\theta_{i+1}); \\
 &\mathcal{D}_{i+1}^{(\alpha)}(\sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} (\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2)f(\zeta_{i+1} - \zeta_i)\theta_{i+1})
 \end{aligned}$$

belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$. From Lemma 5.2.3, we thus conclude that

$$\begin{aligned}
 &\mathcal{I}_{i+1}^{(1)}\left(\partial_{\bar{Y}_i} \bar{X}_{i+1} f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\right); \\
 &\mathcal{I}_{i+1}^{(2)}\left(\sigma'_{Y,i}(\rho_i Z_{i+1}^1 + \sqrt{1 - \rho_i^2} Z_{i+1}^2)f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\right); \\
 &\mathcal{I}_{i+1}^{(2)}\left(\sigma_{Y,i} \frac{\rho'_i}{\sqrt{1 - \rho_i^2}} (\sqrt{1 - \rho_i^2} Z_{i+1}^1 - \rho_i Z_{i+1}^2)f(\zeta_{i+1} - \zeta_i)\theta_{i+1}\right)
 \end{aligned}$$

belong to $\mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$. From the preceding arguments, we eventually deduce that $f(\zeta_{i+1} - \zeta_i)\vec{\theta}_{i+1}^c \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2)$ for any $i \in \{0, \dots, n-1\}$.

Finally, from the very definition of the weights on the last time interval $\vec{\theta}_{n+1}^{e,Y}$ and $\vec{\theta}_{n+1}^{e,X}$ one directly gets that

$$\begin{aligned}
 (1 - F(T - \zeta_n))\vec{\theta}_{n+1}^{e,Y} &= (1 - F(T - \zeta_n))^{-1} \left(m'_n + \sigma'_{Y,n} \left(\rho_n Z_{n+1}^1 + \sqrt{1 - \rho_n^2} Z_{n+1}^2 \right) \right. \\
 &\quad \left. + \sigma_{Y,n} \frac{\rho'_n}{\sqrt{1 - \rho_n^2}} \left(\sqrt{1 - \rho_n^2} Z_{n+1}^1 - \rho_n Z_{n+1}^2 \right) \right)
 \end{aligned}$$

belongs to $\mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 0)$ and that

$$(1 - F(T - \zeta_n)) \vec{\theta}_{n+1}^{e,X} = -\frac{1}{2} a'_{S,n} + \sigma'_{S,n} Z_{n+1}^1$$

belongs to $\mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 1/2)$. The proof is now complete.

5.7 Some technical results

5.7.1 Emergence of jumps in the renewal process N

The first result is used in the proof of the probabilistic representation in Theorem 5.3.1 and is used to express that time integrals add jumps to the renewal process N . In what follows, N is a renewal process in the sense of Definition 5.2.1.

Lemma 5.7.1. *Let $n \in \mathbb{N}$ and $G : \{(t_1, \dots, t_{n+2}) : 0 < t_1 < \dots < t_{n+1} < t_{n+2} := T\} \rightarrow \mathbb{R}_+$ be a measurable function such that $\mathbb{E} \left[\int_{\zeta_n}^T G(\zeta_1, \dots, \zeta_n, s, T) \mathbf{1}_{\{N_T=n\}} ds \right] < \infty$. Then*

$$\begin{aligned} & \mathbb{E} \left[\int_{\zeta_n}^T G(\zeta_1, \dots, \zeta_n, s, T) \mathbf{1}_{\{N_T=n\}} ds \right] \\ &= \mathbb{E} \left[G(\zeta_1, \dots, \zeta_n, \zeta_{n+1}, T) (1 - F(T - \zeta_{n+1}))^{-1} (1 - F(T - \zeta_n)) (f(\zeta_{n+1} - \zeta_n))^{-1} \mathbf{1}_{\{N_T=n+1\}} \right]. \end{aligned}$$

Proof. The proof follows by rewriting the above expectations using (5.2.3). We rewrite the expectation on the right-hand side in integral form. By Fubini's theorem, we obtain

$$\begin{aligned} & \mathbb{E} \left[G(\zeta_1, \dots, \zeta_n, \zeta_{n+1}, T) (1 - F(T - \zeta_{n+1}))^{-1} (1 - F(T - \zeta_n)) (f(\zeta_{n+1} - \zeta_n))^{-1} \mathbf{1}_{\{N_T=n+1\}} \right] \\ &= \int_{\Delta_{n+1}(T)} G(s_1, \dots, s_{n+1}, T) (1 - F(T - s_{n+1}))^{-1} (1 - F(T - s_n)) (f(s_{n+1} - s_n))^{-1} \\ & \quad \times (1 - F(T - s_{n+1})) \prod_{j=0}^n f(s_{j+1} - s_j) ds_{n+1} \\ &= \int_{\Delta_n(T)} \int_{s_n}^T G(s_1, \dots, s_{n+1}, T) ds_{n+1} (1 - F(T - s_n)) \prod_{j=0}^{n-1} f(s_{j+1} - s_j) ds_n. \end{aligned}$$

This completes the proof. \square

Lemma 5.7.2. *Let $n \in \mathbb{N}$. On the set $\{N_T = n\}$, the sequence of weights $(\theta_i)_{1 \leq i \leq n+1}$ defined by (5.3.3) and (5.3.4) satisfy:*

$$\forall i \in \{1, \dots, n\}, f(\zeta_i - \zeta_{i-1}) \theta_i \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1/2), \quad (1 - F(T - \zeta_n)) \theta_{n+1} \in \mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 0).$$

Proof. We investigate each term appearing in the definition of $\theta_i \in \mathbb{S}_{i-1,n}(\bar{X}, \bar{Y})$ and seek to apply Lemma 5.2.3. From the Lipschitz property of a_S and the space-time

inequality (5.1.3), for any $c > 0$ and any $c' > c$, the map $(x_{i-1}, y_{i-1}, x_i, y_i, \mathbf{s}_{n+1}) \mapsto c_S^i(x_{i-1}, y_{i-1}, x_i, y_i, \mathbf{s}_{n+1})$ satisfies

$$\begin{aligned} & |c_S^i(x_{i-1}, y_{i-1}, x_i, y_i, \mathbf{s}_{n+1})|^p \bar{q}_c(s_i - s_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) \\ & \leq C |y_i - m_{i-1}(y_{i-1})|^p \bar{q}_c(s_i - s_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) \\ & \leq C (s_i - s_{i-1})^{p/2} \bar{q}_{c'}(s_i - s_{i-1}, x_{i-1}, y_{i-1}, x_i, y_i) \end{aligned}$$

so that, the random variables $c_S^i \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 1/2)$, for any $i \in \{1, \dots, n+1\}$. Moreover, from the boundedness of the first and second derivatives of a_S , it follows

$$\mathcal{D}_i^{(1)} c_S^i = \partial_{\bar{Y}_i} c_S^i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^{n+1}) = \partial_4 c_S^i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^{n+1}) = \frac{1}{2} a_S'(\bar{Y}_i) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 0)$$

and

$$\mathcal{D}_i^{(1,1)} c_S^i = \partial_{\bar{Y}_i}^2 c_S^i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^{n+1}) = \partial_4^2 c_S^i(\bar{X}_{i-1}, \bar{Y}_{i-1}, \bar{X}_i, \bar{Y}_i, \zeta^{n+1}) = \frac{1}{2} a_S''(\bar{Y}_i) \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, 0)$$

From Lemma 5.2.3, we thus conclude

$$\mathcal{I}_i^{(1)}(c_S^i) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0) \quad \text{and} \quad \mathcal{I}_i^{(1,1)}(c_S^i) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2), \quad i \in \{1, \dots, n\}.$$

In a completely analogous manner, omitting some technical details, we derive

$$\mathcal{I}_i^{(2)}(b_Y^i) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, 0), \quad \text{and} \quad \mathcal{I}_i^{(1,2)}(c_{Y,S}^i), \mathcal{I}_i^{(2,2)}(c_Y^i) \in \mathbb{M}_{i,n}(\bar{X}, \bar{Y}, -1/2).$$

Hence, we obtain $f(\zeta_i - \zeta_{i-1})\theta_i \in \mathbb{M}_{i-1,n}(\bar{X}, \bar{Y}, -1/2)$, for any $i \in \{1, \dots, n\}$. We finally observe that $(1 - F(T - \zeta_n))\theta_{n+1} = 1 \in \mathbb{M}_{n,n}(\bar{X}, \bar{Y}, 0)$. The proof is now complete. \square

Lemma 5.7.3. *Let $T > 0$ and n a positive integer. For any $\mathbf{s}_n = (s_1, \dots, s_n) \in \Delta_n(T)$, for any $(x, y) \in \mathbb{R}^2$, for any positive constant c there exist two positive constants C and $c' > c$ such that the transition density $(s, x, y) \mapsto \bar{q}_c(s, x_0, y_0, x, y)$ satisfies the following semigroup property:*

$$\begin{aligned} & \int_{(\mathbb{R}^2)^n} \bar{q}_c(T - s_n, x_n, y_n, x, y) \times \bar{q}_c(s_n - s_{n-1}, x_{n-1}, y_{n-1}, x_n, y_n) \times \dots \times \bar{q}_c(s_1, x_0, y_0, x_1, y_1) \, d\mathbf{x}_n d\mathbf{y}_n \\ & \leq C^n \bar{q}_{c'}(T, x_0, y_0, x, y). \end{aligned}$$

Proof. The $dx_1 \dots, dx_n$ integrals are treated using the standard semigroup property of Gaussian kernels so that it directly follows

$$\begin{aligned} & \int_{(\mathbb{R}^2)^n} \bar{q}_c(T - s_n, x_n, y_n, x, y) \times \bar{q}_c(s_n - s_{n-1}, x_{n-1}, y_{n-1}, x_n, y_n) \times \dots \times \bar{q}_c(s_1, x_0, y_0, x_1, y_1) \, d\mathbf{x}_n d\mathbf{y}_n \\ & \leq \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-x_0)^2}{2cT}} \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi c(T-s_n)}} e^{-\frac{(y-m_{T-s_n}(y_n))^2}{2c(T-s_n)}} \times \dots \times \frac{1}{\sqrt{2\pi c s_1}} e^{-\frac{(y_1-m_{s_1}(y_0))^2}{2c s_1}} \, d\mathbf{y}_n \end{aligned}$$

We now perform the change of variables $y_1 = m_{s_1}(z_1), y_2 = m_{s_2}(z_2), \dots, y_n = m_{s_n}(z_n)$. Observe that since b_Y admits a bounded first derivative the determinants of the Jacobians $J_{s_1}(z_1) := \partial_x m_{s_1}(z_1), \dots, J_{T-s_n}(z_n) = \partial_x m_{T-s_n}(z_n)$ are (locally)

uniformly bounded for any $(s_1, \dots, s_n) \in \Delta_n(T)$. Remark also that from the semigroup property $m_{s_{i+1}-s_i}(m_{s_i}(z_i)) = m_{s_{i+1}}(z_i)$, for $1 \leq i \leq n$ with the convention $s_{n+1} = T$. Hence, for some positive constants C and $c' > c$ that may change from line to line, we get

$$\begin{aligned}
& \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-x_0)^2}{2cT}} \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi c(T-s_n)}} e^{-\frac{(y-m_{T-s_n}(y_n))^2}{2c(T-s_n)}} \times \dots \times \frac{1}{\sqrt{2\pi cs_1}} e^{-\frac{(y_1-m_{s_1}(y_0))^2}{2cs_1}} dy_n \\
& \leq C^n \frac{1}{\sqrt{2\pi cT}} e^{-\frac{(x-x_0)^2}{2cT}} \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi c(T-s_n)}} e^{-\frac{(y-m_{T-s_n}(z_n))^2}{2c(T-s_n)}} \times \dots \times \frac{1}{\sqrt{2\pi cs_1}} e^{-\frac{(m_{s_1}(z_1)-m_{s_1}(y_0))^2}{2cs_1}} dz_n \\
& \leq C^n \frac{1}{\sqrt{2\pi cT}} e^{-\frac{(x-x_0)^2}{2cT}} \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi c(T-s_n)}} e^{-C^{-1}\frac{(m_T^{-1}(y)-z_n)^2}{2c(T-s_n)}} \times \dots \times \frac{1}{\sqrt{2\pi cs_1}} e^{-C^{-1}\frac{(z_1-y_0)^2}{2cs_1}} dz_n \\
& \leq C^n \frac{1}{2\pi cCT} e^{-\frac{(x-x_0)^2}{2cT}} e^{-\frac{(m_T^{-1}(y)-y_0)^2}{2cCt}} \\
& \leq C^n \bar{q}_{c'}(Tx_0, y_0, x_1, y_1)
\end{aligned}$$

where we first used the bi-Lipschitz property of the flow $(s, x) \mapsto m_s(x)$ which yields

$$\forall t \in [0, T], C^{-1}|x - z|^2 \leq |m_t(x) - m_t(z)|^2 \leq C|x - z|^2$$

for some positive constant $C \geq 1$ and then the semigroup property satisfied by Gaussian kernels. This completes the proof. \square

5.7.2 Formulae for the computations of price, Delta and Vega

The following quantities are needed for the formula of price

$$\begin{aligned}
\mathcal{I}_i^{(1,1)}(c_S^i) &= c_S^i((\mathcal{I}_i^{(1)}(1))^2 - \mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1)), \\
\mathcal{I}_i^{(1)}(c_S^i) &= c_S^i\mathcal{I}_i^{(1)}(1), \\
\mathcal{I}_i^{(2)}(b_Y^i) &= b_Y^i\mathcal{I}_i^{(2)}(1) + \lambda, \\
\mathcal{I}_i^{(1,2)}(c_{Y,S}^i) &= \mathcal{I}_i^{(2)}(c_{Y,S}^i\mathcal{I}_i^{(1)}(1)) = c_{Y,S}^i\mathcal{I}_i^{(1)}(1)\mathcal{I}_i^{(2)}(1) - \mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(2)}(c_{Y,S}^i) - c_{Y,S}^i\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1),
\end{aligned}$$

$$\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{I}_i^{(1)}(c_S^i) + \mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{I}_i^{(1,2)}(c_{Y,S}^i) \right], \quad 1 \leq i \leq N_T.$$

For Delta, we should calculate

$$\begin{aligned}
\mathcal{D}_i^{(1)}(\mathcal{I}_i^{(1,1)}(c_S^i)) &= 2c_S^i\mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1), \\
\mathcal{D}_i^{(1)}(\mathcal{I}_i^{(1)}(c_S^i)) &= c_S^i\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1), \\
\mathcal{D}_i^{(1)}(\mathcal{I}_i^{(2)}(b_Y^i)) &= b_Y^i\mathcal{D}_i^{(1)}\mathcal{I}_i^{(2)}(1), \\
\mathcal{D}_i^{(1)}(\mathcal{I}_i^{(1,2)}(c_{Y,S}^i)) &= c_{Y,S}^i\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1)\mathcal{I}_i^{(2)}(1) + c_{Y,S}^i\mathcal{D}_i^{(1)}\mathcal{I}_i^{(2)}(1)\mathcal{I}_i^{(1)}(1) - \mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(2)}(c_{Y,S}^i).
\end{aligned}$$

$$\mathcal{D}_i^{(1)}\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1,1)}(c_S^i) - \mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(c_S^i) + \mathcal{D}_i^{(1)}\mathcal{I}_i^{(2)}(b_Y^i) + \mathcal{D}_i^{(1)}\mathcal{I}_i^{(1,2)}(c_{Y,S}^i) \right],$$

$$\mathcal{I}_k^{(1)}(\theta_k) = \mathcal{I}_k^{(1)}(1)\theta_k - \mathcal{D}_k^{(1)}\theta_k, \quad k \leq N_T,$$

$$\mathcal{I}_{N_T+1}^{(1)}(\theta_{N_T+1}) = \theta_{N_T+1}\mathcal{I}_{N_T+1}^{(1)}(1) - \mathcal{D}_{N_T+1}^{(1)}\theta_{N_T+1} = \theta_{N_T+1}\mathcal{I}_{N_T+1}^{(1)}(1).$$

Finally, these quantities are needed for the formula of Vega:

$$\mathcal{I}_{i+1}^{(1,1)}(d_S^{i+1}) = m'_i\mathcal{I}_{i+1}^{(1,1)}(c_S^{i+1}),$$

$$\mathcal{I}_{i+1}^{(1)}(e_S^{Y,i+1}) = -m'_i\mathcal{I}_{i+1}^{(1)}(c_S^{i+1}) + \mathcal{D}_i^{(2)}(c_{Y,S}^{i+1})\mathcal{I}_{i+1}^{(1)}(1) - \mathcal{D}_{i+1}^{(1)}\mathcal{D}_i^{(2)}(c_{Y,S}^{i+1}),$$

$$\mathcal{I}_{i+1}^{(2)}(e_Y^{Y,i+1}) = m'_i\mathcal{I}_{i+1}^{(2)}(b_Y^{i+1}),$$

$$\mathcal{I}_{i+1}^{(1,2)}(d_{Y,S}^{i+1}) = m'_i\mathcal{I}_{i+1}^{(1,2)}(c_{Y,S}^{i+1}),$$

$$\mathcal{I}_{i+1}^{(1)}(e_S^{X,i+1}) = e_S^{X,i+1}\mathcal{I}_{i+1}^{(1)}(1) - \mathcal{D}_{i+1}^{(1)}e_S^{X,i+1} = e_S^{X,i+1}\mathcal{I}_{i+1}^{(1)}(1) - \mathcal{D}_{i+1}^{(1)}\mathcal{D}_i^{(2)}(c_S^{i+1}),$$

$$\mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1,1)}(c_S^i)) = \mathcal{D}_i^{(2)}(c_S^i)(\mathcal{I}_i^{(1)}(1))^2 + 2c_S^i\mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1) - \mathcal{D}_i^{(2)}(c_S^i)\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1),$$

$$\mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1)}(c_S^i)) = \mathcal{D}_i^{(2)}(c_S^i)\mathcal{I}_i^{(1)}(1) + c_S^i\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1),$$

$$\mathcal{D}_i^{(2)}(\mathcal{I}_i^{(2)}(b_Y^i)) = b_Y^i\mathcal{D}_i^{(2)}\mathcal{I}_i^{(2)}(1) - \lambda\mathcal{I}_i^{(2)}(1),$$

$$\begin{aligned} \mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1,2)}(c_{Y,S}^i)) &= \mathcal{D}_i^{(2)}(c_{Y,S}^i)\mathcal{I}_i^{(1)}(1)\mathcal{I}_i^{(2)}(1) + c_{Y,S}^i\mathcal{D}_i^{(2)}\mathcal{I}_i^{(2)}(1)\mathcal{I}_i^{(1)}(1) + c_{Y,S}^i\mathcal{I}_i^{(2)}(1)\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1) \\ &\quad - \mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(2,2)}(c_{Y,S}^i) - 2\mathcal{D}_i^{(2)}(c_{Y,S}^i)\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1), \end{aligned}$$

$$\begin{aligned} \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1,1)}(c_S^i)) &= \mathcal{D}_{i-1}^{(2)}(c_S^i)(\mathcal{I}_i^{(1)}(1))^2 + 2c_S^i\mathcal{I}_i^{(1)}(1)\mathcal{D}_{i-1}^{(2)}\mathcal{I}_i^{(1)}(1) \\ &\quad - \mathcal{D}_{i-1}^{(2)}(c_S^i)\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1) - c_S^i\mathcal{D}_{i-1}^{(2)}(\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1)), \end{aligned}$$

$$\mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1)}(c_S^i)) = \mathcal{D}_{i-1}^{(2)}(c_S^i)\mathcal{I}_i^{(1)}(1) + c_S^i\mathcal{D}_{i-1}^{(2)}\mathcal{I}_i^{(1)}(1),$$

$$\mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(2)}(b_Y^i)) = b_Y^i\mathcal{D}_{i-1}^{(2)}\mathcal{I}_i^{(2)}(1) - \mathcal{I}_i^{(2)}(1)\lambda(\partial_{\bar{Y}_{i-1}}\bar{Y}_i - m'_{i-1}),$$

$$\begin{aligned} \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1,2)}(c_{Y,S}^i)) &= \mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{I}_i^{(1)}(1)\mathcal{I}_i^{(2)}(1) + c_{Y,S}^i(\mathcal{I}_i^{(2)}(1)\mathcal{D}_{i-1}^{(2)}\mathcal{I}_i^{(1)}(1) + \mathcal{I}_i^{(1)}(1)\mathcal{D}_{i-1}^{(2)}\mathcal{I}_i^{(2)}(1)) \\ &\quad - \mathcal{I}_i^{(1)}(1)\mathcal{D}_{i-1}^{(2)}(\mathcal{D}_i^{(2)}(c_{Y,S}^i)) - \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1)}(1))\mathcal{D}_i^{(2)}(c_{Y,S}^i) - c_{Y,S}^i\mathcal{D}_{i-1}^{(2)}\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1) \\ &\quad - \mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1). \end{aligned}$$

$$\mathcal{D}_i^{(2)}\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1,1)}(c_S^i)) - \mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1)}(c_S^i)) + \mathcal{D}_i^{(2)}(\mathcal{I}_i^{(2)}(b_Y^i)) + \mathcal{D}_i^{(2)}(\mathcal{I}_i^{(1,2)}(c_{Y,S}^i)) \right],$$

$$\mathcal{D}_{i-1}^{(2)}\theta_i = (f(\zeta_i - \zeta_{i-1}))^{-1} \left[\mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1,1)}(c_S^i)) - \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1)}(c_S^i)) + \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(2)}(b_Y^i)) + \mathcal{D}_{i-1}^{(2)}(\mathcal{I}_i^{(1,2)}(c_{Y,S}^i)) \right],$$

$$\begin{aligned}
\vec{\theta}_i^{e,Y} &= m'_{i-1}\theta_i + (f(\zeta_i - \zeta_{i-1}))^{-1}(\mathcal{I}_i^{(1)}(1)\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i) - \mathcal{D}_i^{(1)}\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)), \\
\mathcal{D}_i^{(2)}\vec{\theta}_i^{e,Y} &= m'_{i-1}\mathcal{D}_i^{(2)}\theta_i + (f(\zeta_i - \zeta_{i-1}))^{-1}\left[\mathcal{D}_i^{(2)}(\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i))\mathcal{I}_i^{(1)}(1) \right. \\
&\quad \left. + \mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{D}_i^{(2)}\mathcal{I}_i^{(1)}(1) - \mathcal{D}_i^{(2)}\mathcal{D}_i^{(1)}\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\right], \\
\mathcal{I}_i^{(2)}(\vec{\theta}_i^{e,Y}) &= \vec{\theta}_i^{e,Y}\mathcal{I}_i^{(2)}(1) - \mathcal{D}_i^{(2)}\vec{\theta}_i^{e,Y}, \\
\mathcal{I}_i^{(2)}(m'_{i-1}\theta_i - \vec{\theta}_i^{e,Y}) &= -(f(\zeta_i - \zeta_{i-1}))^{-1}\mathcal{I}_i^{(1,2)}(\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)) \\
&= -(f(\zeta_i - \zeta_{i-1}))^{-1}\left(\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{I}_i^{(1)}(1)\mathcal{I}_i^{(2)}(1) - \mathcal{I}_i^{(1)}(1)\mathcal{D}_i^{(2)}\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i) \right. \\
&\quad \left. - \mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{D}_i^{(1)}\mathcal{I}_i^{(2)}(1) - \mathcal{D}_i^{(1)}\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\mathcal{I}_i^{(2)}(1) + \mathcal{D}_i^{(2)}\mathcal{D}_i^{(1)}\mathcal{D}_{i-1}^{(2)}(c_{Y,S}^i)\right), \\
\vec{\theta}_i^{e,X} &= (f(\zeta_i - \zeta_{i-1}))^{-1}\mathcal{I}_i^{(1)}(e_S^{X,i}) \\
&= (f(\zeta_i - \zeta_{i-1}))^{-1}(\mathcal{I}_i^{(1)}(1)\mathcal{D}_{i-1}^{(2)}(c_S^i) - \mathcal{D}_i^{(1)}\mathcal{D}_{i-1}^{(2)}(c_S^i)), \\
\mathcal{D}_i^{(1)}\vec{\theta}_i^{e,X} &= (f(\zeta_i - \zeta_{i-1}))^{-1}e_S^{X,i}\mathcal{D}_i^{(1)}\mathcal{I}_i^{(1)}(1), \\
\mathcal{I}_i^{(1)}(\vec{\theta}_i^{e,X}) &= \vec{\theta}_i^{e,X}\mathcal{I}_i^{(1)}(1) - \mathcal{D}_i^{(1)}\vec{\theta}_i^{e,X}, \\
\mathcal{I}_i^{(1)}(\theta_i\mathcal{D}_{i-1}^{(2)}\bar{X}_i) &= (\theta_i\mathcal{I}_i^{(1)}(1) - \mathcal{D}_i^{(1)}\theta_i)\mathcal{D}_{i-1}^{(2)}\bar{X}_i - \mathcal{D}_i^{(1)}(\mathcal{D}_{i-1}^{(2)}\bar{X}_i)\theta_i, \\
\vec{\theta}_i^c &= \mathcal{I}_i^{(2)}(m'_{i-1}\theta_i - \vec{\theta}_i^{e,Y}) + \sigma_{Y,i-1}\frac{\rho'_{i-1}}{\sqrt{1-\rho_{i-1}^2}}\mathcal{I}_i^{(2)}\left(\left(\sqrt{1-\rho_{i-1}^2}Z_i^1 - \rho_{i-1}Z_i^2\right)\theta_i\right) \\
&\quad + \mathcal{I}_i^{(1)}(\theta_i\mathcal{D}_{i-1}^{(2)}\bar{X}_i) - \mathcal{I}_i^{(1)}(\vec{\theta}_i^{e,X}) + \mathcal{D}_{i-1}^{(2)}\theta_i, \\
\mathcal{I}_i^{(2)}\left(\left(\sqrt{1-\rho_{i-1}^2}Z_i^1 - \rho_{i-1}Z_i^2\right)\theta_i\right) &= \left(\sqrt{1-\rho_{i-1}^2}Z_i^1 - \rho_{i-1}Z_i^2\right)\left(\mathcal{I}_i^{(2)}(1)\theta_i - \mathcal{D}_i^{(2)}\theta_i\right) \\
&\quad + \frac{\rho_{i-1}\theta_i}{\sigma_{Y,i-1}\sqrt{1-\rho_{i-1}^2}}, \\
\vec{\theta}_{N_T+1}^{e,Y} &= \theta_{N_T+1}\left(m'_{N_T} + \sigma_{Y,N_T}\frac{\rho'_{N_T}}{\sqrt{1-\rho_{N_T}^2}}\left(\sqrt{1-\rho_{N_T}^2}Z_{N_T+1}^1 - \rho_{N_T}Z_{N_T+1}^2\right)\right), \\
\vec{\theta}_{N_T+1}^{e,X} &= \theta_{N_T+1}\left(-\frac{1}{2}a'_{S,N_T} + \sigma'_{S,N_T}Z_{N_T+1}^1\right), \\
\vec{\theta}_{N_T+1}^c &= 0, \\
\mathcal{I}_{N_T+1}^{(2)}(\vec{\theta}_{N_T+1}^{e,Y}) &= \vec{\theta}_{N_T+1}^{e,Y}\mathcal{I}_{N_T+1}^{(2)}(1) - \mathcal{D}_{N_T+1}^{(2)}\vec{\theta}_{N_T+1}^{e,Y} = \vec{\theta}_{N_T+1}^{e,Y}\mathcal{I}_{N_T+1}^{(2)}(1) + \theta_{N_T+1}\frac{\rho'_{N_T}\rho_{N_T}}{1-\rho_{N_T}^2}, \\
\mathcal{I}_{N_T+1}^{(2)}(\vec{\theta}_{N_T+1}^{e,X}) &= \vec{\theta}_{N_T+1}^{e,X}\mathcal{I}_{N_T+1}^{(2)}(1) - \mathcal{D}_{N_T+1}^{(2)}\vec{\theta}_{N_T+1}^{e,X} = \vec{\theta}_{N_T+1}^{e,X}\mathcal{I}_{N_T+1}^{(2)}(1) - \theta_{N_T+1}\frac{\sigma'_{S,N_T}}{\sigma_{S,N_T}}.
\end{aligned}$$

Part IV

Probabilistic Numerical Methods for Mean-Field Games

Probabilistic Numerical Methods for Mean-Field Games

The content of this chapter is from an article in collaboration with Andrea Angiuli, Christy V. Graves, Jean-François Chassagneux, François Delarue and René Carmona [Ang+19], published in ESAIM: Proceedings and Surveys.

6.1 Introduction

In this chapter we examine numerical methods for solving forward backward stochastic differential equations (FBSDEs) of McKean-Vlasov type. We are particularly interested in equations of this type as they can be used to represent, from the probabilistic viewpoint, solutions to mean field games or, more generally, to mean field stochastic control problems.

Mean field games were developed independently and at about the same time by Lasry and Lions [LL07], and Huang, Caines, and Malhamé [HMC+06]. The goal of this theory is to understand the limit as $N \rightarrow \infty$ of the Nash equilibria for an N player stochastic dynamic game with mean field interaction, meaning that players interact with one another through their collective state. Equivalently, mean field games must be regarded as the continuum limit of games with a large number of symmetric players, each of them having a small effect on the dynamics of the whole group. The applications of mean field games are numerous, and spread across many disciplines, including social science (congestion [LW11][GM15][AL16], cyber attacks [CW16]), biology (flocking [NCM10][NCM11]), and economics (systemic risk [CFS13], production of exhaustible resources [HHS10] [CS17]), just to name a few. As explained in [CD+18; CD13], solutions to mean field games can be characterized through a coupled system of two forward and backward stochastic differential equations of mean field type, like those we address in this note. The forward equation provides the dynamics of one typical player in the population at equilibrium. Generally speaking, the backward equation accounts for the optimality condition in the definition of an equilibrium and the McKean-Vlasov structure is precisely here to stress the fact that the player in hand is representative of the others. As exemplified in [CDL13], other forms of equilibria can be addressed by means of this kind of equations. This includes optimal mean field control problems,

which can be regarded as the continuum limit of a control problem involving a large number of symmetric players obeying a central planner. Below, we mostly focus on examples arising in the theory of mean field games.

Whilst deterministic numerical methods, based upon finite differences or variational approaches, are also conceivable for handling mean field games, see [ACD10; ACCD13; AP16] and [BC15; LST10; Gué12], we here focus on the approach based on FBSDEs. In this regard, we implement (and compare) two different algorithms. The first algorithm, which is based on the paper of Chassagneux, Crisan, and Delarue [CCD+19], relies on a tree structure to represent the pathwise law of the solution. The second algorithm, and main contribution of this chapter, takes the algorithm presented in the paper of Delarue and Menozzi [DM+06] for solving FBSDEs and extends it to the mean field framework in hand. In this algorithm, a grid structure is used to represent the marginal laws of the solution. The serious issue that we are facing in this note is that both methods are based upon a Picard scheme, the first method involving a global Picard scheme upon the whole process and the second one involving a Picard scheme on the sole marginal laws of the process. It is indeed a well-known fact that, because of the strong coupling between the forward and backward equations, Picard schemes for FBSDEs may just converge in small time, even in the classical case without mean field interaction. For sure, this limitation should persist in the mean field setting for the global Picard method; as exemplified below, it turns out that it persists as well when the Picard scheme is applied to the marginal laws. One of our main contribution in this chapter is to apply the time continuation approach presented in [CCD+19] to the grid algorithm and to compare the results with the tree algorithm for which the time continuation approach was originally designed in [CCD+19]. In brief, the time continuation permits to extend, by a continuation argument, the time interval on which the Picard scheme converges. We illustrate both algorithms on a handful of example problems.

Section 6.2 provides a review of Nash equilibria in N player stochastic differential games, and their continuum mean field game counterparts. We review two probabilistic approaches to formulate the solutions of mean field games and provide the general FBSDE system which we would like to solve. In Section 6.3, we describe the algorithms that we implement in this chapter. Some benchmark examples and the corresponding numerical results are presented in Section 6.4. We conclude in Section 6.5.

6.2 Overview of Mean Field Games and FBSDEs

The purpose of this section is to introduce the theoretical material that is needed for our numerical analysis. The objective is purely pedagogical and the text does not contain any new result.

6.2.1 N Player Stochastic Differential Games

We start with the description of the prototype of a finite player game in the theory of mean field games. We consider $N \in \mathbb{Z}^+$ players indexed by $i \in \{1, \dots, N\}$. The dynamic game occurs over a fixed time horizon $[0, T]$ for some $T > 0$. We have N independent m -dimensional Brownian motions $(W_t^i)_{0 \leq t \leq T}$ which are supported by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$. Each player chooses its control $\alpha^i =$

$(\alpha_t^i)_{0 \leq t \leq T}$ from the set \mathbb{A} defined as the set of square integrable \mathbb{F} adapted processes with values in a given set A (typically A is a closed convex subset of a Euclidean space). Each player i has a state X^i which evolves according to the stochastic differential equation:

$$dX_t^i = b^i(t, X_t^i, \bar{\mu}_t, \alpha_t)dt + \sigma^i(t, X_t^i, \bar{\mu}_t, \alpha_t)dW_t^i,$$

where $\bar{\mu}_t$ denotes the empirical distribution of the players' states: $\bar{\mu}_t = \frac{1}{N} \sum_{j=0}^N \delta_{X_t^j} \in \mathcal{P}_2(\mathbb{R}^d)$. Here $\mathcal{P}_2(\mathbb{R}^d)$ is the space of probability measures with a finite second moment, which we equip with the 2-Wasserstein distance, denoted by W_2 . For $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, we call $\Gamma(\mu, \nu)$ the set of all the joint laws with marginals μ and ν . Then, the 2-Wasserstein distance is defined by:

$$W_2(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int |x - y|^2 d\gamma(x, y) \right)^{1/2}.$$

The drift and volatility functions, b^i and σ^i , respectively, are deterministic functions $(b^i, \sigma^i) : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^d \times \mathbb{R}^{d \times m}$. Most of the time, they are assumed to be bounded in time and to be Lipschitz continuous with respect to all the arguments, the Lipschitz property in the measure argument being understood with respect to W_2 . This ensures that, for a given $\alpha = (\alpha^1, \dots, \alpha^N)$, the state dynamics (X^1, \dots, X^N) is well defined.

Given a tuple of controls $\alpha = (\alpha^1, \dots, \alpha^N)$, we associate with player i a cost objective which we take to be of the form:

$$J^i(\alpha) = \mathbb{E} \left[\int_0^T f^i(t, X_t^i, \bar{\mu}_t, \alpha_t^i)dt + g^i(X_T^i, \bar{\mu}_T) \right].$$

Thus each player considers a deterministic running cost $f^i : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}$, and deterministic terminal cost $g^i : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$. Of course, each of them wishes to minimize its own cost by tuning its own control in the most relevant way. Note that we only allow the interaction of the players through their empirical measure, as this will be needed in our formulation of the continuum limit. Still, extensions exist, in which players also interact through the controls, see Subsection 6.2.2.3.

The players are in a *Nash equilibrium* if each player is no better off for switching their strategy when they consider the other players' strategies to be fixed. More precisely, the set of strategies α is a Nash equilibrium if

$$J^i(\alpha) \leq J^i(\alpha^1, \dots, \alpha^{i-1}, \alpha, \alpha^{i+1}, \dots, \alpha^N), \forall \alpha \in \mathbb{A}, \forall i \in \{1, \dots, N\}.$$

6.2.2 Mean Field Games

For games where N is large, the problem quickly becomes of an intractable complexity. Thus we turn to the continuum limit by considering the limit as N tends to infinity. In order for this limit to make sense, we require the players to be symmetric. Precisely, we require $b = b^i$, $\sigma = \sigma^i$, $f = f^i$, and $g = g^i \forall i \in \{1, \dots, N\}$. As the number of players increase, the impact of each player on the empirical distribution decreases, and we expect to have a propagation of chaos such that the players become asymptotically independent of each other. This is the rationale for passing to

the limit: Asymptotically, the influence of one player on the group should be null and the statistical structure of the whole should be pretty simple.

We wish to formulate the analogue of a Nash equilibrium when there is a continuum of players. To this end, we consider the states and actions of the other players to be fixed, and consider the best response for a representative player (as we expect equilibria to inherit the symmetric structure of the game). Thus, the first step is to solve an optimization problem. The next step is to find a fixed point, providing an analogue of a Nash equilibrium for the mean field game.

We again have a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ where the filtration supports an m -dimensional Brownian motion $W = (W_t)_{0 \leq t \leq T}$ and an initial condition $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$.

The strategy for solving the asymptotic game is the following:

1. For a fixed deterministic flow of probability measures $\mu = (\mu_t)_{0 \leq t \leq T} \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$, solve the standard stochastic control problem:

$$\inf_{\alpha \in \mathbb{A}} J^\mu(\alpha) = \mathbb{E} \left[\int_0^T f(t, X_t^\alpha, \mu_t, \alpha_t) dt + g(X_T^\alpha, \mu_T) \right], \quad (6.2.1)$$

subject to

$$\begin{aligned} dX_t^\alpha &= b(t, X_t^\alpha, \mu_t, \alpha_t) dt + \sigma(t, X_t^\alpha, \mu_t, \alpha_t) dW_t \\ X_0^\alpha &= \xi. \end{aligned}$$

2. Find a fixed point, μ , such that $\mathcal{L}(X_t^\alpha) = \mu_t$ for all $0 \leq t \leq T$.

This strategy can be tackled from either the PDE viewpoint (leading to a coupled Hamilton-Jacobi-Bellman and Kolmogorov/Fokker-Plank equations, known as the MFG system in the literature) [LL07] [HMC+06] or the probabilistic viewpoint [CD+18] [CD13], which is the focus of this project. Within the probabilistic viewpoint, there are two approaches, both of which are formulated with FBSDEs. See Chapters 3 and 4 of the manuscript of Carmona and Delarue [CD+18] for reference on the two probabilistic approaches.

For simplicity, from now on we assume m , the dimension of the Brownian motion matches d , the dimension of the state variable. We also assume the diffusion coefficient, σ , is a constant matrix $\sigma \in \mathbb{R}^{d \times d}$. For both approaches, we will utilize the Hamiltonian deriving from the aforementioned stochastic control problem (6.2.1). In fact, since we assume that the drift is uncontrolled, we can just write the reduced Hamiltonian:

$$H(t, x, \mu, \alpha, y) = b(t, x, \mu, \alpha) \cdot y + f(t, x, \mu, \alpha),$$

for $t \in [0, T]$, $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in A$, and an adjoint variable $y \in \mathbb{R}^d$. Then, a key object in order to formulate the solution to (6.2.1) is (whenever it exists):

$$\hat{\alpha}(t, x, \mu, y) = \arg \inf_{\alpha \in A} H(t, x, \mu, \alpha, y).$$

We will provide below explicit examples for $\hat{\alpha}(t, x, \mu, y)$. We now give a brief introduction to the two probabilistic approaches.

6.2.2.1 Weak approach

In the first probabilistic approach, which we will refer to as the weak approach, the optimization problem is solved using a backward SDE for the probabilistic representation of the value function. For a fixed flow of measures $\mu = (\mu_t)_{0 \leq t \leq T}$, let $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ denote the value function:

$$u(t, x) := \inf_{(\alpha_s)_{t \leq s \leq T} \in \mathbb{A}} \mathbb{E} \left[\int_t^T f(s, X_s, \mu_s, \alpha_s) ds + g(X_T, \mu_T) \mid X_t = x \right].$$

The strategy is to evaluate the value function along the solution of the state process $(X_t)_{0 \leq t \leq T}$, namely we let $Y_t = u(t, X_t)$. The weak formulation of the stochastic control problem underpinning u says that, under suitable assumptions that are exemplified below, the pair $(X_t, Y_t)_{0 \leq t \leq T}$ has to solve the following FBSDE:

$$\begin{aligned} dX_t &= b(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1} Z_t)) dt + \sigma dW_t \\ X_0 &= \xi, \\ dY_t &= -f(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, \sigma^{-1} Z_t)) dt + Z_t dW_t \\ Y_T &= g(X_T, \mu_T), \end{aligned}$$

where we assume σ to be invertible. For instance, we take the following set of assumptions from Chapter 3 of the manuscript by Carmona and Delarue [CD+18]: We may assume that the set A for the values of the controls is a bounded closed convex subset of a Euclidean space, the deterministic functions b , f , and g are defined on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, respectively, and there exists a constant $C_0 > 1$ such that:

- For any $t \in [0, T]$, $x, x' \in \mathbb{R}^d$, $\alpha, \alpha' \in A$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$\begin{aligned} |(b, f)(t, x', \mu, \alpha') - (b, f)(t, x, \mu, \alpha)| &+ |\sigma(t, x') - \sigma(t, x)| + \\ &+ |g(x', \mu) - g(x, \mu)| \leq C_0 |(x, \alpha) - (x', \alpha')|. \end{aligned}$$
- The functions b , f , σ and g are bounded by C_0 .
- There exists a function

$$\hat{\alpha} : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \ni (t, x, \mu, y) \rightarrow \hat{\alpha}(t, x, \mu, y)$$

which is C_0 -Lipschitz continuous in (x, y) such that, for each $(t, x, \mu, y) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, $\hat{\alpha}(t, x, \mu, y)$ is the unique minimizer of the Hamiltonian $H(t, x, \mu, y, \alpha)$.

Under this set of assumptions, it is shown in Chapter 3 of the manuscript by Carmona and Delarue [CD+18] that a flow of measures $\mu = (\mu_t)_{0 \leq t \leq T}$ is a mean field game equilibrium if and only if $\mu_t = \mathcal{L}(X_t)$, $\forall t \in [0, T]$, where (X, Y, Z) is a solution of the weak approach FBSDE system in Equation (6.2.2.1), in which case (6.2.2.1) becomes

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1} Z_t)) dt + \sigma dW_t \\ X_0 &= \xi, \\ dY_t &= -f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), \sigma^{-1} Z_t)) dt + Z_t dW_t \\ Y_T &= g(X_T, \mathcal{L}(X_T)), \end{aligned} \tag{6.2.2}$$

where we use the generic notation $\mathcal{L}(\cdot)$ for the law of a random variable. This approach is developed further in the papers and manuscript of Carmona and Delarue [CD+18][CD13].

6.2.2.2 Pontryagin approach

The second probabilistic approach, which we will refer to as the Pontryagin approach, is based on the Pontryagin stochastic maximum principle. In this formulation, the optimization problem is solved using a backward SDE for the probabilistic representation of the spatial derivative of the value function u . Formally, the strategy is thus to evaluate the process $(X_t)_{0 \leq t \leq T}$ along $\nabla_x u$. Hence we let $Y_t = \nabla_x u(t, X_t)$, which makes sense when $\nabla_x u$ is well-defined. In fact the Pontryagin system may be formulated without any further reference to the regularity of u , the Pontryagin formulation having the following general form:

$$\begin{aligned} dX_t &= b(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, Y_t)) dt + \sigma dW_t \\ X_0 &= \xi, \\ dY_t &= - [\nabla_x b(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, Y_t)) \cdot Y_t \\ &\quad + \nabla_x f(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, Y_t))] dt + Z_t dW_t \\ Y_T &= \nabla_x g(X_T, \mu_T), \end{aligned}$$

where we assume b , f and g to be differentiable with respect to x . We may use the following set of assumptions taken from Chapter 3 of the manuscript by Carmona and Delarue [CD+18] to guarantee that the Pontryagin system is both a necessary and a sufficient condition of optimality: We assume the coefficients b , f , and g are defined on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, respectively. We also assume that they satisfy:

- The drift b is an affine function of (x, α) of the form:

$$b(t, x, \mu, \alpha) = b_0(t, \mu) + b_1(t)x + b_2(t)\alpha,$$

where $b_0 : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, \mu) \mapsto b_0(t, \mu)$, $b_1 : [0, T] \ni t \mapsto b_1(t)$ and $b_2 : [0, T] \ni t \mapsto b_2(t)$ are \mathbb{R}^d , $\mathbb{R}^{d \times d}$ and $\mathbb{R}^{d \times d}$ valued, respectively, and are measurable and bounded on bounded subsets of their respective domains.

- There exist two constants $C_1 > 0$ and $C_2 \geq 1$ such that the function $\mathbb{R}^d \times A \ni (x, \alpha) \mapsto f(t, x, \mu, \alpha) \in \mathbb{R}$ is once continuously differentiable with Lipschitz-continuous derivatives (so that $f(t, \cdot, \mu, \cdot)$ is $C^{1,1}$), the Lipschitz constant in x and α being bounded by C_2 (so that it is uniform in t and μ). Moreover, it satisfies the following strong form of convexity:

$$f(t, x', \mu, \alpha') - f(t, x, \mu, \alpha) - (x - x', \alpha - \alpha') \cdot \partial_{(x, \alpha)} f(t, x, \mu, \alpha) \geq C_1 |\alpha' - \alpha|^2.$$

The notation $\partial_{(x, \alpha)} f$ stands for the gradient in the joint variables (x, α) . Finally, f , $\partial_x f$ and $\partial_\alpha f$ are locally bounded over $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$.

- The function $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, \mu) \mapsto g(x, \mu)$ is locally bounded. Moreover, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the function $\mathbb{R}^d \ni x \mapsto g(x, \mu)$ is once continuously differentiable and convex, and has a C_2 -Lipschitz continuous first order derivative.

Under these assumptions, it is shown in Chapter 3 of the manuscript by Carmona and Delarue [CD+18] that a flow of measures $\mu = (\mu_t)_{0 \leq t \leq T}$ is a mean field game equilibrium if and only if $\mu_t = \mathcal{L}(X_t), \forall t \in [0, T]$, where (X, Y, Z) is a solution of the Pontryagin approach FBSDE system in Equation (6.2.2.2), in which case (6.2.2.2) becomes

$$\begin{aligned} dX_t &= b(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t)) dt + \sigma dW_t \\ X_0 &= \xi, \\ dY_t &= -[\nabla_x b((t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t))) \cdot Y_t \\ &\quad + \nabla_x f(t, X_t, \mathcal{L}(X_t), \hat{\alpha}(t, X_t, \mathcal{L}(X_t), Y_t))] dt + Z_t dW_t \\ Y_T &= \nabla_x g(X_T, \mathcal{L}(X_T)). \end{aligned} \tag{6.2.3}$$

This approach is also developed further in the papers and manuscript of Carmona and Delarue [CD+18][CD13].

6.2.2.3 Mean Field Games of Control

In many applications, individuals may interact through their controls, instead of their states. One example is an application of trade crowding which was tackled with a mean field game approach in the paper of Cardaliaguet and Lehalle [CL18]. There is also a model for price impact in the book of Carmona and Delarue which we take as one of our example problems in Section 6.4.4.2. Mean field games where players interact through the law of their controls is sometimes referred to as extended mean field games [GPV14], see also Chapter 4 in [CD+18]. We are interested in testing our numerical methods on a certain class of mean field games of control: those in which the interaction is through the marginal distributions, $\mathcal{L}(X_t)$ and $\mathcal{L}(\alpha_t)$. To design our algorithms to handle such a general framework of mean field interaction, we study numerical methods for solving a general FBSDE system which includes the two approaches detailed above, as well as this class of mean field games of control.

6.2.3 General System

We can address both probabilistic formulations for mean field games and a class of mean field games of control simultaneously by considering the following general FBSDE system. We now take the dimension of the state space to be $d = 1$ since, for simplicity, our algorithms will be just applied to this case. Let $[X] = \mathcal{L}(X)$ denote the law of a process X . With an abuse of notation, let $[X, Y, Z] := (\mathcal{L}(X), \mathcal{L}(Y), \mathcal{L}(Z))$ denote the laws of the individual processes (and not their joint law). The general system is the following:

$$\begin{aligned} dX_t &= B(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + \sigma dW_t \\ X_0 &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}), \\ dY_t &= -F(t, X_t, Y_t, Z_t, [X_t, Y_t, Z_t]) dt + Z_t dW_t \\ Y_T &= G(X_T, [X_T]). \end{aligned} \tag{6.2.4}$$

The assumption that the coefficients depend, at most, upon the marginal laws $\mathcal{L}(X_t)$, $\mathcal{L}(Y_t)$ and $\mathcal{L}(Z_t)$ of the triple (X_t, Y_t, Z_t) (and not upon the full joint law) is tailor made to the applications we have in mind: As explained in the previous paragraph, we want to handle games in which players interact with one another through the law of the control. In order to guess the impact this may have on the FBSDE representation, it is worth recalling that, in the problem (6.2.1), the optimal control may be represented in a quite generic way through the function $\hat{\alpha}$ with the adjoint variable y therein taken as the gradient of the value function. Under the weak formulation approach, this turns the FBSDE system (6.2.2) into an FBSDE system depending on the marginal laws of Z , as Z_t is known to have the representation $Z_t = \nabla_x u(t, X_t)\sigma$. Under the Pontryagin approach, it turns the FBSDE system (6.2.3) into an FBSDE system depending upon the marginal laws of Y . Hence our choice to handle systems of the form (6.2.4). Still the reader should observe that, in order to handle the more general case when the players interact through the joint law of the state and of the control, it is necessary to address systems of the same type as (6.2.4) but with the convention that $[X_t, Y_t, Z_t]$ is understood as the joint law of the triple (X_t, Y_t, Z_t) ; as we just mentioned, we do not address this level of generality here.

In (6.2.4), the diffusion process X is coupled to the diffusion process (Y, Z) through the functions B and F , representing the drift of the forward process and the driver of the backward process, respectively. The functions B and F are assumed to be Lipschitz in each of their arguments on $([0, T], \mathbb{R}^3, (\mathcal{P}_2(\mathbb{R}))^3)$ and G Lipschitz on $(\mathbb{R}, \mathcal{P}_2(\mathbb{R}))$, namely, for $(x, y, z, x', y', z) \in \mathbb{R}^6$ and $(\mu, \nu, \lambda, \mu', \nu', \lambda') \in (\mathcal{P}_2(\mathbb{R}))^6$, we have:

$$\begin{aligned} |B(t', x', y', z', \mu', \nu', \lambda') - B(t, x, y, z, \mu, \nu, \lambda)| &\leq K_B(|t' - t| + |x' - x| + |y' - y| + |z' - z| \\ &\quad + W_2(\mu', \mu) + W_2(\nu', \nu) + W_2(\lambda', \lambda)) \\ |F(t', x', y', z', \mu', \nu', \lambda') - F(t, x, y, z, \mu, \nu, \lambda)| &\leq K_F(|t' - t| + |x' - x| + |y' - y| + |z' - z| \\ &\quad + W_2(\mu', \mu) + W_2(\nu', \nu) + W_2(\lambda', \lambda)) \\ |G(x', \mu') - B(x, \mu)| &\leq K_G(|x' - x| + W_2(\mu', \mu)). \end{aligned}$$

The goal of this project is to study numerical methods for solving this general FBSDE system. At some point in this chapter, we will relax the Lipschitz condition of F in the variables z and λ and address an FBSDE with a quadratic driver F , see our examples in Section 6.4.

6.3 Algorithms

We implement two algorithms for numerically solving the FBSDE system in Equation (6.2.4). In the first algorithm, we represent paths of the stochastic processes (X, Y, Z) using a tree structure, where branches of the tree represent quantization of the Brownian motion. In the second algorithm, we no longer represent the paths of the process, but the marginal laws of the process. In this case, the law of the process is discretized on a fixed temporal and spatial grid.

In both cases, the representation serves as a basis for a Picard scheme for approaching the solution. For the first algorithm, the Picard scheme is implemented in the form of a global Picard scheme on the whole process; for the second one, iterations are just performed on the marginal laws of the process. Although Picard's method sounds

very natural, this strategy suffers, whatever the algorithm, from a serious drawback as Picard iterations for forward backward systems are only expected to converge in small time. Basically, the value of T for which the algorithm will converge depends on the coupling strength of the system; we make this fact clear for the global method by showing how T depends on the Lipschitz coefficients K_B , K_F and K_G . In any case, bifurcations can be observed when T is increased, or equivalently, when the coupling strength between the two equations is increased. For our convenience (since it can be costly to increase T), we will fix T and explore the convergence of the algorithms as we vary the coupling strength. This is one first step in this chapter: Compare how the two algorithms suffer from the coupling strength between the forward and backward equations.

The second key feature of our chapter is to use, for both algorithms, a continuation in time, which allows us to extend the value of the coupling parameter for which the algorithms converge.

In the following sections, we detail our Picard approaches for the two algorithms and the continuation in time method.

6.3.1 Global Picard Iteration on a Small Time Interval

The main difficulty in numerically solving the FBSDE system is the fact that, not only the forward component $X = (X_t)_{0 \leq t \leq T}$ and backward component $(Y, Z) = (Y_t, Z_t)_{0 \leq t \leq T}$ are coupled, but also they run in opposite directions. Thus, neither equation can be solved independently of the other, seemingly requiring us to manage both time directions simultaneously. Several strategies are conceivable to sort out this issue. A first one is to make use of the *decoupling field* of the system in order to work with one time direction instead of two (roughly speaking, the decoupling field is the value function u in the weak approach and its derivative $\nabla_x u$ in the Pontryagin one). We will investigate this method for our second algorithm; its implementation is indeed pretty subtle in the mean field setting and it leads to the aforementioned *Picard method on the marginal laws*. For our first algorithm, however, we limit ourselves to a *brute force* approach. To decouple the equations, we propose a *global* Picard iteration scheme, whose definition is as follows. For the initial and terminal data of the problem (ξ and G), we want to define a mapping $\Phi_{\xi, G}$ that will take the $j - 1$ Picard iterate and produce the j Picard iterate:

$$\Phi_{\xi, G} : (X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}]) \mapsto (X^j, Y^j, Z^j, [X^j, Y^j, Z^j]).$$

We define the decoupled Picard scheme $\Phi_{\xi, G}$ as the following:

1. First, solve

$$\begin{aligned} dX_t^j &= B(t, X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}, [X_t^{j-1}, Y_t^{j-1}, Z_t^{j-1}])dt + \sigma dW_t \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}), \end{aligned}$$

for X^j which gives us $[X^j]$.

2. Next, solve

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^{j-1}, Z_t^{j-1}, [X_t^j, Y_t^{j-1}, Z_t^{j-1}])dt + Z_t^j dW_t \\ Y_T^j &= G(X_T^j, [X_T^j]), \end{aligned}$$

for Y^j and Z^j which gives us $[Y^j]$ and $[Z^j]$.

3. Return $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$.

After initializing $(X^0, Y^0, Z^0, [X^0, Y^0, Z^0])$, we can define a sequence by:

$$(X^j, Y^j, Z^j, [X^j, Y^j, Z^j]) = \Phi_{\xi, G}(X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}]).$$

If the sequence $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])_{j=1, \dots}$ converges to some $(X, Y, Z, [X, Y, Z])$, then:

$$(X, Y, Z, [X, Y, Z]) = \Phi_{\xi, G}(X, Y, Z, [X, Y, Z]),$$

and thus, (X, Y, Z) is a solution to the original FBSDE system in Equation (6.2.4).

Picard schemes such as this one are only guaranteed to converge for a small time horizon, T , depending on the Lipschitz coefficients of the functions B , F and G (and in fact not only the convergence of the Picard sequence but also the solvability of the equation may fail on an arbitrary time interval). We illustrate this idea through the following example (the reader may find the general case in [Del02]): Let the driver function $F = 0$ and define the common Lipschitz coefficient $K = \max(K_B, K_G)$. The FBSDE system becomes:

$$\begin{aligned} dX_t &= B(Y_t)dt + \sigma dW_t, \quad X_0 = \xi \in \mathbb{R} \\ dY_t &= Z_t dW_t, \quad Y_T = G(X_T, [X_T]). \end{aligned}$$

We write the equation above in the integral form and take the expectation conditional to the filtration \mathcal{F}_t in the backward equation:

$$\begin{aligned} X_t &= \xi + \int_0^t B(Y_s)ds + \sigma W_t \\ Y_t &= G(X_T, [X_T]) - \int_t^T Z_s dW_s, \quad \text{i.e., } Y_t = \mathbb{E}(G(X_T, [X_T]) | \mathcal{F}_t). \end{aligned}$$

Instead of one single X , let us now consider two (initial) processes denoted by \hat{X} and \tilde{X} . After one Picard iteration, we arrive at \hat{X}' and \tilde{X}' . The Picard iteration is given by:

$$\begin{aligned} \hat{Y}_t &= \mathbb{E}(G(\hat{X}_T, [\hat{X}_T]) | \mathcal{F}_t) \\ \hat{X}'_t &= \xi + \int_0^t B(\hat{Y}_s)ds + \sigma W_t, \end{aligned}$$

and similarly for \tilde{X}' . From the forward component, we have the following estimate for the difference between \hat{X}' and \tilde{X}' :

$$\begin{aligned}
|\hat{X}'_t - \tilde{X}'_t|^2 &\leq \left| \int_0^t (B(\hat{Y}_s) - B(\tilde{Y}_s)) ds \right|^2 \leq tK^2 \int_0^t |\hat{Y}_s - \tilde{Y}_s|^2 ds, \\
\mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}'_t - \tilde{X}'_t|^2 \right] &\leq T^2 K^2 \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{Y}_t - \tilde{Y}_t|^2 \right].
\end{aligned}$$

Then we write for the backward component:

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{Y}_t - \tilde{Y}_t|^2 \right] &= \mathbb{E} \left[\sup_{0 \leq t \leq T} |\mathbb{E}(G(\hat{X}_T, [\hat{X}_T]) - G(\tilde{X}_T, [\tilde{X}_T]) | \mathcal{F}_t)|^2 \right] \\
&\leq 4\mathbb{E} [|G(\hat{X}_T, [\hat{X}_T]) - G(\tilde{X}_T, [\tilde{X}_T])|^2] \\
&\leq 8K^2 \mathbb{E} [|\hat{X}_T - \tilde{X}_T|^2 + W_2([\hat{X}_T], [\tilde{X}_T])^2].
\end{aligned}$$

In the second to last inequality, we used Doob's martingale inequality for the martingale term $\mathbb{E}(G(\hat{X}_T, [\hat{X}_T]) - G(\tilde{X}_T, [\tilde{X}_T]) | \mathcal{F}_t)_{0 \leq t \leq T}$. Also we used the fact that $W_2([\hat{X}_t], [\tilde{X}_t])^2 \leq \mathbb{E}[|\hat{X}_t - \tilde{X}_t|^2]$.

Combining the inequalities above for both forward and backward components, we obtain the following estimate:

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}'_t - \tilde{X}'_t|^2 \right] \leq 16T^2 K^4 \mathbb{E} [|\hat{X}_T - \tilde{X}_T|^2] \leq 16T^2 K^4 \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}_t - \tilde{X}_t|^2 \right].$$

And then,

$$\sup_{0 \leq t \leq T} W_2([\hat{X}'_t], [\tilde{X}'_t])^2 \leq \sup_{0 \leq t \leq T} \mathbb{E} [|\hat{X}'_t - \tilde{X}'_t|^2] \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}'_t - \tilde{X}'_t|^2 \right].$$

Finally, when $16T^2 K^4 < 1$, i.e. $T \leq 4/K^2$, the mapping $\Phi_{\xi, G}$ realizes a contraction on the forward component and the Picard iteration defined above will converge to the fixed point, providing a solution to the original FBSDE. Thus, we have illustrated that Picard schemes are only expected to converge in small time or for small coupling (i.e. smaller Lipschitz coefficients).

Keeping in mind that we eventually want to describe numerical schemes, we define a solver *picard*, that will implement a finite number, $J_p \in \mathbb{Z}^+$, of Picard iterations. Thus, we define:

- *picard*(ξ, G):

1. Initialize $X_t = \xi$, $Y_t = 0$, and $Z_t = 0, \forall t \in [0, T]$.
2. For $1 \leq j \leq J_p$

$$(X, Y, Z, [X, Y, Z]) = \Phi_{\xi, G}(X, Y, Z, [X, Y, Z])$$

3. Return $(X, Y, Z, [X, Y, Z])$.

6.3.2 Continuation in Time of the Global Method for Arbitrary Interval/Coupling

Of course, we would like the mapping $\Phi_{\xi,G}$ to realize a contraction to make sure that the Picard iteration converges. As we just explained, this is the case when the forward and backward processes have a small enough coupling strength or for small time horizon T . But in practice, this is not always the case: It may happen that the FBSDE system is uniquely solvable, but that we are unable to prove that the Picard sequence converges. In order to overcome this issue, we follow the approach introduced in Chassagneux, Crisan, Delarue [CCD+19]. Basically, the point is to divide the time interval into smaller intervals, called *levels*, and to apply a Picard solver recursively between the levels.

To define the levels, we fix a time mesh: $\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}$. We would like to use the *picard* solver introduced in the previous section to apply the Picard iteration on a given level. Notice that for an arbitrary level $[T_k, T_{k+1}]$, we do not know the initial condition for X_{T_k} or the terminal condition for $Y_{T_{k+1}}$. Thus, our current approximation of these values will be inputs to the *picard* solver in place of ξ and G . We would also like to modify the solver *picard* to also take the current estimate of $(X, [X])$ so that we don't have to start from scratch every time we wish to use *picard*. Thus, we define for level k :

- *picard* $[k](X, Y_{T_{k+1}})$:

1. Initialize $Y_t = 0$, and $Z_t = 0, \forall t \in [T_k, T_{k+1}]$.

2. For $1 \leq j \leq J_p$

$$(X, Y, Z, [X, Y, Z]) = \Phi_{X_{T_k}, Y_{T_{k+1}}}(X, Y, Z, [X, Y, Z])$$

3. Return $(X, Y, Z, [X, Y, Z])$,

where $X = (X_t)_{T_k \leq t \leq T_{k+1}}$, and similarly for Y and Z and their laws. Note that $\Phi_{X_{T_k}, Y_{T_{k+1}}}$ is the same as $\Phi_{\xi,G}$ defined earlier except the time horizon is $[T_k, T_{k+1}]$ instead of $[0, T]$ and the initial and terminal conditions are given by X_{T_k} and $Y_{T_{k+1}}$ instead of ξ and G , respectively. In particular, pay attention that we no longer consider the terminal condition in the form of a mapping (like G) but in the form a random variable (like $Y_{T_{k+1}}$). Implicitly, this requires to store, from one step to another, the full random variable $Y_{T_{k+1}}$; hence our choice below to use a tree.

Now that we have a solver *picard* which will implement the Picard iteration for a given level, next, we want to define a global solver to apply a continuation in time. The global solver, called *solver*, is recursively defined as follows for some $J_s \in \mathbb{Z}^+$. For a given level k , define:

- *solver* $[k](X_{T_k}, [X_{T_k}])$:

1. Initialize $X_t = X_{T_k}$, $Y_t = 0$, and $Z_t = 0, \forall t \in [T_k, T_{k+1}]$.

2. For $1 \leq j \leq J_s$

- (a) $(Y_{T_{k+1}}, [Y_{T_{k+1}}]) = \textit{solver}[k+1](X_{T_{k+1}}, [X_{T_{k+1}}])$

- (b) $(X, Y, Z, [X, Y, Z]) = \textit{picard}[k](X, Y_{T_{k+1}})$

3. Return $(Y_{T_k}, [Y_{T_k}])$.

As before, it is important to notice that the entry X_{T_k} in *solver* is a random variable; $[X_{T_k}]$ is its law. Having the two in our notations is a bit redundant, but we feel it is more transparent for the reader. The break condition of the recursion is given by the terminal condition:

$$\text{solver}[N_\ell](X_{T_{N_\ell}}, [X_{T_{N_\ell}}]) = (Y_{T_{N_\ell}}, [Y_{T_{N_\ell}}]) = (G(X_{T_{N_\ell}}, [X_{T_{N_\ell}}]), \mathcal{L}(G(X_{T_{N_\ell}}, [X_{T_{N_\ell}}])))$$

The goal of the continuation in time is for a Picard iteration scheme to converge even for large coupling parameters or large time horizon. We will see in Section 6.4 that the continuation in time successfully achieves this goal for our benchmark examples. We refer to [CCD+19] for its theoretical analysis.

Thus far, we have described our Picard approach and a continuation in time method. Now we need to provide a scheme for discretizing the Picard iteration mapping $\Phi_{\xi, G}$. In this chapter, we implement a tree algorithm; we also give a variant of it, in the form of a grid algorithm. For both algorithms, we consider the uniform time mesh with time step $h = T/N_t > 0$ with $N_\ell < N_t \in \mathbb{Z}^+$ and $t_i = ih, i = 0, \dots, N_t$. For convenience, we will assume the coarse time mesh used to define the levels, $\{0 = T_0 < T_1, \dots, T_k, \dots, T_{N_\ell-1} < T_{N_\ell} = T\}$, is a subset of the fine time mesh $\{0 = t_0 < t_1, \dots, t_{N_t-1} < t_{N_t} = T\}$.

6.3.3 Tree Algorithm for the Global Method

The first implementation of the Picard iteration, $\Phi_{\xi, G}$, is the tree algorithm presented in Chassagneux, Crisan, Delarue [CCD+19]. We now provide a brief presentation of this algorithm.

6.3.3.1 Time Discretization

Our first step in developing a discretization for the Picard iteration $\Phi_{\xi, G}$ is to discretize the problem in the time domain. We use the decoupled scheme derived in Chassagneux, Crisan, Delarue [CCD+19], and repeated below for convenience. As noted above, we consider the uniform time mesh with time step $h = T/N_t > 0$ with $N_\ell < N_t \in \mathbb{Z}^+$ and $t_i = ih, i = 0, \dots, N_t$.

Step 1) in defining $\Phi_{\xi, G}$ requires solving the forward equation for X^j . We use the classical Euler scheme:

$$X_{t_{i+1}}^j = X_{t_i}^j + h B(t_i, X_{t_i}^{j-1}, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^{j-1}, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) + \sigma \Delta W_i, \quad X_0^j = \xi.$$

Note that when we calculate $X_{t_{i+1}}^j$, the value of $X_{t_i}^j$ and its law is known and could be substituted for $X_{t_i}^{j-1}$ and its law in the drift function, B .

Step 2) in defining $\Phi_{\xi, G}$ requires solving the backward equation for Y^j and Z^j . We derive the discrete-time scheme to approximate the backward component, see e.g. [BT04]. The Y^j component at time t_i in the backward scheme is obtained by taking the expectation conditional to \mathcal{F}_{t_i} , denoted as \mathbb{E}_{t_i} , of the backward equation between t_i and t_{i+1} . Let $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ denote the forward Brownian increment between t_i and t_{i+1} . The driver function F is approximated by its value at time t_i and we use the fact that F at t_i and Z_{t_i} are \mathcal{F}_{t_i} -measurable, and $\mathbb{E}_{t_i}(\Delta W_i) = \mathbb{E}_{t_i}(W_{t_{i+1}} - W_{t_i}) = 0$:

$$\begin{aligned}
 Y_{t_i}^j &= Y_{t_{i+1}}^j + \int_{t_i}^{t_{i+1}} F(t, X_t^j, Y_t^{j-1}, Z_t^{j-1}, [X_t^j, Y_t^{j-1}, Z_t^{j-1}]) dt - \int_{t_i}^{t_{i+1}} Z_t^j dW_t \\
 &\approx Y_{t_{i+1}}^j + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) - Z_{t_i}^j \Delta W_i \\
 Y_{t_i}^j &= \mathbb{E}_{t_i}(Y_{t_{i+1}}^j) + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) \\
 Y_T^j &= G(X_T^j, [X_T^j]).
 \end{aligned}$$

As for the Z^j component, we multiply the approximation of $Y_{t_i}^j$ by the Brownian increment ΔW_i , taking the conditional expectation \mathbb{E}_{t_i} , and using $\mathbb{E}((\Delta W_i)^2) = h$. By noticing that our scheme will never make use of Z_T^j , we can simply set the terminal condition for the Z^j component to 0.

$$\begin{aligned}
 Z_{t_i}^j (\Delta W_i)^2 &\approx Y_{t_{i+1}}^j \Delta W_i + (-Y_{t_i}^j + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}])) \Delta W_i \\
 Z_{t_i}^j &= h^{-1} \mathbb{E}_{t_i}(Y_{t_{i+1}}^j \Delta W_i), \quad Z_T^j = 0.
 \end{aligned}$$

Putting this together, the time-discretized decoupled forward-backward scheme for Picard iteration of our general FBSDE system (6.2.4) is the following:

$$\begin{cases}
 X_{t_{i+1}}^j = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) + \sigma \Delta W_i \\
 X_0^j = \xi, \\
 Y_{t_i}^j = \mathbb{E}_{t_i}(Y_{t_{i+1}}^j) + h F(t_i, X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}, [X_{t_i}^j, Y_{t_i}^{j-1}, Z_{t_i}^{j-1}]) \\
 Y_T^j = G(X_T^{j-1}, [X_T^{j-1}]) \\
 Z_{t_i}^j = h^{-1} \mathbb{E}_{t_i}(Y_{t_{i+1}}^j \Delta W_i) \\
 Z_T^j = 0.
 \end{cases}$$

Other variants of this forward-backward scheme are possible. For example, in the forward scheme we could change j to $j - 1$ and t_i to t_{i+1} on the right hand side. It is easy to observe that the forward-backward system is decoupled because of lagged Picard indices $j - 1$ and j . Thus, given $(X^{j-1}, Y^{j-1}, Z^{j-1}, [X^{j-1}, Y^{j-1}, Z^{j-1}])$ where $X^{j-1} = (X_{t_i}^{j-1})_{0 \leq i \leq N_t}$, and similarly for Y and Z , we can solve the backward scheme autonomously and then the forward scheme to obtain $(X^j, Y^j, Z^j, [X^j, Y^j, Z^j])$. We have not fully provided a discrete scheme for $\Phi_{\xi, G}$ yet, however, because for a given t_i , we have not discretized $(X_{t_i}, Y_{t_i}, Z_{t_i})$. This is the goal of the next section.

6.3.3.2 Spatial Discretization via Tree Structure

The forward-backward decoupled scheme (6.3.3.1) above looks quite simple and explicit. However, it still presents some difficulties for the numerical computation. Firstly, it is difficult to compute the conditional expectation in the backward scheme. Secondly, it is non trivial to compute the law of $X_{t_{i+1}}$ forward in time. Even if there was no drift, the computation would involve the convolution of the law of X_{t_i} and a Gaussian law of the Brownian increment. Ultimately, we will need a spatial discretization.

The approach of this algorithm is to approximate the Brownian increments using a simple binomial approximation: $\Delta W_i = \pm\sqrt{h}$ with probability 1/2. This gives rise to

a binomial tree for the forward scheme. Each node on the tree at depth i represents a value of X_{t_i} , and has two children nodes representing the two possible values of $X_{t_{i+1}}$ (the “up \uparrow ” and the “down \downarrow ” value), conditioned on the value of X_{t_i} . The two values are computed as follows:

$$X_{t_{i+1}}^j(\uparrow\downarrow) = X_{t_i}^j + h B(t_i, X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j, [X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j]) \pm \sigma\sqrt{h}.$$

Suppose that we use M points x_1, \dots, x_M for the approximation of the law ξ of the forward process at the initial time, i.e. $[X_0] = \xi \approx \sum_{k=1}^M p_k^0 \delta_{x_k}(\cdot)$. Then we have M parallel binomial trees at each Picard iteration. For Picard iterate j and time t_i , the number of nodes at depth i is $M \times 2^i$ with values of $(X_{t_i}^j, Y_{t_i}^j, Z_{t_i}^j)$ saved on the nodes of the tree at depth i . The marginal law of each process at time t_i can be determined by looking at all the values on the nodes at depth i . The backward scheme can be easily computed on the binomial tree. At the last time step, $T = t_{N_t}$, we have $Y_T^j = G(X_T^j, [X_T^j])$ for each of the $M \times 2^{N_t}$ nodes. The conditional expectation in the backward scheme at t_i is simply the average of the “up” and “down” branches at t_{i+1} .

To initialize the $j = 0$ Picard iterate as in the definition of the solver *picard*, we want to set $X_{t_i} = \xi, \forall i \in \{0, \dots, N_t\}$. This amounts to taking each initial value x_k and initializing its entire tree to this value, meaning that $X_{t_i}^0 = x_k$ for all nodes at depth i and for all $i \in \{1, \dots, N_t\}$ of the k -th binomial tree. We then begin the Picard iteration by applying the mapping $\Phi_{\xi, G}$ as detailed above. Using the binomial tree and approximation of Brownian increments, the forward-backward decoupled scheme becomes fully implementable.

6.3.4 Picard Iteration on the Marginal Laws: a Grid Algorithm

The complexity of the tree algorithm is exponential with respect to the number of time steps N_t , since the size of the binomial tree is of order 2^{N_t} . The exponential complexity becomes problematic and makes continuation in time much slower when we deal with large time horizons. In order to reduce the size of the tree, a natural idea is to make some “recombination” of the binomial tree. But since the drift function B depends on the value of the process, the two branches “up-down” and “down-up” from the same node at time t_i will not coincide at time t_{i+2} , in general. Instead of recombination, we may fix a spatial grid of a controllable size that the binomial tree can be projected onto, in order to avoid exponential complexity. This will be the first ingredient of our grid algorithm, which is the main novelty in this chapter.

Inspired by the paper of Delarue and Menozzi [DM+06], where the authors used a spatial grid for the approximation of FBSDEs without mean-field interaction, we make an intensive use of the notion of *decoupling field*, which is the second key ingredient of our strategy. Indeed, using the representation result in Proposition 2.2 in [CCD14], we know that there exist deterministic feedback functions $(u, v) : [0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}$ with u a solution to a nonlinear PDE (on the space of measures) such that for a solution (X, Y, Z) to the general FBSDE system in Equation (6.2.4):

$$Y_t = u(t, X_t, [X_t]) \quad \text{and} \quad Z_t = v(t, X_t, [X_t]).$$

Generally speaking, u is called the decoupling field of (6.2.4). Here comes the main observation: The time marginals of the solution to the general FBSDE system in Equation

(6.2.4), $([X_t, Y_t, Z_t])_{0 \leq t \leq T}$, can be equivalently characterized by $(\mu_t, u(t, \cdot), v(t, \cdot))_{0 \leq t \leq T}$ where $[X_t] = \mu_t$. As in [DM+06], our strategy is to thus approximate (u, v) instead of (X, Y, Z) , but this is not so easy as u and v are defined on a space of infinite dimension (because the mean field component lives in $\mathcal{P}(\mathbb{R})$). In order to overcome this difficulty, we propose to freeze the mean field argument in (u, v) . This permits to regard u and v as functions of a finite-dimensional variable and thus to approximate both along the underlying spatial grid. Once an approximation of u and v has been computed for the given proxy of the marginal laws, we can update the value of this proxy by using a Picard method. Hence the name of this subsection. So, as opposed to the tree algorithm, we will no longer keep track of the pathwise laws of the processes. Instead, we will only compute the marginal laws at each time step of the time mesh by means of a Picard iteration.

6.3.4.1 Picard Iteration without Grid

We first give the inputs and outputs of our new Picard mapping without any grid approximation:

$$\Psi_{[\xi], G} : (\mu_t^{j-1}, u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))_{0 \leq t \leq T} \mapsto (\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T},$$

with $\mu_0^{j-1} = \mu_0^j = [\xi]$. Denoting below $\varphi\#\nu$ as the push-forward measure of the measure ν by the function φ , $\Psi_{[\xi], G}((\mu_t^{j-1}, u^{j-1}(t, \cdot), v^{j-1}(t, \cdot))_{0 \leq t \leq T})$ is defined by:

1. Solve the following SDE for X^j :

$$\begin{aligned} dX_t^j &= B(t, X_t^j, u^{j-1}(t, X_t^j), v^{j-1}(t, X_t^j), \mu_t^{j-1}, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot)))\#\mu_t^{j-1})dt + \sigma dW_t \\ X_0^j &= \xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}), \end{aligned}$$

$$\text{Set } \mu_t^j := [X_t^j].$$

2. Next, find $(u^j, v^j)(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^2$ through $(u^j, v^j)(t, X_t^j) := (Y_t^j, Z_t^j)$ by solving:

$$\begin{aligned} dY_t^j &= -F(t, X_t^j, Y_t^j, Z_t^j, \mu_t^j, (u^{j-1}(t, \cdot), v^{j-1}(t, \cdot)))\#\mu_t^j)dt + Z_t^j dW_t, \\ Y_T^j &= G(X_T^j, \mu_T^j). \end{aligned}$$

3. Return $((\mu_t^j, u^j(t, \cdot), v^j(t, \cdot))_{0 \leq t \leq T})$.

Given $\Psi_{[\xi], G}$, we can go through the construction of the global Picard method and define similar routines *picard* and *solver* for our new Picard approach by replacing formally $\Phi_{\xi, G}$ by $\Psi_{[\xi], G}$. The various inputs and outputs in the new routines should be clear. The next step to make the whole algorithm entirely tractable is to show how to compute μ_t^j , u^j , and v^j explicitly in the mapping $\Psi_{[\xi], G}$. Similar to $\Phi_{\xi, G}$, this mapping will be defined explicitly for a discretized scheme on a temporal and spatial grid in the next two sections. As in the tree algorithm, we consider the uniform time mesh with time step $h = T/N_t > 0$ with $N_\ell < N_t \in \mathbb{Z}^+$ and $t_i = ih, i = 0, \dots, N_t$.

6.3.4.2 Grid Approximation of Forward Component and its Law

We begin by fixing a spatial discretization grid. This grid could in principle be defined differently for each time step t_i , but for simplicity, we consider a homogeneous grid Γ fixed for all time steps t_i , $i \in \{0, \dots, N_t\}$ with constant spatial step size Δx . Let Π be the projection function on the grid

$$\Gamma = \{x_k = x_1 + (k-1)\Delta x, k = 1, \dots, N_x\}.$$

Precisely, Π is given by

$$\Pi(x) = x_k \text{ if } x \in [x - \Delta x/2, x + \Delta x/2) \quad \text{and} \quad \Pi(x) = x_0 \text{ if } x < x_0, \quad \Pi(x) = x_{N_x} \text{ if } x \geq x_{N_x}.$$

The initial law ξ of the forward process is approximated as $\xi \approx \mu_0(\cdot)$ on the grid Γ with N_x points. Recall that in the tree algorithm, we cannot choose M , the number of points for the approximation of the initial law, to be too large as we will need M parallel binomial trees. Because the tree algorithm has exponential complexity, we are able to choose N_x to be much larger than M , and in turn, the approximation of the initial law is more accurate in the grid algorithm than in the tree algorithm. We can initialize the Picard iterate $(\mu_i^0, u_i^0, v_i^0)_{0 \leq i \leq N_t}$ similar as before by letting $\mu_0^0 = \mu_0$ and $(u_i^0, v_i^0) = (0, 0)$, for all $i \leq N_t$.

We follow the definition of Step 1) for $\Psi_{[\xi], G}$, but we use the Euler scheme for the forward process between t_i and t_{i+1} :

$$X_{t_{i+1}}^j = X_{t_i}^j + h B(X_{t_i}^j, (u_i^{j-1}, v_i^{j-1})(X_{t_i}^j), \mu_i^{j-1}, (u_i^{j-1}, v_i^{j-1})\#\mu_i^{j-1}) + \sigma \Delta W_i. \quad (6.3.1)$$

Suppose the j Picard iterate at time t_i is given by μ_i^j with $\mu_0^j = \mu_0$:

$$\mu_i^j(\cdot) = \sum_{k=1}^{N_x} p_{i,k}^j \delta_{x_k}(\cdot), \quad p_{i,k}^j \geq 0 \quad \forall k \in \{1, \dots, N_x\} \quad \text{and} \quad \sum_{k=1}^{N_x} p_{i,k}^j = 1.$$

The law of $X_{t_i}^j$ in the Euler scheme of the forward process is μ_i^j . Then we would like to define μ_{i+1}^j as the law of $X_{t_{i+1}}^j$ in the Euler scheme above, but the quantity $X_{t_{i+1}}^j$ may not belong to the grid. Instead of using formula (6.3.1), the natural idea is then to replace it by its projection on the grid Γ :

$$\begin{aligned} X_{t_{i+1}}^j &= \Pi \left(X_{t_i}^j + h B(X_{t_i}^j, (u_i^{j-1}, v_i^{j-1})(X_{t_i}^j), \mu_i^{j-1}, (u_i^{j-1}, v_i^{j-1})\#\mu_i^{j-1}) + \sigma \Delta W_i \right) \\ [X_{t_i}^j] &= \mu_i^j(\cdot) \quad \text{and} \quad \mu_{i+1}^j(\cdot) = [X_{t_{i+1}}^j]. \end{aligned}$$

The law of $X_{t_{i+1}}^j$ is the convolution of μ_i^j with transition density $q^j(t_i, t_{i+1}; x_k, x_n)$ with x_k, x_n two points of the grid Γ and $k, n \in \{1, \dots, N_x\}$. The transition probability is equal to the conditional probability that the process X^j starting at time t_i from point x_k arrives at time t_{i+1} at the point x_n , i.e.

$$q^j(t_i, t_{i+1}; x_k, x_n) = \mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k).$$

The discretized law μ_{i+1}^j is then written as:

$$\begin{aligned}\mu_{i+1}^j(\cdot) &= (\mu_i * q^j(t_i, t_{i+1}))(\cdot) = \sum_{n=1}^{N_x} p_{i+1,n}^j \delta_{x_n}(\cdot) \\ p_{i+1,n}^j &= \sum_{k=1}^{N_x} \left(p_{i,k}^j \times q^j(t_i, t_{i+1}; x_k, x_n) \right) = \sum_{k=1}^{N_x} \left(p_{i,k}^j \times \mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k) \right).\end{aligned}$$

It is worth noticing that if we did not take the projection when computing $X_{t_{i+1}}^j$ in the scheme (6.3.4.2), then its law would be given by convolution of the law μ_i^j with the Gaussian transition density $\bar{q}^j(t_i, t_{i+1}; x_k, y)$ associated to the Euler scheme. The transition densities q^j and \bar{q}^j have the following relation:

$$\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k) = q^j(t_i, t_{i+1}; x_k, x_n) = \int_{\beta(x_n, \Delta x/2)} \bar{q}^j(t_i, t_{i+1}; x_k, y) dy.$$

In fact, for a more tractable implementation of the forward scheme, we use the binomial approximation for the Brownian increments (6.3.3.2) introduced in the previous section. Note that quantization with more points can be easily applied. In this binomial case, with $(\uparrow)/(\downarrow)$ representing the “up” and “down” branches, respectively, the transition probabilities on the grid can be easily computed:

$$\begin{aligned}\mathbb{P}(X_{t_{i+1}}^j = x_n | X_{t_i}^j = x_k) \\ = \frac{1}{2} \left(\mathbf{1}(X_{t_{i+1}}^j(\uparrow) = x_n | X_{t_i}^j = x_k) + \mathbf{1}(X_{t_{i+1}}^j(\downarrow) = x_n | X_{t_i}^j = x_k) \right).\end{aligned}$$

Then we can write μ_{i+1}^j by computing the probabilities:

$$p_{i+1,n}^j = \sum_{k=1}^{N_x} \frac{p_{i,k}^j}{2} \cdot \left(\mathbf{1}(X_{t_{i+1}}^j(\uparrow) = x_n | X_{t_i}^j = x_k) + \mathbf{1}(X_{t_{i+1}}^j(\downarrow) = x_n | X_{t_i}^j = x_k) \right)$$

At Picard iteration $j \geq 1$, the forward scheme finally gives the flow of measures $(\mu_i^j)_{i=0}^{N_t}$ at discrete time steps $(t_i)_{i=0}^{N_t}$ of the discretized law defined on the grid Γ . Thus, we have described an implementation of Step 1) in the definition of the Picard mapping $\Psi_{[\xi], G}$. In the next section, we detail the implementation of the backward components in Step 2).

6.3.4.3 Grid for the approximation of the backward component

Given the current Picard iterates $(\mu_i^j, u_i^{j-1}(\cdot), v_i^{j-1}(\cdot))$, $i = 0, \dots, N_t$, we would like to detail Step 2) in the definition of $\Psi_{[\xi], G}$. Since we have a discrete spatial grid Γ , we wish to compute the values of $u_i^j(x)$ and $v_i^j(x)$ for $x \in \Gamma$. By replacing $Y_{t_i}^j$ and $Z_{t_i}^j$ with their respective feedback functions in the backward component in Equation (6.3.3.1), for $x \in \Gamma$ we have the following backward scheme starting, for $i \leq N_t - 1$, with terminal condition at time $T = t_{N_t}$, $(u_{N_t}^j, v_{N_t}^j) = (G, 0)$:

$$\begin{aligned}u_i^j(x) &= \mathbb{E} \left(u_{i+1}^j(X_{t_{i+1}}^j, \mu_{i+1}^j) + h \cdot F(X_{t_i}^j, u_i^{j-1}(X_{t_i}^j), v_i^j(X_{t_i}^j), \mu_i^j, (u_i^{j-1}, v_i^{j-1}) \# \mu_i^j) \mid X_{t_i}^j = x \right), \\ v_i^j(x) &= \mathbb{E} \left(u_{i+1}^j(X_{t_{i+1}}^j) \cdot \Delta W_i / h \mid X_{t_i}^j = x \right).\end{aligned}$$

The variable $X_{t_{i+1}}^j$ with law $[X_{t_{i+1}}^j] = \mu_{i+1}^j$ is given by the forward scheme presented in the previous section with starting point $X_{t_i}^j = x$. Notice that by construction of the forward scheme, $X_{t_{i+1}}^j \in \Gamma$ and $\text{supp}(\mu_{i+1}^j) = \Gamma$ so $u_{i+1}^j(\cdot)$ has been calculated and saved at time t_{i+1} .

We have a more explicit scheme in the case of binomial approximation of the Brownian increments, which is used for the numerical results for our examples, with $X_{t_{i+1}}^j(\uparrow)$ and $X_{t_{i+1}}^j(\downarrow)$ as defined in the forward scheme, always conditional to $X_{t_i}^j = x$ and given μ_i^j at time t_i :

$$\begin{aligned} u_i^j(x) &= \frac{1}{2} \left(u_{i+1}^j(X_{t_{i+1}}^j(\uparrow)) + u_{i+1}^j(X_{t_{i+1}}^j(\downarrow)) \right) + h \cdot F(x, u_i^{j-1}(x), v_i^j(x), \mu_i^j, (u_i^{j-1}, v_i^{j-1}) \# \mu_i^j), \\ v_i^j(x) &= \frac{h^{-1/2}}{2} \left(u_{i+1}^j(X_{t_{i+1}}^j(\uparrow)) - u_{i+1}^j(X_{t_{i+1}}^j(\downarrow)) \right). \end{aligned}$$

We have now described a scheme for Step 2) in the definition of $\Psi_{[\xi], G}$. Putting this together with the previous section, we have described a fully implementable scheme for $\Psi_{[\xi], G}$. Note that we can use this Picard iteration mapping to define the analogue of the solvers *picard* and *solver*. Importantly, we can also apply the continuation in time method to the grid algorithm as well. In the next section, we apply these two methods, the tree and grid algorithms, to five example problems.

6.4 Examples

We have collected five example problems to test the algorithms presented in Section 6.3.

6.4.1 Linear Example

The first example is a linear model which comes directly from Chassagneux, Crisan, Delarue [CCD+19], in which they implemented the tree algorithm. The system of interest is the following:

$$\begin{aligned} dX_t &= -\rho \mathbb{E}(Y_t) dt + \sigma dW_t \\ X_0 &= x_0 \\ dY_t &= -aY_t dt + Z_t dW_t \\ Y_T &= X_T. \end{aligned}$$

For this problem, the solution is known explicitly:

$$Y_0 = \frac{x_0 e^{aT}}{1 + \frac{\rho}{a}(e^{aT} - 1)}$$

For the numerical results, we let $\rho = 0.1$, $a = 0.25$, $\sigma = 1$, $T = 1$, and $x_0 = 2$. We vary h , the time step size, and for the grid algorithm, we use $\Delta x = h^2$. Figure 6.1 shows the log error between the numerical and true solution values of Y_0 as a function of log number of time steps for the tree and grid algorithm with one level (i.e. without using

continuation in time). The left figure (tree algorithm) repeats the results in [CCD+19] and as expected, it decreases linearly. On the other hand, we also observe a negative trend in the rate of convergence for the grid algorithm. Thus, both algorithms appear to converge to the true solution.

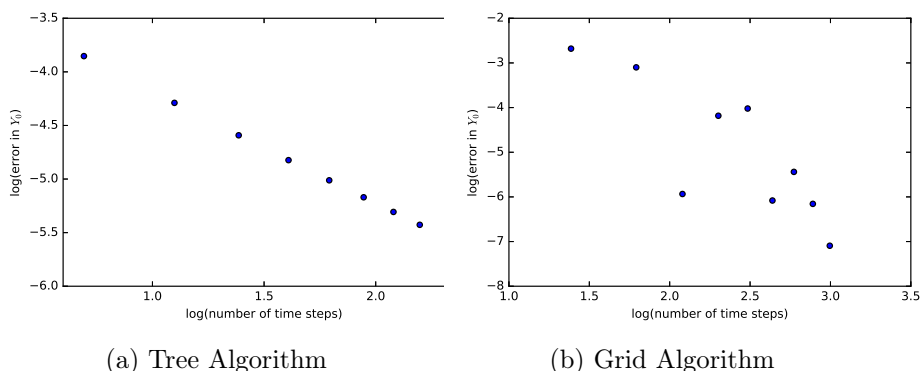


Figure 6.1 – Linear Example: Convergence of the algorithms with one level as the number of time steps increases.

6.4.2 Trigonometric Drivers Example

The second example also comes directly from Chassagneux, Crisan, Delarue [CCD+19]. The system of interest is the following:

$$\begin{aligned} dX_t &= \rho \cos(Y_t) dt + \sigma dW_t \\ X_0 &= x_0 \\ Y_t &= \mathbb{E}_t(\sin(X_T)) \end{aligned}$$

For the numerical results, we let $\sigma = 1$, $T = 1$, $x_0 = 0$, $h = 1/6$ for the tree algorithm and $h = 1/12$ and $\Delta x = h^2$ for the grid algorithm. For this problem, we observe a bifurcation when using the tree algorithm as we increase the coupling parameter, ρ . Figure 6.2 shows the values of Y_0 from the last 5 Picard iterations. Starting at about $\rho = 3.5$, the tree algorithm without continuation in time bifurcates. If the continuation in time method is used for the tree algorithm with two levels, there is no bifurcation for the range of values of ρ showed in the plot. Note that the results from the tree algorithm repeat those in [CCD+19]. The grid algorithm performs quite well in the sense that even with only one level, the algorithm converges for all of the values of ρ in the plot. In particular, it avoids the exponential growth of the data structure characterizing the tree algorithm. Note that even though both the tree method with two levels and the grid method with one level converge, they produce different values for Y_0 for larger values of ρ . We believe this is because the tree algorithm is less accurate since the time step is larger.

For values of ρ for which the tree algorithm bifurcates, we were interested in the effect of changing σ on the convergence. Figure 6.3 shows the last 5 values of Y_0 from the Picard iteration as σ varies for $\rho = 5$. Surprisingly, the value of σ plays no role in

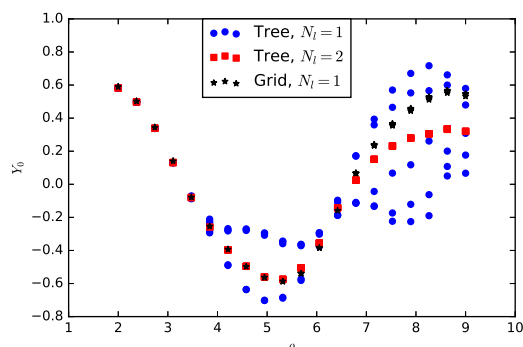


Figure 6.2 – Trigonometric Drivers Example: Bifurcations in the values of Y_0 appear as the coupling parameter, ρ , increases. The tree algorithm with one level is shown in blue circles and with two levels is shown in red squares. The grid algorithm with one level is shown in black asterisks.

the bifurcation for this value of ρ . To see if σ would play a role when ρ is closer to the bifurcation point, we have analogous plots when $\rho = 3.5$ and $\rho = 4$ (see Figure 6.4). In these plots, however, it is clear that σ does affect the bifurcation. Understanding the role of σ on the bifurcation is an open question. It is interesting to note that even though the tree algorithm does not bifurcate for $\rho = 3.5$ when $\sigma = 1$, we still observe a bifurcation when we fix $\rho = 3.5$ and vary σ . This suggests that we check if the grid algorithm bifurcates as we vary σ for a fixed value of ρ where there is no bifurcation when $\sigma = 1$, such as $\rho = 5$. Figure 6.5 shows that the grid algorithm does not bifurcate as we change σ for fixed $\rho = 5$.

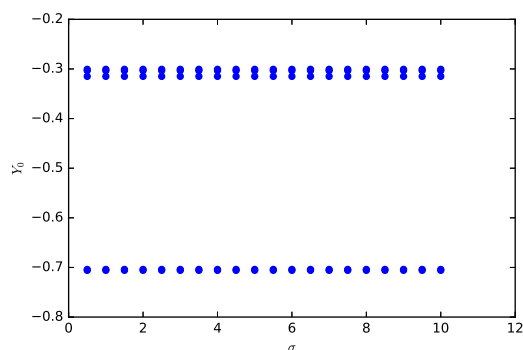


Figure 6.3 – Trigonometric Drivers Example: Tree algorithm with one level for $\rho = 5$. Changing σ has no effect on the bifurcation.

6.4.3 Mixed Model

The third example also comes directly from Chassagneux, Crisan, Delarue [CCD+19]. The system of interest is the following:

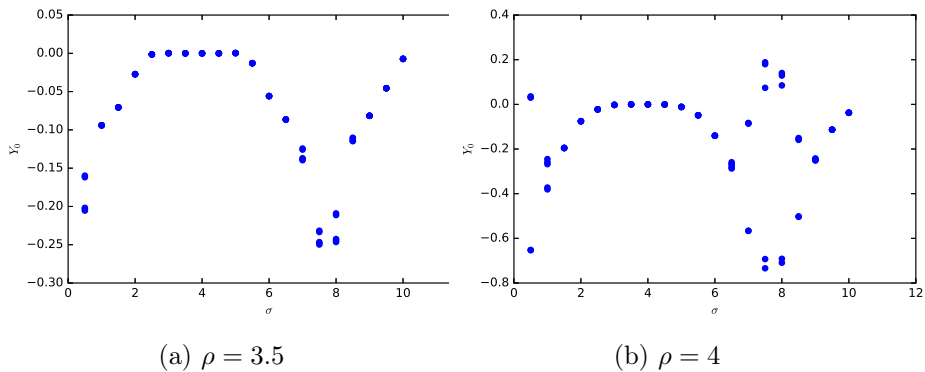


Figure 6.4 – Trigonometric Drivers Example: Tree algorithm with one level for $\rho = 3.5$ and $\rho = 4$. Changing σ produces unexplained results.

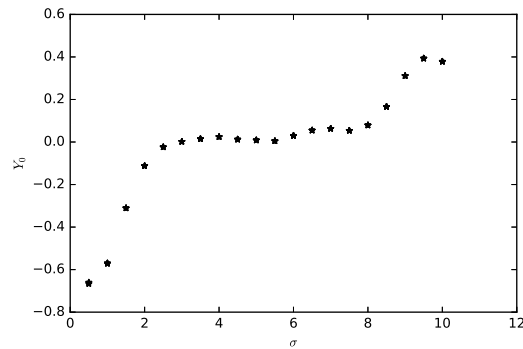


Figure 6.5 – Trigonometric Drivers Example: Grid algorithm with one level for $\rho = 5$. There is no bifurcation as we vary σ .

$$\begin{aligned}
 dX_t &= -\rho Y_t dt + \sigma dW_t \\
 X_0 &= x_0 \\
 dY_t &= \arctan(\mathbb{E}(X_t)) dt + Z_t dW_t \\
 Y_T &= \arctan(X_T).
 \end{aligned}$$

For the numerical results, we let $\sigma = 1$, $T = 1$, $x_0 = 2$, $h = 1/6$ for the tree algorithm and $h = 1/12$ and $\Delta x = h^2$ for the grid algorithm. We also observe a bifurcation for this problem as we increase the coupling parameter, ρ . Figure 6.6 shows the values of Y_0 from the last 5 Picard iterations. Starting at about $\rho = 1.5$, the tree algorithm without continuation in time bifurcates. If the continuation in time method is used for the tree algorithm with two levels, the bifurcation point is pushed back to about $\rho = 3$, and pushed back further when using three levels to about $\rho = 5$. Note that these results for the tree algorithm repeat those in [CCD+19]. The grid algorithm performs quite well again in the sense that it converges for all of the values of ρ shown, even when using only one level. Further, its major attractiveness is the lower complexity compared to the tree algorithm.

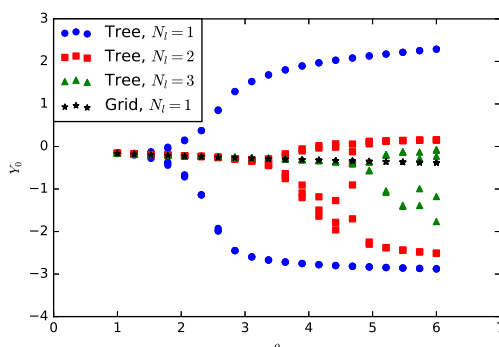


Figure 6.6 – Mixed Model: Bifurcations in the values of Y_0 appear as the coupling parameter, ρ , increases. The tree algorithm with one, two, and three levels is shown in blue circles, red squares, and green triangles, respectively. The grid algorithm with one level is shown in black asterisks.

As in the trigonometric drivers example, we can also investigate the effect of changing σ for a value of ρ where the tree algorithm without continuation in time bifurcates. For $\rho = 2$, Figure 6.7 shows that the tree algorithm converges for large enough values of σ .

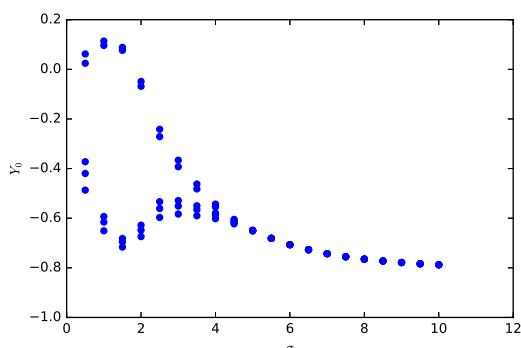


Figure 6.7 – Mixed Model: Tree algorithm with one level for $\rho = 2$. The algorithm converges for large enough values of σ .

Since we have observed that the tree algorithm converges for small value of ρ and large values of σ , this suggests trying a continuation in ρ and/or σ , instead of the continuation in time. Instead of implementing a full continuation method, we used a simpler method of incrementing the parameter of interest.

The incrementation in ρ is performed by starting the algorithm with a small value of ρ and letting it converge. Then ρ is increased by some fixed $\Delta\rho$, and the algorithm is initialized with the solution from the previous value of ρ . Figure 6.8 shows the results from the incrementation in ρ . The incrementation in ρ only increases the bifurcation point by a small amount.

The incrementation in σ is similar, except for each value of ρ , we start the algorithm with a sufficiently large value of σ such that it will converge. Then σ is decreased by a

fixed $\Delta\sigma$. Figure 6.9 shows the results from the incrementation in σ . The incrementation in σ also only increases the bifurcation point by a small amount.

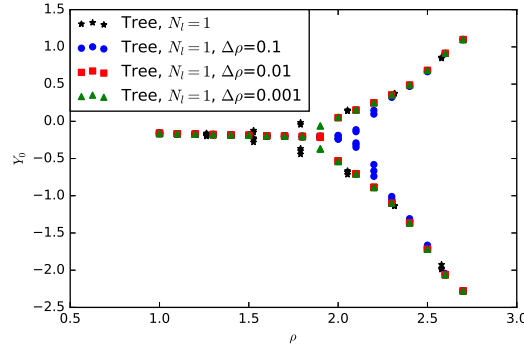


Figure 6.8 – Mixed Model: Tree algorithm with one level. Without continuation or incrementation is shown in black. Incrementation in ρ with $\Delta\rho = 0.1, 0.01$, and 0.001 are shown in blue circles, red squares, and green triangles, respectively.

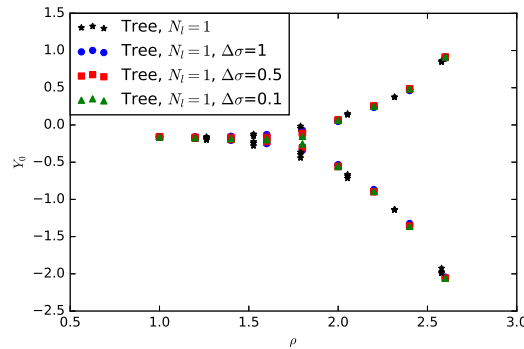


Figure 6.9 – Mixed Model: Tree algorithm with one level. Without continuation or incrementation is shown in black asterisks. Incrementation in σ with $\Delta\sigma = 1, 0.5$, and 0.1 are shown in blue circles, red squares, and green triangles, respectively.

If we take the previous example but change the drift and driver functions to also be in terms of the mean of the processes, then we have the following system, which we will refer to as the mixed model of means:

$$\begin{aligned} dX_t &= -\rho\mathbb{E}(Y_t)dt + \sigma dW_t \\ X_0 &= x_0 = 2 \\ dY_t &= \arctan(\mathbb{E}(X_t)) dt + Z_t dW_t \\ Y_T &= \arctan(\mathbb{E}(X_T)). \end{aligned}$$

For the last example, we noticed that increasing σ allows the tree algorithm to converge. We are thus interested to see if replacing the dynamics with the mean of the process will remove the effect of changing σ on the convergence. We use the same

values of the parameters as before. Figure 6.10 shows the bifurcation with one level of the tree algorithm (i.e. without using continuation in time). The effect of changing σ is shown in Figure 6.11. Our prediction is confirmed: σ no longer affects the convergence when the dynamics are replaced with the mean of the process.

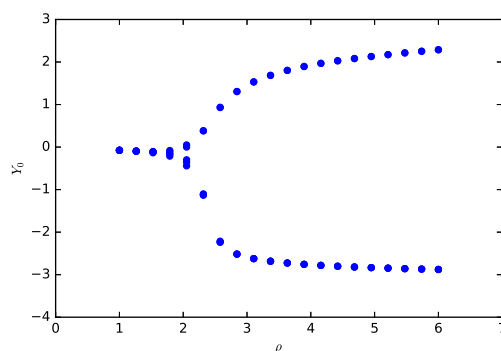


Figure 6.10 – Mixed Model of Means: Bifurcation for the tree algorithm with one level.

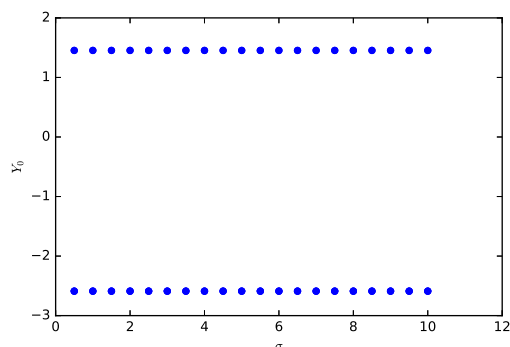


Figure 6.11 – Mixed Model of Means: Tree algorithm with one level with $\rho = 3$. Changing σ has no effect on the bifurcation.

6.4.4 Examples: Linear Quadratic Mean Field Games

The last two examples belong to the family of linear quadratic (LQ) games. In these models, the dynamics of the state are linear in the sense that the drift is defined by an affine function:

$$b(t, x, \alpha) = A_t x + B_t \alpha + \beta_t.$$

Furthermore, the running and the terminal cost are quadratic in the state and control variables. For the sake of simplicity, we choose not to include the cross terms, so that we define f and g as follows:

$$\begin{aligned} f(t, x, \alpha) &= \frac{1}{2} P_t x^2 + \frac{1}{2} Q_t \alpha^2 \\ g(x) &= \frac{1}{2} S x^2. \end{aligned}$$

The FBSDE system derived by approaching the LQ mean field games via the weak approach becomes:

$$\begin{aligned} dX_t &= [A_t x + B_t \alpha + \beta_t] dt + \sigma dW_t, & X_0 &= \xi \\ dY_t &= \left[-\frac{1}{2} P_t X_t^2 + (A_t X_t + \beta_t) \frac{Z_t}{\sigma} + \frac{1}{2} \frac{B_t^2}{Q_t} \frac{Z_t^2}{\sigma^2} \right] dt + dW_t, & Y_T &= \frac{1}{2} S X_T^2. \end{aligned}$$

We present two examples of linear quadratic games: a problem of flocking, and a price impact model of a trader.

Remark 6.4.1. *We observe that the driver of the BSDE is quadratic in Z_t , which makes it seldom solvable. On the other hand, it is possible to obtain explicit solutions for LQ models. If we were to numerically solve equation (6.4.4), we might observe blow up from the quadratic terms in the backward equation and consequently the algorithm to not converge. Luckily, we do not observe such blowup in the examples we consider. But if blowup does occur, one could consider adapting the method presented by Chassagneux and Richou in [CR+16] to numerically approximate a quadratic BSDE.*

6.4.4.1 Flocking Problem

The next example problem models flocking. As in the paper of Nourian, Caines, and Malhamé [NCM10], we consider the spatially homogeneous case where the state, X_t , represents the velocity of a representative player, or bird. Each bird controls its velocity through the process:

$$dX_t = \alpha_t dt + \sigma dW_t,$$

where the control is chosen to minimize:

$$J(\alpha) = \mathbb{E} \left[\int_0^T \left(\frac{1}{2} \alpha_t^2 + \frac{\rho}{2} (X_t - \bar{\mu}_t)^2 \right) dt \right].$$

over $\alpha \in \mathbb{A}$. Above, we let $\bar{\mu}_t$ denote the mean of the distribution μ_t (of the velocities of the birds) at time t . The running cost consists of two components. The first term encourages the birds to minimize their kinetic energy by not choosing a large control. The second term encourages the birds to align their velocities with the mean velocity of the group.

This model falls into the class of linear quadratic games. Assume that the initial condition is given by a constant, $X_0 = x_0$. It can be shown that the solution is Gaussian with mean and variance:

$$\begin{aligned} \mathbb{E}(X_t) &= x_0 \\ \text{Var}(X_t) &= \sigma^2 \int_0^t \exp \left(-2 \int_s^t \eta_u du \right) ds, \end{aligned}$$

where:

$$\eta_t = \sqrt{\rho} \frac{e^{2\sqrt{\rho}(T-t)} - 1}{e^{2\sqrt{\rho}(T-t)} + 1}.$$

Using the weak formulation, the FBSDE system of interest is the following:

$$\begin{aligned} dX_t &= -\frac{Z_t}{\sigma}dt + \sigma dW_t \\ X_0 &= x_0 \\ dY_t &= -\left(\frac{Z_t^2}{2\sigma^2} + \frac{\rho}{2}(X_t - \mathbb{E}X_t)^2\right)dt + Z_t dW_t \\ Y_T &= 0. \end{aligned}$$

If we use the Pontryagin formulation, the FBSDE system becomes:

$$\begin{aligned} dX_t &= -Y_t dt + \sigma dW_t \\ X_0 &= x_0 \\ dY_t &= -\rho(X_t - \mathbb{E}(X_t))dt + Z_t dW_t \\ Y_T &= 0. \end{aligned}$$

The numerical results are presented for $\rho = 1$, $\sigma = 1$, $T = 1$, $x_0 = 0$, $h = 1/20$ for the tree algorithm and $h = 1/130$ and $\Delta x = h^2$ for the grid algorithm. Figure 6.12(a) shows the results for the grid algorithm for both the weak and Pontryagin approaches. The plot shows the weights of the distribution $\mathcal{L}(X_T)$. The results are similar between both approaches and coincide with the true solution.

We can look at the convergence rate by calculating the 2-Wasserstein distance between the numerical results and the true solution as we change the number of time steps. Since our state space is in one dimension, we can calculate the Wasserstein distance explicitly using the representation provided by Prokhorov [Pro56]:

$$W_p(\mu, \nu) = \left(\int_0^1 |F_\mu^{-1}(u) - F_\nu^{-1}(u)|^p du \right)^{1/p},$$

where $F_\mu(x) = \mu([0, x])$, denotes the cumulative distribution function. Figure 6.12(b) presents the convergence rate of the grid algorithm in terms of the 2-Wasserstein distance calculated between the true solution and numerical results with respect to the number of time steps. As expected, the 2-Wasserstein distance decreases towards 0 as we increase the number of time steps, for both the Pontryagin and weak approaches.

The results for the tree algorithm are shown in Figure 6.13 for the weak and Pontryagin approaches. As with the grid algorithm, the weak and Pontryagin solutions are similar to each other and coincide with the true solution.

6.4.4.2 Trader Problem

The last example shows an application of mean field games to finance, such as the trader congestion model in the paper of Cardialaguet and Lehalle, [CL18]. We focus on the *Price Impact Model* presented in the book of Carmona and Delarue in [CD+18]. The interest for this kind of model is motivated by their use in optimal execution problems for high frequency trading. Furthermore, it represents an instance of extended mean field game, also known as mean field game of controls, in which the representative agent interacts with the law of the control instead of the law of their

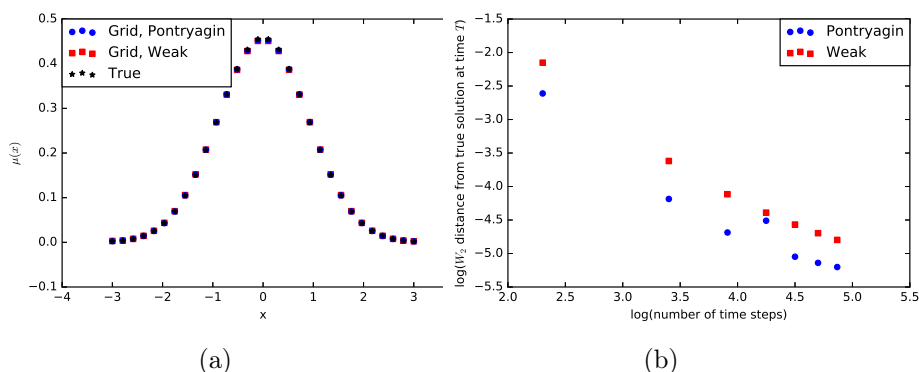


Figure 6.12 – Flocking Problem: (a) Distribution μ_T of the players’ states at time T for the grid algorithm with one level. Pontryagin is in blue circles, weak is in red squares, true solution is shown in black asterisks. (b) 2-Wasserstein distance between true solution and numerical solution for grid algorithm with one level as we increase the number of time steps, plotted as a log-log plot. Pontryagin approach is in blue circles and weak approach is in red squares.

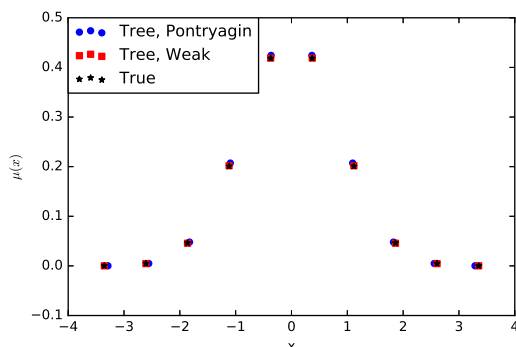


Figure 6.13 – Flocking Problem: Distribution μ_T of the players’ states at time T for the tree algorithm with one level. Pontryagin is shown in blue circles, weak is shown in red squares, and true solution is shown in black asterisks.

state. The problem consists in a group of traders who have to buy or sell a large amount of shares in a given interval of time $[0, T]$. If they trade too fast, they will suffer from market impact. On the other hand, if they trade too slow, they will be affected by a large risk penalization. Approaching this problem as a mean field game, the inventory of the representative trader is modeled by a stochastic process $(X_t)_{0 \leq t \leq T}$ such that

$$dX_t = \alpha_t dt + \sigma dW_t, \quad t \in [0, T],$$

where α_t corresponds to the trading rate. The price of the asset $(S_t)_{0 \leq t \leq T}$ is influenced by the trading strategies of all the traders through the law of the controls $(\theta_t = \mathcal{L}(\alpha_t))_{0 \leq t \leq T}$ as follows:

$$dS_t = \gamma \left(\int_{\mathbb{R}} a d\theta_t(a) \right) dt + \sigma_0 dW_t^0, \quad t \in [0, T],$$

where γ and σ_0 are constants and the Brownian motion W^0 is independent from W . The amount of cash held by the trader at time t is denoted by the process $(K_t)_{0 \leq t \leq T}$. The dynamic of K is modeled by

$$dK_t = -[\alpha_t S_t + c_\alpha(\alpha_t)]dt,$$

where the function $\alpha \mapsto c_\alpha(\alpha)$ is a non-negative convex function satisfying $c_\alpha(0) = 0$, representing the cost for trading at rate α . The wealth V_t of the trader at time t is defined as the sum of the cash held by the trader and the value of the inventory with respect to the price S_t :

$$V_t = K_t + X_t S_t.$$

Applying the self-financing condition of Black-Scholes' theory, the changes over time of the wealth V are given by the equation:

$$\begin{aligned} dV_t &= dK_t + X_t dS_t + S_t dX_t \\ &= \left[-c_\alpha(\alpha_t) + \gamma X_t \int_{\mathbb{R}} a d\theta_t(a) \right] dt + \sigma S_t dW_t + \sigma_0 X_t dW_t^0. \end{aligned}$$

We assume that the trader is subject to a running liquidation constraint modeled by a function c_X of the shares they hold, and to a terminal liquidation constraint at maturity T represented by a scalar function g . Thus, the cost function is defined by:

$$J(\alpha) = \mathbb{E} \left[\int_0^T c_X(X_t) dt + g(X_T) - V_T \right].$$

Applying Equation (6.4.4.2), it follows that

$$J(\alpha) = \mathbb{E} \left[\int_0^T f(t, X_t, \theta_t, \alpha_t) dt + g(X_T) \right],$$

where the running cost is defined by

$$f(t, x, \theta, \alpha) = c_\alpha(\alpha) + c_X(x) - \gamma x \int_{\mathbb{R}} a d\theta(a),$$

for $0 \leq t \leq T$, $x \in \mathbb{R}^d$, $\theta \in \mathcal{P}(A)$ and $\alpha \in A = \mathbb{R}$. We assume that the functions c_X and g are quadratic and that the function c_α is strongly convex in the sense that its second derivative is bounded away from 0. Such a particular case is known as the Almgren-Chriss linear price impact model. Thus, the control is chosen to minimize:

$$J(\alpha) = \mathbb{E} \left[\int_0^T \left(\frac{c_\alpha}{2} \alpha_t^2 + \frac{c_X}{2} X_t^2 - \gamma X_t \int_{\mathbb{R}} a d\theta_t(a) \right) dt + \frac{c_g}{2} X_T^2 \right],$$

over $\alpha \in \mathbb{A}$. To summarize, the running cost consists of three components. The first term represents the cost for trading at rate α . The second term takes into consideration the running liquidation constraint in order to penalize unwanted inventories. The third term defines the actual price impact. Finally, the terminal cost represents the terminal liquidation constraint.

As for the flocking example, this model falls in the class of linear quadratic games. Assume that the initial condition is given by a constant, $X_0 = x_0$. The solution is Gaussian with mean and variance defined as:

$$\mathbb{E}(X_t) = x_0 e^{-\int_0^t \frac{\bar{\eta}_s}{c_\alpha} ds}$$

$$\text{Var}(X_t) = \sigma^2 \int_0^t e^{-\frac{2}{c_\alpha} \int_s^t \eta_r dr} ds$$

where:

$$\bar{\eta}_t = \frac{-C(e^{(\delta^+ - \delta^-)(T-t)} - 1) - c_g(\delta^+ e^{(\delta^+ - \delta^-)(T-t)} - \delta^-)}{(\delta^- e^{(\delta^+ - \delta^-)(T-t)} - \delta^+) - c_g B(e^{(\delta^+ - \delta^-)(T-t)} - 1)}$$

$$\eta_t = -c_\alpha \sqrt{c_X/c_\alpha} \frac{c_\alpha \sqrt{c_X/c_\alpha} - c_g - (c_\alpha \sqrt{c_X/c_\alpha} + c_g) e^{2\sqrt{c_X/c_\alpha}(T-t)}}{c_\alpha \sqrt{c_X/c_\alpha} - c_g + (c_\alpha \sqrt{c_X/c_\alpha} + c_g) e^{2\sqrt{c_X/c_\alpha}(T-t)'}}$$

for $t \in [0, T]$, where $B = 1/c_\alpha$, $C = c_X$, $\delta^\pm = -D \pm \sqrt{R}$, with $D = -\gamma/(2c_\alpha)$ and $R = D^2 + BC$.

Using the weak approach yields the following FBSDEs system:

$$dX_t = -\frac{1}{c_\alpha} \frac{Z_t}{\sigma} dt + \sigma dW_t, \quad X_0 = x_0$$

$$dY_t = -\left[\frac{c_X}{2} X_t^2 + \frac{\gamma}{c_\alpha} \frac{\mathbb{E}[Z_t]}{\sigma} X_t + \frac{1}{2c_\alpha} \left(\frac{Z_t}{\sigma} \right)^2 \right] dt + Z_t dW_t, \quad Y_T = c_g \frac{X_T^2}{2}.$$

Alternatively, the FBSDE system obtained via the Pontryagin approach is:

$$dX_t = -\frac{1}{c_\alpha} Y_t dt + \sigma dW_t, \quad X_0 = x_0$$

$$dY_t = -\left(c_X X_t + \frac{\gamma}{c_\alpha} \mathbb{E}[Y_t] \right) dt + Z_t dW_t, \quad Y_T = c_g X_T.$$

The numerical results focus on the effect of the continuation method for the grid algorithm. In contrast with the previous examples, we show that the grid algorithm is also affected by bifurcation. Figure 6.14 shows the last five Picard iterations of Y_0 for the Pontryagin approach when the number of levels ranges from 1 to 3. Fixing the parameters $x_0 = 1$, $\sigma = 0.7$, $1/c_\alpha = 1.5$, $c_g = 0.3$, $\gamma = 2$, $T = 1$, $h = 1/12$, $\Delta x = h^2$ and increasing the coupling parameter, c_X , we observe that the bifurcation effect can be corrected by increasing the number of levels. In fact, Figure 6.14 shows that the true value of Y_0 matches the value computed numerically when using three levels.

Furthermore, Figure 6.15(a) compares the distribution $\mathcal{L}(X_T)$ obtained by the Pontryagin and the weak approaches using the grid algorithm with parameters $x_0 = 1$, $\sigma = 0.7$, $1/c_\alpha = 0.3$, $c_g = 0.3$, $\gamma = 2$, $c_X = 2$, $T = 1$, $h = 1/130$ and $\Delta x = h^2$. The two approaches produce similar results that coincide with the true solution.

Figure 6.15(b) presents the convergence rate in terms of the 2-Wasserstein distance calculated between the true solution and numerical results with respect to the number of time steps. We again make use of the explicit representation of the Wasserstein distance [Pro56]. The numerical solution is obtained using the grid algorithm with parameters $x_0 = 1$, $\sigma = 0.7$, $1/c_\alpha = 0.3$, $c_g = 0.3$, $\gamma = 2$, $c_X = 2$, $T = 1$, and $\Delta x = h^2$. As expected, the 2-Wasserstein distance decreases towards 0 as we increase the number of time steps.

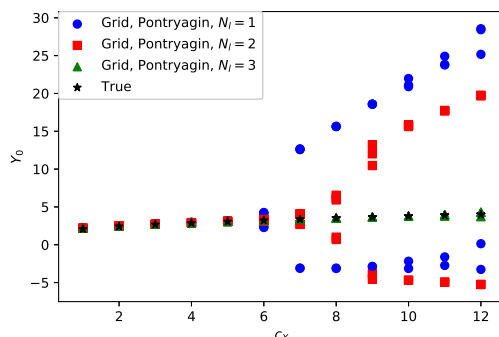


Figure 6.14 – Trader Problem: Bifurcations in the values of Y_0 depending on the coupling parameter c_X for different number of levels in the grid algorithm. One, two, and three levels are shown in blue circles, red squares, and green triangles, respectively. The true value of Y_0 is shown in black asterisks.

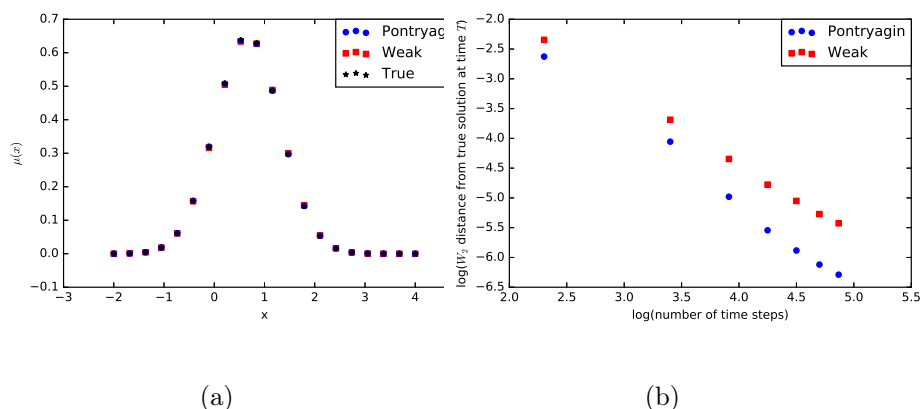


Figure 6.15 – Trader Problem: (a) Distribution μ_T of the players' states at time T for the grid algorithm with one level. Pontryagin is shown in blue circles, weak is shown in red squares, and the true solution is shown in black asterisks. (b) 2-Wasserstein distance between true solution and numerical solution for grid algorithm with one level as we increase the number of time steps, plotted as a log-log plot. Pontryagin approach is shown in blue circles and weak approach is shown in red squares.

The last plot, Figure 6.16, shows the error from the true solution of the control at time 0, α_0 , as we increase the number of time steps. This value is given by $\alpha_0 = -Y_0/c_X$ for the Pontryagin approach and $\alpha_0 = -Z_0/(c_X\sigma)$ for the weak approach. The true value is given by $\alpha_0 = -\bar{\eta}_0 x_0/c_X$. As for the 2-Wasserstein distance, the error in the control decreases towards 0.

6.5 Conclusion

In conclusion, we have provided two algorithms for numerically solving FBSDEs of McKean-Vlasov type, which can be used to formulate the solutions to mean field game

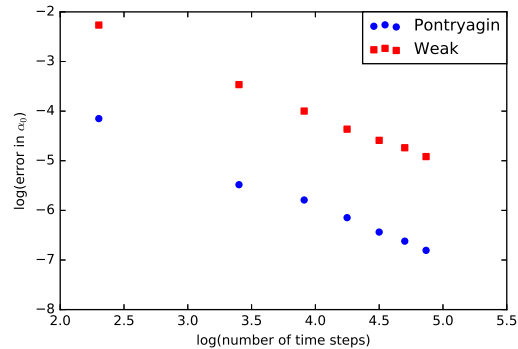


Figure 6.16 – Trader Problem: Error in the control at time 0, α_0 , as we increase the number of time steps, plotted as a log-log plot.

problems. The first algorithm is based on a pathwise tree structure. The second algorithm is based on a marginal grid structure. We have also proposed various refinements to the algorithms, including a continuation in time, and incrementation of a coupling parameter or the diffusion coefficient. The different numerical methods were illustrated on five benchmark examples.

The tree algorithm’s main advantage is that we do not need to project the values of X_t onto a discretized spatial grid, which potentially makes the algorithm more accurate. However, a significant disadvantage of the tree algorithm is the exponential growth of the data structure as the number of time steps is increased. This exponential growth is made worse yet if a higher order of quantization were to be used for approximating the Brownian increments.

The grid algorithm’s main advantage is it avoids the exponential growth of the data structure. A higher order of quantization may be used without drastically changing the algorithm’s complexity. A disadvantage of the grid algorithm is its sensitivity to the spatial step size with respect to the time step size. For the algorithm to be stable, the two step sizes need to be well adjusted to each other.

For both the tree and grid algorithms, we have observed that the continuation in time is able to extend the range of values of the coupling parameters for which the algorithms will converge. The incrementation methods proposed in Subsection 6.4.3, however, were not very successful at avoiding the bifurcations.

This chapter has touched on many things that could be explored deeper. First of all, the extension of the grid algorithm in [DM+06] to the mean field setting has not been studied from a theoretical standpoint. It is an open question to determine if this algorithm converges (meaning that the error decreases as the grid size decreases). The effect of the continuation in time or incrementation of the coupling parameter and/or diffusion coefficient has also yet to be studied. The numerical results also raised questions on the influence of the diffusion coefficient in the bifurcations.

Bibliography

- [ACCD13] Yves Achdou, Fabio Camilli, and Italo Capuzzo-Dolcetta. “Mean field games: convergence of a finite difference method”. In: *SIAM Journal on Numerical Analysis* 51.5 (2013), pp. 2585–2612.
- [ACD10] Yves Achdou and Italo Capuzzo-Dolcetta. “Mean field games: numerical methods”. In: *SIAM Journal on Numerical Analysis* 48.3 (2010), pp. 1136–1162.
- [AG17] Ankush Agarwal and Emmanuel Gobet. “Finite variance unbiased estimation of stochastic differential equations”. In: *2017 Winter Simulation Conference (WSC)*. IEEE. 2017, pp. 1950–1961.
- [AKH+02] Fabio Antonelli, Arturo Kohatsu-Higa, et al. “Rate of convergence of a particle method to the solution of the McKean–Vlasov equation”. In: *The Annals of Applied Probability* 12.2 (2002), pp. 423–476.
- [AKH+17] Patrik Andersson, Arturo Kohatsu-Higa, et al. “Unbiased simulation of stochastic differential equations using parametrix expansions”. In: *Bernoulli* 23.3 (2017), pp. 2028–2057.
- [AL16] Yves Achdou and Mathieu Laurière. “Mean field type control with congestion”. In: *Applied Mathematics & Optimization* 73.3 (2016), pp. 393–418.
- [Ang+19] Andrea Angiuli, Christy V Graves, Houzhi Li, Jean-François Chassagneux, François Delarue, and René Carmona. “Cemracs 2017: numerical probabilistic approach to MFG”. In: *ESAIM: Proceedings and Surveys* 65 (2019), pp. 84–113.
- [AP16] Yves Achdou and Alessio Porretta. “Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games”. In: *SIAM Journal on Numerical Analysis* 54.1 (2016), pp. 161–186.
- [Bal06] Vlad Bally. “Lower bounds for the density of locally elliptic Itô processes”. In: *The Annals of Probability* 34.6 (2006), pp. 2406–2440.
- [Bas98] Richard F Bass. *Diffusions and elliptic operators*. Springer Science & Business Media, 1998.

- [BC15] Jean-David Benamou and Guillaume Carlier. “Augmented Lagrangian methods for transport optimization, mean field games and degenerate elliptic equations”. In: *Journal of Optimization Theory and Applications* 167.1 (2015), pp. 1–26.
- [BFK+05] Adrian D Banner, Robert Fernholz, Ioannis Karatzas, et al. “Atlas models of equity markets”. In: *The Annals of Applied Probability* 15.4 (2005), pp. 2296–2330.
- [BGKH03] Guillaume Bernis, Emmanuel Gobet, and Arturo Kohatsu-Higa. “Monte Carlo evaluation of Greeks for multidimensional barrier and lookback options”. In: *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 13.1 (2003), pp. 99–113.
- [BKH+15] Vlad Bally, Arturo Kohatsu-Higa, et al. “A probabilistic interpretation of the parametrix method”. In: *The Annals of Applied Probability* 25.6 (2015), pp. 3095–3138.
- [BT04] Bruno Bouchard and Nizar Touzi. “Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations”. In: *Stochastic Processes and their applications* 111.2 (2004), pp. 175–206.
- [BT97] Mireille Bossy and Denis Talay. “A stochastic particle method for the McKean-Vlasov and the Burgers equation”. In: *Mathematics of computation* 66.217 (1997), pp. 157–192.
- [CCD14] Jean-François Chassagneux, Dan Crisan, and François Delarue. “A probabilistic approach to classical solutions of the master equation for large population equilibria”. In: *arXiv preprint arXiv:1411.3009* (2014).
- [CCD+19] Jean-François Chassagneux, Dan Crisan, François Delarue, et al. “Numerical method for FBSDEs of McKean–Vlasov type”. In: *The Annals of Applied Probability* 29.3 (2019), pp. 1640–1684.
- [CD13] René Carmona and François Delarue. “Probabilistic analysis of mean-field games”. In: *SIAM Journal on Control and Optimization* 51.4 (2013), pp. 2705–2734.
- [CD+18] René Carmona, François Delarue, et al. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018.
- [CDL13] René Carmona, François Delarue, and Aimé Lachapelle. “Control of McKean–Vlasov dynamics versus mean field games”. In: *Mathematics and Financial Economics* 7.2 (2013), pp. 131–166.
- [CFS13] René Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. “Mean field games and systemic risk”. In: *Available at SSRN 2307814* (2013).
- [CL18] Pierre Cardaliaguet and Charles-Albert Lehalle. “Mean field game of controls and an application to trade crowding”. In: *Mathematics and Financial Economics* 12.3 (2018), pp. 335–363.
- [CLSF+19] Christa Cuchiero, Martin Larsson, Sara Svaluto-Ferro, et al. “Probability measure-valued polynomial diffusions”. In: *Electronic Journal of Probability* 24 (2019).

- [CM+19] Dan Crisan, Eamon McMurray, et al. “Cubature on Wiener space for McKean–Vlasov SDEs with smooth scalar interaction”. In: *The Annals of Applied Probability* 29.1 (2019), pp. 130–177.
- [CR+16] Jean-François Chassagneux, Adrien Richou, et al. “Numerical simulation of quadratic BSDEs”. In: *The Annals of Applied Probability* 26.1 (2016), pp. 262–304.
- [CS17] Patrick Chan and Ronnie Sircar. “Fracking, renewables, and mean field games”. In: *SIAM Review* 59.3 (2017), pp. 588–615.
- [Cuc19] Christa Cuchiero. “Polynomial processes in stochastic portfolio theory”. In: *Stochastic processes and their applications* 129.5 (2019), pp. 1829–1872.
- [CW16] Rene Carmona and Peiqi Wang. “Finite state mean field games with major and minor players”. In: *arXiv preprint arXiv:1610.05408* (2016).
- [Del02] François Delarue. “On the existence and uniqueness of solutions to FB-SDEs in a non-degenerate case”. In: *Stochastic processes and their applications* 99.2 (2002), pp. 209–286.
- [DM+06] François Delarue, Stéphane Menozzi, et al. “A forward–backward stochastic algorithm for quasi-linear PDEs”. In: *The Annals of Applied Probability* 16.1 (2006), pp. 140–184.
- [DPF07] Giuseppe Da Prato and H elene Frankowska. “Stochastic viability of convex sets”. In: *Journal of mathematical analysis and applications* 333.1 (2007), pp. 151–163.
- [DS94] Freddy Delbaen and Walter Schachermayer. “A general version of the fundamental theorem of asset pricing”. In: *Mathematische annalen* 300.1 (1994), pp. 463–520.
- [EK81] Stewart N Ethier and Thomas G Kurtz. “The infinitely-many-neutral-alleles diffusion model”. In: *Advances in Applied Probability* 13.3 (1981), pp. 429–452.
- [EK93] Stewart N Ethier and Thomas G Kurtz. “Fleming–Viot processes in population genetics”. In: *SIAM Journal on Control and Optimization* 31.2 (1993), pp. 345–386.
- [Fer02] E Robert Fernholz. “Stochastic portfolio theory”. In: *Stochastic Portfolio Theory*. Springer, 2002, pp. 1–24.
- [Fer99] Robert Fernholz. “Portfolio generating functions”. In: *Quantitative Analysis in Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar*. World Scientific, 1999, pp. 344–367.
- [FK05] Robert Fernholz and Ioannis Karatzas. “Relative arbitrage in volatility-stabilized markets”. In: *Annals of Finance* 1.2 (2005), pp. 149–177.
- [FK09] Robert Fernholz and Ioannis Karatzas. “Stochastic portfolio theory: an overview”. In: *Handbook of numerical analysis* 15 (2009), pp. 89–167.
- [FK10] Daniel Fernholz and Ioannis Karatzas. “On optimal arbitrage”. In: *The Annals of Applied Probability* (2010), pp. 1179–1204.

- [FKHL+19] Noufel Frikha, Arturo Kohatsu-Higa, Libo Li, et al. “Integration by parts formula for killed processes: a point of view from approximation theory”. In: *Electronic Journal of Probability* 24 (2019).
- [FKK05] Robert Fernholz, Ioannis Karatzas, and Constantinos Kardaras. “Diversity and relative arbitrage in equity markets”. In: *Finance and Stochastics* 9.1 (2005), pp. 1–27.
- [FL16] Damir Filipović and Martin Larsson. “Polynomial diffusions and applications in finance”. In: *Finance and Stochastics* 20.4 (2016), pp. 931–972.
- [Fou+01] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, and Pierre-Louis Lions. “Applications of Malliavin calculus to Monte-Carlo methods in finance. II”. In: *Finance and Stochastics* 5.2 (2001), pp. 201–236.
- [Fou+99] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. “Applications of Malliavin calculus to Monte Carlo methods in finance”. In: *Finance and Stochastics* 3.4 (1999), pp. 391–412.
- [Fri08] Avner Friedman. *Partial differential equations of parabolic type*. Courier Dover Publications, 2008.
- [Fri10] Avner Friedman. “Stochastic differential equations and applications”. In: *Stochastic Differential Equations*. Springer, 2010, pp. 75–148.
- [FS82] Robert Fernholz and Brian Shay. “Stochastic portfolio theory and stock market equilibrium”. In: *The Journal of Finance* 37.2 (1982), pp. 615–624.
- [Fun84] T. Funaki. “A certain class of diffusion processes associated with nonlinear parabolic equations”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 67.3 (1984), pp. 331–348. ISSN: 1432-2064. DOI: [10.1007/BF00535008](https://doi.org/10.1007/BF00535008). URL: <https://doi.org/10.1007/BF00535008>.
- [Gär88] J. Gärtner. “On the McKean-Vlasov Limit for Interacting Diffusions”. In: *Mathematische Nachrichten* 137.1 (1988), pp. 197–248. DOI: [10.1002/mana.19881370116](https://doi.org/10.1002/mana.19881370116). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mana.19881370116>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mana.19881370116>.
- [GJ06] Christian Gourieroux and Joann Jasiak. “Multivariate Jacobi process with application to smooth transitions”. In: *Journal of Econometrics* 131.1-2 (2006), pp. 475–505.
- [GKH+03] Emmanuel Gobet, Arturo Kohatsu-Higa, et al. “Computation of Greeks for barrier and lookback options using Malliavin calculus”. In: *Electronic communications in probability* 8 (2003), pp. 51–62.
- [GM15] Diogo A Gomes and Hiroyoshi Mitake. “Existence for stationary mean-field games with congestion and quadratic Hamiltonians”. In: *Nonlinear Differential Equations and Applications NoDEA* 22.6 (2015), pp. 1897–1910.

- [GPV14] Diogo A Gomes, Stefania Patrizi, and Vardan Voskanyan. “On the existence of classical solutions for stationary extended mean field games”. In: *Nonlinear Analysis: Theory, Methods & Applications* 99 (2014), pp. 49–79.
- [Gué12] Olivier Guéant. “New numerical methods for mean field games with quadratic costs”. In: *Networks & Heterogeneous Media* 7.2 (2012), p. 315.
- [Hai16] Martin Hairer. “Convergence of Markov processes”. In: *Lecture notes*, <http://www.hairer.org/notes/Convergence.pdf> (2016).
- [HHS10] Chris Harris, Sam Howison, and Ronnie Sircar. “Games with exhaustible resources”. In: *SIAM Journal on Applied Mathematics* 70.7 (2010), pp. 2556–2581.
- [HLTT+17] Pierre Henry-Labordere, Xiaolu Tan, Nizar Touzi, et al. “Unbiased simulation of stochastic differential equations”. In: *The Annals of Applied Probability* 27.6 (2017), pp. 3305–3341.
- [HMC+06] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. “Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.
- [Jou97] B. Jourdain. “Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized Burgers’ equations”. In: *ESAIM: PS* 1 (1997), pp. 339–355. DOI: [10.1051/ps:1997113](https://doi.org/10.1051/ps:1997113). URL: <https://doi.org/10.1051/ps:1997113>.
- [JR15] Benjamin Jourdain and Julien Reygner. “Capital distribution and portfolio performance in the mean-field Atlas model”. In: *Annals of Finance* 11.2 (2015), pp. 151–198.
- [KR17] Ioannis Karatzas and Johannes Ruf. “Trading strategies generated by Lyapunov functions”. In: *Finance and Stochastics* 21.3 (2017), pp. 753–787.
- [KS85] Shigeo Kusuoka and Daniel Stroock. “Applications of the Malliavin calculus, Part II”. In: *J. Fac. Sci. Univ. Tokyo* 32.1-76 (1985), p. 18.
- [KS88] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate texts in mathematics. World Publishing Company, 1988. ISBN: 9783540965350. URL: <https://books.google.fr/books?id=uxnvAAAAMAAJ>.
- [Lip+01] Robert S Liptser, Robert Ševilevič Lipcer, Albert Nikolaevič Širjaev, Robert Sevilevič Lipcer, Albert Nikolaevich Shiriaev, and Albert N Shiryaev. *Statistics of random processes: I. General theory*. Vol. 1. Springer Science & Business Media, 2001.
- [LL07] Jean-Michel Lasry and Pierre-Louis Lions. “Mean field games”. In: *Japanese journal of mathematics* 2.1 (2007), pp. 229–260.
- [LST10] Aime Lachapelle, Julien Salomon, and Gabriel Turinici. “Computation of mean field equilibria in economics”. In: *Mathematical Models and Methods in Applied Sciences* 20.04 (2010), pp. 567–588.

- [LW11] Aimé Lachapelle and Marie-Therese Wolfram. “On a mean field game approach modeling congestion and aversion in pedestrian crowds”. In: *Transportation research part B: methodological* 45.10 (2011), pp. 1572–1589.
- [Lö13] Eva Löcherbach. “Ergodicity and speed of convergence to equilibrium for diffusion processes”. In: *Lecture notes*, <https://eloecherbach.u-cergy.fr/cours.pdf> (2013).
- [MPS11] Eberhard Mayerhofer, Oliver Pfaffel, and Robert Stelzer. “On strong solutions for positive definite jump diffusions”. In: *Stochastic processes and their applications* 121.9 (2011), pp. 2072–2086.
- [MT06] Paul Malliavin and Anton Thalmaier. *Stochastic calculus of variations in mathematical finance*. Springer Science & Business Media, 2006.
- [MU12] Aleksandar Mijatović and Mikhail Urusov. “On the martingale property of certain local martingales”. In: *Probability Theory and Related Fields* 152.1-2 (2012), pp. 1–30.
- [MV18] Y. S. Mishura and A. Y. Veretennikov. *Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations*. Preprint Arxiv arXiv:1603.02212. 2018.
- [Nag02] Hideo Nagai. “Optimal strategies for ergodic control problems arising from portfolio optimization”. In: *Stochastic Theory and Control*. Springer, 2002, pp. 353–368.
- [NCM10] Mojtaba Nourian, Peter E Caines, and Roland P Malhamé. “Synthesis of Cucker-Smale type flocking via mean field stochastic control theory: Nash equilibria”. In: *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2010, pp. 814–819.
- [NCM11] Mojtaba Nourian, Peter E Caines, and Roland P Malhamé. “Mean field analysis of controlled Cucker-Smale type flocking: Linear analysis and perturbation equations”. In: *IFAC Proceedings Volumes* 44.1 (2011), pp. 4471–4476.
- [Nua06] David Nualart. *The Malliavin calculus and related topics*. Vol. 1995. Springer, 2006.
- [Pal+11] Soumik Pal et al. “Analysis of market weights under volatility-stabilized market models”. In: *The Annals of Applied Probability* 21.3 (2011), pp. 1180–1213.
- [Pic14] Radka Picková. “Generalized volatility-stabilized processes”. In: *Annals of Finance* 10.1 (2014), pp. 101–125.
- [Pro56] Yu V Prokhorov. “Convergence of random processes and limit theorems in probability theory”. In: *Theory of Probability & Its Applications* 1.2 (1956), pp. 157–214.
- [RF19] Paul-Eric Chaudru de Raynal and Noufel Frikha. “Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space”. In: *Journal de Mathématiques Pures et Appliquées (Forthcoming)* (2019).

- [RST18] Goncalo dos Reis, Greig Smith, and Peter Tankov. “Importance sampling for mckean-vlasov sdes”. In: *arXiv preprint arXiv:1803.09320* (2018).
- [Ruf13] Johannes Ruf. “Hedging under arbitrage”. In: *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 23.2 (2013), pp. 297–317.
- [Ruf+15] Johannes Ruf et al. “The martingale property in the context of stochastic differential equations”. In: *Electronic Communications in Probability* 20 (2015).
- [RX19] Johannes Ruf and Kangjiana Xie. “Generalised Lyapunov functions and functionally generated trading strategies”. In: *Applied Mathematical Finance* 26.4 (2019), pp. 293–327.
- [SS91] Elias M Stein and Jeremy C Stein. “Stock price distributions with stochastic volatility: an analytic approach”. In: *The review of financial studies* 4.4 (1991), pp. 727–752.
- [STT+19] Lukasz Szpruch, Shuren Tan, Alvin Tse, et al. “Iterative multilevel particle approximation for McKean–Vlasov SDEs”. In: *The Annals of Applied Probability* 29.4 (2019), pp. 2230–2265.
- [Szn91] Alain-Sol Sznitman. “Topics in propagation of chaos”. In: *Ecole d’été de probabilités de Saint-Flour XIX—1989*. Springer, 1991, pp. 165–251.
- [Ver15] Alexander Vervuurt. “Topics in stochastic portfolio theory”. In: *arXiv preprint arXiv:1504.02988* (2015).
- [Ver81] Alexander Ju Veretennikov. “On strong solutions and explicit formulas for solutions of stochastic integral equations”. In: *Mathematics of the USSR-Sbornik* 39.3 (1981), p. 387.

Titre : Étude de méthodes numériques pour certaines équations différentielles stochastiques en finance et modélisation de la distribution du capital dans le marché financier

Résumé : Nous apportons dans cette thèse quelques contributions à la modélisation du marché financier dans le cadre de la théorie stochastique du portefeuille et à l'étude des méthodes numériques pour quelques équations différentielles stochastiques en modélisation financière et en théorie des jeux. Nous modélisons le marché par des poids relatifs des actifs et nous étudions un schéma probabiliste pour les lois marginales des solutions des EDS de McKean-Vlasov. Nous proposons également une représentation probabiliste et des formules d'intégration par parties à des modèles à volatilité stochastique pour obtenir des estimateurs Monte-Carlo sans biais du prix et des sensibilités. Enfin nous présentons deux algorithmes pour la résolution numérique des EDSPRs issues des jeux à champ moyen.

Mots clés : Théorie stochastique du portefeuille; Arbitrage relatif; Poids du marché; Processus ergodiques; EDS de McKean-Vlasov; Équations de Fokker-Planck non-linéaires; Analyse numérique probabiliste; Itération de Picard; Modèles à volatilité stochastique; Formules d'intégration par parties; Jeux à champ moyen; EDSPRs

Title : Study of numerical methods for some stochastic differential equations in finance and modeling of capital distribution in financial market

Abstract : In this thesis, we make some contributions to the modeling of the financial market in the context of stochastic portfolio theory, as well as to the study of numerical methods for some stochastic differential equations in financial modeling and game theory. The market is modeled by relative weights of assets and we study a probabilistic scheme for marginal laws of solutions of McKean-Vlasov SDEs. We also develop a probabilistic representation and integration by parts formulae for some stochastic volatility models to obtain unbiased Monte-Carlo estimators of price and sensitivities. Finally we present two algorithms for numerical resolution of FBSDEs arising from mean-field games.

Keywords : Stochastic Portfolio Theory; Relative Arbitrage; Market Weights; Ergodic Processes; McKean-Vlasov SDEs; Non-linear Fokker-Planck equations; Probabilistic numerical analysis; Picard iteration; Stochastic volatility models; Integration by parts formulae; Mean-field games; FBSDEs