Numerical Computations for Backward Doubly Stochastic Differential Equations and Nonlinear Stochastic PDEs

Achref Bachouch

L'UNAM l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

Université de Tunis El-Manar

Ecole Nationale d'ingénieurs de Tunis

Université du Maine
Le Mans - Laval

Thèse de Doctorat

Achref BACHOUCH

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université du Maine
sous le label de L'Université Nantes Angers Le Mans
et de l'Ecole Nationale d'Ingénieurs de Tunis
sous le label de L'Université de Tunis EL Manar

École doctorale : *STIM*

Discipline : *26*
Spécialité : *Mathématiques*
Unité de recherche : *LMM*

Soutenue le *1er octobre 2014*

Numerical Computations for Backward Doubly Stochastic Differential Equations and Non-linear Stochastic PDEs

JURY

Rapporteurs : **Bruno BOUCHARD**, Professeur, *ENSAE et Université Paris Dauphine*
Dan CRISAN, Professeur, Imperial College London

Examinateurs : **Laurent DENIS**, Professeur, Université du Maine, Le Mans
Nabil GMATI, Professeur, Ecole Nationale d'Ingénieurs de Tunis, Tunis
Emmanuel GOBET, Professeur, Ecole Polytechnique, Paris
Annie Millet, Professeur, Université Paris 1
Nizar TOUZI, Professeur, Ecole Polytechnique, Paris

Directeur de Thèse : **Anis MATOUSSI**, Professeur, Université du Maine, Le Mans

Co-directeur de Thèse : **Mohamed MNIF**, Professeur, Ecole Nationale d'Ingénieurs de Tunis, Tunis

Remerciements

Tout d'abord, je tiens à remercier particulièrement mon directeur de thèse, Anis Matoussi, pour ses conseils, ses remarques, la générosité avec laquelle il partageait son savoir, sa disponibilité et surtout sa grande patience. Il m'a motivé dès mon premier cours de master. Son soutien a toujours été constant et déterminant, autant sur le plan scientifique que sur le plan psychologique. L'accomplissement de cette thèse lui doit énormément.

Ensuite, je ne manquerais pas de remercier profondément mon co-directeur de thèse, Mohamed Mnif, qui a dirigé une bonne partie de ce travail. Il a été présent par ses conseils et ses remarques, pour accompagner mes premiers pas en rehcerche depuis le master.

Je remercie chaleureusement Bruno Bouchard et Dan Crisan qui, bien qu'ayant des emplois du temps très chargés, ont accepté de me faire l'honneur de rapporter cette thèse. Je les remercie également pour tout l'intérêt qu'ils ont montré pour mes travaux de recherche. Leurs corrections et leurs remarques pertinentes ont beaucoup contribué à l'amélioration de ce manuscrit.

Mes remerciements s'adressent très particulièrement à Emmanuel Gobet, pour avoir accepté de faire partie de mon jury de thèse, et surtout pour tout le temps qu'il m'a consacré, tous les conseils qu'il m'a prodigués et toutes ses remarques. J'ai beaucoup appris en sa compagnie, autant sur le plan scientifique que sur le plan humain.

Je suis très reconnaissant à Annie Millet, Laurent Denis et Nizar Touzi pour m'avoir fait l'honneur de faire partie du jury et pour l'intérêt qu'ils portent à mon travail.

Je ne pourrais manquer d'exprimer ma gratitude pour Nabil Gmati, d'une part, pour avoir accepté de faire partie de mon jury, et d'autre part pour le soutien qu'il nous a toujours apporté en tant que directeur du laboratoire Lamsin à l'ENIT. Il a toujours su écouter les doctorants, et a tout fait pour nous offrir les meilleures conditions de travail.

J'ai une pensée très particulière à tous les participants au CEMRACS 2013, en particulier à mes deux coéquipiers du projet Redvar, François et Lionel.

Je ne saurais oublier de remercier tous les membres du laboratoire LMM, qui ont beaucoup facilité ma vie de doctorant, avec une pensée particulière à Brigitte et Irène.

Je ne pourrais également manquer de remercier tous les membres du laboratoire LAMSIN, avec une pensée particulière à Raoudha.

Mes remerciements les plus profonds vont naturellement à tous les membres de ma famille, qui m'ont soutenu constamment durant toutes ces longues années d'études : Papa, maman, Ines, Imène et Yosri.

Je remercie tous mes amis et proches, qui ont contribué de près ou de loin à l'accomplissement de cette thèse. Je remercie Mhamed Gaigi, mon "binôme" depuis le master, tout simplement pour son amitié. Je suis également très redevable à mon cher ami Sofiène Bacha, à qui cette thèse doit beaucoup.

Enfin, je ne me lasserais jamais de remercier Héla qui, malgré la distance, a toujours été présente à mes côtés .

Résumé

L'objectif de cette thèse est l'étude d'un schéma numérique pour l'approximation des solutions d'équations différentielles doublement stochastiques rétrogrades (EDDSR). Durant les deux dernières décennies, plusieurs méthodes ont été proposées afin de permettre la résolution numérique des équations différentielles stochastiques rétrogrades standards. Dans cette thèse, on propose une extension de l'une de ces méthodes au cas doublement stochastique. Notre méthode numérique nous permet d'attaquer une large gamme d'équations aux dérivées partielles stochastiques (EDPS) nonlinéaires. Ceci est possible par le biais de leur représentation probabiliste en termes d'EDDSRs. Dans la dernière partie, nous étudions une nouvelle méthode des particules dans le cadre des études de protection en neutroniques.
Cette thèse contient quatre chapitres. Dans le second chapitre, on propose un schéma numérique pour les EDDSRs. Nous étudions l'erreur de discrétisation en temps issue de notre schéma puis nous donnons la vitesse de convergence associée. Ensuite, nous déduisons un schéma numérique pour l'approximation des solutions faibles d'EDPS semilinéaires et nous donnons la vitesse de convergence en temps pour ce dernier schéma. Nous finissons le chapitre par des tests numériques. Dans le troisième chapitre, nous étendons notre méthode numérique aux équations différentielles doublement stochastiques rétrogrades généralisées (EDDSRG). Nous étudions l'erreur de discrétisation en temps et donnons la vitesse de convergence associée. Ensuite, nous déduisons un schéma numérique pour l'approximation des solution des EDPS quasilinéaires associés aux EDDSRG, en donnant la vitesse de convergence en temps. Nous finissons ce chapitre par des tests numériques. Dans le quatrième chapitre, on propose un schéma pour l'approximation par projections et simulations de Monte-Carlo des solutions d'EDDSRs discrètes. Ces EDDSRs discrètes apparaissent naturellement suite à la discrétisation temporelle des EDDSRs. On étudie l'erreur de régression dans un cas particulier mais très instructif. Afin d'éviter le problème de grandes dimensions dû au bruit auxilière, on procède à une analyse conditionnelle de l'erreur sachant les trajectoires de ce bruit extérieur. On obtient des bornes supérieures presques sûres non asymptotiques mais explicites de l'erreur de régression conditionnelle qui assurent la convergence de notre schéma. Ces bornes nous permettent de choisir les paramètres pour atteindre une précision donnée.
Dans le dernier chapitre, on étudie un problème d'estimation de probabilités faibles dans le cadre des études de protection en neutroniques. On adapte une méthode récente d'estimatiom de faibles probabilités par un système de particules en interaction, se basant sur l'algorithme de Hastings-Metropolis et qui est proposée initialement pour les variables aléatoires, au cas des chaînes de Markov. La formulation en termes de chaînes de Markov est très naturelle dans le cadre des neutroniques. On montre la convergence de notre algorithme. Enfin, l'implémentation de la méthode est donnée en détails dans le cas unidimensionnel ainsi que dans le cas bidimensionnel, avec des résultats numériques.

Mots-clés : Equations différentielles doublement stochastiques rétrogrades, Equations aux dérivées partielles stochastiques semilinéaires, Equations différentielles doublement stochastiques rétrogrades généralisées, Equations aux dérivées partielles stochastiques quasilinéaires, Projections, Simulations de Monte-Carlo, régression, Système de particules en interaction, Algorithme Hastings-Metropolis, Chaînes de Markov.

Abstract

The purpose of this thesis is to study a numerical method for backward doubly stochastic differential equations (BDSDEs in short). In the last two decades, several methods were proposed to approximate solutions of standard backward stochastic differential equations. In this thesis, we propose an extension of one of these methods to the doubly stochastic framework. Our numerical method allows us to tackle a large class of nonlinear stochastic partial differential equations (SPDEs in short), thanks to their probabilistic interpretation. In the last part, we study a new particle method in the context of shielding studies.

This thesis contains four chapters. In the second chapter, we propose a numerical scheme for BDSDEs. We study the error arising from the time discretization of these BDSDEs and we give the rate of convergence in time. Then, we deduce a numerical scheme to approximate the weak solutions of the associated semilinear SPDEs and we give the rate of convergence in time. Numerical tests are also given.

In the third chapter, we extend our numerical scheme to generalized backward doubly stochastic differential equations (GBDSDEs in short). We study the time discretization error and we give the rate of convergence in time. Then, we deduce a numerical scheme to approximate the solutions of the associated quasilinear SPDEs, with a divergence term. We deduce the rate of convergence and we give numerical tests.

In the fourth chapter, we propose an algorithm based on projections and Monte-Carlo simulations to approximate solutions of discrete BDSDEs, arising from the time discretization of BDSDEs. We study the regression error in a particular but very instructive case. In order to avoid the curse of dimension induced by the auxilary noise, we proceed to a conditional analysis of the error given this exterior noise. We obtain non asymptotic but explicit almost sure upper bounds for the regression conditional error. This insures the convergence of our scheme and allows us to choose parameters to achieve a given accuracy.

In the last chapter, we study a problem of small probability estimation in the context of shielding studies in neutron transport. Thus, we adapt a recent interacting particle method for small probability estimation, based on Hastings-Metropolis algorithm and given initially for random variables, to the case of Markov chains, which is a natural formulation for neutron transport problems. We show the convergence of our algorithm. Then, the practical implementation is given in details in the one and two-dimensional cases, with numerical results.

Keywords : Backward Doubly Stochastic Differential Equations, Semilinear Stochastic PDEs, Generalized Backward Doubly Stochastic Differential Equations, Quasilinear Stochastic PDEs, Projections, Monte-Carlo simulations, regression, Interacting particle systems, Hastings-Metropolis algorithm, Markov chains.

Table des matières

Chapitre 1

Introduction

The aim of this thesis is to study a numerical method for backward doubly stochastic differential equations (BDSDEs in short). In the same time, using the probabilistic interpretation of non linear stochastic partial differential equations (SPDEs in short), we deduce a numerical probabilistic method for the latter SPDEs, which numerical resolution was mainly done by analytic methods in the literature.

In the second chapter, we extend the numerical method given by Zhang [70] and Bouchard and Touzi [16] for standard backward stochastic differential equations (BSDEs in short) to the doubly stochastic framework. We study the time discretization error of our numerical scheme and we prove the convergence and the rate of convergence for this scheme under standard assumptions on our model. Afterthat, we deduce a numerical probabilistic scheme for weak solutions of semi linear SPDEs using their probabilistic interpretation in terms of BDSDEs given by Bally and Matoussi [9]. Finally, we deduce the convergence and the rate of convergence for the deduced numerical probabilistic method for the weak soltuions of SPDEs.

In the third chapter, we extend our numerical method for BDSDEs to generalized backward stochastic differential equations (GBDSDEs in short). We prove the convergence in time and the rate of convergence under stander assumptions. Then, we deduce a numerical probabilistic scheme to approximate weak solutions of quasilinear SPDEs (with a divergence term) using their probabilistic interpretation in terms of GBDSDEs given by Matoussi and Stoica [54]. Finally, we deduce the convergence in time and the rate of convergence for the latter scheme for the weak solutions of quasilinear SPDEs.

In the fourth chapter, we propose an algorithm to approximate the conditional expectations involved by our probabilistic scheme. This algorithm is based on a least-squares regression approach. In fact, we extend the algorithms studied for BSDEs by Gobet, Lemor and Warin [34] and more recently by Gobet and Turkedjiev [35] to the case of BDSDEs. We proceed to a conditionnal analysis of the regression error and give the rate of convergence in a particular but very instrucive case of BDSDEs.

In the fifth chapter, we study an interacting particle method for small-probability estimation in the context of neutronic shielding. This method is based on the Hastings-Metropolis algorithm and was first presented by Guyader and al [36] in the case of random variables. We show how to adapt the Hastings-Metropolis algorithm to the case of Markov chains and we prove a convergence result for this algorithm in this case. Finally, we give the practical implementation of the resulting method for small-probability estimation, for an academic one-dimensional problem, and for a two-dimensional shielding study.We deduce, for these two cases, that the proposed interacting-particle method beats a simple-Monte Carlo method, when the probability to estimate is small.

We begin by recalling some preliminaries on BSDEs. At first, such equations were introduced in the linear case in 1973 by J.M. Bismut[11] in Stochastic Optimal Control theory. He studied linear BSDEs in order to give a probabilistic interpretation of the Pontryagin maximum principle. In 1990, Pardoux and Peng [61] proved the first result of Well-posedness for the more general (non-linear) case. A non linear BSDE is defined as follows :
For a finite horizon time T and a given filtered probability space $\Big(\Omega, \mathcal{F}^W, \{\mathcal{F}_t^W\}_{0 \leqslant t \leqslant T}), \mathbb{P}\Big)$

generated by an $\mathbb{R}^d$-valued standard Brownian motion W, we call a BSDE the following equation :

$$Y_t = \xi + \int_t^T f(t, Y_t, Z_t)dt - \int_t^T Z_t dW_t, \forall t \in [0,T], \mathbb{P} - a.s., \tag{1.0.1}$$

where ξ is a given $\mathbb{R}^k$-valued and $\mathcal{F}_T$-measurable random variable and f is a given $\mathbb{R}^k$-valued progressively-measurable function defined on $[0,T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$. ξ is called the terminal condition and the function f is called the generator to the BSDE.
Resolving the BSDE (1.0.1) remains to find a pair of progressively-measurable processes (Y,Z) satisfying this equation and such that the process Y is continuous, $E[\sup_{0 \leqslant t \leqslant} |Y_t|^2] < \infty$ and $E[\int_0^T \|Z_t\|^2] < \infty$. We note here that the role of the process Z is to guarantee the adaptability of the process Y.
Pardoux and Peng [61] proved the existence and the uniqueness for the solution of such equation when $E\Big[\xi + \int_0^T |f(t,0,0)|^2 dt\Big]$ is finite.
As soon as this result was proved, these standard BSDEs were largely studied in the literature, since they have many applications in Mathematical Finance, Stochastic Control, Game theory etc ... (see [31] for an overview of BSDEs and their application).
Another application of BSDEs is in PDEs theory. Indeed, in the particular case of Markovian BSDEs, these equations are linked to PDEs. A BSDE is called Markovian when the randomness of the terminal condition and the generator is completely generated by a diffusion process $\{(X_s^{t,x})_{t \leqslant s \leqslant T}, (t,x) \in [0,T] \times \mathbb{R}^d\}$ which is the strong solution of the following standard SDE :

$$X_s^{t,x} = x + \int_s^T b(X_r^{t,x})dr + \int_s^T \sigma(X_r^{t,x})dW_r, \quad t \leqslant s \leqslant T, \tag{1.0.2}$$

b and σ are two functions on $\mathbb{R}^d$ with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d \times d}$, satisfying the standard Lipschitz continuous and linear growth assumptions.
The BSDE has the following form :

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(r, X_r, Y_r^{t,x}, Z_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r, t \leqslant s \leqslant T, \tag{1.0.3}$$

where Φ and f are two deterministic functions on $\mathbb{R}^d$ and $[0,T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$ with values in $\mathbb{R}^k$.
Now let Us consider the following PDE :

$$du_t(x) + \big(\mathcal{L}u_t(x) + f(t, x, u_t(x), \nabla u_t \sigma(t,x))\big)\, dt = 0, \tag{1.0.4}$$

over the time interval $[0,T]$, with a given final condition $u_T = \Phi$ and non-linear deterministic coefficient f. $\mathcal{L}$ is the infinitesimal generator associated to the diffusion which solution is X given by $\mathcal{L}u(t,x) = \frac{1}{2}\sum_{i,j=1}^d ((\sigma^*(x)\sigma(x))_{i,j}\partial_{i,j}^2 u(t,x) + \sum_{i=1}^d b_i(x)\partial_i u(t,x)$.
If we assume that this PDE has a solution which is regular enough, we obtain by applying the Itô formula that $(u(s, X_s^{t,x}), \nabla u(s, X_s^{t,x})\sigma(X_s^{t,x}))_{t \leqslant s \leqslant T}$ is a the solution of (1.0.3), which is a generalization of the Feynman-Kac formula to a semilinear case. The latter relation allows the numerical resolution of semilinear PDEs by probabilistic methods. This will be of special interest in this thesis.

1.1 Numerical scheme for semilinear SPDEs via BDSDEs

The aim of the second chapter is to investigate a numerical method for BDSDEs (a class of BSDEs) and then to deduce a numerical probabilistic scheme for the associated semilinear SPDEs. This chapter is based on the work [5].

1.1.1 Preliminaries on BDSDEs and Semilinear Stochastic PDEs

BDSDEs was introduced by Pardoux and Peng [63] in 1994 as a generalisation of the standard BSDEs. They are called doubly stochastic because they involve, in addition to the stochastic forward Itô integral w.r.t. the initial Brownian motion W, another stochastic Itô backward integral w.r.t. another Brownian motion B which is independent from W. More precisely, let $(\Omega, \mathcal{F}, P)$ be a probability space and $T > 0$ be a fixed horizon time. Then, let $\{W_t, 0 \leqslant t \leqslant T\}$ and $\{B_t, 0 \leqslant t \leqslant T\}$ be two mutually independent standard Brownian motions with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^l$, defined on $(\Omega, \mathcal{F}, P)$. We fix $t \in [0, T]$. For each $s \in [t, T]$, we define the family of $\sigma - Algebras$

$$\mathcal{F}_s^t := \mathcal{F}_{t,s}^W \vee \mathcal{F}_{s,T}^B$$

where $\mathcal{F}_{t,s}^W = \sigma\{W_r - W_t, t \leqslant r \leqslant s\}$, and $\mathcal{F}_{s,T}^B = \sigma\{B_r - B_s, s \leqslant r \leqslant T\}$. We take $\mathcal{F}^W = \mathcal{F}_{0,T}^W$, $\mathcal{F}^B = \mathcal{F}_{0,T}^B$ and $\mathcal{F} = \mathcal{F}^W \vee \mathcal{F}^B$. When $t = 0$, we denote the $\sigma - algebra$ $\mathcal{F}_s^0$ by $\mathcal{F}_s$ for simplicity. Without loss of generality, we assume that $\mathcal{F}^W$ and $\mathcal{F}^B$ are complete.
We stress that the collection $(\mathcal{F}_s^t)_{t \leqslant s \leqslant T}$ is neither increasing nor decreasing, and it does not constitute a filtration.
The BDSDE of our interest, is defined by

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr + \int_t^T h(r, Y_r, Z_r)\overleftarrow{dB_r} - \int_t^T Z_r dW_r, \forall t \in [0, T], \tag{1.1.1}$$

where ξ is an $\mathbb{R}^k$-valued and $\mathcal{F}_T$-measurable random variable, f and h are two given random functions on $[0, T] \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$ with values respectively in $\mathbb{R}^k$ and $\mathbb{R}^l$ and such that for all (y, z) in $\mathbb{R}^k \times \mathbb{R}^{k \times d}$, for all t in $[0, T]$, $f(t, y, z)$ and $h(t, y, z)$ are $\mathcal{F}_t$-measurable. A solution to such BDSDE is a pair of processes (Y, Z) such that for all t in $[0, T]$, Y_t and Z_t are $\mathcal{F}_t$-measurable, (Y, Z) are satisfying this equation and such that the process Y is continuous, $E[\sup_{0 \leqslant t \leqslant T} |Y_t|^2] < \infty$ and $E[\int_0^T \|Z_t\|^2] < \infty$.
In 1994, Pardoux and Peng [63] proved the Well-posedness result of the last equation when $E\Big[\xi + \displaystyle\int_0^T |f(t, 0, 0)|^2 dt + \int_0^T |h(t, 0, 0)|^2 dt\Big]$ is finite.

In the Markovian case, analogously to BSDEs, the BDSDE (1.1.1) becomes : for a given (t, x) in $[0, T] \times \mathbb{R}^d$,

$$\begin{aligned} Y_s^{t,x} &= \Phi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\overleftarrow{dB_r} \\ &- \int_s^T Z_r^{t,x} dW_r, \forall s \in [t, T], \end{aligned} \tag{1.1.2}$$

where $(X_s^{t,x})_{t \leqslant s \leqslant T}$ is the process introduced in (1.0.4). Pardoux and Peng [63] introduced Markovian BDSDEs to give a probabilistic interpretation to the classical solutions of the following Semi-linear SPDE :

$$du_t(x) + \big(\mathcal{L}u_t(x) + f(t, x, u_t(x), \nabla u_t \sigma(x))\big)\, dt + h(t, x, u_t(x), \nabla u_t \sigma(x)) \cdot \overleftarrow{dB_t} = 0, \tag{1.1.3}$$

where $\mathcal{L}$ is the second order differential operator defined in (1.0.4). We note here that $\{u_t(x) = u(t,x), (t,x) \text{ in } [0,T] \times \mathbb{R}^d\}$ is a random field such that for each (t,x) in $[0,T] \times \mathbb{R}^d$, $u(t,x)$ is $\mathcal{F}^B_{t,T}$-measurable.

1.1.1.1 Notations and Assumptions

In order to introduce classical then weak Sobolev solutions of SPDEs, we need to introduce some notations :
• $C^k(\mathbb{R}^p, \mathbb{R}^q)$ and $C^k_p(\mathbb{R}^p, \mathbb{R}^q)$ denote respectively the set of functions of class C^k from $\mathbb{R}^p$ to $\mathbb{R}^q$ and the set of functions of class C^k from $\mathbb{R}^p$ to $\mathbb{R}^q$, which, together with all their partial derivatives of order less or equal to k, grow at most like a polynomial function of the variable x at infinity.
• $C^k_b(\mathbb{R}^p, \mathbb{R}^q)$ (respectively $C^k_b([0,T] \times \mathbb{R}^p, \mathbb{R}^q)$) denotes the set of functions of class C^k from $\mathbb{R}^p$ (repectively from $[0,T] \times \mathbb{R}^p$) to $\mathbb{R}^q$ whose partial derivatives of order less or equal to k are bounded.
• $C^\infty_b(\mathbb{R}^p, \mathbb{R}^q)$ denotes the set of functions of class C^∞ from $\mathbb{R}^p$ to $\mathbb{R}^q$ whose partial derivatives are bounded.

The assumptions we will use in this chapter are :
Assumption (H1) There exist a positive constant K such that

$$|b(x) - b(x')| + \|\sigma(x) - \sigma(x')\| \leqslant K|x - x'|, \forall x, x' \in \mathbb{R}^d.$$

Assumption (H2) There exist two constants $K > 0$ and $0 \leqslant \alpha < 1$ such that for any $(t_1, x_1, y_1, z_1), (t_2, x_2, y_2, z_2) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$,

(i) $|f(t_1, x_1, y_1, z_1) - f(t_2, x_2, y_2, z_2)| \leqslant K\big(\sqrt{|t_1 - t_2|} + |x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\|\big)$,

(ii) $\|h(t_1, x_1, y_1, z_1) - h(t_2, x_2, y_2, z_2)\|^2 \leqslant K\big(|t_1 - t_2| + |x_1 - x_2|^2 + |y_1 - y_2|^2\big) + \alpha^2\|z_1 - z_2\|^2$,

(iii) $|\Phi(x_1) - \Phi(x_2)| \leqslant K|x_1 - x_2|$,

(iv) $\sup_{0 \leqslant t \leqslant T}(|f(t,0,0,0)| + ||h(t,0,0,0)||) \leqslant K$.

Assumption (H3)

(i) $b \in C^2_b(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in C^2_b(\mathbb{R}^d, \mathbb{R}^{d \times d})$

(ii) $\Phi \in C^2_b(\mathbb{R}^d, \mathbb{R}^k), f \in C^2_b([0,T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{d \times k}, \mathbb{R}^k)$
and $h \in C^2_b([0,T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{d \times k}, \mathbb{R}^{k \times l})$.

1.1.1.2 Solutions of SPDEs and Probabilistic representations

First, we state the the Well-posedness result for the Markovian BDSDE (1.1.2) :

Theorem 1.1.1. *Assume that* (**H1**) *and* (**H2**) *hold. Then there exist a unique solution* $(Y^{t,x}_s, Z^{t,x}_s)_{t \leqslant s \leqslant T}$ *to the BDSDE* (1.1.2) *such that*

$$E\Big[\sup_{t \leqslant s \leqslant T}|Y^{t,x}_s|^2 + \Big(\int_t^T \|Z^{t,x}_s\|^2 ds\Big)\Big] < +\infty. \tag{1.1.4}$$

The following theorems given in [63] introduce the classical solution of the SPDE (1.1.3) and give its probabilistic representation in terms of BDSDEs

Theorem 1.1.2. *Assume that* (**H1**) *and* (**H2**) *hold, that* $\Phi \in C^2(\mathbb{R}^d, \mathbb{R}^k)$. *Let* $\{u(t,x),(t,x) \in [0,T] \times \mathbb{R}^d\}$ *be a random field such that* $u(t,x)$ *is* $\mathcal{F}^B_{t,T}$*-measurable for each* (t,x), $u \in \mathcal{C}^{0,2}([0,T]\times \mathbb{R}^d, \mathbb{R}^k)$ *a.s. and* u *satisfies equation* (1.1.3). *Then* $u(t,x) = Y^{t,x}_t$, *where* $\{(Y^{t,x}_s, Z^{t,x}_s)_{t \leqslant s \leqslant T}, 0 \leqslant t \leqslant T, x \in \mathbb{R}^d\}$ *is the unique solution of the BDSDE* (1.1.2).

Conversely, we have

Theorem 1.1.3. *Let* $\{(Y^{t,x}_s, Z^{t,x}_s)_{t \leqslant s \leqslant T}, 0 \leqslant t \leqslant T, x \in \mathbb{R}^d\}$ *be the unique solution of the BDSDE* (1.1.2). *Assume that* (**H2**) *holds,* $b \in C^2(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C^3(\mathbb{R}^d, \mathbb{R}^{d\times d})$, $\Phi \in C^3_p(\mathbb{R}^d, \mathbb{R}^k)$ *and such that for each* $s \in [0,T]$, $f(s,.,.,.)$ *belongs to* $C^3(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{d\times k}, \mathbb{R}^k)$ *and* $h(s,.,.,.)$ *belongs to* $C^3(\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{d\times k}, \mathbb{R}^{k\times l})$. *Then* $\{u(t,x) := Y^{t,x}_t, 0 \leqslant t \leqslant T\}$ *is the unique classical solution of the SPDE* (1.1.3).

For weak solutions, Bally and Matoussi [9] showed that the previous representation of the SPDE's solution remains true under weaker assumptions, namely when the terminal condition Φ is only meaurable in x and the coefficients f and g are only measurable in (t,x). They considered weak sobolev solutions of SPDE (1.1.3) and their approach was based on flow technics. Their result is crucial for us, since our numerical method for BDSDEs solves also SPDEs by using the representation they proved in the case of weak solutions.

First, let us give the definition of the weak Sobolev solution of the SPDE (1.1.3). Since we work on the whole space $\mathbb{R}^d$, we introduce a weight function ρ which is positive, satisfying $\int_{\mathbb{R}^d}(1+|x|^2)\rho(x)dx < \infty$ and such that ρ and $\frac{1}{\rho}$ are locally integrable. For example, we can take $\rho(x) = e^{-\frac{x^2}{2}}$ or $\rho(x) = e^{-|x|}$. As a consequence of (**H3**), we have $\int_{\mathbb{R}^d} |\Phi(x)|^2\rho(x)dx < \infty$, $\int_0^T \int_{\mathbb{R}^d} |f(t,x,0,0)|^2 \rho(x)dxdt < \infty$ and $\int_0^T \int_{\mathbb{R}^d} |h(t,x,0,0)|^2\rho(x)dxdt < \infty$.

We need also to introduce the following spaces :

· $L^2(\mathbb{R}^d, \rho(x)dx)$ is the weighted Hilbert space and we employ the following notation for its scalar product and its norm is : $(u,v)_\rho = \int_{\mathbb{R}^d} u(x)v(x)\rho(x)dx$ and $\|u\|_\rho = (u,u)^{\frac{1}{2}}_\rho$.

· $\mathrm{H}^1_\sigma(\mathbb{R}^d)$ is the associated weighted first order Dirichlet space and its norm is $\|u\|_{\mathrm{H}^1_\sigma(\mathbb{R}^d)} = (\|u\|^2_\rho + \|\nabla u\sigma\|^2_\rho)^{\frac{1}{2}}$.

· $\mathcal{D} := \mathcal{C}^\infty_c([0,T]) \otimes \mathcal{C}^2_c(\mathbb{R}^d)$ is the space of test functions where $\mathcal{C}^\infty_c([0,T])$ denotes the space of all real valued infinite differentiable functions with compact support in $[0,T]$ and $\mathcal{C}^2_c(\mathbb{R}^d)$ the set of C^2-functions with compact support in $\mathbb{R}^d$.

· $\mathcal{H}_T$ is the space of predictable processes $(u_t)_{t\geqslant 0}$ with values in $\mathrm{H}^1_\sigma(\mathbb{R}^d)$ such that

$$\|u\|_T = \Big(E\Big[\sup_{0\leqslant t\leqslant T}\|u_t\|^2_\rho\Big] + E\Big[\int_0^T \|\nabla u_t\sigma\|^2_\rho dt\Big]\Big)^{\frac{1}{2}} < \infty.$$

Definition 1.1.1. *We say that* $u \in \mathcal{H}_T$ *is a weak solution of the equation (1.1.3) associated with the terminal condition* Φ *and the coefficients* (f,g), *if the following relation holds almost surely, for each* $\varphi \in \mathcal{D}$

$$\begin{aligned}\int_t^T (u(s,.),\partial_s\varphi(s,.))ds + \int_t^T \mathcal{E}(u(s,.),\varphi(s,.))ds + (u(t,.),\varphi(t,.)) - (\Phi(.),\varphi(T,.)) \quad (1.1.5)\\ = \int_t^T (f(s,.,u(s,.),(\nabla u\sigma)(s,.)),\varphi(s,.))ds + \sum_{i=1}^l \int_t^T (h(s,.,u(s,.),(\nabla u\sigma)(s,.)),\varphi(s,.))\overleftarrow{dB^i_s},\end{aligned}$$

where

$$\begin{cases} (.,.) \text{ denotes the usual scalar product in } L^2(\mathbb{R}^d, dx) \\ \text{and} \\ \mathcal{E}(u,\varphi) = (Lu,\varphi) = \int_{\mathbb{R}^d}((\nabla u\sigma)(\nabla\varphi\sigma) + \varphi\nabla((\frac{1}{2}\sigma^*\nabla\sigma + b)u))(x)dx \text{ is} \\ \text{the energy associated to the diffusion operator.}\end{cases}$$

From Bally and Matoussi [9], we have the following representation result :

Theorem 1.1.4. *Assume Assumptions* $(\mathbf{H1})-(\mathbf{H3})$ *hold, there exists a unique weak solution* $u \in \mathcal{H}_T$ *of the SPDE* (1.1.3). *Moreover,* $u(t,x) = Y_t^{t,x}$ *and* $Z_t^{t,x} = \nabla u_t \sigma$, $dt \otimes dx \otimes dP$ *a.e. where* $(Y_s^{t,x}, Z_s^{t,x})_{t \leqslant s \leqslant T}$ *is the solution of the BDSDE* (2.1.2). *Furthermore, we have for all* $s \in [t,T]$, $u(s, X_s^{t,x}) = Y_s^{t,x}$ *and* $(\nabla u \sigma)(s, X_s^{t,x}) = Z_s^{t,x}$ $dt \otimes dx \otimes dP$ *a.e.*

1.1.2 Numerical methods : Different approaches

We recall that our aim is two folds : first we present a method for solving numerically the BDSDE (1.1.2) and using the probabilistic representation of solutions of the SPDE (1.1.3), we will also deduce a new probabilistic method for solving this last equation.
Let us begin by an overview for existing probabilistic methods for solving standard BSDEs.

1.1.2.1 Numerical methods for BSDEs

In the deterministic PDE's case i.e. $h \equiv 0$, the numerical approximation of standard BSDEs has already been widely studied in the literature. We can cite for example Bally [8], Zhang [70], Bouchard and Touzi [16], Gobet, Lemor and Warin[34], Bouchard and Elie [15] and Crisan and Manolarakis [22]. Zhang [70] proposed a discrete-time numerical approximation, by step processes, for a class of decoupled F-BSDEs with path-dependent terminal values. He proved an L^2-type regularity result for the control process Z under Lipschitz assumptions. This result allowed him to derive a rate of convergence for his scheme of order the time step of the square of the L^2-error. A similar numerical scheme was proposed by Bouchard and Touzi [16] for decoupled F-BSDEs. In order to approximate the conditional expectations arising from the time discretization, they use the Malliavin approach and the Monte carlo method. Afterthat, Crisan, Manolarakis and Touzi [20] proposed a betterment of this method on the Malliavin weights. A completely explicit numerical scheme was proposed by Gobet, Lemor and Warin [34]. They also introduced an algorithm based on the least-squares regression approach and using the Monte Carlo method to compute the conditional expectations. Finally, another algorithm based on the curbature method and the approach of Bouchard and Touzi [16] and Zhang [70] was introduced by Crisan and Manolarakis [22]. These authors proposed also a second order discretization of BSDEs in [21].
When $h \neq 0$ and it does not depend on the control variable z, Aman [3] suggested a numerical scheme following the approach of Bouchard and Touzi [16]. Aboura [1] studied the same numerical scheme but following Gobet et al. [33]. Both of the two authors obtained a convergence of order the time discretization step of the square of the L^2- error.
In the case when $h \neq 0$, the numerical resolution of BDSDEs offers a probabilistic approach to resolve semilinear SPDEs which was mainly resolved by analytic methods in the literature.

1.1.2.2 Analytic methods for SPDEs

Analytic methods are based on time-space discretization of the SPDEs. Mainly, the discretization on space is done by three methods : the finite differences method, finite elements and the spectral Galerkin method. The Euler finite-difference scheme was studied mainly by Gyongy [39], Gyongi and Nualart [38], Gyongy and Krylov [37] and Gerencsér and Gyongy [32]. In [38], Gyongi and Nualart proved the convergence of this scheme. Then, Gyongy [39] gave the rate of convergence. After that, Gyongy and Krylov [37] gave a rate of convergence for a symmetric finite difference scheme for a class of linear SPDE driven by infinite dimensional brownian motion. They also proved that this rate can be improved by Richardson acceleration method.

Recently, Gerencsér and Gyongy [32] considered weak sobolev solutions for linear SPDEs, where the smoothness of coefficients are dimension-invariant. Then, they used the Richardson method to accelerate their rate of convergence. The finite element method was investigated by J.B. Walsh in [69]. He proved the convergence of this scheme and obtained similar rate to those of the finite difference scheme. The third method was mainly studied by Jentzen and Kloeden [44]. It is based on the spectral Galerkin approximation, which consists on deriving Taylor expansions for the solution of the SPDE and consequently adding more regularity assumptions on the coefficients of the SPDE.
Other methods was also explored in order to resolve numerically SPDEs by an analytic approach. We can cite the spectral approach used by Lototsky, Mikulevicius and Rozovskii [51] to approximate numerically the conditional law of the solution of the Zakai equation. The method of Characteristics, which is based on the averaging characteristic formula, was used by Milstein and Tretyakov [57] to solve a linear SPDE. They based their numerical scheme on Monte Carlo technics. Finally, Crisan [19] studied a particle approximation for a class of nonlinear stochastic partial differential equations.

1.1.3 Our main results and contributions

1.1.3.1 Numerical scheme and main results

In order to solve the F-BDSDE (1.1.2), we we introduce the following discretized version. Let

$$\pi : t_0 = 0 < t_1 < \ldots < t_N = T, \tag{1.1.6}$$

be a partition of the time interval $[0,T]$. For simplicity we take an equidistant partition of $[0,T]$ i.e. $\Delta_N = \frac{T}{N}$ and $t_n = n\Delta_N$, $0 \leqslant n \leqslant N$. Throughout the rest, we will use the notations $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ and $\Delta B_n = B_{t_{n+1}} - B_{t_n}$, for $n = 1, \ldots, N$.
The forward component X will be approximated by the classical Euler scheme :

$$\begin{cases} X^N_{t_0} &= X_{t_0}, \\ X^N_{t_n} &= X^N_{t_{n-1}} + b(X^N_{t_{n-1}})(t_n - t_{n-1}) + \sigma(X^N_{t_{n-1}})(W_{t_n} - W_{t_{n-1}}), \text{ for } n = 1, \ldots, N. \end{cases} \tag{1.1.7}$$

It is known that as N goes to infinity, one has $\sup_{0 \leqslant n \leqslant N} E|X_{t_n} - X^N_{t_n}|^2 \to 0$.
Then, the solution (Y, Z) of (1.1.2) is approximated by (Y^N, Z^N) defined by :

$$Y^N_{t_N} = \Phi(X^N_T) \text{ and } Z^N_{t_N} = 0, \tag{1.1.8}$$

and for $n = N-1, \ldots, 0$, we set

$$Y^N_{t_n} = E_{t_n}[Y^N_{t_{n+1}} + h(t_{n+1}, \Theta^N_{n+1})\Delta B_n] + \Delta_N f(t_n, \Theta^N_n), \tag{1.1.9}$$

$$\Delta_N Z^N_{t_n} = E_{t_n}\left[Y^N_{t_{n+1}} \Delta W^*_n + h(t_{n+1}, \Theta^N_{n+1})\Delta B_n \Delta W^*_n \right], \tag{1.1.10}$$

where

$$\Theta^N_n := (X^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}), \forall n = 0, \ldots, N.$$

$*$ denotes the transposition operator and E_{t_n} denotes the conditional expectation over the σ-algebra $\mathcal{F}^0_{t_n}$.
For all $n = 0, \ldots, N-1$, we define the pair of processes $(Y^N_t, Z^N_t)_{t_n \leqslant t < t_{n+1}}$ as the solution of the following BDSDE :

$$Y^N_t = Y^N_{t_{n+1}} + \int_t^{t_{n+1}} f(t_n, \Theta^N_n) ds + \int_t^{t_{n+1}} h(t_{n+1}, \Theta^N_{n+1}) \overleftarrow{dB_s} - \int_t^{t_{n+1}} Z^N_s dW_s, \; t_n \leqslant t < t_{n+1}. \tag{1.1.11}$$

First, we show two results concerning the time discretization error on our scheme for resolving BDSDEs. The first is an upper bound result :

Theorem 1.1.5. *Assume that Assumptions **(H1)** and **(H2)** hold, define the error*

$$Error_N(Y,Z) := \sup_{0 \leqslant s \leqslant T} E[|Y_s - Y_s^N|^2] + \sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds], \tag{1.1.12}$$

where Y^N and Z^N are given by (1.1.11) and the process $\bar{Z}$ is defined by $\bar{Z}_t = \frac{1}{\Delta_N} E_{t_n}[\int_{t_n}^{t_{n+1}} Z_s ds], \forall t \in [t_n, t_{n+1}), \forall n \in \{0, \ldots, N-1\}$ *and* $\bar{Z}_N = 0$.
Then

$$\begin{aligned} Error_N(Y,Z) &\leqslant C\Delta_N(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds. \end{aligned} \tag{1.1.13}$$

Then, we obtain the rate of convergence under more regularity by adding the Assumption (**H3**) :

Theorem 1.1.6. *Under Assumptions* **(H1)**-**(H3)**, *there exists a positive constant C (depending only on T, K, α, $|b(0)|$, $||\sigma(0)||$, $|f(t,0,0,0)|$ and $||h(t,0,0,0)||$) such that*

$$Error_N(Y,Z) \leqslant C\Delta_N(1+|x|^2). \tag{1.1.14}$$

In the second chapter, we show under Assumption (**H3**) a representation result for Z that we use to prove the rate of convergence. However, the same rate of convergence is shown only under (**H1**) and (**H2**), which are natural assumptions in BDSDE's setting (see Remark 1.1.1).
Finally, we deduce the numerical scheme to resolve SPDE (1.1.3) by using the following lemma which comes from the Markov property of Y^N and Z^N and the flow property of X^N :

Lemma 1.1.1. *Let $x \in \mathbb{R}^d$ and $t_n \in \pi$. Define*

$$u_{t_n}^N(x) := Y_{t_n}^{N,t_n,x} \text{ and } v_{t_n}^N(x) := Z_{t_n}^{N,t_n,x}. \tag{1.1.15}$$

Then $u_{t_n}^N$ (resp. $v_{t_n}^N$) is $\mathcal{F}_{t_n,T}^B$-measurable and we have for all $x \in \mathbb{R}^d$ and for all $t, t_n \in \pi$ such that $t \leqslant t_n$:

$$u_{t_n}^N(X_{t_n}^{t,x}) = Y_{t_n}^{N,t,x} \quad (resp.\ v_{t_n}^N(X_{t_n}^{t,x}) = Z_{t_n}^{N,t,x}).$$

After that, thanks to a norm equivalence result which was already proved by Barles and Lesigne [10] and Bally and Matoussi [9] when $b \in C_b^1(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in C_b^2(\mathbb{R}^d, \mathbb{R}^{d\times d})$, we deduce the rate of convergence for the numerical scheme for the SPDE (1.1.3). We stress that here we do not assume that σ satisfies the uniform ellipticity condition. We recall that $u(t,x) = Y_t^{t,x}$ and $v(t,x) = Z_t^{t,x}$ $dt \otimes dx \otimes dP$ a.e. We define the process (u_s^N, v_s^N) as follows :

$$u_s^N(x) := Y_s^{N,s,x} \text{ and } v_s^N(x) := Z_s^{N,s,x}, \forall s \in [t_n, t_{n+1}). \tag{1.1.16}$$

Thus, we obtain

$$u_s^N(X_s^{t,x}) = Y_s^{N,t,x} \text{ and } v_s^N(X_s^{t,x}) = Z_s^{N,t,x}, \forall t \leqslant s, t, s \in [t_n, t_{n+1}). \tag{1.1.17}$$

We define the error between the solution of the SPDE and the numerical scheme as follows :

$$\begin{aligned} Error_N(u,v) &:= \sup_{0 \leqslant s \leqslant T} E[\int_{\mathbb{R}^d} |u_s^N(x) - u(s,x)|^2 \rho(x)dx] \\ &+ \sum_{n=0}^{N-1} E[\int_{\mathbb{R}^d} \int_{t_n}^{t_{n+1}} \|v_s^N(x) - v(s,x)\|^2 ds\rho(x)dx], \end{aligned} \tag{1.1.18}$$

where ρ is the weight function defined in subsection 1.1.1.2.
The following theorem shows the convergence of the numerical scheme 1.1.15 of the solution of the SPDE (1.1.3).

Theorem 1.1.7. *Assume that* **(H1)**-**(H3)** *hold. Then, the error* $Error_N(u,v)$ *converges to 0 as* $N \to \infty$ *and there exists a positive constant* C *(depending only on* T*,* K*,* α*,* $|b(0)|$*,* $||\sigma(0)||$*,* $|f(t,0,0,0)|$ *and* $||h(t,0,0,0)||$*) such that*

$$Error_N(u,v) \leqslant C\Delta_N. \tag{1.1.19}$$

Remark 1.1.1. *The Assumptions needed to prove the rate of convergence* (1.1.14) *and consequently* (1.1.19) *can be relaxed. Indeed, we can obtain this rate of convergence* (1.1.14) *only under Assumptions* **(H1)** *and* **(H2)**. *This can be done by using the* L_2*-regularity for the process* Z *proved in section* (2.7).

1.1.3.2 Our Contributions

In our work, we extended the approach of Bouchard-Touzi-Zhang for standard BSDEs to the case of BDSDEs. We wish to emphasize that this generalization is not obvious, especially when the function h depends also on the control variable z, because of the strong impact of the backward stochastic integral term on the numerical approximation scheme. It is known that in the associated Stochastic PDE's (2.1.1), the term $h(u, \nabla u)$ leads to a second order perturbation type which explains the contraction condition assumed on h with respect to the variable z (see [63], [59]). Our scheme is implicit in Y and explicit in Z. We prove the convergence of our numerical scheme in time and we give the rate of convergence. The square of the L^2- error has an upper bound of order the discretization step in time. As a consequence, we obtain a numerical scheme for the weak solution of the associated semi linear SPDE. We give also a rate of convergence result for the later weak solution. Then, we propose a path-dependent algorithm based on iterative regression functions which are approximated by projections on vector space of functions with coefficients evaluated using Monte Carlo simulations. More precisely, we fix one path of the discretized Brownian motion B and we approximate the conditional expectations arising in our numerical scheme (1.1.9)-(1.1.10) by performing regressions using simulations of the forward process X. The analysis of the regression error will be handeled in chapter 4. Finally, we present some numerical tests.
Compared to the deterministic numerical method developed by Gyongy and Krylov [37], the probabilistic approach could tackle the semilinear SPDEs which could be degenerate and needs less regularity conditions on the coefficients than the finite difference scheme. However, the rate of convergence obtained (as the classical Monte Carlo method) is clearly slower than the results obtained by finite difference and finite element schemes, but of course more available in higher dimension. Indeed, our method has all the the known advantages of Monte Carlo methods. These latter methods are tractable especially when the dimension of the state process is very large unlike the finite difference method. Furthermore, their parallel nature provides another advantage to the probabilistic approach : each processor of a parallel computer can be assigned the task of making a random trial and doing the calculus independently.

1.2 Numerical computations for Quasilinear Stochastic PDE's

In the third chapter, we are interested in the numerical resolution of quasilinear SPDEs by a probabilistic method, using their probabilistic interpretation in terms of Generalized Backward Doubly Stochastic Differential Equations (GBDSDEs for short). This chapter is based on the work [7].

1.2.1 GBDSDEs and Quasilinear SPDEs

We are interested in the following SPDE on $\mathbb{R}^d$ and over the time interval $[0,T]$,

$$du_t(x)+\big[\frac{1}{2}\Delta u_t(x)+f_t(x,u_t(x),\nabla u_t(x))+\mathrm{div} g_t(x,u_t(x),\nabla u_t(x))\big]dt+h_t(x,u_t(x),\nabla u_t(x))\cdot\overleftarrow{dB}_t=0, \tag{1.2.1}$$

with a given final condition $u_T=\Phi$ and where f,h and g are non-linear random functions. When h is identically null, equation (1.2.1) becomes a Quasilinear PDE. This equation was studied by Stoica [68]. More precisely, he studied the solution $u:\mathbb{R}^d:\to\mathbb{R}$ of the following equation

$$(\partial_t+\mathcal{L}_0)u+f-\sum_{i,j}\partial_i(a^{i,j}g_j)=0, \tag{1.2.2}$$

where f and g are given real-valued functions on $[0,T]\times\mathbb{R}^d$, $\mathcal{L}_0$ is the elliptic divergence form operator given by :

$$\mathcal{L}_0:=\sum_{i,j}\partial_i(a_j^{i,j})+\sum_i b^i\partial_i,$$

$b(x):=(b^1(x),\dots,b^d(x))$ is a vector field and for all i, j, $a^{i,j}$ are bounded measurable functions on $\mathbb{R}^d$ satisfying the uniform ellipticity condition : for some positive constant λ,

$$\lambda^{-1}|\xi|^2\leqslant\sum_{i,j}a^{i,j}(x)\xi^i\xi^j\leqslant\lambda|\xi|^2,\forall\xi,x\in\mathbb{R}^d.$$

Denoting by $(\Omega,\mathcal{F},(\mathcal{F}_t)_t,X_t,\theta_t,P^x)$ the diffusion process generated by $\mathcal{L}_0$ in $\mathbb{R}^d$, Stoica [68] proved that the following relation holds (see Theorem 3.2 [68]), when f, g and the terminal condition Φ are square integrable :

$$u_t(X_t)-u_s(X_s)=\sum_{i=1}^d\int_s^t\partial_i u_r(X_r)dM_r^i-\int_s^t f_r(X_r)dr-\frac{1}{2}\int_s^t g_r*dX_r, \tag{1.2.3}$$

where M^i is the martingale part of the component X^i of the process and the integral denoted with $*$ is a stochastic martingale expressed in terms of forward and backward martingales.
We denote that if $\mathcal{L}_0$ is symmetric under the probability measure P^m (see section 3.2.2 for the definition of P^m), the term $\int_s^t g_r*dX_r$ becomes :

$$\int_s^t g_r*dX_r=\sum_{i=1}^d\left(\int_s^t g_i(r,X_r)dM_r^i+\int_s^t g_i(r,X_r)\overleftarrow{dM_r^i}\right).$$

Since the function g is assumed only measurable, the term $\sum_{i,j=1}^d\partial_i(a^{i,j}g_j)$ in equation (1.2.2) is a distribution. Hence, the stochastic integral $\int_s^t g_r*dX_r$ gives a probabilistic interpretation for a distribution. In other words, the solution u is represented in (1.2.3) in terms of a stochastic process, of the function f and the field.

In the general case (h and g are non-null functions), the SPDE (1.2.1) was given a probabilistic interpretation by Matoussi and Stoica [54] in term of the following GBDSDE :

$$u_t(W_t) - u_s(W_s) = \sum_{i=1}^{d} \int_s^t \partial_i u_r(W_r) dW_r^i - \int_s^t f_r(W_r)dr - \frac{1}{2}\int_s^t g_r * dW_r + hf_r(W_r)\overleftarrow{dB}_r. \quad (1.2.4)$$

We will see in chapter 3, that the study of the SPDE (1.2.1) covers a more general case. Indeed, if instead of $\frac{1}{2}\Delta$ in (1.2.1) we had $\mathcal{L} := \sum_{i,j} \partial_i(a_j^{i,j})$ where in addition a is symmetric, we show by the mean of a change of variable that the case of the SPDE (1.2.1) with the operator $\mathcal{L}$ is covered by our framework.

The aim of chapter 3 is to use the probabilistic interpretation (1.2.4) to give a probabilistic numerical scheme in order to resolve numerically the associated quasilinear SPDE (1.2.1).

1.2.2 Our Contribution : Numerical resolution of Quasilinear SPDEs - Time discretization error

We first present our numerical scheme to approximate solutions of (1.2.1).

1.2.2.1 Numerical scheme

We use the same notations of the numerical scheme presented in subsection 1.1.3.1. Let W^N denote the discrete time approximation of the Brownian motion W. The solution (Y, Z) of the GBDSDE (1.2.4) will be approximated by (Y^N, Z^N) defined in the following :

$$Y_{t_N}^N = \Phi(W_T^N) \text{ and } Z_{t_N}^N = 0, \quad (1.2.5)$$

and for $n = N-1, \ldots, 0$, we set

$$\begin{aligned} Y_{t_n}^N &= E_{t_n}[Y_{t_{n+1}}^N + \Delta_N f(t_n, \Theta_n^N) + \frac{1}{2}g(t_{n+1}, \Theta_{n+1}^N)\Delta W_n] \\ &+ E_{t_n}[h(t_{n+1}, \Theta_{n+1}^N)\Delta B_n], \end{aligned} \quad (1.2.6)$$

$$\begin{aligned} \Delta_N Z_{t_n}^N &= E_{t_n}[Y_{t_{n+1}}^N \Delta W_n^* + h(t_{n+1}, \Theta_{n+1}^N)\Delta B_n \Delta W_n^*] \\ &+ \frac{1}{2}E_{t_n}[\{g(t_n, \Theta_n^N) + g(t_{n+1}, \Theta_{n+1}^N)\}\Delta W_n \Delta W_n^*], \end{aligned} \quad (1.2.7)$$

where

$$\Theta_n^N := (W_{t_n}^N, Y_{t_n}^N, Z_{t_n}^N), \forall n = 0, \ldots, N.$$

$*$ denotes the transposition operator and E_{t_n} denotes the conditional expectation over the σ-algebra $\mathcal{F}_{t_n}^0$.

1.2.2.2 Contribution and main results

The probabilistic interpretation given below allows us to give a numerical probabilistic scheme for the quasilinear SDPE (1.2.1) based on a time-discretization of the GBDSDE (1.2.4).
First, we extend the Ito Formula for GBDSDEs. Then we give an an upper bound for the time discretization error. Afterthat, we extend the result concerning the Zhang L^2-regularity of the martingale integrand Z proved by Zhang [70] for standard BSDEs to our GBDSDE's case. This result is very important to derive the rate of convergence of our numerical scheme, which is of order the time discretization step of the square of the L^2- error. Finally, we give some numerical experiments to test statically our scheme.

1.3 An empirical regression method for BDSDEs

In the fourth chapter, which is based on the work [6], we present an algorithm to approximate the solutions of the discrete BDSDE arising from the time discretization of BDSDEs in the two previous chapters. Afterthat, we give a conditional analysis of the regression error involved by our algorithm. Let Us start by recalling some preleminaries on the regression method.

1.3.1 Empirical regression method for BSDEs

We recall that the empirical least-squares approach was developed in the standard BSDE's case by Gobet, Lemor and Warin [34] and more recently by Gobet and Turkedjiev [35]. This approach is based on iterative regression functions which are approximated by projections on vector space of functions with coefficients evaluated using Monte Carlo simulations. The BDSDE's of our interest is of the following form

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x})dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x})\overleftarrow{dB_r} - \int_s^T Z_r^{t,x} dW_r. \tag{1.3.1}$$

We study the case when the generators are independent from the variable z. For such BDSDE, Aboura [1] proposed an algorithm based on the empirical least-squares regression approach, following [33]. He considered the solution of the BDSDE at time t_i as a measurable deterministic function of $\Big(X_{t_i}, (B_{t_{k+1}} - B_k)_{i \leqslant k \leqslant N}\Big)$, where $t_i \in \pi := \{t_0, \dots, t_N\}$ and π is a discrete time grid of the time interval $[0, T]$. His approach imply a high-dimensionality problem, since he is dealing with a dimension of $d + l \times N$, where d is the dimension of the state process X and l is the dimension of the Brownian motion B.

1.3.2 Our contribution : Conditional approach for the empirical regression method for BDSDEs

We follow a conditional approach given the trajectories of the auxiliary noise B. We give first our numerical scheme.

1.3.2.1 Numerical scheme

We introduce the following notations :
· π is a discrete time grid of the time interval $[0, T]$ and $t_i \in \pi := \{t_0, \dots, t_N\}$.Then, $\Delta_i := t_{i+1} - t_i$.
· $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ and $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, for $i = 1, \dots, N-1$.
· We define the filtration $\mathcal{G}_t := \mathcal{F}_{0,t}^W \vee \mathcal{F}_{0,T}^B$ and we note by $E_{t_i}[.]$ the conditional expectation over $\mathcal{G}_{t_i}$.

We study the Multi step Dynamic Programming equation involved by the time discretizatin of the BDSDE (1.3.1) :
The forward component X will be approximated by the classical Euler scheme and the approximation is denoted by X^N.
The solution Y of (1.3.1) is approximated by (Y^N) defined by the following Multi step-forward Dynamic Programming (**MDP** for short) equation instead of the classical One step-forward Dynamic Programming (**ODP** for short) equation given in the previous chapters :
For $i = N-1, \dots, 0$, we set

$$Y_{t_i}^N = E_{t_i}\Big[\Phi(X_T^N) + \sum_{k=i}^{N-1} \Delta_k f(Y_{t_{k+1}}^N) + h(Y_{t_{k+1}}^N)\Delta B_k\Big]. \tag{1.3.2}$$

1.3.2.2 Main result

The main result of chapter 4 is non asymptotic and explicit error estimates for our algorithm, given in Theorem 4.4.1. These estimates are non asymptotic since the constants of the upper bound for the error depend on the time discretization number N and the fixed path $(\Delta B_k)_{0 \leqslant k \leqslant N-1}$. A direct consequence of the estimates in Theorem 4.4.1 is the following convergence result. We denote for all $i \in \{0, \dots, N-1\}$ by $\mathcal{K}_{Y,i}$ the finite dimensional approximation space and by M_i the number of Monte Carlo simulations of the forward process X used at time t_i to perform the regression. We assume that **Assumptions ($H1$-$H2$)** hold and that $\Phi(0), f(0)$ and $h(0)$ are bounded. Then, for a fixed time discretization number N and a fixed path $(\Delta B_k)_{0 \leqslant k \leqslant N-1}$, we obtain the convergence of our least-squares MDP algorithm by taking $(M_k)_{0 \leqslant k \leqslant N-1}$ and $(card(\mathcal{K}_{Y,k}))_{0 \leqslant k \leqslant N-1}$ large enough.

1.3.2.3 Our Contributions

First, we proceed to a conditional analysis of the error, given the trajectories of the Brownian motion B. To achieve this goal, we use the tools for the regression error analysis which were developed recently in [35] for standard BSDEs but in a very general context. Thus, we reduce the dimension of the regression problem from $d + l \times N$ in [1] to d. In this sense, our approach is better than the unconditional one.
Second, the MDP scheme studied leads to averaged local error terms, which is better than the sum as in the ODP's case. We obtain convergence results for fixed time discretization number N and a fixed path $(\Delta B_k)_{0 \leqslant k \leqslant N-1}$.

1.4 Variance reduction for small probability estimation : a Hastings-Metropolis algorithm on Markov chains

The fifth chapter is devoted to a work which has been done at CEMRACS 2013 (Centre d'été Mathématiques de Recherche Avancée en Calcul Scientifique). The CEMRACS is a scientific event of the SMAI (the french Society of Applied and Industrial Mathematics). The CEMRACS 2013 consisted of six weeks. In the first week, a summer school on numerical methods and algorithms for high performance computing was proposed. The remaining five weeks were intensive long research sessions on different research projects. These projects was proposed by an industrial or an academic partner. This event took place at CIRM, Luminy, Marseille, from 22 July to 30 August 2013. It was devoted to " Modelling and simulation of complex systems : stochastic and deterministic approaches". The project described in this chapter has the acronym REDVAR. Its main motivation is variance reduction technics for the estimation of rare events in the context of neutronic shielding. This work was concretized by a preprint [4] submitted to ESAIM Proceedings.

1.4.1 Monte Carlo method for rare events and Hastings-Metropolis algorithm on random variables

In this last part, we study an interacting particle method for small probability estimation developed in [36] and we adapt it to the context of neutronic shielding. Neutronics is the study of neutron population in fissile media that can be modeled using the linear Boltzmann equation called also the transport equation. The study of neutronics began in the 40's, when nuclear energy was starting to be used either for setting up nuclear devices like bombs or for civil purposes like the production of energy. Neutronics can be divided in two different sub-domains. The first sub-domain aim at understanding the neutron population dynamics due to the branching process that mimics fission reaction (see for instance [71] for a recent survey on branching process in neutronics). The second sub-domain deals with the propagation of neutrons through media where fission reactions do not occur, or can safely be neglected. In this case, the neutron transport can be modeled by simple exponential flights [72] : between each collisions, neutrons travel along straight path distributed exponentially.

For this last category, national nuclear authorities require shielding studies of nuclear systems before giving their agreement for the design of nuclear systems like reactor core. The study of such structure is complicated by 3-dimensional effects due to the geometry and by non-trivial energetic spectrum that can hardly be modeled. The Monte-Carlo transport codes (like MCNP [50], Geant4 [2], Tripoli-4 [29]) are often used for shielding studies since they require very few hypotheses. However, these studies remain a big numerical challenge for Monte-Carlo codes. Indeed, the shielding studies require to evaluate the proportion of neutrons that pass through the shielding disposal and this proportion is by construction very small. Consequently, the Monte Carlo code has to evaluate a small probability, which is the main motivation of this work.

Classical techniques for variance reduction in these small probability estimation problems often rely on a zero-variance scheme [43, 42, 12] adapted to the Boltzmann equation allied with weight-watching techniques [13]. The particular forms that this scheme takes when concretely developed in various transport codes range from the use of weight windows [17, 42, 43, 50], like in MCNP, to the use of the exponential transform [56, 55] like in Tripoli-4. Now, all these techniques have proven to be limited since the requirements made by national nuclear authorities have been progressively strengthen. As a consequence, new variance reduction techniques have been recently proposed (see for instance [30] for the use of neural networks for evaluating the importance function).

Recently in [36], the authors developed an interacting-particle method for small probability estimation in the case of random variables. In our work, we apply this method to a neutronic shielding's Monte Carlo code. This application to shielding study with Monte Carlo codes is not straightforward. In fact, a Monte Carlo code consists in sampling the trajectory of a neutron which can, depending of the complexity of the physical modeling, be the realization of a branching process, or of a stochastic process. Indeed, since a neutron travels along straight paths between collisions, there is no loss of information in considering only the characteristic of the collisions (dates, positions, energies, subparticle creations) as random.

In order to simplify the matter, the subparticle creation phenomena are not taken into account in this work. Similarly, we do not take energy dependence into account. As a result, we consider here the simplified but realistic case of monokinetic particle, that is a particle that has a constant speed and that can not give birth to other subparticles. For this model of a monokinetic particle, the set of the successive collision point positions constitute thus a Markov chain. Furthermore, with probability one, the monokinetic particle is absorbed after a finite number of collisions. The small probability we are interested in, in this work, is thus the probability that a Markov chain, that is almost surely stationary, in finite time, "pass" through a shielding system and reach a domain of interest before absorption.

The method proposed in [36] relies on the Hasting Metropolis algorithm [55, 41] for practical application. This algorithm is clearly a textbook method when applied to probability distributions on the Euclidean space. Nevertheless, we have discussed below that small probability estimation problems in Monte Carlo codes are defined as involving Markov Chains instead of random vectors. It is thus not straightforward to apply the method of [36] to these kind of problems. Let us begin by recalling the interacting particle method and the Hastings-Metropolis algorithm as presented in [36]

1.4.2 Preliminaries on Interacting Particle Method and Hastings-Metropolis algorithm for random variables

We present the interacting-particle method [36] and highlight its need of the Hastings-Metropolis (HM) algorithm for practical application.

We consider a probability space $(\Omega, \mathcal{F}, P)$, and a measurable space $(S, \mathcal{S}, Q)$. We consider a random variable X from $(\Omega, \mathcal{F}, P)$ to $(S, \mathcal{S}, Q)$. We assume that we are able to sample realizations of X.

We consider an objective function $\Phi : S \to \mathbb{R}$, for which we only assume that $\Phi(X)$ has a continuous cumulative distribution function F. The interacting-particle method aims at estimating the probability of the event $\Phi(X) \geqslant l$, for a given level $l \in \mathbb{R}$. We denote this probability by p.

The method can be presented in two steps. First, we assume that an ideal, or a theoretical, method can be implemented exactly. In this case, the finite-sample distribution of the corresponding estimator of the probability p is known exactly, so that exact finite-sample confidence intervals are available. Furthermore, the limit, for large number of sampling from X, of the probability estimation error, has attractive properties as shown in [36]. Nevertheless, this ideal method can not be implemented exactly for a large range of practical problems. Thus, it is proposed in [36] to approximate the ideal method by using a HM algorithm.

1.4.2.1 Theoretical version of the interacting-particle method

We assume that we are able to sample realizations of X, conditionally to the event $\Phi(X) \geqslant t$, for any $t \in \mathbb{R}$. This is a strong assumption and that is why the corresponding method that we present is called the ideal method.

The ideal algorithm for estimating p is then parameterized by a number of particle N and is as follows.

Algorithm 1.4.2.1

- Generate an *iid* sample $(X_1, .., X_N)$, from the distribution of X, and initialize $m = 1$, $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ and $X_1^1 = X_1, ..., X_N^1 = X_N$.
- While $L_m \leqslant l$ do
  - For $i = 1, ..., N$
    - Set $X_i^{m+1} = X_i^m$ if $\Phi(X_i^m) > L_m$, and else $X_i^{m+1} = X^*$, where X^* follows the distribution of X conditionally to $\Phi(X) \geqslant L_m$, and is independent of any other random variables involved in the algorithm.
  - Set $m = m + 1$.
  - Set $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$.
- The natural estimator of the probability p is $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

For each finite N, the ideal estimator $\hat{p}_{ipm}$ obtained from the algorithm 1.4.2.1 has an explicit distribution that is detailed in [36]. Here, we just consider two properties of $\hat{p}_{ipm}$. First, the estimator is unbiased : $\mathbb{E}(\hat{p}_{ipm}) = p$. Second, asymptotic 95% confidence intervals, for N large, are of the form

$$I_{\hat{p}_{ipm}} = \left[\hat{p}_{ipm} \exp\left(-1.96 \sqrt{\frac{-\log \hat{p}_{ipm}}{N}} \right), \hat{p}_{ipm} \exp\left(1.96 \sqrt{\frac{-\log \hat{p}_{ipm}}{N}} \right) \right]. \quad (1.4.1)$$

1.4.2.2 Practical implementation of the interacting-particle method with the Hastings-Metropolis algorithm

For practical implementation of the previous algorithm, the only problem we have to solve is the conditional sampling, with the distribution of X, conditionally to $\Phi(X) \geqslant t$, for any $t \in \mathbb{R}$.

An application of the HM algorithm is proposed in [36]. For this, the following is assumed

- The distribution of X has a probability distribution function (pdf) f with respect to $(S, \mathcal{S}, Q)$. For any $x \in S$ we can compute $f(x)$.
- We dispose of a transition kernel on $(S, \mathcal{S}, Q)$ with conditional pdf $\kappa(x, y)$ (pdf of y conditionally to x). We are able to sample from $\kappa(x, .)$ for any $x \in S$ and we can compute $\kappa(x, y)$ for any $x, y \in S$.

Let $t \in \mathbb{R}$ and $x \in S$ so that $\Phi(x) \geqslant t$. Then, the following algorithm enables to, starting from x, sample approximately with the distribution of X, conditionally to $\Phi(X) \geqslant t$. The algorithm is parameterized by a number of iterations $T \in \mathbb{N}^*$.

Algorithm 1.4.2.2.1

- Let $X = x$.
- For $i = 1, ..., T$
  - Independently from any other random variable, generate X^* following the $\kappa(X, .)$ distribution.
  - If $\Phi(X^*) \geqslant t$
    - Let $r = \frac{f(X^*)\kappa(X^*,X)}{f(X)\kappa(X,X^*)}$.
    - With probability $min(r, 1)$, let $X = X^*$.
- Return X.

The random variable returned by algorithm 1.4.2.2.1 is denoted $X_T(x)$.

For consistency, we now give the actual interacting-particle method, involving algorithm 1.4.2.2.1. This method is parameterized by the number of particles N and the number of HM iterations T.

Algorithm 1.4.2.2.2

- Generate an *iid* sample $(X_1, .., X_N)$ from the distribution of X and initialize $m = 1$, $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ and $X_1^1 = X_1, ..., X_N^1 = X_N$.
- While $L_m \leqslant l$ do
  - For $i = 1, ..., N$
    - Set $X_i^{m+1} = X_i^m$ if $\Phi(X_i^m) > L_m$, and else pick at random an integer J among the integers $1 \leqslant j \leqslant N$ so that $\Phi(X_j^m) > L_m$. Then, let $X_i^{m+1} = X_T(X_J^m)$, with the notation of algorithm 1.4.2.2.1.
  - Set $m = m + 1$.
  - Set $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$.
- The estimate of the probability p is $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

The estimator $\hat{p}_{ipm}$ of algorithm 1.4.2.2.2 is the practical estimator that we studied in the numerical results of section 5.5.

In [36], it is shown that, when the space S is a subset of $\mathbb{R}^d$, under mild assumptions, the distribution of the estimator of Algorithm converges, as $T \to +\infty$, to the distribution of the ideal estimator of Algorithm 1.4.2.1. For this reason, we call the estimator of ideal Algorithm 1.4.2.1 the estimator corresponding to the case $T = +\infty$. We also call the confidence intervals (1.4.1) the confidence intervals of the case $T = +\infty$.

Nevertheless, the space S we are interested in is a space of sequences that are killed after a finite time. Thus, it is not straightforward that the convergence, as $T \to +\infty$, discussed above, hold in our case. Furthermore, even the notion of pdf on this space of sequences has to be defined.

This is the object of the section 5.3, that defines the notion of pdf, on a space of sequences that are killed after a finite time, and that gives a convergence result for the HM algorithm.

1.4.3 Our contribution : Hastings-Metropolis algorithm on Markov chains for rare events

The contribution of this work is two-fold. First it is shown how the Hasting Metropolis algorithm can be extended to sampling of Markov chains that are stationary after finite time. This enables to use the Hastings-Metropolis algorithm, which is necessary to implement the method [36] in practice. A convergence result has also been shown for the Hastings-Metropolis algorithm in this setting.

The second contribution of the paper is to give the actual probability density function equations, for implementing the interacting-particle method in an academic one-dimensional problem, and a simplified but realistic two-dimensional problem. In both cases, the method is shown to be valid. Furthermore, the method outperforms a simple-Monte Carlo estimator, for estimating a small probability.

Prospects are possible for both contributions. First, the proof of the convergence of the Hastings-Metropolis could be extended under more general assumptions. Second, several possibilities for practical improvement of the interacting-particle method are presented in section 5.5.3.

Chapitre 2

Numerical scheme for semilinear SPDEs via BDSDEs

2.1 Introduction

Stochastic partial differential equations (SPDEs) combine the features of partial differential equations and Itô equations. Such equations play important roles in many applied fields such as the filtering of partially observable diffusion processes, genetic population and other areas. We study the following stochastic partial differential equation (in short SPDE) for a system-valued of predictable random field $u_t(x) = u(t,x)$, satisfying the following equation :

$$du_t(x) + \big(\mathcal{L}u_t(x) + f(t,x,u_t(x),\nabla u_t\sigma(x))\big)\,dt + g(t,x,u_t(x),\nabla u_t\sigma(x))\cdot \overleftarrow{dB}_t = 0, \qquad (2.1.1)$$

over the time interval $[0,T]$, with a given final condition $u_T = \Phi$ and non-linear deterministic coefficients f and g. $\mathcal{L}u = \big(Lu_1, \cdots, Lu_k\big)$ is a second order differential operator and σ is the diffusion coefficient. The differential term with $\overleftarrow{dB}_t$ refers to the backward stochastic integral with respect to a l-dimensional brownian motion on $\big(\Omega, \mathcal{F}, \mathbb{P}, (B_t)_{t\geqslant 0}\big)$. We use the backward stochastic integral in the SPDE because we will employ the framework of Backward Doubly Stochastic Differential Equation (BDSDE) introduced first by Pardoux and Peng [63]. They gave a probabilistic representation for the classical solution $u_t(x)$ of the SPDE (2.1.1) (written in the integral form) in term of the following class of BDSDE's :

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x})dr + \int_s^T g(r,X_r^{t,x},Y_r^{t,x},Z_r^{t,x})\overleftarrow{dB}_r - \int_s^T Z_r^{t,x}dW_r, \quad (2.1.2)$$

where $(X_s^{t,x})_{t\leqslant s\leqslant T}$ is a diffusion process starting from x at time t driven by the finite dimensional brownian motion $(W_t)_{t\geqslant 0}$ and with infinitesimal generator L. More precisely, under some regularity assumptions on the final condition Φ and coefficients f and g , they have proved that $u_t(x) = Y_t^{t,x}$ and $\nabla u_t\sigma(x) = Z_t^{t,x}$, $\forall (t,x) \in [0,T]\times\mathbb{R}^d$. Then, Bally and Matoussi [9] (see also [53]) showed that the same representation remains true in the case when the final condition (respectively the coefficients f and g) is only measurable in x (resp. are jointly measurable in (t,x) and Lispchitz in u and ∇u). In this paper, weak Sobolev solution of the equation (2.1.1) has been considered, and the approach was based on stochastic flow technics (see also [47, 48]). Moreover their results were generalized in [53] in the case of a larger class of SPDE's (2.1.1) driven by a Kunita-Ito non-linear noise (see [47, 48, 49] for more details). In particular, the Kunita-Ito non-linear noise covers a class of infinite dimensional space-time colored-white noise (see [37], [66], [44]). Generally, the explicit resolution of semi-linear SPDEs is not possible, so it is then necessary to resort to numerical methods.

The first approach used to solve numerically nonlinear SPDEs is analytic methods, based on time-space discretization of the SPDEs. The discretization on space can be achieved either by finite differences, or finite elements and spectral Galerkin methods. But most numerical works

on SPDEs have concentrated on the Euler finite-difference scheme. Gyongi and Nualart [38] have proved that these schemes converge, and Gyongy [39] determined the order of convergence. J.B. Walsh [69] investigated schemes based on the finite elements methods. He studied the rate of convergence of these schemes for parabolic SPDEs, including the Forward and Backward Euler and the Crank-Nicholson schemes. He found substantially similar rate of convergence to those found for finite difference schemes. The spectral Galerkin approximation was used by Jentzen and Kloeden [44]. They based their method on Taylor expansions derived for the solution of the SPDE, under some regularity conditions.
Lototsky, Mikulevicius and Rozovskii in 1997 [51] used the spectral approach for the numerical estimation of the conditional distribution solution of a linear SPDE known as the Zakai equation. Further developments on spectral methods can be found in Lototsky [52]. Milstein and Tretyakov [57] solved a linear Stochastic Partial Differential Equation by using the method of characteristics (the averaging over the characteristic formula). They proposed a numerical scheme based on Monte Carlo technique. Layer methods for linear and semilinear SPDEs are constructed. Picard [65] considered a filtering problem where the observation is a function of a diffusion corrupted by an independent white noise. He estimated the error caused by a discretization of the time interval. He obtained some approximations of the optimal filter which can be computed with Monte-Carlo methods. Crisan [19] studied a particle approximation for a class of nonlinear stochastic partial differential equations. Very interesting results have been obtained by Gyongy and Krylov [37] where they considered a symmetric finite difference scheme for a class of linear SPDE driven by infinite dimensional brownian motion. They have proved that the approximation error is proportional to k^2 where k is the discretization step in space and by the Richardson acceleration method they have even got the error proportional to k^4.

The other alternative for resolving numerically SPDEs is the probabilistic approach by using Monte Carlo methods. These latter methods are tractable especially when the dimension of the state process is very large unlike the finite difference method. Furthermore, their parallel nature provides another advantage to the probabilistic approach : each processor of a parallel computer can be assigned the task of making a random trial and doing the calculus independently. The probabilistic approach requires weaker assumptions on the SPDE's coefficients. In the deterministic PDE's case i.e. $g \equiv 0$, the numerical approximation of the BSDE has already been studied in the literature by Bally [8], Zhang [70], Bouchard and Touzi [16], Gobet, Lemor and Warin[34] and Bouchard and Elie [15]. Zhang [70] proposed a discrete-time numerical approximation, by step processes, for a class of decoupled FBSDEs with possible path-dependent terminal values. He proved an L^2-type regularity of the BSDE's solution, the convergence of his scheme and he derived its rate of convergence. Bouchard and Touzi [16] suggested a similar numerical scheme for decoupled FBSDEs. The conditional expectations involved in their discretization scheme were computed using the Malliavin approach and the Monte carlo method. Crisan, Manolarakis and Touzi [20] proposed an improvement on the Malliavin weights. Gobet, Lemor and Warin in [34] proposed an explicit numerical scheme. In the case when $g \neq 0$ and it does not depend on the control variable z, Aman [3] proposed a numerical scheme following the idea used by Bouchard and Touzi [16] and obtained a convergence of order h of the square of the L^2- error (h is the discretization step in time). Aboura [1] studied the same numerical scheme under the same kind of hypothesis, but following Gobet et al. [33]. He obtained a convergence of order h in time and used the regression Monte Carlo method to implement his scheme, following always [33].
In our work, we extend the approach of Bouchard-Touzi-Zhang in the general case when g depends also on the control variable z. We wish to emphasize that this generalization is not obvious because of the strong impact of the backward stochastic integral term on the numerical approximation scheme. It is known that in the associated Stochastic PDE's (2.1.1), the term $g(u, \nabla u)$

leads to a second order perturbation type which explains the contraction condition assumed on g with respect to the variable z (see [63], [59]). Our scheme is implicit in Y and explicit in Z. We prove the convergence of our numerical scheme and we give the rate of convergence. The square of the L^2- error has an upper bound of order the discretization step in time. As a consequence, we get a numerical scheme for the weak solution of the associated semi linear SPDE. We give also a rate of convergence result for the later weak solution. Then, we propose a numerical scheme based on iterative regression functions which are approximated by projections on vector space of functions with coefficients evaluated using Monte Carlo simulations. Finally, we present some numerical tests. Compared to the deterministic numerical method developed by Gyongy and Krylov [37], the probabilistic approach could tackle the semilinear SPDE which could be degenerate and needs less regularity conditions on the coefficients than the finite difference scheme. However, the rate of convergence obtained (as the classical Monte Carlo method) is clearly slower than the results obtained by finite difference and finite element schemes, but of course more available in higher dimension.

This paper is organized as follows. In section 2 we introduce preliminaries and assumptions and we describe the approximation scheme for the BDSDE. In section 3 we show an upper bound result for the time discretization error. In section 4 we give a Malliavin regularity result for the solution of our Forward-Backward Doubly SDE's. Then, we show a L^2-regularity result for the Z-component of the solution of the BDSDE (2.1.2) which is crucial to obtain the rate of convergence of our numerical scheme. Section 5 is devoted to the numerical scheme of the SPDE's weak solution. In section 6, we test statistically the convergence of this scheme by using a path dependent algorithm based on the regression Monte Carlo Method. Finally, we give some technical results in the Appendix

2.2 Preliminaries and notations

2.2.1 Forward Backward Doubly Stochastic Differential Equation

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $T > 0$ be a fixed horizon time. Then, let $\{W_t, 0 \leqslant t \leqslant T\}$ and $\{B_t, 0 \leqslant t \leqslant T\}$ be two mutually independent standard Brownian motions with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^l$, defined on $(\Omega, \mathcal{F}, P)$. We fix $t \in [0, T]$. For each $s \in [t, T]$, we define the family of $\sigma - Algebras$

$$\mathcal{F}_s^t := \mathcal{F}_{t,s}^W \vee \mathcal{F}_{s,T}^B$$

where $\mathcal{F}_{t,s}^W = \sigma\{W_r - W_t, t \leqslant r \leqslant s\}$, and $\mathcal{F}_{s,T}^B = \sigma\{B_r - B_s, s \leqslant r \leqslant T\}$. We take $\mathcal{F}^W = \mathcal{F}_{0,T}^W$, $\mathcal{F}^B = \mathcal{F}_{0,T}^B$ and $\mathcal{F} = \mathcal{F}^W \vee \mathcal{F}^B$. When $t = 0$, we denote the $\sigma - algebra$ $\mathcal{F}_s^0$ by $\mathcal{F}_s$ for simplicity. Without loss of generality, we assume that $\mathcal{F}^W$ and $\mathcal{F}^B$ are complete.

We stress that the collection $(\mathcal{F}_s^t)_{t \leqslant s \leqslant T}$ is neither increasing nor decreasing, and it does not constitute a filtration.

After that, we introduce the following spaces :

• $C_b^k(\mathbb{R}^p, \mathbb{R}^q)$ (respectively $C_b^\infty(\mathbb{R}^p, \mathbb{R}^q)$) denotes the set of functions of class C^k from $\mathbb{R}^p$ to $\mathbb{R}^q$ whose partial derivatives of order less or equal to k are bounded (respectively the set of functions of class C^∞ from $\mathbb{R}^p$ to $\mathbb{R}^q$ whose partial derivatives are bounded).

For any $m \in \mathbb{N}$, we introduce the following notations :

• $\mathbb{H}_m^2([0, T])$ denotes the set of (classes of $dP \times dt$ a.e. equal) $\mathbb{R}^m$-valued jointly measurable processes $\{\psi_u; u \in [0, T]\}$ satisfying :

(i) $||\psi||^2_{\mathbb{H}_m^2([0,T])} := E[\int_0^T |\psi_u|^2 du] < \infty$,

(ii) ψ_u is $\mathcal{F}_u$-measurable, for a.e. $u \in [0, T]$.

• $\mathbb{S}^2_m([0,T])$ denotes similary the set of $\mathbb{R}^m$-valued continuous processes satisfying :
(i) $||\psi||^2_{\mathbb{S}^2_m([0,T])} := E[\sup_{0 \leqslant u \leqslant T} |\psi_u|^2] < \infty$,
(ii) ψ_u is $\mathcal{F}_u$-measurable, for any $u \in [0,T]$.
• $\mathbb{S}$ the set of random variables F of the form :

$$F = \hat{f}(W(h_1), \ldots, W(h_{m_1}), B(k_1), \ldots, B(k_{m_2}))$$

with $\hat{f} \in C_b^\infty(\mathbb{R}^{m_1+m_2}, \mathbb{R})$, $h_1, \ldots, h_{m_1} \in L^2([0,T], \mathbb{R}^d), k_1, \ldots, k_{m_2} \in L^2([0,T], \mathbb{R}^l)$, where

$$W(h_i) := \int_0^T h_i(s) dW_s, \quad B(k_j) := \int_0^T k_j(s) \overleftarrow{dB_s}.$$

For any random variable $F \in \mathbb{S}$, we define its Malliavin derivative $(D_s F)_s$ with respect to the brownian motion W by

$$D_s F := \sum_{i=1}^{m_1} \nabla_i \hat{f}\Big(W(h_1), \ldots, W(h_{m_1}); B(k_1), \ldots, B(k_{m_2})\Big) h_i(s),$$

where $\nabla_i \hat{f}$ is the derivative of $\hat{f}$ with respect to its i-th argument.
We define a norm on $\mathbb{S}$ by :

$$\|F\|_{1,2} := \big\{E[F^2] + E\big[\int_0^T |D_s F|^2 ds\big]\big\}^{\frac{1}{2}}.$$

• $\mathbb{D}^{1,2} \triangleq \overline{\mathbb{S}}^{\|\cdot\|_{1,2}}$ is then a Sobolev space.
• $\mathcal{S}^2_k([0,T], \mathbb{D}^{1,2})$ is the set of processes $Y = (Y_u, 0 \leqslant u \leqslant T)$ such that $Y \in \mathbb{S}^2_k([0,T])$, $Y^i_u \in \mathbb{D}^{1,2}$, $1 \leqslant i \leqslant k$, $0 \leqslant u \leqslant T$ and

$$\|Y\|_{1,2} := \{E[\int_0^T |Y_u|^2 du] + E[\int_0^T \int_0^T ||D_\theta Y_u||^2 du d\theta]\}^{\frac{1}{2}} < \infty.$$

• $\mathcal{M}^2_{k\times d}([0,T], \mathbb{D}^{1,2})$ is the set of processes $Z = (Z_u, 0 \leqslant u \leqslant T)$ such that $Z \in \mathbb{H}^2_{k\times d}([0,T])$, $Z^{i,j}_u \in \mathbb{D}^{1,2}$,$1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant d$, $0 \leqslant u \leqslant T$ and

$$\|Z\|_{1,2} := \{E[\int_0^T \|Z_u\|^2 du] + E[\int_0^T \int_0^T ||D_\theta Z_u||^2 du d\theta]\}^{\frac{1}{2}} < \infty.$$

• $\mathcal{B}^2([0,T], \mathbb{D}^{1,2}) := \mathcal{S}^2_k([0,T], \mathbb{D}^{1,2}) \times \mathcal{M}^2_{k\times d}([0,T], \mathbb{D}^{1,2})$.
We define also for a given $t \in [0,T]$:
• $L^2([t,T], \mathbb{D}^{1,2})$ is the set of progressively measurable processes $(v_s)_{t \leqslant s \leqslant T}$ such that :
(i) $v(s,.) \in \mathbb{D}^{1,2}$, for a.e. $s \in [t,T]$,
(ii) $(s,w) \longrightarrow Dv(s,w) \in L^2([t,T] \times \Omega)$,
(iii) $E[\int_t^T |v_s|^2 ds] + E[\int_t^T \int_t^T |D_u v_s|^2 du ds] < \infty$.
• $L^2([t,T], \mathbb{D}^{1,2} \times \mathbb{D}^{1,2}) := L^2([t,T], \mathbb{D}^{1,2}) \times L^2([t,T], \mathbb{D}^{1,2})$.

For all $(t,x) \in [0,T] \times \mathbb{R}^d$, let $(X^{t,x}_s)_s$ be the unique strong solution of the following stochastic differential equation :

$$dX^{t,x}_s = b(X^{t,x}_s) ds + \sigma(X^{t,x}_s) dW_s, \quad s \in [t,T], \qquad X^{t,x}_s = x, \quad 0 \leqslant s \leqslant t, \tag{2.2.1}$$

where b and σ are two functions on $\mathbb{R}^d$ with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$. We will omit the dependance of the forward process X in the initial condition if it starts at time $t = 0$.
We consider the following BDSDE : For all $t \leqslant s \leqslant T$,

$$\begin{cases} dY^{t,x}_s &= -f(s, X^{t,x}_s, Y^{t,x}_s, Z^{t,x}_s) ds - g(s, X^{t,x}_s, Y^{t,x}_s, Z^{t,x}_s) \overleftarrow{dB_s} + Z^{t,x}_s dW_s, \\ Y^{t,x}_T &= \Phi(X^{t,x}_T), \end{cases} \tag{2.2.2}$$

where f and Φ are two functions respectively on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$ and $\mathbb{R}^d$ with values in $\mathbb{R}^k$ and g is a function on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$ with values in $\mathbb{R}^{k\times l}$.
We note that the integral with respect to $(B_s, t\leqslant s\leqslant T)$ is a "backward Itô integral" (see Kunita [49] and Nualart and Pardoux [59] for the definition) and the integral with respect to $(W_s, t\leqslant s\leqslant T)$ is a standard forward Itô integral.
Finally, for each real matrix A, we denote by $\|A\|$ its Frobenius norm defined by $\|A\| = (\sum_{i,j} a_{i,j}^2)^{1/2}$.
For a vector x, $|x|$ stands for its Euclidean norm defined by $|x| = (\sum_i |x_i|^2)^{1/2}$.
The following assumptions will be needed in our work :

Assumption (H1) There exist a positive constant K such that

$$|b(x)-b(x')|+\|\sigma(x)-\sigma(x')\|\leqslant K|x-x'|, \forall x,x'\in\mathbb{R}^d.$$

Assumption (H2) There exist two constants $K>0$ and $0\leqslant\alpha<1$ such that for any $(t_1,x_1,y_1,z_1),(t_2,x_2,y_2,z_2)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$,

(i) $|f(t_1,x_1,y_1,z_1)-f(t_2,x_2,y_2,z_2)|\leqslant K\big(\sqrt{|t_1-t_2|}+|x_1-x_2|+|y_1-y_2|+\|z_1-z_2\|\big)$,

(ii) $\|g(t_1,x_1,y_1,z_1)-g(t_2,x_2,y_2,z_2)\|^2\leqslant K\big(|t_1-t_2|+|x_1-x_2|^2+|y_1-y_2|^2\big)+\alpha^2\|z_1-z_2\|^2$,

(iii) $|\Phi(x_1)-\Phi(x_2)|\leqslant K|x_1-x_2|$,

(iv) $\sup_{0\leqslant t\leqslant T}(|f(t,0,0,0)|+||g(t,0,0,0)||)\leqslant K$.

Assumption (H3)

(i) $b\in C_b^2(\mathbb{R}^d,\mathbb{R}^d)$ and $\sigma\in C_b^2(\mathbb{R}^d,\mathbb{R}^{d\times d})$

(ii) $\Phi\in C_b^2(\mathbb{R}^d,\mathbb{R}^k)$, $f\in C_b^2([0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{d\times k},\mathbb{R}^k)$

and $g\in C_b^2([0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{d\times k},\mathbb{R}^{k\times l})$.

Pardoux and Peng [63] proved that there exists a unique solution $(Y,Z)\in\mathbb{S}_k^2([t,T])\times\mathbb{H}_{k\times d}^2([t,T])$ to the BDSDE (2.2.2).

Remark 2.2.1. *Pardoux and Peng [63] assumed the contraction condition $0\leqslant\alpha<1$ to prove the existence and the uniqueness results for the BDSDE's solution.*

From [31], [63] and [46], the standard estimates for the solution of the Forward-Backward Doubly SDE (2.2.1)-(2.2.2) hold and we remind the following theorem :

Theorem 2.2.1. *Under Assumptions **(H1)** and **(H2)**, there exist, for any $p\geqslant 2$, two positive constants C and C_p and an integer q such that :*

$$E[\sup_{t\leqslant s\leqslant T}|X_s^{t,x}|^2]\leqslant C(1+|x|^2), \tag{2.2.3}$$

$$E\Big[\sup_{t\leqslant s\leqslant T}|Y_s^{t,x}|^p+\Big(\int_t^T\|Z_s^{t,x}\|^2ds\Big)^{p/2}\Big]\leqslant C_p(1+|x|^q), \tag{2.2.4}$$

2.2.2 Numerical Scheme for decoupled Forward-BDSDE

In order to approximate the solution of the BDSDE (2.2.2), we introduce the following discretized version. Let

$$\pi: t_0=0<t_1<\ldots<t_N=T, \tag{2.2.5}$$

be a partition of the time interval $[0,T]$. For simplicity we take an equidistant partition of $[0,T]$ i.e. $h = \frac{T}{N}$ and $t_n = nh$, $0 \leqslant n \leqslant N$. Throughout the rest, we will use the notations $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ and $\Delta B_n = B_{t_{n+1}} - B_{t_n}$, for $n = 1, \dots, N$.
The forward component X will be approximated by the classical Euler scheme :

$$\begin{cases} X^N_{t_0} &= X_{t_0}, \\ X^N_{t_n} &= X^N_{t_{n-1}} + b(X^N_{t_{n-1}})(t_n - t_{n-1}) + \sigma(X^N_{t_{n-1}})(W_{t_n} - W_{t_{n-1}}), \text{ for } n = 1, \dots, N. \end{cases} \tag{2.2.6}$$

It is known that as N goes to infinity, one has $\sup_{0 \leqslant n \leqslant N} E|X_{t_n} - X^N_{t_n}|^2 \to 0$.
Quite naturally, the solution (Y,Z) of (2.2.2) is approximated by (Y^N, Z^N) defined by :

$$Y^N_{t_N} = \Phi(X^N_T) \text{ and } Z^N_{t_N} = 0, \tag{2.2.7}$$

and for $n = N-1, \dots, 0$, we set

$$Y^N_{t_n} = E_{t_n}[Y^N_{t_{n+1}} + g(t_{n+1}, \Theta^N_{n+1})\Delta B_n] + hf(t_n, \Theta^N_n), \tag{2.2.8}$$

$$hZ^N_{t_n} = E_{t_n}\left[\{Y^N_{t_{n+1}} + g(t_{n+1}, \Theta^N_{n+1})\Delta B_n\}\Delta W^*_n\right], \tag{2.2.9}$$

where

$$\Theta^N_n := (X^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}), \forall n = 0, \dots, N.$$

$*$ denotes the transposition operator and E_{t_n} denotes the conditional expectation over the σ-algebra $\mathcal{F}^0_{t_n}$.
For all $n = 0, \dots, N-1$, we define the pair of processes $(Y^N_t, Z^N_t)_{t_n \leqslant t < t_{n+1}}$ as the solution of the following BDSDE :

$$Y^N_t = Y^N_{t_{n+1}} + \int_t^{t_{n+1}} f(t_n, \Theta^N_n)ds + \int_t^{t_{n+1}} g(t_{n+1}, \Theta^N_{n+1})\overleftarrow{dB_s} - \int_t^{t_{n+1}} Z^N_s dW_s, \ t_n \leqslant t < t_{n+1}. \tag{2.2.10}$$

Remark 2.2.2. *Equation (2.2.10) is the continuous approximation of the solution of BDSDE (2.2.2). The sequences $(Y^N_{t_n})_{0 \leqslant n \leqslant N}$ given by (2.2.8) and (2.2.10) coincide. In Lemma 2.3.1 we will give the relation between $(Z^N_{t_n})_{0 \leqslant n \leqslant N-1}$ and $(Z^N_s)_{t_n \leqslant s < t_{n+1}}$.*

Remark 2.2.3. *For the approximation of $Y^N_{t_n}$, (2.2.8) is well-defined, indeed $Y^N_{t_n}(\omega)$ is a fixed point of*

$$\varphi(x) = hf(t_n, X^N_{t_n}(\omega), x, Z^N_{t_n}(\omega)) + E_{t_n}[Y^N_{t_{n+1}} + g(t_{n+1}, X^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Delta B_n](\omega),$$

which exists and is unique as soon as $Kh < 1$.

Remark 2.2.4. *The superscript (t,x) indicates the dependence of the solution (X,Y,Z) on the initial date (t,x). To alleviate notations, we omit the dependence on (t,x) of $(Y^{t,x}, Z^{t,x})$ and $(Y^{N,t,x}, Z^{N,t,x})$ when the context is clear.*
We note also that in the next computations, the constant C denotes a generic constant that may change from line to line. It depends on K, T, α, $|b(0)|$, $||\sigma(0)||$, $|f(t,0,0,0)|$ and $||g(t,0,0,0)||$.

2.3 Upper bound for the discrete time approximation error

Fisrt, we define the process $\bar{Z}$ by

$$\begin{cases} \bar{Z}_t = \frac{1}{h} E_{t_n}[\int_{t_n}^{t_{n+1}} Z_s ds], \forall t \in [t_n, t_{n+1}), \forall n \in \{0, \ldots, N-1\}, \\ \bar{Z}_{t_N} = 0. \end{cases} \tag{2.3.1}$$

Then we give the following property of the continuous approximation Z^N which shows that $Z^N_{t_n}$ is the best $L^2(\mathcal{F}_{t_n})$-estimate of $(Z^N_s)_s$.

Lemma 2.3.1. *For all $n = 0, \ldots, N-1$, we have*

$$Z^N_{t_n} = \frac{1}{h} E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s ds]. \tag{2.3.2}$$

Proof. From (2.2.10) we have

$$\begin{aligned} \int_{t_n}^{t_{n+1}} Z^N_s dW_s \Delta W_n &= Y^N_{t_{n+1}} \Delta W_n + \int_{t_n}^{t_{n+1}} f(t_n, \Theta^N_n) ds \Delta W_n \\ &+ \int_{t_n}^{t_{n+1}} g(t_{n+1}, \Theta^N_{n+1}) \overleftarrow{dB_s} \Delta W_n - Y^N_{t_n} \Delta W_n. \end{aligned}$$

then

$$\begin{aligned} E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s dW_s \Delta W_n] &= E_{t_n}[Y^N_{t_{n+1}} \Delta W_n] + \mathbb{E}_{t_n}[\int_{t_n}^{t_{n+1}} f(t_n, \Theta^N_n) ds \Delta W_n] \\ &+ E_{t_n}[\int_{t_n}^{t_{n+1}} g(t_{n+1}, \Theta^N_{n+1}) \overleftarrow{dB_s} \Delta W_n] - E_{t_n}[Y^N_{t_n} \Delta W_n] \\ &= E_{t_n}[Y^N_{t_{n+1}} \Delta W_n] + h E_{t_n}[f(t_n, \Theta^N_n) \Delta W_n] \\ &+ E_{t_n}[g(t_{n+1}, \Theta^N_{n+1}) \Delta B_n \Delta W_n] - E_{t_n}[Y^N_{t_n} \Delta W_n] \\ &= E_{t_n}[Y^N_{t_{n+1}} \Delta W_n] + E_{t_n}[g(t_{n+1}, \Theta^N_{n+1}) \Delta B_n \Delta W_n] \\ &= h Z^N_{t_n}. \end{aligned} \tag{2.3.3}$$

Here we used the fact that $Y^N_{t_n}$ and $f(t_n, \Theta^N_n)$ are $\mathcal{F}_{t_n}$-measurable and then we have

$$E_{t_n}[f(t_n, \Theta^N_n) \Delta W_n] = E_{t_n}[Y^N_{t_n} \Delta W_n] = 0.$$

Now by using the integration by parts formula we have

$$\begin{aligned} E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s dW_s \Delta W_n] &= E_{t_n}[\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_u Z^N_s dW_s] + E_{t_n}[\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} Z^N_u dW_u dW_s] \\ &+ E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s ds]. \end{aligned}$$

Then

$$E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s dW_s \Delta W_n] = E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s ds]. \tag{2.3.4}$$

Equations (2.3.3) and (2.3.4) give that

$$Z^N_{t_n} = \frac{1}{h} E_{t_n}[\int_{t_n}^{t_{n+1}} Z^N_s ds].$$

□

The next theorem states an upper bound result regarding the time discretization error.

Theorem 2.3.1. *Assume that Assumptions **(H1)** and **(H2)** hold, define the error*

$$Error_N(Y,Z) := \sup_{0\leqslant s\leqslant T} E[|Y_s - Y_s^N|^2] + \sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds], \tag{2.3.5}$$

where Y^N and Z^N are given by (2.2.10). Then

$$\begin{aligned} Error_N(Y,Z) \quad &\leqslant \quad Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ &+ \quad C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds. \end{aligned} \tag{2.3.6}$$

Proof. For all $t \in [t_n, t_{n+1})$, $n = 0, \ldots, N-1$ we define the following quantities :

$$\begin{cases} \theta_t := (X_t, Y_t, Z_t)\ , \delta Y_t^N := Y_t - Y_t^N,\ \delta Z_t^N := Z_t - Z_t^N, \\ \delta f_t := f(t,\theta_t) - f(t_n, \Theta_n^N), \\ \delta g_t := g(t,\theta_t) - g(t_{n+1}, \Theta_{n+1}^N). \end{cases} \tag{2.3.7}$$

We have :

$$\delta Y_t^N = \delta Y_{t_{n+1}}^N + \int_t^{t_{n+1}} \delta f_s ds + \int_t^{t_{n+1}} \delta g_s \overleftarrow{dB_s} - \int_t^{t_{n+1}} \delta Z_s^N dW_s, \forall t \in [t_n, t_{n+1}).$$

Using the Generalized Itô's Lemma (see Lemma 1.3, [63]), we obtain

$$\begin{aligned} |\delta Y_t^N|^2 + \int_t^{t_{n+1}} \|\delta Z_s^N\|^2 ds - |\delta Y_{t_{n+1}}^N|^2 \quad &= \quad 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta f_s)ds + 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta g_s)\overleftarrow{dB_s} \\ &+ \quad \int_t^{t_{n+1}} \|\delta g_s\|^2 ds - 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta Z_s^N)dW_s, \forall t \in [t_n, t_{n+1}). \end{aligned}$$

Then taking the expectation, we have

$$\begin{aligned} A_t^n := E[|\delta Y_t^N|^2] + \int_t^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds - E[|\delta Y_{t_{n+1}}^N|^2] \quad &= \quad 2\int_t^{t_{n+1}} E[(\delta Y_s^N, \delta f_s)]ds \\ &+ \quad \int_t^{t_{n+1}} E[\|\delta g_s\|^2]ds. \end{aligned} \tag{2.3.8}$$

From Assumption **(H2)-(ii)**,

$$\begin{aligned} &\int_t^{t_{n+1}} E[\|\delta g_s\|^2]ds \leqslant Kh^2 + K\int_t^{t_{n+1}} E[|X_s - X_{t_{n+1}}^N|^2]ds \\ + \quad &K\int_t^{t_{n+1}} E[|Y_s - Y_{t_{n+1}}^N|^2]ds + \alpha^2 E\big[\int_t^{t_{n+1}} ||Z_s - Z_{t_{n+1}}^N||^2 ds\big]. \end{aligned} \tag{2.3.9}$$

Using the Young's inequality, for a positive constant ε, we obtain for all $n = 0, \ldots, N-2$,

$$\begin{aligned} E\big[\int_t^{t_{n+1}} ||Z_s - Z_{t_{n+1}}^N||^2 ds\big] \quad &\leqslant \quad (1+\frac{1}{\varepsilon})E\big[\int_t^{t_{n+1}} ||Z_s - \bar{Z}_{t_{n+1}}||^2 ds\big] \\ &+ \quad (1+\varepsilon)E\big[\int_t^{t_{n+1}} ||\bar{Z}_{t_{n+1}} - Z_{t_{n+1}}^N||^2 ds\big]. \end{aligned} \tag{2.3.10}$$

Let Us note that the last inequality is not needed for $n = N-1$, since $Z^N_{t_N} = 0$ and then $E\big[\int_t^{t_N} ||Z_s - Z^N_{t_N}||^2 ds\big] = E\big[\int_t^{t_N} ||Z_s||^2 ds\big], \forall t \in [t_{N-1}, t_N)$.
For all $n = 0, \dots, N-2$, we use Lemma 2.3.1, the definition of $\bar{Z}$ and the Jensen's inequality to get

$$\begin{aligned} E\big[||\bar{Z}_{t_{n+1}} - Z^N_{t_{n+1}}||^2\big] &= E\Big[||\frac{1}{h}E_{t_{n+1}}\big[\int_{t_{n+1}}^{t_{n+2}} \delta Z^N_r dr\big]||^2\Big]. \\ &\leqslant \frac{1}{h^2}E\Big[E_{t_{n+1}}\big[||\int_{t_{n+1}}^{t_{n+2}} \delta Z^N_r dr||^2\big]\Big]. \end{aligned}$$

By using Cauchy Schwartz inequality, we obtain for all $n = 0, \dots, N-2$

$$E\big[||\bar{Z}_{t_{n+1}} - Z^N_{t_{n+1}}||^2\big] \leqslant \frac{1}{h}E\Big[\int_{t_{n+1}}^{t_{n+2}} \|\delta Z^N_r\|^2 dr\Big]. \tag{2.3.11}$$

Plugging (2.3.11) in (2.3.10) then (2.3.10) in (2.3.9), we get for all $n = 0, \dots, N-2$

$$\begin{aligned} &\int_t^{t_{n+1}} E[\|\delta g_s\|^2]ds \leqslant Kh^2 + K\int_t^{t_{n+1}} E[|X_s - X^N_{t_{n+1}}|^2]ds + K\int_t^{t_{n+1}} E[|Y_s - Y^N_{t_{n+1}}|^2]ds \\ +\ &(1+\frac{1}{\varepsilon})\alpha^2 \int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + (1+\varepsilon)\alpha^2 \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z^N_s\|^2]ds. \end{aligned} \tag{2.3.12}$$

The previous inequality becomes trivially for $n = N-1$,

$$\begin{aligned} &\int_t^{t_N} E[\|\delta g_s\|^2]ds \leqslant Kh^2 + K\int_t^{t_N} E[|X_s - X^N_{t_N}|^2]ds + K\int_t^{t_N} E[|Y_s - Y^N_{t_N}|^2]ds \\ +\ &\alpha^2 E\big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\big], \forall t \in [t_{N-1}, t_N). \end{aligned} \tag{2.3.13}$$

We mention that in the rest of the proof, we will omit to treat the case $n = N-1$. This case remains only to replace inequality (2.3.12) by inequality (2.3.13) in the following estimations, which is simpler to handle.
We set $\alpha' := (1+\varepsilon)\alpha^2$. We choose ε such that $\alpha' \in (0,1)$. This is possible since $\alpha^2 \in (0,1)$. Then, we use the inequality $2ab \leqslant \frac{1-\alpha'}{4K}a^2 + \frac{4K}{1-\alpha'}b^2$ and equation (2.3.12) to obtain for all $n = 0, \dots, N-2$

$$\begin{aligned} A^n_t &\leqslant \frac{4K}{1-\alpha'}\int_t^{t_{n+1}} E[|\delta Y^N_s|^2]ds + \frac{1-\alpha'}{4K}\int_t^{t_{n+1}} E[|\delta f_s|^2]ds + Kh^2 \\ &+ K\int_t^{t_{n+1}} E[|X_s - X^N_{t_{n+1}}|^2]ds + K\int_t^{t_{n+1}} E[|Y_s - Y^N_{t_{n+1}}|^2]ds \\ &+ (1+\frac{1}{\varepsilon})\alpha^2 \int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha' \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z^N_s\|^2]ds \end{aligned}$$

Now using Assumption **(H2)-(i)** in the last inequality, we get

$$\begin{aligned} A^n_t &\leqslant \frac{4K}{1-\alpha'}\int_t^{t_{n+1}} E[|\delta Y^N_s|^2]ds + \frac{1-\alpha'}{4K}K\{h^2 + \int_t^{t_{n+1}} E[|X_s - X^N_{t_n}|^2]ds \\ &+ \int_t^{t_{n+1}} E[|Y_s - Y^N_{t_n}|^2]ds + \int_t^{t_{n+1}} E[||Z_s - Z^N_{t_n}||^2]ds\} \\ &+ Kh^2 + K\int_t^{t_{n+1}} E[|X_s - X^N_{t_{n+1}}|^2]ds + K\int_t^{t_{n+1}} E[|Y_s - Y^N_{t_{n+1}}|^2]ds \\ &+ (1+\frac{1}{\varepsilon})\alpha^2 \int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha' \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z^N_s\|^2]ds. \end{aligned}$$

Then, by plugging $\bar{Z}_{t_n}$ in the last inequality

$$\begin{aligned} A_t^n &\leqslant \frac{4K}{1-\alpha'}\int_t^{t_{n+1}} E[|\delta Y_s^N|^2]ds + \frac{1-\alpha'}{4K}K\{h^2 + \int_t^{t_{n+1}} E[|X_s - X_{t_n}^N|^2]ds \\ &+ \int_t^{t_{n+1}} E[|Y_s - Y_{t_n}^N|^2]ds + 2\int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + 2\int_{t_n}^{t_{n+1}} E[||\delta Z_s^N||^2]ds\} \\ &+ Kh^2 + K\int_t^{t_{n+1}} E[|X_s - X_{t_{n+1}}^N|^2] + K\int_t^{t_{n+1}} E[|Y_s - Y_{t_{n+1}}^N|^2]ds \\ &+ (1+\frac{1}{\varepsilon})\alpha^2\int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha'\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds. \end{aligned}$$

It is well known from Kloeden and Platen [45] that for all $s \in [t_n, t_{n+1})$ and for all $n = 0, \ldots, N-1$

$$E[|X_s - X_{t_n}^N|^2] \leqslant Ch \text{ and } E[|X_s - X_{t_{n+1}}^N|^2] \leqslant Ch, \tag{2.3.14}$$

where C is a positive constant independent of x and depending on K,T, $|b(0)|$ and $\|\sigma(0)\|$.
On the other hand, it is easy to check that $\sup_{t_n \leqslant s \leqslant t_{n+1}}(|Y_s - Y_{t_{n+1}}|^2 + |Y_s - Y_{t_n}|^2) \leqslant Ch(1+|x|^2)$.
This implies that

$$\begin{aligned} E[|Y_s - Y_{t_{n+1}}^N|^2] &\leqslant C\{E[|Y_s - Y_{t_{n+1}}|^2] + E[|Y_{t_{n+1}} - Y_{t_{n+1}}^N|^2]\} \\ &\leqslant C\{h(1+|x|^2) + E[|\delta Y_{t_{n+1}}^N|^2]\} \end{aligned} \tag{2.3.15}$$

and similarly we have

$$E[|Y_s - Y_{t_n}^N|^2] \leqslant C\{h(1+|x|^2) + E[|\delta Y_{t_n}^N|^2]\}, \tag{2.3.16}$$

where C is a positive constant independent of x.
From (2.3.14), (2.3.15) and (2.3.16), we obtain

$$\begin{aligned} A_t^n &\leqslant C\int_t^{t_{n+1}} E[|\delta Y_s^N|^2]ds + ChE[|\delta Y_{t_{n+1}}^N|^2] + ChE[|\delta Y_{t_n}^N|^2] + Ch^2(1+|x|^2) \\ &+ C\int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + (\frac{1-\alpha'}{2})\int_{t_n}^{t_{n+1}} E[||\delta Z_s^N||^2]ds \\ &+ (1+\frac{1}{\varepsilon})\alpha^2\int_t^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha'\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds. \end{aligned} \tag{2.3.17}$$

where C is a generic positive constant depending on α' and independent of x.

From (2.3.8) and (2.3.17), we get

$$\begin{aligned} E[|\delta Y_t^N|^2] &\leqslant A_t^n + E[|\delta Y_{t_{n+1}}^N|^2] \\ &= E[|\delta Y_t^N|^2] + \int_t^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \\ &\leqslant C\int_t^{t_{n+1}} E[|\delta Y_s^N|^2]ds + B_n, \ \forall t \in [t_n, t_{n+1}), \end{aligned} \tag{2.3.18}$$

where we set for all $n = 0, \ldots, N-2$:

$$\begin{aligned} B_n &:= E[|\delta Y_{t_{n+1}}^N|^2] + ChE[|\delta Y_{t_{n+1}}^N|^2] + ChE[|\delta Y_{t_n}^N|^2] + Ch^2(1+|x|^2) \\ &+ C\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + (\frac{1-\alpha'}{2})\int_{t_n}^{t_{n+1}} E[||\delta Z_s^N||^2]ds \\ &+ (1+\frac{1}{\varepsilon})\alpha^2\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha'\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds. \end{aligned} \tag{2.3.19}$$

Using Gronwall Lemma, we have

$$E[|\delta Y_t^N|^2] \leqslant B_n e^{Ch}, \ \forall t \in [t_n, t_{n+1}). \tag{2.3.20}$$

From inequalities (2.3.20) and (2.3.18), we get for h small enough

$$\begin{aligned} E[|\delta Y_t^N|^2] + \int_t^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds &\leqslant (1 + Che^{Ch})B_n \\ &\leqslant (1 + Ch)B_n, \ \forall t \in [t_n, t_{n+1}). \end{aligned} \tag{2.3.21}$$

By taking $t = t_n$ in the last inequality, we obtain

$$\begin{aligned} E[|\delta Y_{t_n}^N|^2] + \int_{t_n}^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \leqslant (1 + Ch)\Big\{ E[|\delta Y_{t_{n+1}}^N|^2] + ChE[|\delta Y_{t_n}^N|^2] \\ + \ Ch^2(1 + |x|^2) + C\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + (\frac{1-\alpha'}{2})\int_{t_n}^{t_{n+1}} E[||\delta Z_s^N||^2]ds \\ + \ (1 + \frac{1}{\varepsilon})\alpha^2 \int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + \alpha' \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds \Big\}. \end{aligned}$$

Then

$$\begin{aligned} (1 - Ch)E[|\delta Y_{t_n}^N|^2] + [1 - (1 + Ch)\frac{1-\alpha'}{2}] \int_{t_n}^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \leqslant (1 + Ch)\Big\{ E[|\delta Y_{t_{n+1}}^N|^2] \\ + \ Ch^2(1 + |x|^2) + C\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + (1 + \frac{1}{\varepsilon})\alpha^2 \int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds \\ + \ \alpha' \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds \Big\}. \end{aligned}$$

For h small enough, we get

$$\begin{aligned} E[|\delta Y_{t_n}^N|^2] + \frac{1+\alpha'}{2}\int_{t_n}^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \leqslant (1 + Ch)\Big\{ E[|\delta Y_{t_{n+1}}^N|^2] + \alpha' \int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds \\ + \ Ch^2(1 + |x|^2) + C\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + (1 + \frac{1}{\varepsilon})\alpha^2 \int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds \Big\}. \end{aligned} \tag{2.3.22}$$

Iterating the last inequality, we obtain for all $n = 0, \ldots, N-1$

$$\begin{aligned} E[|\delta Y_{t_n}^N|^2] + \frac{1+\alpha'}{2}\int_{t_n}^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \leqslant (1 + Ch)^{N-1}\Big\{ E[|\delta Y_T^N|^2] + Ch(1 + |x|^2) \\ + \ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds \\ + \ \alpha^2 \int_{t_{N-1}}^{t_N} E[||Z_s||^2]ds \Big\}. \end{aligned}$$

Using the Assumption **(H2)-(iii)**, we get

$$\begin{aligned} E[|\delta Y_{t_n}^N|^2] + \frac{1+\alpha'}{2}\int_{t_n}^{t_{n+1}} E[\|\delta Z_s^N\|^2]ds \leqslant (1 + Ch)^{N-1}\Big\{ Ch(1 + |x|^2) \\ + \ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds \\ + \ \alpha^2 \int_{t_{N-1}}^{t_N} E[||Z_s||^2]ds \Big\}. \end{aligned} \tag{2.3.23}$$

Now we sum up inequality (2.3.22) over n, we get

$$\begin{aligned}
&\sum_{n=0}^{N-1} E[|\delta Y_{t_n}^N|^2] + \frac{1+\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant (1+Ch)\Big\{\sum_{n=0}^{N-1} E[|\delta Y_{t_{n+1}}^N|^2] \\
+ & \sum_{n=0}^{N-1} Ch^2(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds \\
+ & \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]\Big\} + (1+Ch)\alpha' \sum_{n=0}^{N-2}\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds.
\end{aligned}$$

Using that $Nh = T$ and $\sum_{n=0}^{N-2}\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds = \int_{t_1}^T E[\|\delta Z_s^N\|^2]ds$, we get

$$\begin{aligned}
&\sum_{n=0}^{N-1} E[|\delta Y_{t_n}^N|^2] + \frac{1+\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant (1+Ch)\Big\{\sum_{n=0}^{N-1} E[|\delta Y_{t_{n+1}}^N|^2] \\
+ & Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds \\
+ & \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]\Big\} + (1+Ch)\alpha' \int_0^T E[\|\delta Z_s^N\|^2]ds.
\end{aligned}$$

Then

$$\begin{aligned}
&\sum_{n=0}^{N-1} E[|\delta Y_{t_n}^N|^2] + \Big[\frac{1+\alpha'}{2} - (1+Ch)\alpha'\Big]\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant (1+Ch)\Big\{\sum_{n=0}^{N-1} E[|\delta Y_{t_{n+1}}^N|^2] \\
+ & Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds \\
+ & \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]\Big\}.
\end{aligned}$$

We obtain for h small enough

$$\begin{aligned}
&\sum_{n=0}^{N-1} E[|\delta Y_{t_n}^N|^2] + \frac{1-\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant (1+Ch)\Big\{\sum_{n=0}^{N-1} E[|\delta Y_{t_{n+1}}^N|^2] \\
+ & Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds \\
+ & \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]\Big\}.
\end{aligned}$$

Hence, we get

$$\begin{aligned}
&\frac{1-\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant (1+Ch)E[|\delta Y_T^N|^2] + [(1+Ch)-1]\sum_{n=1}^{N-1} E[|\delta Y_{t_n}^N|^2] \\
- & E[|\delta Y_{t_0}^N|^2] + Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_n}||^2]ds \\
+ & C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds + \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big].
\end{aligned}$$

Using Assumption **(H2)-(iii)** on Φ, we get

$$\frac{1-\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant Ch(1+|x|^2) + Ch\sum_{n=1}^{N-1} E[|\delta Y_{t_n}^N|^2]$$
$$+ \quad C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds$$
$$+ \quad \alpha^2 E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]. \tag{2.3.24}$$

Summing up (2.3.23) over n, we have

$$h\sum_{n=0}^{N-1} E[|\delta Y_{t_n}^N|^2 \leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds$$
$$+ \quad C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + C\int_{t_{N-1}}^{t_N} E[||Z_s||^2]ds.$$

Plugging the last inequality in (2.3.24), we obtain

$$\frac{1-\alpha'}{2}\int_0^T E[\|\delta Z_s^N\|^2]ds \leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds$$
$$+ \quad C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]. \tag{2.3.25}$$

Now, turning Back to equation (2.3.21), we have for all $n = 0, \ldots, N-2$

$$\begin{aligned}
E[|\delta Y_t^N|^2] &\leqslant (1+Ch)B_n \\
&\leqslant (1+Ch)\Big\{E[|\delta Y_{t_{n+1}}^N|^2] + \alpha'\int_{t_{n+1}}^{t_{n+2}} E[\|\delta Z_s^N\|^2]ds \\
&+ ChE[|\delta Y_{t_n}^N|^2] + (\frac{1-\alpha'}{2})\int_{t_n}^{t_{n+1}} E[||\delta Z_s^N||^2]ds \\
&+ Ch^2(1+|x|^2) + C\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds \\
&+ (1+\frac{1}{\varepsilon})\alpha^2\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds\Big\}, \forall t \in [t_n, t_{n+1}).
\end{aligned}$$

Using inequality (2.3.23), we get

$$\begin{aligned}
E[|\delta Y_t^N|^2] &\leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds \\
&+ C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big].
\end{aligned}$$

By taking the supremum over t in the last inequality, we obtain

$$\sup_{0\leqslant t\leqslant T} E[|\delta Y_t^N|^2] \leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_n}||^2]ds$$
$$+ \quad C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]. \tag{2.3.26}$$

Equations (2.3.26) and (2.3.25) give together

$$\sup_{0\leqslant t\leqslant T}E[|\delta Y_t^N|^2]+\int_0^T E[\|\delta Z_s^N\|^2]ds\leqslant Ch(1+|x|^2)+C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_n}||^2]ds$$

$$+C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds+CE\Big[\int_{t_{N-1}}^{t_N}||Z_s||^2ds\Big]. \qquad (2.3.27)$$

Plugging $\bar{Z}_{t_n}$, we deduce from Lemma 2.3.1 that

$$E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}||Z_s-Z_{t_n}^N||^2ds] \leqslant CE[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}||Z_s-\bar{Z}_{t_n}||^2ds] + CE[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}||\delta Z_s^N||^2ds].$$

Using the last inequality in (2.3.27), we get

$$\sup_{0\leqslant t\leqslant T}E[|\delta Y_t^N|^2]+E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}||Z_s-Z_{t_n}^N||^2ds]\leqslant Ch(1+|x|^2)$$

$$+\ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_n}||^2]ds+C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds+CE\Big[\int_{t_{N-1}}^{t_N}||Z_s||^2ds\Big]$$

which can be written, if we set $\bar{Z}_{t_N}:=0$

$$\sup_{0\leqslant t\leqslant T}E[|\delta Y_t^N|^2]+E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}||Z_s-Z_{t_n}^N||^2ds]\leqslant Ch(1+|x|^2)$$

$$+\ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_n}||^2]ds+C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}E[||Z_s-\bar{Z}_{t_{n+1}}||^2]ds. \qquad (2.3.28)$$

□

Remark 2.3.1. *BDSDEs theory works well in the L^2 framework on the full probability space (Existence and uniqueness for the theorotical solution). Thus, since the time discretization error is studied on the full space (under the expectation), the analysis works in an analogous way to the BSDEs case. We need the Itô Lemma for BDSDEs and to use the contraction condition in Assumption* ***(H2)****. After that, we need to prove Zhang L^2-regularity results for this kind of equations. However, for the regression error analysis, we need almost sure estimates for the solution. We send the reader to chapter 4 for more details on this question.*

2.4 Path regularity of the process Z

The purpose of this section is to prove L^2-regularity of the Z component of the solution of the BDSDE (2.1.2). Such result is crucial to obtain the rate of convergence of our numerical scheme. For this end, we need to introduce the Malliavin derivatives of the solution. This will allow us to provide representation and regularity results for Y and Z that will immediately imply the rate of our scheme.

We recall that the tools on the Malliavin calculus in the context of BDSDEs were introduced in

Pardoux and Peng [63]. Pardoux and Peng have skipped details of this part considering that it is just a natural extension of the works on standard BSDEs (The reader can see [62, 31, 64] for the BSDEs case). For the sake of completeness, we give some details which are crucial to obtain regularity result of the process Z and we give some technical proofs in the Appendix.

2.4.1 Malliavin calculus on the Forward SDE's

In this section, we recall some properties on the differentiability in the Malliavin sense of the forward process $(X_s^{t,x})$. Under **(H3(i))**, Nualart [58] stated that $X_s^{t,x} \in \mathbb{D}^{1,2}$ for any $s \in [t,T]$ and for $l \leqslant k$ the derivative $D_r^l X_s^{t,x}$ is given by :
(i) $D_r^l X_s^{t,x} = 0, \quad for \quad s < r \leqslant T,$
(ii) For any $t < r \leqslant T$, a version of $\{D_r^l X_s^{t,x}, r \leqslant s \leqslant T\}$ is the unique solution of the linear SDE

$$D_r^l X_s^{t,x} = \sigma^l(X_r^{t,x}) + \int_r^s \nabla b(X_u^{t,x}) D_r^l X_u^{t,x} du + \sum_{i=1}^d \int_r^s \nabla \sigma^i(X_u^{t,x}) D_r^l X_u^{t,x} dW_u^i,$$

where $(\sigma^i)_{i=1,\ldots,k}$ denotes the i-th column of the matrix σ.
Moreover, $D_r^l X_s^{t,x} \in \mathbb{D}^{1,2}$ for all $r, s \leqslant T$. For all $v \leqslant T$ and $l' \leqslant k$, we have

$$D_v^{l'} D_r^l X_s^{t,x} = 0 \text{ if } s < v \vee r,$$

and for all $s \geqslant v \vee r$ a version of $D_v^{l'} D_r^l X_s^{t,x}$ is the unique solution of the SDE :

$$\begin{aligned} D_v^{l'} D_r^l X_s^{t,x} &= \nabla \sigma^l(X_r^{t,x}) D_v^{l'} X_r^{t,x} + \sum_{i=1}^d \nabla \sigma^i(X_v^{t,x}) D_r^l X_v^{t,x} 1_{\{t \leqslant v \leqslant s\}} \\ &+ \int_r^s \Big[\sum_{j=1}^k \nabla((\nabla b)^j(X_u^{t,x})) D_v^{l'} X_u^{t,x} (D_r^l X_u^{t,x})^j + \nabla b(X_u^{t,x}) D_v^{l'} D_r^l X_u^{t,x}\Big] du \\ &+ \sum_{i=1}^d \int_r^s \Big[\sum_{j=1}^k \nabla(\nabla \sigma^i(X_u^{t,x}))^j D_v^{l'} X_u^{t,x} (D_r^l X_u^{t,x})^j + \nabla \sigma^i(X_u^{t,x}) D_v^{l'} D_r^l X_u^{t,x}\Big] dW_u^i, \end{aligned}$$

where $((\nabla b)^j)_{j=1,\ldots,k}$ (resp.$((\nabla \sigma^i(X_u^{t,x}))^j)_{j=1,\ldots,k}$) denotes the j-th column of the matrix (∇b) (resp. $(\nabla \sigma^i(X_u^{t,x}))$) and $((D_r^l X_u^{t,x})^j)_{j=1,\ldots,k}$ denotes the j-th component of the vector $(D_r^l X_u^{t,x})$. The following inequalities will be useful later. For the proofs, we refer to Nualart [58] for example. From Lemma 2.7 in [58] applied to X and $D_s X$, there exists a positive constant C_p, depending on p, such that : for all $0 \leqslant r \leqslant s \leqslant T$, we have the following inequalities

$$E\Big[\sup\nolimits_{0 \leqslant u \leqslant T} ||D_s X_u||^p\Big] \leqslant C_p(1 + |x|^p), \tag{2.4.1}$$

$$E\Big[\sup\nolimits_{s \vee r \leqslant u \leqslant T} ||D_s X_u - D_r X_u||^p\Big] \leqslant C_p |s - r|(1 + |x|^p), \tag{2.4.2}$$

The same argument used for $D_r D_s X$ shows that there exists $C_p > 0$ such that

$$E\Big[\sup\nolimits_{0 \leqslant u \leqslant T} ||D_r D_s X_u||^p\Big] \leqslant C_p(1 + |x|^{2p}). \tag{2.4.3}$$

2.4.2 Malliavin calculus for the solution of BDSDE's

Now, our aim is to study the differentiability in the Malliavin sense of the solution of the BDSDE (2.2.2). We start with the following lemma which shows that a backward Itô integral is differentiable in the Malliavin sense if and only if its integrand is so. We recall that Pardoux and Peng [62] proved that the result holds for the classical Itô integral.

Lemma 2.4.1. *Let $U \in \mathbb{H}_1^2([t,T])$ and $I_i(U) = \int_t^T U_r dW_r^i, i = 1, \ldots, d$. Then, for each $\theta \in [0,T]$ we have $U_\theta \in \mathbb{D}^{1,2}$. if and only if $I_i(U) \in \mathbb{D}^{1,2}, i = 1, \ldots, d$ and for all $\theta \in [0,T]$, we have*

$$
\begin{aligned}
D_\theta I_i(U) &= \int_\theta^T D_\theta U_r dW_r^i + U_\theta, \ \theta > t, \\
D_\theta I_i(U) &= \int_t^T D_\theta U_r dW_r^i, \ \theta \leqslant t.
\end{aligned}
$$

For backward Itô integral, and since the Malliavin derivative is with respect to the brownian motion W, we have the following result :

Lemma 2.4.2. *Let $U \in \mathbb{H}_1^2([t,T])$ and $I_i(U) = \int_t^T U_r \overleftarrow{dB_r^i}, i = 1, \ldots, l$. Then for each $\theta \in [0,T]$ we have $U_\theta \in \mathbb{D}^{1,2}$ if and only if $I_i(U) \in \mathbb{D}^{1,2}, i = 1, \ldots, l$ and for all $\theta \in [0,T]$, we have*

$$
\begin{aligned}
D_\theta I_i(U) &= \int_\theta^T D_\theta U_r \overleftarrow{dB_r^i}, \ \theta > t, \\
D_\theta I_i(U) &= \int_t^T D_\theta U_r \overleftarrow{dB_r^i}, \ \theta \leqslant t.
\end{aligned}
$$

For later use, we need to prove the a priori estimates for the solution of the BDSDE (see [31] for similar estimates for a standard BSDE).

Proposition 2.4.1. *Let (ϕ^1, f^1, g^1) and (ϕ^2, f^2, g^2) be two standard parameters of the BDSDE (2.2.2) and (Y^1, Z^1) and (Y^2, Z^2) the associated solutions. Assume that Assumption* **(H2)** *holds. For $s \in [t,T]$, set $\delta Y_s := Y_s^1 - Y_s^2$, $\delta_2 f_s := f^1(s, X_s, Y_s^2, Z_s^2) - f^2(s, X_s, Y_s^2, Z_s^2)$ and $\delta_2 g_s := g^1(s, X_s, Y_s^2, Z_s^2) - g^2(s, X_s, Y_s^2, Z_s^2)$. Then, we have*

$$
||\delta Y||^2_{\mathbb{S}_d^2([t,T])} + ||\delta Z||^2_{\mathbb{H}^2_{d\times k}([t,T])} \leqslant CE[|\delta Y_T|^2 + \int_t^T |\delta_2 f_s|^2 ds + \int_t^T \|\delta_2 g_s\|^2 ds], \tag{2.4.4}
$$

where C is a positive constant depending only on K, T and α.

Proof. Using the same argument as in the classical BSDE's setting, one can prove this stability result for BDSDEs (see El Karoui et al.[31] for the BSDE's case).

□

Now, we study the differentiability in the Malliavin sense of the solution of the BDSDE which is technical. To our knowledge, it does not exist in the literature. We have to precise that Pardoux and Peng [63] have skipped details considering that it was just an easy extension of the work on standard BSDEs [62]. We show that the derivative is a solution of a linear BDSDE as Pardoux and Peng [62] did for standard BSDEs, see also El Karoui Peng and Quenez ([31], Proposition 5.3)).

Proposition 2.4.2. *Assume that* **(H1)-(H3)** *hold. For any $t \in [0,T]$ and $x \in \mathbb{R}^d$, let $\{(Y_s, Z_s), t \leqslant s \leqslant T\}$ denotes the unique solution of the BDSDE :*

$$
Y_s = \Phi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r, Z_r) dr + \int_s^T g(r, X_r^{t,x}, Y_r, Z_r) \overleftarrow{dB_r} - \int_s^T Z_r dW_r, \ t \leqslant s \leqslant T.
$$

Then, $(Y,Z) \in \mathcal{B}^2([t,T], \mathbb{D}^{1,2})$ and $\{D_\theta Y_s, D_\theta Z_s; t \leqslant s, \theta \leqslant T\}$ is given by :
(i) $D_\theta Y_s = 0, D_\theta Z_s = 0$ for all $t \leqslant s < \theta \leqslant T$

(ii) for any fixed $\theta \in [t,T]$, $\theta \leqslant s \leqslant T$ and $1 \leqslant i \leqslant d$, a version of $(D^i_\theta Y_s, D^i_\theta Z_s)$ is the unique solution of the BDSDE :

$$\begin{aligned} D^i_\theta Y_s &= \nabla\Phi(X^{t,x}_T)D^i_\theta X^{t,x}_T + \int_s^T \Big(\nabla_x f(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta X^{t,x}_r\Big)dr \\ &+ \int_s^T \Big(\nabla_y f(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta Y_r + \sum_{j=1}^d \nabla_{z^j} f(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta Z^j_r\Big)dr \\ &+ \sum_{n=1}^l \int_s^T \Big(\nabla_x g^n(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta X^{t,x}_r + \nabla_y g^n(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta Y_r\Big)\overleftarrow{dB^n_r} \\ &+ \sum_{n=1}^l \int_s^T \sum_{j=1}^d \Big(\nabla_{z^j} g^n(r,X^{t,x}_r,Y_r,Z_r)D^i_\theta Z^j_r\Big)\overleftarrow{dB^n_r} - \int_s^T \sum_{j=1}^d D^i_\theta Z^j_r dW^j_r, \end{aligned} \tag{2.4.5}$$

where $(z^j)_{1 \leqslant j \leqslant d}$ denotes the j-th column of the matrix z, $(g^n)_{1 \leqslant n \leqslant l}$ denotes the n-th column of the matrix g and $B = (B^1, \ldots, B^l)$.

Proof. See Appendix.

□

The second order differentiability in the Malliavin sense of the solution of the BDSDE will be given in Appendix.

2.4.3 Representation results for BDSDEs

In this subsection, we will prove a representation result of (Z, DZ) which will be useful to prove the rate of convergence of our numerical scheme.

Proposition 2.4.3. *Assume that* **(H1)**-**(H3)** *hold. Then : For $t \leqslant s \leqslant T$, we have*

$$D_s Y_s = Z_s, \tag{2.4.6}$$

and

$$\|Z\|^2_{\mathbb{S}^2_{k\times d}([t,T])} \leqslant C(1+|x|^2). \tag{2.4.7}$$

For $l_1, l_2 \leqslant d$, $t \leqslant s \leqslant T$, we have

$$D^{l_2}_s D^{l_1}_t Y_s = D^{l_2}_t Z^{l_1}_s, \tag{2.4.8}$$

and

$$\|D^{l_1}_s Z\|^2_{\mathbb{S}^2_{k\times d}([t,T])} \leqslant C(1+|x|^4). \tag{2.4.9}$$

Proof. To simplify the notations, we restrict ourselves to the case $k = d = 1$.
1. Notice that for $t \leqslant s$

$$Y_s = Y_t - \int_t^s f(r,\Sigma_r)dr - \int_t^s g(r,\Sigma_r)\overleftarrow{dB_r} + \int_t^s Z_r dW_r,$$

where $\Sigma_r := (X^{t,x}_r, Y_r, Z_r)$.
It follows from Lemma 2.4.1 and Lemma 2.4.2 that, for $t < \theta \leqslant s$

$$\begin{aligned} D_\theta Y_s &= Z_\theta - \int_\theta^s \Big(\nabla_x f(r,\Sigma_r)D_\theta X_r + \nabla_y f(r,\Sigma_r)D_\theta Y_r + \nabla_z f(r,\Sigma_r)D_\theta Z_r\Big)dr \\ &- \int_\theta^s \Big(\nabla_x g(r,\Sigma_r)D_\theta X_r + \nabla_y g(r,\Sigma_r)D_\theta Y_r + \nabla_z g(r,\Sigma_r)D_\theta Z_r\Big)\overleftarrow{dB_r} + \int_\theta^s D_\theta Z_r dW_r. \end{aligned}$$

Then by taking $\theta = s$, it follows that equality (2.4.6) holds. From (2.8.1), we deduce that (2.4.7) holds.
2. Notice that for $\theta \leqslant t \leqslant s$

$$\begin{aligned} D_\theta Y_s &= D_\theta Y_t - \int_t^s \Big(\nabla_x f(r,\Sigma_r) D_\theta X_r + \nabla_y f(r,\Sigma_r) D_\theta Y_r + \nabla_z f(r,\Sigma_r) D_\theta Z_r\Big) dr \\ &- \int_t^s \Big(\nabla_x g(r,\Sigma_r) D_\theta X_r + \nabla_y g(r,\Sigma_r) D_\theta Y_r + \nabla_z g(r,\Sigma_r) D_\theta Z_r\Big) \overleftarrow{dB_r} + \int_t^s D_\theta Z_r dW_r. \end{aligned}$$

It follows from Lemma 2.4.1 and Lemma 2.4.2 that, for $\theta \leqslant t < v \leqslant s$

$$\begin{aligned} D_v D_\theta Y_s &= D_\theta Z_v - \int_v^s D_v(\Sigma_r)^*[Hf](r,\Sigma_r) D_\theta(\Sigma_r) dr - \int_v^s \nabla f(r,\Sigma_r) D_v D_\theta(\Sigma_r) dr \\ &- \int_v^s D_v(\Sigma_r)^*[Hg](r,\Sigma_r) D_\theta(\Sigma_r) \overleftarrow{dB_r} - \int_v^s \nabla g(r,\Sigma_r) D_v D_\theta(\Sigma_r) \overleftarrow{dB_r} \\ &+ \int_v^s D_v D_\theta Z_r dW_r. \end{aligned}$$

Then by taking $v = s$ and $t = \theta$, it follows that equality (2.4.8) holds. We have from estimate (2.2.4) and inequality (2.4.3), that for each $v \leqslant T$ and $\theta \leqslant T$

$$E[\sup_{t \leqslant s \leqslant T} |D_v D_\theta Y_s|^2] + E[\int_t^T |D_v D_\theta Z_s|^2 ds] \leqslant C(1+|x|^4). \tag{2.4.10}$$

and then by taking $v = s$ and $t = \theta$ we deduce that (2.4.9) holds.

□

2.4.4 Zhang L^2-regularity

In this subsection, we extend the result of Zhang [70] which concerns the L^2-regularity of the martingale integrand Z. Such result is crucial to derive the rate of convergence of our numerical scheme. We start with the following proposition which gives an upper bound for

$$E\Big[\sup_{r\in[s,u]} |Y_r - Y_s|^2\Big] \quad \text{and} \quad E\Big[||Z_u - Z_s||^2\Big], \quad t \leqslant s \leqslant u \leqslant T.$$

Proposition 2.4.4. *Assume that* **(H1)-(H3)** *hold. Then for $t \leqslant s \leqslant u \leqslant T$, we have*

$$E\Big[\sup_{r\in[s,u]} |Y_r - Y_s|^2\Big] \leqslant C(1+|x|^2)|u-s|, \tag{2.4.11}$$

$$E\Big[||Z_u - Z_s||^2\Big] \leqslant C(1+|x|^2)|u-s|. \tag{2.4.12}$$

Proof. To simplify the notations, we restrict ourselves to the case $k = d = l = 1$.
(i) Plugging inequality (2.4.7) in the estimate (2.8.12), the result (2.4.11) holds.
(ii) From Proposition 2.4.3, we have

$$E\Big[|Z_u - Z_s|^2\Big] \leqslant CE[|D_u Y_u - D_s Y_u|^2] + CE[|D_s Y_u - D_s Y_s|^2]. \tag{2.4.13}$$

From the definition of the BDSDE (2.4.5), we have

$$D_uY_u - D_sY_u = \nabla\Phi(X_T)(D_uX_T - D_sX_T) + \int_u^T \Big(\nabla_x f(r,\Sigma_r)(D_uX_r - D_sX_r)\Big)dr$$
$$+ \int_u^T \Big(\nabla_y f(r,\Sigma_r)(D_uY_r - D_sY_r) + \nabla_z f(r,\Sigma_r)(D_uZ_r - D_tZ_r)\Big)dr$$
$$+ \int_u^T \Big(\nabla_x g(r,\Sigma_r)(D_uX_r - D_sX_r) + \nabla_y g(r,\Sigma_r)(D_uY_r - D_sY_r)\Big)\overleftarrow{dB_r}$$
$$+ \int_u^T \Big(\nabla_z g(r,\Sigma_r)(D_uZ_r - D_sZ_r)\Big)\overleftarrow{dB_r} - \int_u^T (D_uZ_r - D_sZ_r)dW_r.$$

Applying the generalized Itô's formula (see [63], Lemma 1.3), we obtain

$$|D_uY_T - D_sY_T|^2 - |D_uY_u - D_sY_u|^2 =$$
$$- 2\int_u^T \nabla_x f(r,\Sigma_r)(D_uX_r - D_sX_r)(D_uY_r - D_sY_r)dr - 2\int_u^T \nabla_y f(r,\Sigma_r)(D_uY_r - D_sY_r)^2dr$$
$$- 2\int_u^T \nabla_z f(r,\Sigma_r)(D_uZ_r - D_uZ_r)(D_uY_r - D_sY_r)dr$$
$$- 2\int_u^T \nabla_x g(r,\Sigma_r)(D_uX_r - D_sX_r)(D_uY_r - D_sY_r)\overleftarrow{dB_r}$$
$$- 2\int_u^T \nabla_y g(r,\Sigma_r)(D_uY_r - D_sY_r)^2\overleftarrow{dB_r}$$
$$- 2\int_u^T \nabla_z g(r,\Sigma_r)(D_uZ_r - D_sZ_r)(D_uY_r - D_sY_r)\overleftarrow{dB_r}$$
$$+ 2\int_u^T (D_uZ_r - D_sZ_r)(D_uY_r - D_sY_r)dW_r$$
$$- \int_u^T \Big|\nabla_x g(r,\Sigma_r)(D_uX_r - D_sX_r) + \nabla_y g(r,\Sigma_r)(D_uY_r - D_sY_r) + \nabla_z g(r,\Sigma_r)(D_uZ_r - D_sZ_r)\Big|^2dr$$
$$+ \int_u^T |D_uZ_r - D_sZ_r|^2dr.$$

From inequalities (2.8.1) and (2.4.1), using the Burkholder-Davis-Gundy's inequality and Assumption **(H2)**, the stochastic integrals which appear in the last equation disappear when we take the expectation. By Young inequality, we obtain, for $\varepsilon' > 0$

$$E[|D_uY_u - D_sY_u|^2] + E[\int_u^T |D_uZ_r - D_sZ_r|^2]dr \leqslant E[|\nabla\Phi(X_T)(D_uX_T - D_sX_T)|^2]$$
$$+ 2E[\int_u^T \nabla_x f(r,\Sigma_r)(D_uX_r - D_sX_r)(D_uY_r - D_sY_r)dr]$$
$$+ 2E[\int_u^T \nabla_y f(r,\Sigma_r)(D_uY_r - D_sY_r)^2dr]$$
$$+ 2E[\int_u^T \nabla_z f(r,\Sigma_r)(D_uZ_r - D_uZ_r)(D_uY_r - D_sY_r)dr]$$
$$+ C(1+\frac{1}{\varepsilon'})E[\int_u^T \nabla_x g(r,\Sigma_r)^2|D_uX_r - D_sX_r|^2dr]$$
$$+ C(1+\frac{1}{\varepsilon'})E[\int_u^T \nabla_y g(r,\Sigma_r)^2|D_uY_r - D_sY_r|^2dr]$$
$$+ (1+\varepsilon')E[\int_u^T \nabla_z g(r,\Sigma_r)^2|D_uZ_r - D_sZ_r|^2dr].$$

Hence by using Assumption **(H2)** and Young inequality, we have for $\varepsilon, \varepsilon' > 0$ and $C > 0$,

$$\begin{aligned} & E[|D_uY_u - D_sY_u|^2] + E[\int_u^T |D_uZ_r - D_sZ_r|^2 dr] \leqslant K^2 E[|D_uX_T - D_sX_T|^2] \\ + & 2KE[\int_u^T |D_uX_r - D_sX_r|^2 dr] + 4KE[\int_u^T |D_uY_r - D_sY_r|^2 dr] \\ + & K\varepsilon E[\int_u^T |D_uY_r - D_sY_r|^2 dr] + \frac{K}{\varepsilon} E[\int_u^T |D_uZ_r - D_sZ_r|^2 dr] \\ + & CK^2(1+\frac{1}{\varepsilon'})E[\int_u^T |D_uX_r - D_sX_r|^2 dr] + CK^2(1+\frac{1}{\varepsilon'})E[\int_u^T |D_uY_r - D_sY_r|^2 dr] \\ + & (1+\varepsilon')\alpha^2 E[\int_u^T |D_uZ_r - D_sZ_r|^2 dr]. \end{aligned}$$

Then, we obtain

$$\begin{aligned} & E[|D_uY_u - D_sY_u|^2] + E[\int_u^T |D_uZ_r - D_sZ_r|^2 dr] \leqslant K^2 E[|D_uX_T - D_sX_T|^2] \\ + & K(2 + KC(1+\frac{1}{\varepsilon'}))E[\int_u^T |D_uX_r - D_sX_r|^2 dr] \\ + & (K^2C(1+\frac{1}{\varepsilon'}) + (4+\varepsilon)K)E[\int_u^T |D_uY_r - D_sY_r|^2 dr] \\ + & ((1+\varepsilon')\alpha^2 + \frac{K}{\varepsilon})E[\int_u^T |D_uZ_r - D_sZ_r|^2 dr]. \end{aligned}$$

For ε large enough and ε' small enough, we have $(1+\varepsilon')\alpha^2 + \frac{K}{\varepsilon} < 1$. From inequality (2.4.2), we deduce that

$$E[|D_uY_u - D_sY_u|^2] \leqslant C\Big((1+|x|^2)|u-s| + E[\int_u^T |D_uY_r - D_sY_r|^2 dr]\Big),$$

where C is a positive constant. From Gronwall's lemma we have

$$E[|D_uY_u - D_sY_u|^2] \leqslant C(1+|x|^2)|u-s|. \tag{2.4.14}$$

Since $(D_sY_u)_{s \leqslant u \leqslant T}$ satisfies the BDSDE (2.4.5), inequalities (2.8.12)-(2.4.7) hold for $(D_sY_u, D_sZ_u)_{s \leqslant u \leqslant T}$ and yield

$$E[|D_sY_u - D_sY_s|^2] \leqslant C(1+|x|^2)|u-s|. \tag{2.4.15}$$

Plugging (2.4.14) and (2.4.15) into (2.4.13), we obtain (2.4.12).

□

The following theorem states the rate of convergence of our numerical scheme.

Theorem 2.4.1. *Under Assumptions* **(H1)-(H3)**, *there exists a positive constant C (depending only on T, K, α, $|b(0)|$, $||\sigma(0)||$, $|f(t,0,0,0)|$ and $||g(t,0,0,0)||$) such that*

$$Error_N(Y,Z) \leqslant Ch(1+|x|^2). \tag{2.4.16}$$

Proof. From the definition (3.3.2), $\bar{Z}_{t_n}$ is the best approximation of $(Z_t)_{t_n \leqslant t < t_{n+1}}$ by $\mathcal{F}_{t_n}$-measurable random variable in the following sense

$$E\Big[\int_{t_n}^{t_{n+1}} \|Z_s - \bar{Z}_{t_n}\|^2 ds\Big] = \inf_{Z_n \in L^2(\Omega, \mathcal{F}_{t_n})} E\Big[\int_{t_n}^{t_{n+1}} \|Z_s - Z_n\|^2 ds\Big]$$

which implies

$$E[\|Z_s - \bar{Z}_{t_n}\|^2] \leqslant E[\|Z_s - Z_{t_n}\|^2].$$

From Proposition 2.4.4, we have

$$E\Big[||Z_s - Z_{t_n}||^2\Big] \leqslant C(1+|x|^2)|s-t_n| \leqslant Ch(1+|x|^2),$$

for all $s \in [t_n, t_{n+1}]$ and $0 \leqslant n \leqslant N-1$ where C depends only on T, K, $b(0)$, $\sigma(0)$, $f(t,0,0,0)$ and $g(t,0,0,0)$. Then

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} ||Z_s - \bar{Z}_{t_n}||^2 ds\Big] \leqslant Ch(1+|x|^2).$$

Similarly, from Proposition 2.4.4 we get

$$\sum_{n=0}^{N-2} E\Big[\int_{t_n}^{t_{n+1}} ||Z_s - \bar{Z}_{t_{n+1}}||^2 ds\Big] \leqslant Ch(1+|x|^2).$$

Finally, using the same argument again, we obtain

$$E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big] \leqslant Ch(1+|x|^2).$$

Then, from Theorem 2.3.1

$$Error_N(Y,Z) \leqslant Ch(1+|x|^2).$$

□

Remark 2.4.1. *One could define the error as follows*

$$\widetilde{Error}_N(Y,Z) := \sup_{0 \leqslant s \leqslant T} E[|Y_s - Y_s^N|^2] + \sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} ||Z_s - Z_s^N||^2 ds]. \qquad (2.4.17)$$

Then, we have

$$\widetilde{Error}_N(Y,Z) \leqslant Ch(1+|x|^2). \qquad (2.4.18)$$

2.5 Numerical scheme for the weak solution of the SPDE

Most numerical work on SPDEs has concentrated on the Euler finite-difference scheme (see [38], [39] , [37]), on finite element method (see [69]) and also on spectral Galerkin methods (see [44] and the references therein). Here, we follow a probabilistic method based on the Feynman-Kac's formula for the weak solution of the semilinear SPDE's (2.1.1) based on BSDE's approach (see [9], [53]). We consider a weak Sobolev solution of such SPDE in the sense that u shall be considered as a predictable process in some first order Sobolev space. Therefore, we shall improve the convergence and the rate of convergence of the L^2-norm error of such solution by using the convergence results on BDSDEs proved in section 4 and an equivalence norm result given in Barles and Lesigne [10] and Bally and Matoussi [9].

2.5.1 Weak solution for SPDE

Since we work on the whole space $\mathbb{R}^d$, we introduce a weight function ρ satisfying the following conditions : ρ is a positive locally integrable function , $\frac{1}{\rho}$ are locally integrable and $\int_{\mathbb{R}^d}(1+|x|^2)\rho(x)dx < \infty$. For example, we can take $\rho(x) = e^{-\frac{x^2}{2}}$ or $\rho(x) = e^{-|x|}$. As a consequence of (**H3**), we have $\displaystyle\int_{\mathbb{R}^d} |\Phi(x)|^2\rho(x)dx < \infty$, $\int_0^T\int_{\mathbb{R}^d} |f(t,x,0,0)|^2\rho(x)dxdt < \infty$ and $\int_0^T\int_{\mathbb{R}^d} |g(t,x,0,0)|^2\rho(x)dxdt < \infty$.

We denote by $L^2(\mathbb{R}^d,\rho(x)dx)$ the weighted Hilbert space and we employ the following notation for its scalar product and its norm : $(u,v)_\rho = \int_{\mathbb{R}^d} u(x)v(x)\rho(x)dx$ and $\|u\|_\rho = (u,u)_\rho^{\frac{1}{2}}$. Then, we define by $\mathrm{H}^1_\sigma(\mathbb{R}^d)$ the associated weighted first order Dirichlet space and its norm $\|u\|_{\mathrm{H}^1_\sigma(\mathbb{R}^d)} = (\|u\|^2_\rho + \|\nabla u\sigma\|^2_\rho)^{\frac{1}{2}}$. Finally, $(.,.)$ denotes the usual scalar product in $L^2(\mathbb{R}^d,dx)$.
We define also $\mathcal{D} := \mathcal{C}_c^\infty([0,T])\otimes\mathcal{C}_c^2(\mathbb{R}^d)$ the space of test functions where $\mathcal{C}_c^\infty([0,T])$ denotes the space of all real valued infinite differentiable functions with compact support in $[0,T]$ and $\mathcal{C}_c^2(\mathbb{R}^d)$ the set of C^2-functions with compact support in $\mathbb{R}^d$.
We introduce $\mathcal{H}_T$ the space of predictable processes $(u_t)_{t\geqslant 0}$ with values in $\mathrm{H}^1_\sigma(\mathbb{R}^d)$ such that

$$\|u\|_T = \Big(E\Big[\sup_{0\leqslant t\leqslant T}\|u_t\|^2_\rho\Big] + E\Big[\int_0^T \|\nabla u_t\sigma\|^2_\rho dt\Big]\Big)^{\frac{1}{2}} < \infty.$$

We say that $u\in\mathcal{H}_T$ is a weak solution of the equation (2.1.1) associated with the terminal condition Φ and the coefficients (f,g), if the following relation holds almost surely, for each $\varphi\in\mathcal{D}$

$$\int_t^T (u(s,.),\partial_s\varphi(s,.))ds + \int_t^T \mathcal{E}(u(s,.),\varphi(s,.))ds + (u(t,.),\varphi(t,.)) - (\Phi(.),\varphi(T,.)) \quad (2.5.1)$$
$$= \int_t^T (f(s,.,u(s,.),(\nabla u\sigma)(s,.)),\varphi(s,.))ds + \sum_{i=1}^l\int_t^T (g(s,.,u(s,.),(\nabla u\sigma)(s,.)),\varphi(s,.))\overleftarrow{dB^i_s},$$

where $\mathcal{E}(u,\varphi) = (Lu,\varphi) = \int_{\mathbb{R}^d}((\nabla u\sigma)(\nabla\varphi\sigma) + \varphi\nabla((\frac{1}{2}\sigma^*\nabla\sigma + b)u))(x)dx$ is the energy associated to the diffusion operator.

From Bally and Matoussi [9], we have the following result :

Theorem 2.5.1. *Assume Assumptions* (**H1**) $-$ (**H3**) *hold, there exists a unique weak solution* $u\in\mathcal{H}_T$ *of the SPDE* (2.1.1). *Moreover,* $u(t,x) = Y_t^{t,x}$ *and* $Z_t^{t,x} = \nabla u_t\sigma$, $dt\otimes dx\otimes dP$ *a.e. where* $(Y_s^{t,x},Z_s^{t,x})_{t\leqslant s\leqslant T}$ *is the solution of the BDSDE* (2.1.2). *Furthermore, we have for all* $s\in[t,T]$, $u(s,X_s^{t,x}) = Y_s^{t,x}$ *and* $(\nabla u\sigma)(s,X_s^{t,x}) = Z_s^{t,x}$ $dt\otimes dx\otimes dP$ *a.e.*

2.5.2 Numerical Scheme for SPDE

Let Us first recall that (X^N,Y^N,Z^N) denotes the numerical Euler scheme of the FBDSDE's (2.1.2) given in (2.2.6)-(2.2.7)-(2.2.8)-(2.2.9). The numerical approximation of the SPDE (2.1.1) will be presented in the following lemma :

Lemma 2.5.1. *Let* $x\in\mathbb{R}^d$ *and* $t_n\in\pi$. *Define*

$$u^N_{t_n}(x) := Y^{N,t_n,x}_{t_n} \text{ and } v^N_{t_n}(x) := Z^{N,t_n,x}_{t_n} \quad (2.5.2)$$

Then $u^N_{t_n}$ *(resp.* $v^N_{t_n}$*) is* $\mathcal{F}^B_{t_n,T}$*-measurable and we have for all* $x\in\mathbb{R}^d$ *and* $t,t_n\in\pi$ *such that* $t\leqslant t_n$

$$u^N_{t_n}(X^{t,x}_{t_n}) = Y^{N,t,x}_{t_n} \quad (\text{resp. } v^N_{t_n}(X^{t,x}_{t_n}) = Z^{N,t,x}_{t_n}).$$

Proof. From the Markov property of Y^N and Z^N, the random variables $u^N_{t_n}$ and $v^N_{t_n}$ are $\mathcal{F}^B_{t_n,T}$ measurable. From the definition of $u^N_{t_n}$ and $v^N_{t_n}$, we have

$$u^N_{t_n}(X^{t,x}_{t_n}) = Y^{N,t_n,X^{t,x}_{t_n}}_{t_n} \text{ and } v^N_{t_n}(X^{t,x}_{t_n}) = Z^{N,t_n,X^{t,x}_{t_n}}_{t_n}.$$

From (2.2.8), (2.2.9) and by taking $(t,x) = (t_n, X^{t,x}_{t_n})$, we obtain :

$$\begin{aligned} Y^{N,t_n,X^{t,x}_{t_n}}_{t_n} &= E_{t_n}[Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}] + hE_{t_n}[f(t_n, X^{N,t_n,X^{t,x}_{t_n}}_{t_n}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_n}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_n})] \\ &+ E_{t_n}[g(t_{n+1}, X^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}})\Delta B_n] \\ hZ^{N,t_n,X^{t,x}_{t_n}}_{t_n} &= E_{t_n}[Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}\Delta W^*_n + g(t_{n+1}, X^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}})\Delta B_n \Delta W^*_n]. \end{aligned}$$

From the flow property, we have $X^{N,t_n,X^{t,x}_{t_n}}_{t_n} = X^{N,t,x}_{t_n}$, then we obtain

$$\begin{aligned} Y^{N,t_n,X^{t,x}_{t_n}}_{t_n} &= E_{t_n}[Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}] + hE_{t_n}[f(t_n, X^{N,t,x}_{t_n}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_n}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_n})], \\ &+ E_{t_n}[g(t_{n+1}, X^{N,t,x}_{t_{n+1}}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}})\Delta B_n] \\ hZ^{N,t_n,X^{t,x}_{t_n}}_{t_n} &= E_{t_n}[Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}\Delta W^*_n + g(t_{n+1}, X^{N,t,x}_{t_{n+1}}, Y^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}}, Z^{N,t_n,X^{t,x}_{t_n}}_{t_{n+1}})\Delta B_n \Delta W^*_n]. \end{aligned}$$

Then from the uniqueness of the solution of (2.2.8)-(2.2.9) we obtain the result.

□

2.5.3 Rate of convergence for the weak solution of SPDEs

We give a norm equivalence result which was already proved by Barles and Lesigne [10] and Bally and Matoussi [9] when $b \in C^1_b(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in C^2_b(\mathbb{R}^d, \mathbb{R}^{d\times d})$. We note that σ can be degenerated and so we do not assume ellipticity condition.

Proposition 2.5.1. *Under Assumptions* (**H1**) − (**H3**)*, there exist two positive constants C_1 and C_2 such that for every $t \leqslant s \leqslant T$ and $\phi \in L^1(\mathbb{R}^d \times \Omega_B, \rho(x)dx \otimes dP_B)$, we have*

$$C_1 \int_{\mathbb{R}^d} E[|\phi(x)|]\rho(x)dx \leqslant \int_{\mathbb{R}^d} E[|\phi(X^{t,x}_s)|]\rho(x)dx \leqslant C_2 \int_{\mathbb{R}^d} E[|\phi(x)|]\rho(x)dx. \qquad (2.5.3)$$

Moreover, for every $\Psi \in L^1(\mathbb{R}^d \times (0,T) \times \Omega_B, \rho(x)dx \otimes dt \otimes dP_B)$

$$\begin{aligned} C_1 \int_{\mathbb{R}^d}\int_t^T E[|\Psi(s,x)|]ds\rho(x)dx &\leqslant \int_{\mathbb{R}^d}\int_t^T E[|\Psi(s,X^{t,x}_s)|]ds\rho(x)dx \\ &\leqslant C_2 \int_{\mathbb{R}^d}\int_t^T E[|\Psi(x)|]ds\rho(x)dx. \end{aligned} \qquad (2.5.4)$$

We recall that $u(t,x) = Y^{t,x}_t$ and $v(t,x) = Z^{t,x}_t$ $dt\otimes dx \otimes dP$ a.e. We define the process (u^N_s, v^N_s) as follows :

$$u^N_s(x) := Y^{N,s,x}_s \text{ and } v^N_s(x) := Z^{N,s,x}_s, \forall s \in [t_n, t_{n+1}). \qquad (2.5.5)$$

Using (2.2.10) and following the proof of Lemma 2.5.1, we obtain

$$u^N_s(X^{t,x}_s) = Y^{N,t,x}_s \text{ and } v^N_s(X^{t,x}_s) = Z^{N,t,x}_s, \forall t \leqslant s, t, s \in [t_n, t_{n+1}). \qquad (2.5.6)$$

As in Gyongy and Krylov [37], we define the error between the solution of the SPDE and the numerical scheme as follows :

$$\begin{aligned} Error_N(u,v) \quad &:= \quad \sup_{0 \leqslant s \leqslant T} E[\int_{\mathbb{R}^d} |u_s^N(x) - u(s,x)|^2 \rho(x)dx] \\ &+ \quad \sum_{n=0}^{N-1} E[\int_{\mathbb{R}^d} \int_{t_n}^{t_{n+1}} \|v_s^N(x) - v(s,x)\|^2 ds\rho(x)dx]. \end{aligned} \tag{2.5.7}$$

The following theorem shows the convergence of the numerical scheme (2.5.2) of the solution of the SPDE (2.1.1).

Theorem 2.5.2. *Assume that* **(H1)**-**(H3)** *hold. Then, the error* $Error_N(u,v)$ *converges to 0 as* $N \to \infty$ *and there exists a positive constant* C *(depending only on* T, K, α, $|b(0)|$, $||\sigma(0)||$, $|f(t,0,0,0)|$ *and* $||g(t,0,0,0)||$*) such that*

$$Error_N(u,v) \leqslant Ch. \tag{2.5.8}$$

Proof. We take $t = t_0$. From the norm equivalence result (see inequality (2.5.3)), for all $s \in [t_n, t_{n+1})$ such that $s \geqslant t$, we have

$$E[\int_{\mathbb{R}^d} |u_s^N(x) - u(s,x)|^2 \rho(x)dx] \leqslant CE[\int_{\mathbb{R}^d} |u_s^N(X_s^{t,x}) - u(s,X_s^{t,x})|^2 \rho(x)dx],$$

where C is positive generic constant. From equation (2.5.6), we get

$$E[\int_{\mathbb{R}^d} |u_s^N(x) - u(s,x)|^2 \rho(x)dx] \leqslant C \int_{\mathbb{R}^d} E[|Y_s^{N,t,x} - Y_s^{t,x}|^2]\rho(x)dx.$$

Therefore Remark 2.4.1 implies that

$$\sup_{0 \leqslant s \leqslant T} E[\int_{\mathbb{R}^d} |u_s^N(x) - u(s,x)|^2 \rho(x)dx] \leqslant Ch \int_{\mathbb{R}^d} (1+|x|^2)\rho(x)dx \leqslant Ch. \tag{2.5.9}$$

From the norm equivalence result (see inequality (2.5.4)), we have

$$\begin{aligned} &\sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} ||v_s^N(x) - v(s,x)||^2 \rho(x)dxds] \\ \leqslant \quad & C \sum_{n=0}^{N-1} E[\int_{\mathbb{R}^d} \int_{t_n}^{t_{n+1}} ||v_s^N(X_s^{t,x}) - v(s,X_s^{t,x})||^2 \rho(x)dxds]. \end{aligned}$$

From equation (2.5.6), we get

$$\begin{aligned} &\sum_{n=0}^{N-1} E[\int_{\mathbb{R}^d} \int_{t_n}^{t_{n+1}} ||v_s^N(X_s^{t,x}) - v(s,X_s^{t,x})||^2 \rho(x)dxds] \\ = \quad & \sum_{n=0}^{N-1} E[\int_{\mathbb{R}^d} \int_{t_n}^{t_{n+1}} ||Z_s^{N,t,x} - Z_s^{t,x}||^2 \rho(x)dxds], \end{aligned}$$

and so from Remark 2.4.1 we deduce that

$$\sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} ||v_s^N(x) - v(s,x)||^2 \rho(x)dxds] \leqslant Ch \int_{\mathbb{R}^d} (1+|x|^2)\rho(x)dx \leqslant Ch. \tag{2.5.10}$$

From inequalities (2.5.9) and (2.5.10), we deduce that (2.5.8) holds.

□

Remark 2.5.1. *Gyongy and Krylov [37] considered the following linear SPDE on* $[0,T]\times\mathbb{R}^d$,

$$\begin{cases} du(t,x) &= (\mathcal{L}_1 u(t,x) + f(t,x))dt + \sum_{i=1}^{\infty}(\mathcal{L}_{2,i}u(t,x) + g(t,x)_i)dw_t^i \\ u(0,x) &= u_0 \in L^2(\Omega, P), \end{cases}$$

where $\mathcal{L}_1 u(t,x) = \sum_{q,l=1}^{d} a(t,x)_{lq}\frac{\partial^2 u(t,x)}{\partial x_l \partial x_q}$, $\mathcal{L}_{2,i}u(t,x) = \sum_{q=1}^{d} b_{iq}(t,x)\frac{\partial u(t,x)}{\partial x_q}, 1 \leqslant i \leqslant \infty$ *and* $(b(t,x)_{i,q})_{i=0}^{\infty} \in \ell^2$, $(t,x)\in[0,T]\times\mathbb{R}^d$, $1\leqslant q\leqslant d$. *They approximate the SPDE by*

$$du^h(t,x) = (\mathcal{L}_1^h u^h(t,x) + f(t,x))dt + \sum_{i=1}^{\infty}(\mathcal{L}_{2,i}^h u^h(t,x) + g(t,x)_i)dw_t^i,$$

$\mathcal{L}_1^h$, $\mathcal{L}_{2,i}^h$ *are the approximation of* $\mathcal{L}_1$, $\mathcal{L}_{2,i}$ *by using finite difference scheme on the space grid* $\mathbb{G}_h$. *Their results revolve to prove the existence of the random process* $u^{(j)}(t,x)$, $j=1,\dots,k$ *for some* $k\geqslant 0$ *s.t.*

$$u^h(t,x) = u^{(0)}(t,x) + \sum_{j=1}^{k}\frac{h^j}{j!}u^{(j)}(t,x) + R^h(t,x),$$

where $u^{(0)}$ *is the solution of the SPDE. They assumed that the SPDE is non degenerate and for* $m > k+1+\frac{d}{2}$, *the coefficients are m-times continuously differentiable in x. When they used a symmetric finite difference scheme and* $d=2$, *the* L^2*-error is proportional to* h^2 *where h is the discretization step in space and by the Richardson acceleration, the error is proportional to* h^4. *Compared to their work, our scheme is more general. It converges in the non linear case. Our convergence is of order* $\sqrt{h}$ *where h is the discretization step in time. However, our scheme is expected, as usual for Monte Carlo methods, to deal better than analytic methods with high dimensional problems.*

Remark 2.5.2. *If we assume more regularity conditions on the coefficients and the final condition as in Pardoux and Peng [63], namely,* $\Phi\in C_b^3(\mathbb{R}^d,\mathbb{R}^k)$, $f\in C_b^3([0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{d\times k},\mathbb{R}^k)$ *and* $g\in C_b^3([0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{d\times k},\mathbb{R}^{k\times l})$. *If* $(Y_s^{t,x}, Z_s^{t,x})_{t\leqslant s\leqslant T}$ *is the solution the BDSDE (2.1.2). Then,* $u_t(x) = Y_t^{t,x}$, $\forall(t,x)\in[0,T]\times\mathbb{R}^d$ *is the unique classical solution of the SPDE (2.1.1) in the integral sense (see [63]). Therefore, we can obtain a stronger result. In fact, the estimation on the error (2.5.7) obtained in the previous theorem can be replaced by :*

$$E[\sup_{0\leqslant t\leqslant T}|u_t^N(x) - u(t,x)|^2] + E[\int_0^T ||v_t^N(x) - v(t,x)||^2 dt] \leqslant Ch.$$

This last equation gives an estimation which holds for all $x\in\mathbb{R}^d$ *and which is not only almost sure anymore. For the Monte Carlo method, we estimate the solution for one point x at time t, and by varying x and t we obtain the solution* $u(t,x)$ *on the whole domain.*

2.6 Implementation and numerical tests

In this part, we are interested in implementing our numerical scheme. Our aim is only to test statically its convergence. Further analysis of the convergence of the used method and of the error bounds will be accomplished in a future work.

2.6.1 Notations and algorithm

We use a path-dependent algorithm, for every fixed path of the brownian motion B, we approximate by a regression method the solution of the associated PDE. Then, we replace the conditional expectations which appear in (2.6.1) and (2.6.2) by $L^2(\Omega, \mathcal{P})$ projections on the function basis approximating $L^2(\Omega, \mathcal{F}_{t_n})$. We compute $Z^N_{t_n}$ in an explicit manner and we use I Picard iterations to compute $Y^N_{t_n}$ in a implicit way. Actually, we proceed as in [34], except that in our case the solutions $Y^N_{t_n}$ and $Z^N_{t_n}$ are measurable functions of $(X^N_{t_n}, (\Delta B_i)_{n \leqslant i \leqslant N-1})$. So, each solution given by our algorithm depends on the fixed path of B.

2.6.1.1 Numerical scheme

For each fixed path of B, the solution of (2.2.1)-(2.2.2) is approximated by (Y^N, Z^N) defined by the following algorithm, given in the multidimensional case.
For $0 \leqslant n \leqslant N-1$:
$\forall j_1 \in \{1, \dots, k\}$,

$$Y^N_{t_n,j_1} = E_{t_n}\Big[Y^N_{t_{n+1},j_1} + h f_{j_1}(X^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}) + \sum_{j=1}^{l} g_{j_1,j}(X^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Delta B_{n,j}\Big], \tag{2.6.1}$$

$\forall j_1 \in \{1, \dots, k\}$ and $\forall j_2 \in \{1, \dots, d\}$

$$hZ^N_{t_n,j_1,j_2} = E_{t_n}\Big[Y^N_{t_{n+1},j_1}\Delta W_{n,j_2} + \sum_{j=1}^{l} g_{j_1,j}(X^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Delta B_{n,j}\Delta W_{n,j_2}\Big]. \tag{2.6.2}$$

We stress that at each discretization time, the solution of the algorithm depends on the fixed path of the brownian motion B.

2.6.1.2 Vector spaces of functions

At every t_n, we select $k(d+1)$ deterministic functions bases $(p_{i,n}(.))_{1 \leqslant i \leqslant k(d+1)}$ and we look for approximations of $Y^N_{t_n}$ and $Z^N_{t_n}$, which will be denoted respectively by y^N_n and z^N_n, in the vector spaces spanned respectively by the basis $(p_{j_1,n}(.))_{1 \leqslant j_1 \leqslant k}$ and the basis $(p_{j_1,j_2,n}(.))_{1 \leqslant j_1 \leqslant k, 1 \leqslant j_2 \leqslant d}$. Each basis $p_{i,n}(.)$ is considered as a vector of functions of dimension $L_{i,n}$. In other words, $P_{i,n}(.) = \{\alpha.p_{i,n}(.), \alpha \in \mathbb{R}^{L_{i,n}}\}$.
As an example, we cite the hypercube basis (**HC**) used in [34]. In this case, $p_{i,n}(.)$ does not depend nor on i neither on n and its dimension is simply denoted by L. A domain $D \subset \mathbb{R}^d$ centered on $X_0 = x$, that is $D = \prod_{i=1}^{d}(x_i - a, x_i + a]$, can be partitionned on small hypercubes of edge δ. Then, $D = \bigcup_{i_1,\dots,i_d} D_{i_1,\dots,i_d}$ where $D_{i_1,\dots,i_d} = (x_i - a + i_1\delta, x_i - a + i_1\delta] \times \dots \times (x_i - a + i_d\delta, x_i - a + i_d\delta]$. Finally we define $p_{i,n}(.)$ as the indicator functions of this set of hypercubes.

2.6.1.3 Monte Carlo simulations

To compute the projection coefficients α, we will use M independent Monte Carlo simulations of $X_{t_n}^N$ and ΔW_n which will be respectively denoted by $X^{N,m}_{t_n}$ and $\Delta W^m_n, m = 1, \dots, M$.

2.6.1.4 Description of the algorithm

$\rightarrow$ Initialization : For $n = N$, take $(y^{N,m,I}_N) = (\Phi(X^{N,m}_{t_N}))$ and $(z^{N,m}_N) = 0$.
$\rightarrow$ Iteration : For $n = N-1, \dots, 0$:

• We approximate (2.6.2) by computing for all $j_1 \in \{1, \ldots, k\}$ and $j_2 \in \{1, \ldots, d\}$

$$\begin{aligned} \alpha^M_{j_1,j_2,n} &= \operatorname*{arginf}_{\alpha} \frac{1}{M} \sum_{m=1}^{M} \Big| y^{N,M,I}_{n+1,j_1}(X^{N,m}_{t_{n+1}}) \frac{\Delta W^m_{n,j_2}}{h} \\ &+ \sum_{j=1}^{l} g_{j_1,j}\Big(X^{N,m}_{t_{n+1}}, y^{N,M,I}_{n+1}(X^{N,m}_{t_{n+1}}), z^{N,M}_{n+1}(X^{N,m}_{t_{n+1}})\Big) \frac{\Delta B_{n,j} \Delta W^m_{n,j_2}}{h} - \alpha.p^m_{j_1,j_2,n}\Big|^2. \end{aligned}$$

Then we set $z^{N,M}_{n,j_1,j_2}(.) = (\alpha^M_{j_1,j_2,n}.p_{j_1,j_2,n}(.)),\ j_1 \in \{1, \ldots, k\},\ j_2 \in \{1, \ldots, d\}$.

• We use I Picard iterations to obtain an approximation of Y_{t_n} in (2.6.1) :

· For $i = 0 : \forall j_1 \in \{1, \ldots, k\},\ \alpha^{M,0}_{j_1,n} = 0$.

· For $i = 1, \ldots, I$: We approximate (2.6.1) by calculating $\alpha^{M,i}_{j_1,n}$, $\forall j_1 \in \{1, \ldots, k\}$, as the minimizer of :

$$\begin{aligned} &\frac{1}{M} \sum_{m=1}^{M} \Big| y^{N,M,I}_{n+1,j_1}(X^{N,m}_{t_{n+1}}) + h f_{j_1}\Big(X^{N,m}_{t_n}, y^{N,M,i-1}_n(X^{N,m}_{t_n}), z^{N,M}_n(X^{N,m}_{t_n})\Big) \\ &\qquad + \sum_{j=1}^{l} g_{j_1,j}\Big(X^{N,m}_{t_{n+1}}, y^{N,M,I}_{n+1}(X^{N,m}_{t_{n+1}}), z^{N,M}_{n+1}(X^{N,m}_{t_{n+1}})\Big) \Delta B_{n,j} - \alpha p^m_{j_1,k}\Big|^2. \end{aligned}$$

Finally, we define $y^{N,M,I}_n(.)$ as :

$$y^{N,M,I}_{n,j_1}(.) = (\alpha^{M,I}_{j_1,n}.p_{j_1,n}(.)), \forall j_1 \in \{1, \ldots, k\}.$$

2.6.2 One-dimensional case (Case when $d = k = l = 1$)

2.6.2.1 Function bases

We use the basis (**HC**) defined above. So we set :

$$d_1 = \min_{n,m} X^m_{t_n}, \quad d_2 = \max_{n,m} X^m_{t_n} \text{ and } L = \frac{d_2 - d_1}{\delta}$$

where δ is the edge of the hypercubes $(D_j)_{1 \leqslant j \leqslant L}$ defined by $D_j = \Big[d + (j-1)\delta, d + j\delta\Big), \forall j$. We take at each time t_n

$$1_{D_j}(X^{N,m}_{t_n}) = 1_{[d+(j-1)\delta, d+j\delta)}(X^{N,m}_{t_n}), j = 1, \ldots, L$$

and

$$(p^m_{i,n}(.)) = \Big\{\sqrt{\frac{M}{card(D_j)}} 1_{D_j}(X^{N,m}_{t_n}), 1 \leqslant j \leqslant L\Big\}, i = 0, 1.$$

$Card(D_j)$ denotes the number of simulations of $X^N_{t_n}$ which are in our cube D_j.
This system is orthonormal with respect to the empirical scalar product defined by

$$< \psi_1, \psi_2 >_{n,M} := \frac{1}{M} \sum_{m=1}^{M} \psi_1(X^{N,m}_{t_n}) \psi_2(X^{N,m}_{t_n}).$$

In this case, the solutions of our least squares problems are given by :

$$\begin{aligned}
\alpha_{1,n}^{M} &= \frac{1}{M}\sum_{m=1}^{M} p_{1,n}(X_{t_n}^{N,m})\Big\{ y_{n+1}^{N,M,I}(X_{t_{n+1}}^{N,m})\frac{\Delta W_n^m}{h} \\
&+ g\Big(X_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(X_{t_{n+1}}^{N,m}), z_{n+1}^{N,M,}(X_{t_{n+1}}^{N,m})\Big)\frac{\Delta B_n^m \Delta W_n^m}{h}\Big\}, \\
\alpha_{0,n}^{M,i} &= \frac{1}{M}\sum_{m=1}^{M} p_{0,n}(X_{t_n}^{N,m})\Big\{ y_{n+1}^{N,M,I}(X_{t_{n+1}}^{N,m}) + hf\Big(X_{t_n}^{N,m}, y_n^{N,M,i-1}(X_{t_n}^{N,m}), z_n^{N,M}(X_{t_n}^{N,m})\Big) \\
&+ g\Big(X_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(X_{t_{n+1}}^{N,m}), z_{n+1}^{N,M}(X_{t_{n+1}}^{N,m})\Big)\Delta B_n^m\Big\}.
\end{aligned}$$

Remark 2.6.1. *We note that for each value of M, N and δ, we launch the algorithm 50 times and we denote by $(Y_{0,m'}^{0,x,N,M,I})_{1 \leqslant m' \leqslant 50}$ the set of collected values. Then we calculate the empirical mean $\overline{Y}_0^{0,x,N,M,I}$ and the empirical standard deviation $\sigma^{N,M,I}$ defined by :*

$$\overline{Y}_0^{0,x,N,M,I} = \frac{1}{50}\sum_{m'=1}^{50} Y_{0,m'}^{0,x,N,M,I} \ \textit{and}\ \sigma^{N,M,I} = \sqrt{\frac{1}{49}\sum_{m'=1}^{50} |Y_{0,m'}^{0,x,N,M,I} - \overline{Y}_0^{0,x,N,M,I}|^2}. \tag{2.6.3}$$

We also note before starting the numerical examples that our algorithm converges after at most three Picard iterations. Finally, we stress that (2.6.3) gives us an approximation of $u(0,x)$ the solution of the SPDE (2.1.1) at time $t = 0$.

2.6.2.2 Case when f and g are linear in y and independent of z

$$\begin{cases} dX_t = X_t(\mu dt + \sigma dW_t), \\ \Phi(x) = -x + K,\ f(y) = a_0 y,\ g(y) = b_0 y \end{cases}$$

and we set $K = 115$, $r = 0.01$, $R = 0.06$, $X_0 = 100$, $\mu = 0.05$, $\sigma = 0.2$, $T = 0.25$, $d_1 = 60$, $d_2 = 200$, a_0 and b_0 are fixed constants.
Let $Y_{explicit}$ be the solution of our BDSDE in this particular case. By the integration by parts formula, we get

$$Y_{t,explicit}^{t,x} = E[\Phi(X_T^{t,x})e^{a_0(T-t)+b_0(B_T-B_t)-\frac{1}{2}b_0^2(T-t)}/\mathcal{F}_{t,T}^B].$$

At t=0, we have

$$\begin{aligned}
Y_{0,explicit}^{0,x} &= E[\Phi(X_T^{0,x})e^{(a_0-\frac{1}{2}b_0^2)T+b_0B_T}/\mathcal{F}_{0,T}^B] \\
&= e^{(a_0-\frac{1}{2}b_0^2)T+b_0B_T}E[\Phi(X_T^{0,x})].
\end{aligned}$$

Then, we define $\overline{Y}_0^{0,x,N,M,I}$ as the numerical approximation of the solution of the BDSDE in this case (computed by our algorithm) and $\sigma^{N,M,I}$ as its standard deviation. In the other hand, we compute the solution $Y_{0,explicit}^{0,x}$ in this linear case by using the explicit formula of the expectation of $\Phi(X_T^{0,x})$, as follows

$$Y_{explicit}^{0,x} = e^{(a_0-\frac{1}{2}b_0^2)T+b_0B_T}E[\Phi(X_T^{0,x})] = e^{(a_0-\frac{1}{2}b_0^2)T+b_0B_T}(K - xe^{\mu T}).$$

For $a_0 = 0.5$, $b_0 = 0.5$ and $\delta = 1$

N=20, $Y^{0,x}_{explicit} = 13.724$

M	$\overline{Y}^{0,x,N,M,I}_0(\sigma^{N,M,I})$	$\frac{\lvert Y^{0,x}_{explicit} - \overline{Y}^{0,x,N,M,I}_0 \rvert}{Y^{0,x}_{explicit}}$
100	13.910(1.178)	0.013
1000	13.792(0.309)	0.004
5000	13.847(0.117)	0.008

For $a_0 = 0.5$, $b_0 = 0.5$ and $\delta = 0.5$

N=30, $Y^{0,x}_{explicit} = 14.115$

M	$\overline{Y}^{0,x,N,M,I}_0(\sigma^{N,M,I})$	$\frac{\lvert Y^{0,x}_{explicit} - \overline{Y}^{0,x,N,M,I}_0 \rvert}{Y^{0,x}_{explicit}}$
100	14.246(1.045)	0.009
1000	14.195(0.337)	0.005
5000	14.236(0.129)	0.008

2.6.2.3 Comparison of numerical approximations of the solutions of the FBDSDE and the FBSDE : the general case

Now we take

$$\begin{cases} \Phi(x) = -x + K, \\ f(t,x,y,z) = -\theta z - ry + (y - \frac{z}{\sigma})^-(R - r), \\ g_1(t,x,y,z) = 0.1z + 0.5y + log(x) \end{cases}$$

and we set $\theta = (\mu - r)/\sigma$, $K = 115$, $X_0 = 100$, $\mu = 0.05$, $\sigma = 0.2$, $r = 0.01$, $R = 0.06$, $\delta = 1$, $N = 20$, $T = 0.25$ and we fix $d_1 = 60$ and $d_2 = 200$ as in citeG05. The functions g_1,g_2 and g_3 taken in the following are examples of the function g. They are sufficiently regular and Lipschitz on $[60, 200] \times \mathbb{R} \times \mathbb{R}$ and could be extended to regular Lipschitz functions on $\mathbb{R}^3$. In this case, Assumptions (**H1**)-(**H3**) are satisfied.
We compare the numerical solution of our BDSDE (noted again $\overline{Y}^{t,x,N,M,I}_t$) and the BSDE's one (noted here by $\overline{Y}^{0,x,N,M}_{t,BSDE}$), without g and B.

When t is close to maturity

M	$\overline{Y}^{0,x,N,M}_{t_{19},BSDE}(\sigma^{N,M})$	$\overline{Y}^{0,x,N,M,I}_{t_{19}}(\sigma^{N,M,I})$
128	13.748(0.879)	15.452(0.948)
512	13.827(0.384)	15.534(0.409)
2048	13.762(0.223)	15.464(0.240)
8192	13.781(0.091)	15.484(0.097)
32768	13.796(0.054)	15.501(0.058)

M	$\overline{Y}^{0,x,N,M}_{t_{15},BSDE}(\sigma^{N,M})$	$\overline{Y}^{0,x,N,M,I}_{t_{15}}(\sigma^{N,M,I})$
128	14.168(0.905)	17.894(1.096)
512	14.113(0.388)	17.774(0.429)
2048	13.988(0.226)	17.607(0.270)
8192	13.985(0.093)	17.623(0.104)
32768	13.994(0.055)	17.627(0.064)

When $t = 0$

M	$\overline{Y}_{0,BSDE}^{0,x,N,M}(\sigma^{N,M})$	$\overline{Y}_{0}^{0,x,N,M,I}(\sigma^{N,M,I})$
128	15.431(1.005)	13.571(1.146)
512	15.029(0.428)	13.173(0.500)
2048	14.763(0.243)	12.885(0.280)
8192	14.718(0.098)	12.825(0.106)
32768	14.715(0.060)	12.804(0.064)

For $g_2(y,z) = 0.1z + 0.5y$

M	$\overline{Y}_{t_{19}}^{0,x,N,M,I}(\sigma^{N,M,I})$
128	14.767(0.949)
512	14.850(0.410)
2048	14.781(0.240)
8192	14.801(0.097)
32768	14.818(0.058)

M	$\overline{Y}_{t_{15}}^{0,x,N,,M,I}(\sigma^{N,M,I})$
128	16.267(1.093)
512	16.166(0.428)
2048	16.007(0.270)
8192	16.024(0.104)
32768	16.029(0.064)

When $t = 0$

M	$\overline{Y}_{0}^{0,x,N,M,I}(\sigma^{N,M,I})$
128	13.821(0.063)
512	14.555(1.132)
2048	14.176(0.495)
8192	13.899(0.277)
32768	13.842(0.105)

For $g_3(x,y) = logx + 0.5y$:
When t is close to maturity

M	$\overline{Y}_{t_{19}}^{0,x,N,M,I}(\sigma^{N,M,I})$
128	15.452(0.948)
512	15.534(0.409)
2048	15.464(0.240)
8192	15.484(0.097)
32768	15.501(0.058)

M	$\overline{Y}_{t_{15}}^{0,x,N,M,I}(\sigma^{N,M,I})$
128	18.253(1.068)
512	18.166(0.453)
2048	18.010(0.266)
8192	18.006(0.109)
32768	18.017(0.065)

When $t = 0$

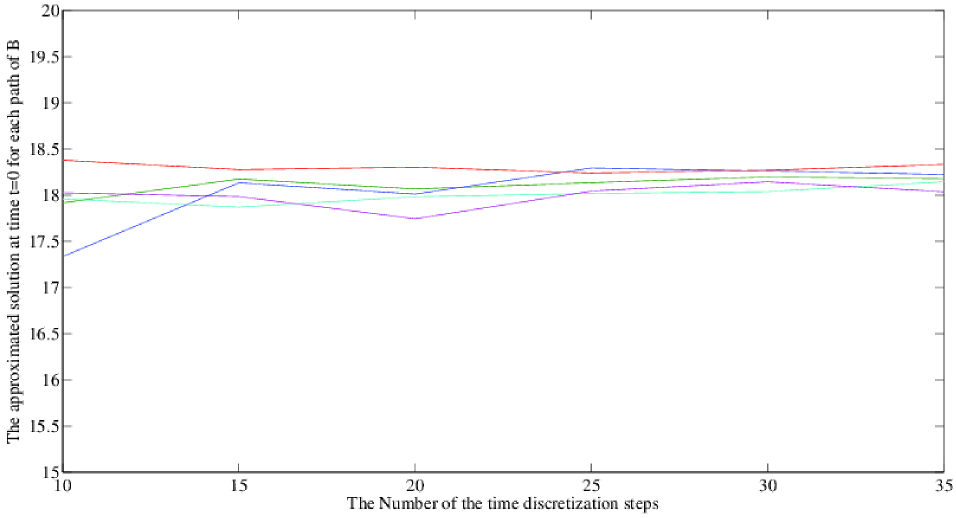

FIGURE 2.1 – The BDSDE's solution with respect to the number of time discretization steps for five different paths of B. The figure is obtained for $M = 2000$ and $\delta = 1$.

M	$\overline{Y}_0^{0,x,N,M,I}(\sigma^{N,M,I})$
128	12.071(0.054)
512	12.075(0.088)
2048	12.122(0.218)
8192	12.384(0.381)
32768	12.791(0.903)

In the previous tables, we test our algorithm for different examples of the function g (g_1 and g_2 are dependent in z, g_3 is independent of z). We see the convergence of the BDSDE's solution when we increase the number of simulations M.

In figure 2.1, we study statically the main result of this paper. So, we fix all the parameters ($\delta = 1,$ and $M = 2000$) and we draw the map of the BDSDE's solution, for the function g_1, with respect to the number of time discretization steps N. The solution is computed for five different paths of the brownian motion B. We can examine there the convergence of our scheme.

We see on Figure 2.2 the impact of the function g on the solution ; we variate N, M and δ as in [34], by taking these quantities as follows : First we fix $d_1 = 40$ and $d_2 = 180$ (which means that $x \in [d_1, d_2] = [40, 180]$ and in this case our assumptions (**H1**)-(**H3**) are satisfied). Let $j \in \mathbb{N}$, we take $\alpha_M = 3$, $\beta = 1$, $N = 2(\sqrt{2})^{(j-1)}$, $M = 2(\sqrt{2})^{\alpha_M(j-1)}$ and $\delta = 50/(\sqrt{2})^{(j-1)(\beta+1)/2}$. Then, we draw the map of each solution at $t = 0$ with respect to j.

2.7 Zhang L^2-Regularity results under Lipschitz assumptions

In order to derive the rate of convergence for our numerical scheme, we proved in the subsection 2.4.4 the Zhang L^2-Regularity for the martingale integrand Z under strong assumptions on the parameters. Indeed, to use the Malliavin calculus tools, we assumed that the coefficients of the BDSDE (2.2.2) are of class $\mathcal{C}^2$. The aim of this section is to prove the Zhang L^2-regularity results

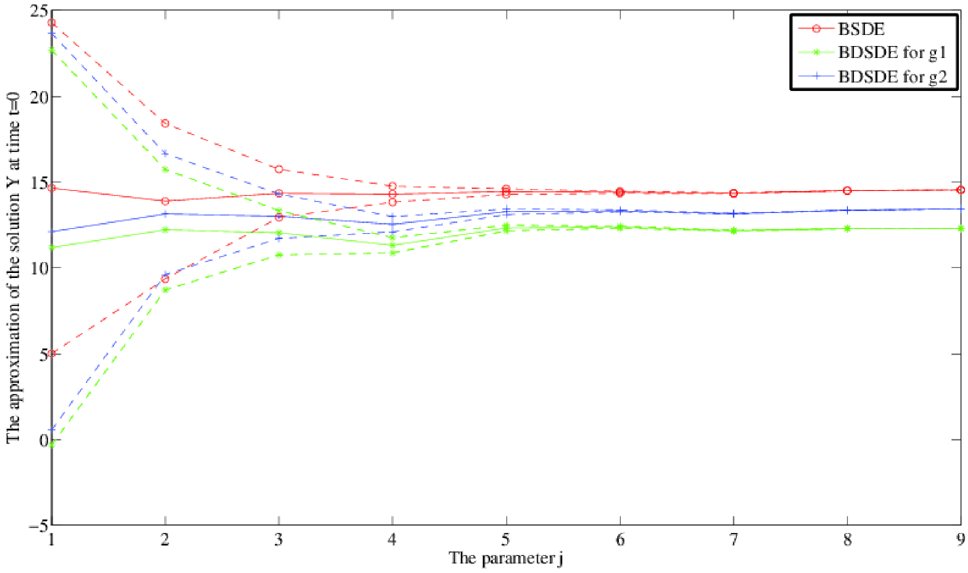

FIGURE 2.2 – Comparison of the BSDE's solution and the BDSDE's one : The solution of the BSDE is with circle markers, the solution of the BDSDE for $g_1(x, y, z) = 0.1z + 0.5y + log(x)$ is with star markers and the one for $g_2(y, z) = 0.1z + 0.5y$ is with cross markers. Confidence intervals are with dotted lines.

only under standard Lipschitz assumptions on the coefficients of the BDSDE (2.2.2). Thus, we will be able to derive a rate of convergence for our numerical scheme only under Lipschitz assumptions.

We recall the assumptions :

Assumption (H1) There exist a positive constant K such that

$$|b(x) - b(x')| + \|\sigma(x) - \sigma(x')\| \leqslant K|x - x'|, \forall x, x' \in \mathbb{R}^d.$$

Assumption (H2) There exist two constants $K > 0$ and $0 \leqslant \alpha < 1$ such that for any $(t_1, x_1, y_1, z_1), (t_2, x_2, y_2, z_2) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k\times d}$,

(i) $|f(t_1, x_1, y_1, z_1) - f(t_2, x_2, y_2, z_2)| \leqslant K\big(\sqrt{|t_1 - t_2|} + |x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\|\big),$

(ii) $\|g(t_1, x_1, y_1, z_1) - g(t_2, x_2, y_2, z_2)\|^2 \leqslant K\big(|t_1 - t_2| + |x_1 - x_2|^2 + |y_1 - y_2|^2\big) + \alpha^2\|z_1 - z_2\|^2,$

(iii) $|\Phi(x_1) - \Phi(x_2)| \leqslant K|x_1 - x_2|,$

(iv) $\sup_{0 \leqslant t \leqslant T}(|f(t, 0, 0, 0)| + ||g(t, 0, 0, 0)||) \leqslant K.$

We introduce the following assumption

Assumption (H2') There exist two constants $K > 0$ and $0 \leqslant \alpha < 1$ such that for any $(t_1, x_1, y_1, z_1), (t_2, x_2, y_2, z_2) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k\times d}$,

(i) $|f(t_1, x_1, y_1, z_1) - f(t_2, x_2, y_2, z_2)| \leqslant K\big(\sqrt{|t_1 - t_2|} + |x_1 - x_2| + |y_1 - y_2| + \|z_1 - z_2\|\big),$

(ii) $\|g(t_1, x_1, y_1, z_1) - g(t_2, x_2, y_2, z_2)\|^2 \leqslant K\big(|t_1 - t_2| + |x_1 - x_2|^2 + |y_1 - y_2|^2\big) + \alpha^2\|z_1 - z_2\|^2,$

(iii) $|\Phi(x_1) - \Phi(x_2)| \leqslant K|x_1 - x_2|,$

(iv) $f(t, 0, 0) \in \mathbb{H}^2_k([0, T])$ and $g(t, 0, 0) \in \mathbb{H}^2_{k\times l}([0, T]).$

The following lemma will be needed later in our estimations.

Lemma 2.7.1. *Assume that Assumptions* **(H1)** *and* **(H2')** *hold and* $\xi \in L^2(\mathcal{F}_T)$. *Let* $\Theta := (X^{t,x}, Y, Z)$ *denote the solution of the following F-BDSDE*

$$X_s^{t,x} = x + \int_t^s b(X_u^{t,x})ds + \int_t^s \sigma(X_u^{t,x})dW_u, \quad s \in [t,T] \tag{2.7.1}$$

$$Y_s = \xi + \int_t^s f(u, Y_u, Z_u)ds + \int_t^s g(u, Y_u, Z_u)\overleftarrow{dB_u} - \int_t^s Z_u dW_u. \tag{2.7.2}$$

Then we have the following
(i)for all $p \geqslant 2$, *there exists a constant* C_p, *depending on* T, K, α *and* p *such that*

$$E\Big[\sup_{t \leqslant s \leqslant T}|Y_s|^p + \Big(\int_t^T \|Z_s\|^2 ds\Big)^{\frac{p}{2}}\Big] \leqslant C_p E\Big\{|\xi|^p + \int_t^T |f(s,0,0)|^p ds + \int_t^T |g(s,0,0)|^p ds\Big\} \tag{2.7.3}$$

and

$$\begin{aligned} E\Big[|Y_s - Y_t|^p\Big] &\leqslant C_p E\Big\{|\xi|^p + \sup_{0 \leqslant s \leqslant T}|f(s,0,0)|^p + \sup_{0 \leqslant s \leqslant T}|g(s,0,0)|^p ds\Big\}|s-t|^{p-1} \\ &+ E\Big[\Big(\int_t^s \|Z_u\|^2 du\Big)^{\frac{p}{2}}\Big]. \end{aligned} \tag{2.7.4}$$

(ii)Let $\Theta^\varepsilon := (X^{\varepsilon,t,x}, Y^\varepsilon, Z^\varepsilon)$ *denote the solution of the perturbed F-BDSDE* (2.7.1) *and* (2.7.2) *with coefficients replaced by* $b^\varepsilon, \sigma^\varepsilon, f^\varepsilon$ *and* g^ε, *initial condition replaced by* x^ε *and* ξ^ε *as a terminal value. Assume that* $b^\varepsilon, \sigma^\varepsilon, f^\varepsilon$ *and* g^ε *satisfy Assumptions* **(H1)** *and* **(H2')**, *that* $\lim_{\varepsilon \longrightarrow 0} x^\varepsilon = x$ *and that for fixed (x,y,z) in* $\mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{k\times d}$,

$$\lim_{\varepsilon \longrightarrow 0} |b^\varepsilon(x) - b(x)|^2 + |\sigma^\varepsilon(x) - \sigma(x)|^2 = 0,$$
$$\lim_{\varepsilon \longrightarrow 0} E\Big\{|\xi^\varepsilon - \xi|^2 + \int_t^T |g^\varepsilon(s,y,z) - g(s,y,z))|^2 ds + \int_t^T |f^\varepsilon(s,y,z) - f(s,y,z)|^2 ds\Big\} = 0.$$

Then we have

$$\lim_{\varepsilon \longrightarrow 0} E\Big\{\sup_{t \leqslant s \leqslant T}|X_s^{\varepsilon,t,x} - X_s^{t,x}|^2 + \sup_{t \leqslant s \leqslant T}|Y_s^\varepsilon - Y_s|^2 + \int_t^T |Z_s^\varepsilon - Z_s|^2 ds\Big\} = 0. \tag{2.7.5}$$

□

The following lemma will be useful for the upper bound result on our time discretization error and for proving the Zhang L^2- regularity under Lipschitz assumptions.

Lemma 2.7.2. *Assume that Assumptions* **(H1)** *and* **(H2)** *hold. Then for all* $p \geqslant 2$, *there exists a constant* $C_p > 0$ *depending only on* T, K, α *and* p *such that*

$$\Big(E[\|Z_s^{t,x}\|^p]\Big)^{\frac{1}{p}} \leqslant C_p(1 + |x|^2) \ a.e. \ s \in [t,T]. \tag{2.7.6}$$

In addition, there exist a positive constant C *independent from* h *the time step of our uniform time-grid such that*

$$\begin{aligned} \max_{0 \leqslant n \leqslant N-1} \Big\{\sup_{t_n \leqslant s \leqslant t_{n+1}} E[|Y_s^{t,x} - Y_{t_n}^{t,x}|^2] &+ \sup_{t_n \leqslant s \leqslant t_{n+1}} E[|Y_s^{t,x} - Y_{t_{n+1}}^{t,x}|^2]\Big\} \\ &\leqslant Ch(1 + |x|^2). \end{aligned} \tag{2.7.7}$$

Proof. Fisrt, we consider the case when b, σ, f, g and $\Phi \in C_b^1$ and satisfying assumptions **(H1)** and **(H2)**. Let $\nabla\Theta^{t,x} := (\nabla X^{t,x}, \nabla Y^{t,x}, \nabla Z^{t,x})$ be the solution of the following equations

$$\nabla X_s^{t,x} = I_d + \int_t^s \nabla_x b(X_u^{t,x})\nabla X_u^{t,x} du + \int_t^s \nabla_x \sigma(X_u^{t,x})\nabla X_u^{t,x} dW_u \tag{2.7.8}$$

and

$$\begin{aligned}\nabla Y_s^{t,x} &= \nabla\Phi(X_T^{t,x})\nabla X_T^{t,x} + \int_s^T \Big(\nabla_x f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla X_r^{t,x} + \nabla_y f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla Y_r^{t,x}\\ &+ \nabla_z f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla Z_r^{t,x}\Big)dr + \int_s^T \Big(\nabla_x g(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla X_r^{t,x}\\ &+ \nabla_y g(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla Y_r^{t,x} + \nabla_z g(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\nabla Z_r^{t,x}\Big)\overleftarrow{dB_r^n} - \int_s^T \nabla Z_r^{t,x} dW_r.\end{aligned} \tag{2.7.9}$$

Since $\nabla X^{t,x}$ is the solution of the SDE (2.7.8), $[\nabla X^{t,x}]^{-1}$ is also the solution of an SDE and we have the following estimation

$$E[\sup_{0 \leqslant t \leqslant T} |[\nabla X_s^{t,x}]^{-1}|^p] \leqslant C_p. \tag{2.7.10}$$

On the other hand, $\nabla Y^{t,x}$ is the solution of the linear BDSDE (2.7.9). Using estimation (2.7.3), we get

$$E[\sup_{0 \leqslant t \leqslant T} |\nabla Y_s^{t,x}|^p] \leqslant C_p. \tag{2.7.11}$$

Now, let Us recall the following representation result (see [63], Proposition 2.3),

$$Z_s^{t,x} = \nabla Y_s^{t,x}[\nabla X_s^{t,x}]^{-1}\sigma(X_u^{t,x}), P-a.s.\ \forall s \in [t,T]. \tag{2.7.12}$$

Using the Hölder inequality, we get

$$\begin{aligned}\|Z_s^{t,x}\|_p &\leqslant \|\nabla Y_s^{t,x}\|_{3p}\|[\nabla X_s^{t,x}]^{-1}\|_{3p}\|\sigma(X_u^{t,x})\|_{3p}\\ &\leqslant C_p(1+|x|^2),\ \forall s \in [t,T].\end{aligned} \tag{2.7.13}$$

Now the aim is to generalize the previous estimation to Lipschitz coefficients case. So let b, σ, Φ, f and g satisfying Assumptions (**H1**) and (**H2**) and let $b^k, \sigma^k, \Phi^k, f^k$ and g^k smooth molifiers of these functions. Denoting $Z^{t,x,k}$ the solution of the F-BDSDE associated to the regular coefficients, we deduce from (2.7.13) that $\|Z_s^{k,t,x}\|_p \leqslant C_p(1+|x|^2)$, $\forall s \in [t,T]$, where C_p is independent from k. Using the stability result (2.7.5) , we get

$$\lim_{k\longrightarrow+\infty} E\Big[\int_t^T |Z_s^{k,t,x} - Z_s^{t,x}|^2 ds\Big] = 0. \tag{2.7.14}$$

We deduce that for a.e. $s \in [t,T]$, there exist a subsequence of $(Z^{k,t,x})_k$ such that $\lim_{k\longrightarrow+\infty} Z_s^{k,t,x} = Z_s^{t,x}$ in probability. By the Fatou's Lemma, we get $\|Z_s^{t,x}\|_p \leqslant C_p(1+|x|^2)$. Inserting the latter inequality in estimation (2.7.4), we get the estimation (2.7.7). □

The following theorem states the main result of this section, which is the extension of the Zhang L^2-regularity to our case.

Theorem 2.7.1. *Under Assumptions* (**H1**) *and* (**H2**), *we have the following estimation*

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} \Big\{\|Z_s - Z_{t_n}\|^2 + \|Z_s - Z_{t_{n+1}}\|^2\Big\}ds\Big] \leqslant Ch(1+|x|^2) \tag{2.7.15}$$

Proof. Let $\pi' : 0 = t_0, \dots, t_N = T$ denote our uniform partitipon of $[0,T]$ with time step h and $\pi : 0 = s_0, \dots, s_m = T$ any partition finer than π. Without loss of generality, we assume that $s_{l_i} = t_i$ for $i = 1, \dots, N$. We will prove the theorem for π'. Let Φ^π, f^π, h^π and $g^\pi \in C_b^1$ smooth molifiers of Φ, f, h and g, such that all the derivatives are bounded by K. We denote by $\Theta^\pi = (W, Y^\pi, Z^\pi)$ the solution of the following F-BDSDE :

$$X_s^\pi = x + \int_t^s b^\pi(X_u^\pi)ds + \int_t^s \sigma^\pi(X_u^\pi)dW_u, \quad s \in [t,T] \tag{2.7.16}$$

$$Y_s^\pi = \Phi(X_T^\pi) + \int_t^s f^\pi(u, \Theta^\pi)ds + \int_t^s g^\pi(u, \Theta^\pi)\overleftarrow{dB_u} - \int_t^s Z_u^\pi dW_u. \tag{2.7.17}$$

By the stability result (2.7.5), we have

$$\lim_{|\pi| \longrightarrow 0} E\Big\{\sup_{t \leqslant s \leqslant T}|X_s^\pi - X_s|^2 + \sup_{t \leqslant s \leqslant T}|Y_s^\pi - Y_s|^2 + \int_t^T |Z_s^\pi - Z_s|^2 ds\Big\} = 0. \tag{2.7.18}$$

By (2.7.18), there exists a subsequence denoted again by π such that $\lim_{|\pi| \longrightarrow 0} E|Z_s^\pi - Z_s|^2 = 0$ for a.e. s $\in [t,T]$. Let Us note that, for $s \in [t_n, t_{n+1})$, we have

$$\begin{aligned} E|Z_s - Z_{t_n}|^2 + E|Z_s - Z_{t_{n+1}}|^2 &\leqslant CE\Big\{|Z_s - Z_s^\pi|^2 + |Z_s^\pi - Z_{t_n}^\pi|^2 \\ &+ |Z_s^\pi - Z_{t_{n+1}}^\pi|^2\Big\}. \end{aligned} \tag{2.7.19}$$

By (2.7.18), proving the theorem remains to estimate $E|Z_s^\pi - Z_{t_n}^\pi|^2$ and $E|Z_s^\pi - Z_{t_{n+1}}^\pi|^2$ for $s \in [t_n, t_{n+1})$.

To this end, we denote by $(\nabla^\pi X, \nabla^\pi Y)$ the solution of the linear equations (2.7.8) – (2.7.9) with coefficients replaced by $\Phi^\pi, b^\pi, \sigma^\pi, f^\pi$ and g^π.

Using the representaion result (2.7.12) for (Z^π), we have

$$Z_s^\pi - Z_{s'}^\pi = \nabla Y_s^\pi[\nabla X_s^\pi]^{-1}\sigma^\pi(X_s^\pi) - \nabla Y_{s'}^\pi[\nabla X_{s'}^\pi]^{-1}\sigma^\pi(X_{s'}^\pi), s, s' \in [t_n, t_{n+1}). \tag{2.7.20}$$

Then

$$\begin{aligned} |Z_s^\pi - Z_{s'}^\pi|^2 &\leqslant 3|\nabla Y_s^\pi - \nabla Y_{s'}^\pi|^2|[\nabla X_s^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi)|^2 \\ &+ 3|\nabla Y_{s'}^\pi|^2|[\nabla X_s^\pi]^{-1} - [\nabla X_{s'}^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi)|^2 \\ &+ 3|\nabla Y_{s'}^\pi|^2|[\nabla X_{s'}^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi) - \sigma^\pi(X_{s'}^\pi)|^2. \end{aligned} \tag{2.7.21}$$

Thus, we get

$$\begin{aligned} |Z_s^\pi - Z_{t_n}^\pi|^2 &\leqslant C\Big\{|\nabla Y_s^\pi - \nabla Y_{t_n}^\pi|^2|[\nabla X_s^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi)|^2 \\ &+ |\nabla Y_{t_n}^\pi|^2|[\nabla X_s^\pi]^{-1} - [\nabla X_{t_n}^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi)|^2 \\ &+ |\nabla Y_{t_n}^\pi|^2|[\nabla X_{t_n}^\pi]^{-1}|^2|\sigma^\pi(X_s^\pi) - \sigma^\pi(X_{t_n}^\pi)|^2\Big\}. \end{aligned} \tag{2.7.22}$$

We conclude by using the Hölder's inequality and the estimation (2.7.7) that

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} |Z_s^\pi - Z_{t_n}^\pi|^2 ds\Big] \leqslant Ch(1 + |x|^2),$$

here we used also the same kind of estimation as (2.7.4) but for $[\nabla X^\pi]^{-1}$ (instead of ∇Y_s^π) as it is a solution of an SDE.

By the same arguments used on $E|Z_s^\pi - Z_{t_{n+1}}^\pi|^2$ (taking $s' = t_{n+1}$), we conclude that

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} |Z_s^\pi - Z_{t_{n+1}}^\pi|^2 ds\Big] \leqslant Ch(1 + |x|^2).$$

This gives

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s^\pi - Z_{t_n}^\pi|^2 + |Z_s^\pi - Z_{t_{n+1}}^\pi|^2\Big\}ds\Big] \leqslant Ch(1+|x|^2). \tag{2.7.23}$$

Using (2.7.19), we obtain

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s - Z_{t_n}|^2 + |Z_s - Z_{t_{n+1}}|^2\Big\}ds\Big] \leqslant CE\Big[\int_t^T |Z_s - Z_s^\pi|^2 ds\Big] + Ch(1+|x|^2). \tag{2.7.24}$$

Recalling (2.7.18) and letting $|\pi| \longrightarrow 0$, we finally get

$$\sum_{n=0}^{N-1} E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s - Z_{t_n}|^2 + |Z_s - Z_{t_{n+1}}|^2\Big\}ds\Big] \leqslant Ch(1+|x|^2). \tag{2.7.25}$$

□

Now, we are able to derive the rate of convergence of our scheme under Lipschitz assumptions.

Corollary 2.7.1. *Under Assumptions* (**H1**) *and* (**H2**), *we have*

$$\begin{aligned} Error_N(Y,Z) &:= \sup_{0 \leqslant s \leqslant T} E[|Y_s - Y_s^N|^2] + \sum_{n=0}^{N-1} E[\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds] \\ &\leqslant Ch(1+|x|^2). \end{aligned} \tag{2.7.26}$$

Proof. First we recall that under (**H1**) and (**H2**), we have by (2.3.6)

$$\begin{aligned} Error_N(Y,Z) &\leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[|Z_s - \bar{Z}_{t_n}|^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds. \end{aligned} \tag{2.7.27}$$

Then, as the conditional expectation minimizes the conditional mean square error, we have

$$\int_{t_n}^{t_{n+1}} E|Z_s - \bar{Z}_{t_n}|^2 ds \quad \leqslant \quad \int_{t_n}^{t_{n+1}} E|Z_s - Z_{t_n}|^2 ds.$$

On the other hand, plugging $Z_{t_{n+1}}$ in the following, we get

$$\int_{t_n}^{t_{n+1}} E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds \leqslant C\int_{t_n}^{t_{n+1}} E[|Z_s - Z_{t_{n+1}}|^2]ds + ChE[|Z_{t_{n+1}} - \bar{Z}_{t_{n+1}}|^2]$$

By the definition of $\bar{Z}_{t_{n+1}}$, Jensen's inequality and Cauchy-Schwarz inequality, we have for all $n = 0, \ldots, N-2$

$$\begin{aligned} hE[|Z_{t_{n+1}} - \bar{Z}_{t_{n+1}}|^2] &= \frac{1}{h}E\Big\{\Big|E_{t_{n+1}}\Big[\int_{t_{n+1}}^{t_{n+2}}\{Z_{t_{n+1}} - Z_s\}ds\Big]\Big|^2\Big\} \\ &\leqslant \frac{1}{h}E\Big\{\Big|\Big[\int_{t_{n+1}}^{t_{n+2}}\{Z_{t_{n+1}} - Z_s\}ds\Big]\Big|^2\Big\} \\ &\leqslant E\int_{t_{n+1}}^{t_{n+2}} |Z_{t_{n+1}} - Z_s|^2 ds. \end{aligned}$$

Hence, the inequality (2.7.27) becomes

$$\begin{aligned} Error_N(Y,Z) &\leqslant Ch(1+|x|^2) + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[|Z_s - Z_{t_n}|^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E[|Z_s - Z_{t_{n+1}}|^2]ds. \end{aligned}$$

We conclude the rate of convergence by using Theorem 2.7.1.

2.8 Appendix

2.8.1 Proof of Proposition 2.4.2.

To simplify the notations, we restrict ourselves to the case $k = d = l = 1$. $(D_\theta Y, D_\theta Z)$ is well defined and from inequalities (2.2.4) and (2.4.1), we deduce that for each $\theta \leqslant T$

$$E[\sup_{t \leqslant s \leqslant T}|D_\theta Y_s|^2] + E[\int_t^T |D_\theta Z_s|^2 ds] \leqslant C(1+|x|^2). \tag{2.8.1}$$

We define recursively the sequence (Y^m, Z^m) as follows. First we set $(Y^0, Z^0) = (0,0)$. Then, given (Y^{m-1}, Z^{m-1}), we define (Y^m, Z^m) as the unique solution in $\mathbb{S}^2_k([t,T]) \times \mathbb{H}^2_{k\times d}([t,T])$ of

$$Y^m_s = \Phi(X^{t,x}_T) + \int_s^T f(r, X^{t,x}_r, Y^{m-1}_r, Z^{m-1}_r)dr + \int_s^T g(r, X^{t,x}_r, Y^{m-1}_r, Z^{m-1}_r)\overleftarrow{dB_r} - \int_s^T Z^m_r dW_r.$$

We recursively show that $(Y^m, Z^m) \in \mathcal{B}^2([t,T], \mathbb{D}^{1,2})$. Suppose that $(Y^m, Z^m) \in \mathcal{B}^2([t,T], \mathbb{D}^{1,2})$ and let us show that $(Y^{m+1}, Z^{m+1}) \in \mathcal{B}^2([t,T], \mathbb{D}^{1,2})$.
From the induction assumption, we have $\Phi(X_T) + \int_s^T f(r, \Sigma^m_r)dr \in \mathbb{D}^{1,2}$.
We have $g(r, \Sigma^m_r) \in \mathbb{D}^{1,2}$ for all $r \in [t,T]$. From Lemma 2.4.2, we have $\int_t^T g(r, \Sigma^m_r)\overleftarrow{dB_r} \in \mathbb{D}^{1,2}$. then

$$Y^{m+1}_s = E\big[\Phi(X^{t,x}_T) + \int_s^T f(r, \Sigma^m_r)dr + \int_s^T g(r, \Sigma^m_r)\overleftarrow{dB_r}|\mathcal{F}^W_{t,s} \vee \mathcal{F}^B_{t,T}\big] \in \mathbb{D}^{1,2},$$

where $\Sigma^m_r := (X^{t,x}_r, Y^m_r, Z^m_r)$.
Hence

$$\int_t^T Z^{m+1}_r dW_r = \Phi(X^{t,x}_T) + \int_t^T f(r, \Sigma^m_r)dr + \int_t^T g(r, \Sigma^m_r)\overleftarrow{dB_r} - Y^{m+1}_t \in \mathbb{D}^{1,2}.$$

It follows from Lemma 2.4.1 that $Z^{m+1} \in \mathcal{M}^2_{k\times d}([t,T], \mathbb{D}^{1,2})$ and we have $D_\theta Y^{m+1}_s = D_\theta Z^{m+1}_s = 0$ for $t \leqslant s \leqslant \theta$ and for $\theta \leqslant s \leqslant T$

$$\begin{aligned} D_\theta Y^{m+1}_s &= \nabla\Phi(X^{t,x}_T)D_\theta X^{t,x}_T \\ &+ \int_s^T \Big(\nabla_x f(r, \Sigma^m_r)D_\theta X_r + \nabla_y f(r, \Sigma^m_r)D_\theta Y^m_r + \nabla_z f(r, \Sigma^m_r)D_\theta Z^m_r\Big)dr \\ &+ \int_s^T \Big(\nabla_x g(r, \Sigma^m_r)D_\theta X_r + \nabla_y g(r, \Sigma^m_r)D_\theta Y^m_r + \nabla_z g(r, \Sigma^m_r)D_\theta Z^m_r\Big)\overleftarrow{dB_r} \\ &- \int_s^T D_\theta Z^{m+1}_r dW_r. \end{aligned} \tag{2.8.2}$$

From inequality (2.2.4), we deduce that for each $\theta \leqslant T$

$$E[\sup_{t \leqslant s \leqslant T}|D_\theta Y^{m+1}_s|^2] + E[\int_t^T |D_\theta Z^{m+1}_s|^2 ds] \leqslant C(1+|x|^2).$$

It is known that inequality (2.2.4) holds for (Y^{m+1}, Z^{m+1}) and so we deduce that

$$\|Y^{m+1}\|_{1,2} + \|Z^{m+1}\|_{1,2} < \infty,$$

which shows that $(Y^{m+1}, Z^{m+1}) \in \mathcal{B}^2([t,T],\mathbb{D}^{1,2})$. Using the contraction mapping argument as in El Karoui Peng and Quenez [31], we deduce that (Y^{m+1}, Z^{m+1}) converges to (Y,Z) in $\mathbb{S}^2([t,T]) \times \mathbb{H}^2([t,T])$. We will show that $(D_\theta Y^m, D_\theta Z^m)$ converges to (Y^θ, Z^θ) in $L^2(\Omega \times [t,T] \times [t,T], dP \otimes dt \otimes dt)$, where $Y_s^\theta = Z_s^\theta = 0$ for all $t \leqslant s \leqslant \theta$ and $(Y_s^\theta, Z_s^\theta, \theta \leqslant s \leqslant T)$ is the solution of the BDSDE.

$$\begin{aligned} Y_s^\theta &= \nabla\Phi(X_T^{t,x}) D_\theta X_T^{t,x} \\ &+ \int_s^T \Big(\nabla_x f(r,\Sigma_r) D_\theta X_r + \nabla_y f(r,\Sigma_r) Y_r^\theta + \nabla_z f(r,\Sigma_r) Z_r^\theta\Big) dr \\ &+ \int_s^T \Big(\nabla_x g(r,\Sigma_r) D_\theta X_r + \nabla_y g(r,\Sigma_r) Y_r^\theta + \nabla_z g(r,\Sigma_r) Z_r^\theta\Big) \overleftarrow{dB_r} \\ &- \int_s^T Z_r^\theta dW_r. \end{aligned} \tag{2.8.3}$$

From equations (2.8.2) and (2.8.3), we have

$$\begin{aligned} D_\theta Y_s^{m+1} - Y_s^\theta &= \int_s^T \Big((\nabla_x f(r,\Sigma_r^m) - \nabla_x f(r,\Sigma_r)) D_\theta X_r^{t,x} \\ &+ \nabla_y f(r,\Sigma_r^m) D_\theta Y_r^m - \nabla_y f(r,\Sigma_r) Y_r^\theta + \nabla_z f(r,\Sigma_r^m) D_\theta Z_r^m - \nabla_z f(r,\Sigma_r) Z_r^\theta \Big) dr \\ &+ \int_s^T \Big((\nabla_x g(r,\Sigma_r^m) - \nabla_x g(r,\Sigma_r)) D_\theta X_r^{t,x} + \nabla_y g(r,\Sigma_r^m) D_\theta Y_r^m - \nabla_y g(r,\Sigma_r) Y_r^\theta \Big) \overleftarrow{dB_r} \\ &+ \int_s^T \Big(\nabla_z g(r,\Sigma_r^m) D_\theta Z_r^m - \nabla_z g(r,\Sigma_r) Z_r^\theta \Big) \overleftarrow{dB_r} \\ &- \int_s^T (D_\theta Z_r^{m+1} - Z_r^\theta) dW_r. \end{aligned}$$

From Proposition 2.4.1, we have

$$\begin{aligned} & E[\sup_{\theta \leqslant s \leqslant T} |D_\theta Y_s^{m+1} - Y_s^\theta|^2] + E[\int_s^T |D_\theta Z_r^{m+1} - Z_r^\theta|^2 dr] \\ &\leqslant CE\Big[\int_s^T \Big|(\nabla_x f(r,\Sigma_r^m) - \nabla_x f(r,\Sigma_r)) D_\theta X_r^{t,x} + \nabla_y f(r,\Sigma_r^m) Y_r^\theta - \nabla_y f(r,\Sigma_r) Y_r^\theta \\ &+ \nabla_z f(r,\Sigma_r^m) Z_r^\theta - \nabla_z f(r,\Sigma_r) Z_r^\theta\Big|^2 dr\Big] \\ &+ CE\Big[\int_s^T \Big|(\nabla_x g(r,\Sigma_r^m) - \nabla_x g(r,\Sigma_r)) D_\theta X_r + \nabla_y g(r,\Sigma_r^m) Y_r^\theta - \nabla_y g(r,\Sigma_r) Y_r^\theta \\ &+ \nabla_z g(r,\Sigma_r^m) Z_r^\theta - \nabla_z g(r,\Sigma_r) Z_r^\theta\Big|^2 dr\Big]. \end{aligned} \tag{2.8.4}$$

Therefore, we obtain

$$\begin{aligned} & E[\int_t^T \int_t^T |D_\theta Y_s^{m+1} - Y_s^\theta|^2 ds d\theta] + E[\int_t^T \int_t^T |D_\theta Z_s^{m+1} - Z_s^\theta|^2 ds d\theta] \\ \leqslant \; & CE[\int_t^T \int_t^T |\delta_{r,\theta}^m|^2 dr d\theta] + CE[\int_t^T \int_t^T |\rho_{r,\theta}^m|^2 dr d\theta], \end{aligned} \tag{2.8.5}$$

where

$$\begin{aligned} \delta_{r,\theta}^m &= (\nabla_x f(r,\Sigma_r^m) - \nabla_x f(r,\Sigma_r)) D_\theta X_r^{t,x} + \nabla_y f(r,\Sigma_r^m) Y_r^\theta - \nabla_y f(r,\Sigma_r) Y_r^\theta \\ &+ \nabla_z f(r,\Sigma_r^m) Z_r^\theta - \nabla_z f(r,\Sigma_r) Z_r^\theta, \end{aligned} \tag{2.8.6}$$

and

$$\begin{aligned} \rho^m_{r,\theta} &= (\nabla_x g(r,\Sigma^m_r) - \nabla_x g(r,\Sigma_r))D_\theta X^{t,x}_r + \nabla_y g(r,\Sigma^m_r)Y^\theta_r - \nabla_y g(r,\Sigma_r)Y^\theta_r \\ &+ \nabla_z g(r,\Sigma^m_r)Z^\theta_r - \nabla_z g(r,\Sigma_r)Z^\theta_r. \end{aligned} \tag{2.8.7}$$

From the definition of $(\delta^m_{r,\theta})_{t\,\leqslant\, r,\theta\,\leqslant\, T}$, we have $E[\int_t^T\int_t^T|\delta^m_{r,\theta}|^2drd\theta]\leqslant C\int_t^T(A_m(\theta,t,T)+B_m(\theta,t,T))d\theta$, where

$$\begin{aligned} A_m(\theta,t,T) &= E\Big[\int_t^T |(\nabla_x f(r,\Sigma^m_r) - \nabla_x f(r,\Sigma_r))D_\theta X^{t,x}_r|^2dr\Big] \\ B_m(\theta,t,T) &= E\Big[\int_t^T \big|(\nabla_y f(r,\Sigma_r) - \nabla_y f(r,\Sigma^m_r))Y^\theta_r\big|^2dr\Big] \\ &+ E\Big[\int_t^T \big|(\nabla_z f(r,\Sigma_r) - \nabla_z f(r,\Sigma^m_r))Z^\theta_r\big|^2dr\Big] \end{aligned}$$

Moreover, since $\nabla_x f$ is bounded and continuous with respect to (x,y,z), it follows by the dominated convergence theorem and inequality (2.2.3) that

$$\lim_{m\to\infty}\int_t^T A_m(\theta,t,T)d\theta = 0. \tag{2.8.8}$$

Furthermore, since $\nabla_y f$ and $\nabla_z f$ are bounded and continuous with respect to (x,y,z), it follows, also, by the dominated convergence theorem and inequality (2.2.4) that

$$\lim_{m\to\infty}\int_t^T B_m(\theta,t,T)d\theta = 0. \tag{2.8.9}$$

From the definition of $(\rho^m_{r,\theta})_{s\,\leqslant\, r,\theta\,\leqslant\, T}$, we have

$$E[\int_t^T\int_t^T|\rho^m_{r,\theta}|^2drd\theta]\leqslant C\int_t^T(A'_m(\theta,t,T)+B'_m(\theta,t,T))d\theta$$

with

$$\begin{aligned} A'_m(\theta,t,T) &= E\Big[\int_t^T |(\nabla_x g(r,\Sigma^m_r) - \nabla_x g(r,\Sigma_r))D_\theta X^{t,x}_r|^2dr\Big] \\ B'_m(\theta,t,T) &= E\Big[\int_t^T \big|(\nabla_y g(r,\Sigma_r) - \nabla_y g(r,\Sigma^m_r))Y^\theta_r\big|^2dr\Big] \\ &+ E\Big[\int_t^T \big|(\nabla_z g(r,\Sigma_r) - \nabla_z g(r,\Sigma^m_r))Z^\theta_r\big|^2dr\Big]. \end{aligned}$$

Similarly as shown above, since $\nabla_y g$ and $\nabla_z g$ are bounded and continuous with respect to (x,y,z) we can show that :

$$\lim_{m\to\infty}\int_t^T A'_m(\theta,t,T)d\theta = \lim_{m\to\infty}\int_t^T B'_m(\theta,t,T)d\theta = 0. \tag{2.8.10}$$

Plugging (2.8.8), (2.8.9) and (2.8.10) into inequality (2.8.5), we deduce that

$$\lim_{m\to\infty} E[\int_t^T\int_t^T|D_\theta Y^{m+1}_s - Y^\theta_s|^2dsd\theta] + E[\int_t^T\int_t^T|D_\theta Z^{m+1}_s - Z^\theta_s|^2dsd\theta] = 0.$$

It follows that (Y^m,Z^m) converges to (Y,Z) in $L^2([t,T],\mathbb{D}^{1,2}\times\mathbb{D}^{1,2})$ and a version of $(D_\theta Y, D_\theta Z)$ is given by (Y^θ,Z^θ) which is the desired result. □

2.8.2 Second order Malliavin derivative of the solution of BDSDE's

We apply know similar computation to get the second order Malliavin derivatives representation of the solution of BDSDE 's, so we will omit the proof.

Proposition 2.8.1. *Assume that assumptions* **(H2)** *and* **(H3)** *hold. We fix* $t \in [0,T]$*. Then for each* $t \leqslant \theta \leqslant T$*,* $(D_\theta Y, D_\theta Z)$ *belongs to* $\mathcal{B}^2([t,T],\mathbb{D}^{1,2})$*. For each* $t \leqslant v \leqslant T$ *and* $1 \leqslant i,j \leqslant d$*,*

$$D^j_v D^i_\theta Y_s = D^j_v D^i_\theta Z^n_s = 0,\ 1 \leqslant n \leqslant d,\ \textit{if } s < \theta \vee v,$$

and a version of $(D^j_v D^i_\theta Y_s, D^j_v D^i_\theta Z_s)_{v\vee\theta \leqslant s \leqslant T}$ *is the unique solution of the equation :*

$$D^j_v D^i_\theta Y_s = T_1(\Phi) + T_2(f) + T_3(g) + T_4(W),$$

where

$$T_1(\Phi) = \sum_{n_1=1}^{k} \nabla((\nabla\Phi)^{n_1}(X^{t,x}_T))D^j_v X^{t,x}_T (D^i_\theta X^{t,x}_T)^{n_1} + \nabla\Phi(X^{t,x}_T)D^j_v D^i_\theta X^{t,x}_T,$$

$$\begin{aligned}
T_2(f) &= \int_s^T \sum_{n_1=1}^{k} \Big(\nabla_x((\nabla_x f)^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v X^{t,x}_r (D^i_\theta X^{t,x}_r)^{n_1} \\
&+ \nabla_x f(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta X^{t,x}_r\Big)dr \\
&+ \int_s^T \Big(\sum_{n_1=1}^{k} \nabla_y((\nabla_y f)^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v Y_r (D^i_\theta Y_r)^{n_1} \\
&+ \nabla_y f(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta Y_r\Big)dr \\
&+ \sum_{n_2=1}^{d} \int_s^T \sum_{n_1=1}^{k} \nabla_{z^{n_2}}((\nabla_{z^{n_2}} f)^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v Z^{n_2}_r (D^i_\theta Z^{n_2}_r)^{n_1} dr \\
&+ \sum_{n_2=1}^{d} \int_s^T \nabla_{z^{n_2}} f(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta Z^{n_2}_r dr,
\end{aligned}$$

$$\begin{aligned}
T_3(g) &= \sum_{n_3=1}^{l} \int_s^T \sum_{n_1=1}^{k} \nabla_x((\nabla_x g^{n_3})^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v X^{t,x}_r (D^i_\theta X^{t,x}_r)^{n_1} \overleftarrow{dB^{n_3}_r} \\
&+ \sum_{n_3=1}^{l} \int_s^T \nabla_x g^{n_3}(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta X^{t,x}_r \overleftarrow{dB^{n_3}_r} \\
&+ \sum_{n_3=1}^{l} \int_s^T \sum_{n_1=1}^{k} \nabla_y((\nabla_y g^{n_3})^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v Y_r (D^i_\theta Y_r)^{n_1} \overleftarrow{dB^{n_3}_r} \\
&+ \sum_{n_3=1}^{l} \int_s^T \nabla_y g^{n_3}(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta Y_r \overleftarrow{dB^{n_3}_r} \\
&+ \sum_{n_3=1}^{l} \sum_{n_2=1}^{d} \int_s^T \sum_{n_1=1}^{k} \nabla_{z^{n_2}}((\nabla_{z^{n_2}} g^{n_3})^{n_1}(r,X^{t,x}_r,Y_r,Z_r))D^j_v Z^{n_2}_r (D^i_\theta Z^{n_2}_r)^{n_1} \overleftarrow{dB^{n_3}_r} \\
&+ \sum_{n_3=1}^{l} \sum_{n_2=1}^{d} \int_s^T \nabla_{z^{n_2}} g^{n_3}(r,X^{t,x}_r,Y_r,Z_r)D^j_v D^i_\theta Z^{n_2}_r \overleftarrow{dB^{n_3}_r},
\end{aligned}$$

$$T_4(W) = -\sum_{n_2=1}^{d} \int_s^T D_v^j D_\theta^i Z_r^{n_2} dW_r^{n_2},$$

$(z^j)_{1 \leqslant j \leqslant d}$ *denotes the j-th column of the matrix z,* $(g^{n_3})_{1 \leqslant n_3 \leqslant l}$ *denotes the* n_3*-th column of the matrix g,* $B = (B^1, \dots, B^l)$, $(D_\theta^i X_r^{t,x})^{n_1}$ *is the* n_1*-th component of the vector* $(D_\theta^i X_r^{t,x})$, $(D_\theta^i Y_r)^{n_1}$ *is the* n_1*-th component of the vector* $(D_\theta^i Y_r)$ *and* $(D_\theta^i Z_r^{n_2})^{n_1}$ *is the* n_1*-th component of the vector* $(D_\theta^i Z_r^{n_2})$.

2.8.3 Some estimates on the solution of the FBDSDE

Lemma 2.8.1. *Let* (b^1, σ^1) *and* (b^2, σ^2) *be the standard parameters of the SDE* (2.2.1) *with initial condition* x^1 *(resp.* x^2*). We assume that* **(H1)** *holds. Put* $\delta X_s = X_s^1 - X_s^2$, $\delta b_s = (b^1 - b^2)(X_s^1)$ *and* $\delta\sigma_s = (\sigma^1 - \sigma^2)(X_s^1)$*. Then*

$$||X^1||_{\mathbb{S}_d^2} \leqslant C(1 + |x|^2).$$

For all $s_1, s_2 \in [0, T]$*, we have*

$$E\Big[\sup_{s_1 \leqslant u \leqslant s_2} |X_u^1 - X_{s_1}^1|\Big] \leqslant C(1 + |x|^2)|s_2 - s_1|,$$

and for all $s_1 \leqslant s \leqslant s_2$*, we have*

$$||\delta X||_{\mathbb{S}_d^2([s_1,s_2])} \leqslant C\Big(|x^1 - x^2|^2 + |s_2 - s_1| + E\big[\int_{s_1}^{s_2} |\delta b_s|^2 + |\delta\sigma_s|^2 ds\big]\Big),$$

where C is a generic constant depending only on K, T, $(b^1(0), \sigma^1(0))$ *and* $(b^2(0), \sigma^2(0))$.

Lemma 2.8.2. *Let* $(X^{t,x}, Y^{t,x}, Z^{t,x})$ *be the solution of the FBDSDE* (2.2.1)-(2.2.2)*. We assume that Assumptions* **(H1)** *and* **(H2)** *hold. Then, we have*

$$||Y^{t,x}||_{\mathbb{S}_d^2} + ||Z^{t,x}||_{\mathbb{H}_{d\times k}^2} \leqslant C(1 + |x|^2), \tag{2.8.11}$$

and for all $s', s \in [t, T], s' \leqslant s$*, we have*

$$E\Big[\sup_{s' \leqslant u \leqslant s} |Y_u^{t,x} - Y_{s'}^{t,x}|^2\Big] \leqslant C\Big((1 + |x|^2)|s - s'| + ||Z^{t,x}||_{M_{k\times d}^2[s',s]}\Big). \tag{2.8.12}$$

Proof. The technics used to prove these estimates are classical in the BSDE's theory (see El Karoui et al.[31]) so we omit it.

□

Chapitre 3

Numerical computations for Quasilinear Stochastic PDEs

3.1 Introduction

We consider the following stochastic PDE, in $\mathbb{R}^d$,

$$du_t(x) + \Big[\, \frac{1}{2}\Delta u_t(x) + f_t(x, u_t(x), \nabla u_t(x)) + \mathrm{div} g_t\left(x, u_t\left(x\right), \nabla u_t\left(x\right)\right) \Big]\, dt \\ + h_t(x, u_t(x), \nabla u_t(x)) \cdot \overleftarrow{dB}_t = 0, \tag{3.1.1}$$

over the time interval $[0, T]$, with a given final condition $u_T = \Phi$ and $f, g = \big(g_1, \cdots, g_d\big)$, $h = \big(h_1, \cdots, h_{d^1}\big)$ non-linear random functions. The differential term with $\overleftarrow{dB}_t$ refers to the backward stochastic integral with respect to a d^1-dimensional Brownian motion on $\big(\Omega, \mathcal{F}, \mathbb{P}, (B_t)_{t\,\geqslant\,0}\big)$. We use the backward notation because in the proof we will employ the doubly stochastic framework introduced by Pardoux and Peng [63].

When h is identically null and g is not, we obtain a Quasilinear deterministic PDE. The latter equation was studied by Stoica [68], who gave the probabilistic interpretation for such equation. In fact, Stoica studied a more general case. He considered the equation 3.1.1, but with the elliptic divergence form operator $\mathcal{L}_0$ instead of the operator $\frac{1}{2}\Delta$. The equation he studied was of the form :

$$(\partial_t + \mathcal{L}_0)u + f - \sum_{i,j} \partial_i(a^{i,j} g_j) = 0, \tag{3.1.2}$$

where f and g are given real-valued functions on $[0, T] \times \mathbb{R}^d$, $\mathcal{L}_0$ is the given by :

$$\mathcal{L}_0 := \sum_{i,j} \partial_i(a^{i,j}_j) + \sum_i b^i \partial_i,$$

$b(x) := (b^1(x), \ldots, b^d(x))$ is a vector field and for all i, j, $a^{i,j}$ are bounded measurable functions on $\mathbb{R}^d$. He proved that under uniform ellipticity assumption on the matrix a and if the coefficients f, g, h and Φ are square integrable, the solution of (3.1.2) satisfy the following relation :

$$u_t(X_t) - u_s(X_s) = \sum_{i=1}^{d} \int_s^t \partial_i u_r(X_r) dM^i_r - \int_s^t f_r(X_r)dr - \frac{1}{2}\int_s^t g_r * dX_r, \tag{3.1.3}$$

where $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, X_t, \theta_t, P^x)$ is the diffusion process generated by $\mathcal{L}_0$ in $\mathbb{R}^d$, M^i is the martingale part of the component X^i of the process and the integral denoted with $*$ is a stochastic martingale expressed in terms of forward and backward martingales.

When $\mathcal{L}_0$ is symmetric under the probability measure P^m (see subsection **??** for the rigorous definition of P^m), we have :

$$\int_s^t g_r * dX_r = \sum_{i=1}^{d} \left(\int_s^t g_i(r, X_r) dM^i_r + \int_s^t g_i(r, X_r) \overleftarrow{dM^i_r} \right).$$

Hence, Stoica generalized the BSDE's method for semilinear PDEs to a quasilinear case, by giving the interpretation of the term $\sum_{i,j} \partial_i(a^{i,j} g_j)$ in terms of the stochastic integral $*$.

When h and g are non-null functions, equation (3.1.1) is a quasilinear stochastic PDE and will be of special interest in this chapter. The probabilistic interpretation for the SPDE (3.1.1) was given by Matoussi and

Stoica [54]. They proved also by the mean of a useful change of variable that, when the matrix a is assumed symmetric, the case when we have $\mathcal{L} := \sum_{i,j} \partial_i(a_j^{i,j})$ instead of the operator $\frac{1}{2}\Delta$ in (3.1.1) is covered by their framework.
Since we are mainly interested in the numerical resolution of SPDE (3.1.1) by a probabilistic method, let Us give an overview of the probabilistic methods for PDEs in the literature. In the deterministic PDE's case when $g \equiv 0$ and $h \equiv 0$, the numerical approximation of the BSDE has already been studied in the literature by Bally [8], Zhang [70], Bouchard and Touzi [16], Gobet, Lemor and Warin[34] and Bouchard and Elie [15]. Zhang [70] proposed a discrete-time numerical approximation, by step processes, for a class of decoupled FBSDEs with possible path-dependent terminal values. He proved an L^2-type regularity of the BSDE's solution, the convergence of his scheme and he derived its rate of convergence. Bouchard and Touzi [16] suggested a similar numerical scheme for decoupled FBSDEs. The conditional expectations involved in their discretization scheme were computed by using the Malliavin approach and the Monte carlo method. Crisan, Manolarakis and Touzi [20] proposed an improvement on the Malliavin weights. Gobet, Lemor and Warin in [34] proposed an explicit numerical scheme. When $g \equiv 0$, $h \neq 0$ and it does not depend on the control variable z, Aman [3] proposed a numerical scheme following the idea used by Bouchard and Touzi [16] and obtained a convergence of order the time discretization step of the square of the L^2- error. Aboura [1] studied the same numerical scheme under the same kind of hypothesis, but following Gobet et al. [33]. He obtained a convergence of order the time discretization step and used the regression Monte Carlo method to implement his scheme, following always [33].

In this work, we explore a probabilistic numerical method to approximate the solution of equation (3.1.1), which have been studied in particular by Denis and Stoica [28], Denis-Matoussi-Stoica [25, 26, 27] and Matoussi -Stoica [54].

3.2 Preliminaries and notations

3.2.1 Quasilinear SPDEs : Theorotical aspect

The basic Hilbert space of our framework is $\mathbf{L}^2(\mathbb{R}^d)$ and we employ the usual notation for its scalar product and its norm,

$$(u,v) = \int_{\mathbb{R}^d} u(x)\, v(x)\, dx,\ \|u\|_2 = \left(\int_{\mathbb{R}^d} u^2(x)\, dx\right)^{\frac{1}{2}}.$$

In general, we shall use the notation

$$(u,v) = \int_{\mathbb{R}^d} u(x)v(x)\, dx,$$

where u, v are measurable functions defined in $\mathbb{R}^d$ and $uv \in \mathbf{L}^1(\mathbb{R}^d)$.

Our evolution problem will be considered over a fixed time interval $[0,T]$ and the norm for a function $L^2([0,T]\times\mathbb{R}^d)$ will be denoted by

$$\|u\|_{2,2} = \left(\int_0^T \int_{\mathbb{R}^d} |u(t,x)|^2 dxdt\right)^{\frac{1}{2}}.$$

Another Hilbert space that we use is the first order Sobolev space $H^1(\mathbb{R}^d) = H^1_0(\mathbb{R}^d)$. Its natural scalar product and norm are

$$(u,v)_{H^1(\mathbb{R}^d)} = (u,v) + (\nabla u, \nabla v),\ \|u\|_{H^1(\mathbb{R}^d)} = \left(\|u\|_2^2 + \|\nabla u\|_2^2\right)^{\frac{1}{2}}$$

where we denote the gradient by $\nabla u(t,x) = \big(\partial_1 u(t,x), \cdots, \partial_d u(t,x)\big)$.

Of special interest is the subspace $\widetilde{F} \subset \mathbf{L}^2([0,T]; H^1(\mathbb{R}^d))$ consisting of all functions $u(t,x)$ such that $t \longmapsto u_t = u(t,\cdot)$ is continuous in $\mathbf{L}^2(\mathbb{R}^d)$. The natural norm on $\widetilde{F}$ is

$$\|u\|_T = \sup_{0 \leqslant t \leqslant T} \|u_t\|_2 + \left(\int_0^T \|\nabla u_t\|_2 dt\right)^{\frac{1}{2}}.$$

The Lebesgue measure in $\mathbb{R}^d$ will be denoted by m. The space of test functions which we employ in the definition of weak solutions of the evolution equations (3.1.1) or (3.2.1) is $\mathcal{D}_T = \mathcal{C}^\infty([0,T]) \otimes \mathcal{C}_c^\infty(\mathbb{R}^d)$, where $\mathcal{C}^\infty([0,T])$ denotes the space of real functions which can be extended as infinite differentiable functions in the neighborhood of $[0,T]$ and $\mathcal{C}_c^\infty(\mathbb{R}^d)$ is the space of infinite differentiable functions with compact support in $\mathbb{R}^d$. Finally, for all m in $\mathbb{N}$, $\mathbb{H}_m^2([0,T])$ will denote the set of (classes of $dP \times dt$ a.e. equal) $\mathbb{R}^m$-valued jointly measurable processes $(\psi_u)_{0 \leqslant u \leqslant T}$ such that $||\psi||^2_{\mathbb{H}^2_m([0,T])} := E[\int_0^T |\psi_u|^2 du]$ is finite and ψ_u is $\mathcal{F}_u$-measurable, for a.e. $u \in [0,T]$, while $\mathbb{S}_m^2([0,T])$ will denote the set of $\mathbb{R}^m$-valued continuous processes $(\psi_u)_{0 \leqslant u \leqslant T}$ such that $||\psi||^2_{\mathbb{S}^2_m([0,T])} := E[\sup_{0 \leqslant u \leqslant T} |\psi_u|^2]$ is finite and such that ψ_u is $\mathcal{F}_u$-measurable, for any $u \in [0,T]$.

3.2.2 The probabilistic interpretation of the divergence term

The operator $\partial_t + \frac{1}{2}\Delta$, which represents the main linear part in the equation (3.1.1), is probabilistically interpreted by the Bownian motion in $\mathbb{R}^d$. We shall view the Brownian motion as a Markov process and therefore we next introduce some detailed notation for it. The sample space is $\Omega' = \mathcal{C}([0,\infty);\mathbb{R}^d)$, the canonical process $(W_t)_{t \geqslant 0}$ is defined by $W_t(\omega) = \omega(t)$, for any $\omega \in \Omega'$, $t \geqslant 0$ and the shift operator, $\theta_t : \Omega' \longrightarrow \Omega'$, is defined by $\theta_t(\omega)(s) = \omega(t+s)$, for any $s \geqslant 0$ and $t \geqslant 0$. The canonical filtration $\mathcal{F}_t^0 = \sigma(W_s; s \leqslant t)$ is completed by the standard procedure with respect to the probability measures produced by the transition function

$$P_t(x,dy) = q_t(x-y)dy, \quad t>0, \quad x \in \mathbb{R}^d,$$

where $q_t(x) = (2\pi t)^{-\frac{d}{2}} \exp(-|x|^2/2t)$ is the gaussian density. Thus we get a continuous Hunt process $(\Omega', W_t, \theta_t, \mathcal{F}, \mathcal{F}_t^0, \mathbb{P}^x)$. We shall also use the backward filtration of the future events $\mathcal{F}_t' = \sigma(W_s; \ s \geqslant t)$ for $t \geqslant 0$. $\mathbb{P}^0$ is the Wiener measure, which is supported by the set $\Omega_0' = \{\omega \in \Omega', \ w(0) = 0\}$. We also set $\Pi_0(\omega)(t) = \omega(t) - \omega(0)$, $t \geqslant 0$, which defines a map $\Pi_0 : \Omega' \to \Omega_0'$. Then $\Pi = (W_0, \Pi_0) : \Omega' \to \mathbb{R}^d \times \Omega_0'$ is a bijection. For each probability measure on $\mathbb{R}^d$, the probability $\mathbb{P}^\mu$ of the Brownian motion started with the initial distribution μ is given by

$$\mathbb{P}^\mu = \Pi^{-1}(\mu \otimes \mathbb{P}^0).$$

In particular, for the Lebesgue measure in $\mathbb{R}^d$, which we denote by $m = dx$, we have

$$\mathbb{P}^m = \Pi^{-1}(dx \otimes \mathbb{P}^0).$$

These relations are saying that W_0 is independent of Π_0. It is known that each component $(W_t^i)_{t \geqslant 0}$ of the Brownian motion, $i = 1, \cdots, d$, is a martingale under any of the measures $\mathbb{P}^\mu$. The next lemma shows that $(W_{t-r}^i, \mathcal{F}_{t-r}')$, $r \in (0,t]$ is a backward local martingale under $\mathbb{P}^m$ (see [54] for the proof).

Lemma 3.2.1. *Let $0 < s < t$. If $A \in \sigma(W_t)$ is such that $\mathbb{E}^m[|W_t|; A] < \infty$, then one has $\mathbb{E}^m[|W_s|; A] < \infty$. Moreover, for each $B \in \mathcal{F}_t'$, and $i = 1, \cdots, d$, one has*

$$\mathbb{E}^m[W_s^i; A \cap B] = \mathbb{E}^m[W_t^i; A \cap B].$$

Now let us assume that f and $|g|$ belong to $\mathbf{L}^2([0,T] \times \mathbb{R}^d)$ and $u \in \widetilde{F}$ is a solution of the following deterministic equation :

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + f(t,x) + \mathrm{div} g(t,x) = 0. \tag{3.2.1}$$

. Let us denote by

$$\int_s^t g_r * dW_r = \sum_{i=1}^d \left(\int_s^t g_i(r,W_r) dW_r^i + \int_s^t g_i(r,W_r) d\overleftarrow{W_r^i} \right). \tag{3.2.2}$$

Then one has the following representation (Theorem 3.2 in [68])

Theorem 3.2.1. *The following relation holds $\mathbb{P}^m$-a.s. for each $0 \leqslant s \leqslant t \leqslant T$,*

$$u_t(W_t) - u_s(W_s) = \sum_{i=1}^d \int_s^t \partial_i u_r(W_r) dW_r^i - \int_s^t f_r(W_r) dr - \frac{1}{2} \int_s^t g_r * dW_r. \tag{3.2.3}$$

In [68] one uses the backward martingale $\overleftarrow{M}^{\mu,i}$ defined under an arbitrary $\mathbb{P}^{\mu}$, with μ a probability measure in $\mathbb{R}^d$, in order to express the integral $\int_s^t g_r * dW_r$. Though formally the definition looks different, one easily sees that it is the same object. After that, we introduce the quasicontinuity notion, which will be useful for us to obtain the continuity of the process $u_t(W_t)$ in Theorem 3.2.4 :

Definition 3.2.1. *A function $\psi : [0,T] \times \mathbb{R}^d \to \overline{\mathbb{R}}$ is called quasicontinuous provided that for each $\varepsilon > 0$, there exists an open set, $D_\varepsilon \subset [0,T] \times \mathbb{R}^d$, such that ψ is finite and continuous on D_ε^c and*

$$\mathbb{P}^m \left(\{ \omega \in \Omega' / \exists t \in [0,T] \;\; s.t. \, (t, W_t(\omega)) \in D_\varepsilon \} \right) < \varepsilon.$$

We note that if a function ψ is quasicontinuous, then the process $(\psi_t(W_t))_{t \in [0,T]}$ is continuous.

3.2.2.1 Hypotheses

Let $B = (B_t)_{t \geqslant 0}$ be a standard d^1-dimentional Brownian motion on a probability space $\left(\Omega, \mathcal{F}^B, \mathbb{P}\right)$. So $B_t = \left(B_t^1, \cdots, B_t^{d_1}\right)$ takes values in $\mathbb{R}^{d^1}$. Over the time interval $[0,T]$ we define the backward filtration $\left(\mathcal{F}_{s,T}^B\right)_{s \in [0,T]}$ where $\mathcal{F}_{s,T}^B$ is the completion in $\mathcal{F}^B$ of $\sigma(B_r - B_s; s \leqslant r \leqslant T)$.

We denote by $\mathcal{H}_T$ the space of $H^1(\mathbb{R}^d)$-valued predictable and $\mathcal{F}_{t,T}^B$-adapted processes $(u_t)_{0 \leqslant t \leqslant T}$ such that the trajectories $t \to u_t$ are in $\widetilde{F}$ a.s. and

$$E \left\| u \right\|_T^2 < \infty.$$

In the remainder of this paper we assume that the final condition Φ is a given function in $\mathbf{L}^2(\mathbb{R}^d)$ and the functions appearing in the equation (3.1.1)

$$\begin{aligned} f &: [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R} \,, \\ g &= (g_1, ..., g_d) \,:\, [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d \\ h &= (h_1, ..., h_{d^1}) \,:\, [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^{d^1} \end{aligned}$$

are measurable functions . We set

$$f(\cdot,\cdot,\cdot,0,0) := f^0, \quad g(\cdot,\cdot,\cdot,0,0) := g^0 = (g_1^0, ..., g_d^0) \quad \text{and} \quad h(\cdot,\cdot,\cdot,0,0) := h^0 = (h_1^0, ..., h_{d^1}^0).$$

and assume the following hypotheses :

Assumption (H) : There exist non-negative constants C, α, β such that

(i) $|f_t(x,y,z) - f(t,x,y'z')| \leqslant C\big(|y-y'| + |z-z'|\big)$

(ii) $\left(\sum_{j=1}^{d_1} |h_{j,t}(,x,y,z) - h_j(t,,x,y',z')|^2 \right)^{\frac{1}{2}} \leqslant C\,|y-y'| + \beta\,|z-z'|$,

(iii) $\left(\sum_{i=1}^{d} |g_{i,t}(x,y,z) - g_i(t,,x,y',z')|^2 \right)^{\frac{1}{2}} \leqslant C\,|y-y'| + \alpha\,|z-z'|$.

(iv) the contraction property (as in [28]) : $\alpha + \frac{\beta^2}{2} < \frac{1}{2}$.

Assumption (HD2)

$$\left\| f^0 \right\|_{2,2}^2 + \left\| g^0 \right\|_{2,2}^2 + \left\| h^0 \right\|_{2,2}^2 < \infty.$$

Assumption (HD2')

(i) $|\Phi(x) - \Phi(x')| \leqslant K|x-x'|$,

(ii) $\sup_{0 \leqslant t \leqslant T}(|f(t,0,0,0)| + |h(t,0,0,0)| + |g(t,0,0,0)|) \leqslant K$.

(iii) the contraction property (as in [28]) : $\alpha + \frac{\beta^2}{2} + \frac{\alpha^2}{8} < \frac{1}{2}$.

We recall that a solution of the equation (3.1.1) with final condition $u_T = \Phi$, is a processus $u \in \mathcal{H}_T$ such that for each test function $\varphi \in \mathcal{D}_T$ and any $\forall t \in [0,T]$, we have a.s.

$$\begin{aligned}&\int_t^T \left[(u_s, \partial_s \varphi_s) + \frac{1}{2} (\nabla u_s, \nabla \varphi_s) + \big(g_s, \nabla \varphi_s\big) \right] ds - \big(\Phi, \varphi_T\big) + \big(u_t, \varphi_t\big) \\ &= \int_t^T (f_s, \varphi_s)\, ds + \int_t^T (h_s, \varphi_s) \cdot \overleftarrow{dB}_s.\end{aligned} \tag{3.2.4}$$

By Theorem 8 in [28] we have existence and uniqueness of the solution. Moreover, the solution belongs to $\mathcal{H}_T$. We denote by $\mathcal{U}(\Phi, f, g, h)$ this solution. Let $\mathcal{L} = \sum_{ij} \partial_i a^{ij} \partial_j$ be an elliptic operator in divergence form, with the matrix $a = \left(a^{ij}\right) : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ being symmetric, measurable and such that

$$\lambda \left|\xi\right|^2 \leqslant \sum_{ij} a^{ij}\left(x\right) \xi^i \xi^j \leqslant \Lambda \left|\xi\right|^2,$$

for any $x, \xi \in \mathbb{R}^d$. If instead of the operator $\frac{1}{2}\Delta$ in our equation (1) we had the operator $\mathcal{L}$, then the contraction condition (iv) of hypothesis (**H**) would be replaced by $\alpha + \frac{\beta^2}{2} < \lambda$ (this ensures the contraction condition as formulated in [8]). Then the time change $t \to \frac{1}{2\Lambda} t'$ yields a one to one correspondence betwen the solutions u of the equation

$$du_t + \left[\mathcal{L} u_t + f_t\left(u_t, \nabla u_t\right) + div g_t\left(u_t, \nabla u_t\right)\right] dt + h_t\left(u_t, \nabla u_t\right) \cdot \overleftarrow{dB}_t = 0,$$

over $[0,T]$ and the solutions $\widehat{u}_t = u_{\frac{1}{2\Lambda} t}$ satisfying the equation

$$d\widehat{u}_t + \left[\frac{1}{2}\Delta \widehat{u}_t + \widehat{f}_t\left(\widehat{u}_t, \nabla \widehat{u}_t\right) + div \widehat{g}_t\left(\widehat{u}_t, \nabla \widehat{u}_t\right)\right] dt + \widehat{h}_t\left(\widehat{u}_t, \nabla \widehat{u}_t\right) \cdot \overleftarrow{d\widehat{B}}_t = 0,$$

over the interval $[0, 2\Lambda T]$, with the transformed coeficients

$$\widehat{f}\left(t,x,y,z\right) = \frac{1}{2\Lambda} f\left(\frac{1}{2\Lambda} t, x, y, z\right), \widehat{h}\left(t,x,y,z\right) = \frac{1}{(2\Lambda)^{\frac{1}{2}}} h\left(\frac{1}{2\Lambda} t, x, y, z\right),$$

$$\widehat{g}_i\left(t,x,y,z\right) = \frac{1}{2\Lambda}\left(g_i\left(\frac{1}{2\Lambda} t, x, y, z\right) + \sum_j a^{ij}\left(x\right) z_j - \Lambda z_i \right),\ i = 1, ..., d,$$

and the transformed Brownian motion $\widehat{B}_t = (2\Lambda)^{\frac{1}{2}} B_{\frac{1}{2\Lambda} t}, t \in [0, 2\Lambda T]$. This can be checked just by direct calculations using the above definition of a solution. Moreover, if one writes $\mathcal{L}$ in the form $\mathcal{L} u = \Lambda \Delta u - div\left(\gamma \nabla u\right)$, where $\gamma = \left(\gamma^{ij}\right)$ is a matrix with the enties $\gamma^{ij}\left(x\right) = \Lambda \delta^{ij} - a^{ij}\left(x\right), i, j = 1, ..., d$, then one has

$$0 \leqslant \gamma = \Lambda I - a \leqslant \left(\Lambda - \lambda\right) I,$$

in the sense of the order induced by the cone of non-negative definite matrices. This implies that one has

$$\left|\gamma\left(x\right)\xi\right| \leqslant \left(\Lambda - \lambda\right)\left|\xi\right|,$$

for any $x, \xi \in \mathbb{R}^d$. Then it easy to deduce that $\widehat{g}_t\left(x,y,z\right) = \frac{1}{2\Lambda}\left(g_{\frac{1}{2\Lambda} t}\left(x,y,z\right) + \gamma\left(x\right) z\right)$ fulfils condition (iii) of assumption (**H**) with a constant $\widehat{\alpha} = \frac{1}{2\Lambda}\left(\alpha + \left(\Lambda - \lambda\right)\right)$. On the other hand one can see that $\widehat{h}$ satisfies condition (ii) with $\widehat{\beta} = \frac{1}{(2\Lambda)^{\frac{1}{2}}} \beta$, so that the condition $\alpha + \frac{\beta^2}{2} < \lambda$, ensures $\widehat{\alpha} + \frac{\widehat{\beta}^2}{2} < \frac{1}{2}$, which is condition (iv) of our assumption (**H**). Therefore we conclude that our framework covers the case of an equation that involves an elliptic operator like $\mathcal{L}$, because the properties of the solution u are immediately obtained from those of the solution $\widehat{u}$. From now on, our aim will be the numerical approximation of the solution of SPDE (3.1.1) using a Monte Carlo method. Let us stress that the probabilistic numerical method described here allows us to approximate solutions of equation (3.1.1) by using M Monte Carlo simulations of the brownian motion W, instead of using a classical Euler scheme to discretize the Laplacien operator Δ. Then, we can benefit from advantages of the Monte Carlo method, which is more tractable in high dimensions and practical for parallel computing.

3.2.2.2 Itô formula and quasicontinuity property for the solution of the SPDE

In this part, we state some properties of the solution of equation of (3.1.1). These properties will be useful for us to study the time discretization error (3.3.3). We start by recalling the following result from [28] (stated for linear SPDE i.e. when f, g, h do not depend on u and ∇u) :

Theorem 3.2.2. *Let $u \in \mathcal{H}_T$ be a solution of the equation*

$$du_t + \frac{1}{2}\Delta u_t dt + \big(f_t + divg_t\big)\, dt + h_t \overleftarrow{dB}_t = 0,$$

where f, g, h are predictable processes such that

$$\int_0^T \big[\|f_t\|_2^2 + \|g_t\|_2^2 + \|h_t\|_2^2\big]\, dt < \infty \quad \textit{and} \quad \|\Phi\|_2^2 < \infty.$$

Then, for any $0 \leqslant s \leqslant t \leqslant T$, one has the following stochastic representation, $\mathbb{P}^m$-a.s.,

$$u\left(t, W_t\right) - u\left(s, W_s\right) = \sum_i \int_s^t \partial_i u\left(r, W_r\right) dW_r^i - \int_s^t f_r\left(W_r\right) dr - \frac{1}{2}\int_s^t g * dW - \int_s^t h_r\left(W_r\right) \cdot \overleftarrow{dB}_r. \quad (3.2.5)$$

Setting for all t in $[0,T]$, $Y_t := u_t(W_t)$ and $Z_t := \nabla u_t(W_t)$, the stochastic representation (3.2.5) can be written

$$Y_t = Y_T + \int_t^T f_r(W_r)dr + \frac{1}{2}\int_t^T g_r(W_r) * dW_r + \int_t^T h_r(W_r)\overleftarrow{dB}_r - \int_t^T Z_r dW_r. \quad (3.2.6)$$

We remark that $\mathcal{F}_T$ and $\mathcal{F}_{0,T}^B$ are independent under $\mathbb{P} \otimes \mathbb{P}^m$ and therefore in the above formula the stochastic integrals with respect to dW_t and $\overleftarrow{dW}_t$ act independently of $\mathcal{F}_{0,T}^B$ and similarly the integral with respect to $\overleftarrow{dB}_t$ acts independently of $\mathcal{F}_T$.
In particular the process $(u_t(W_t))_{t\in[0,T]}$ admits a continuous version. For this continuous version, we keep the same notation used in (3.2.6) and we denote it by $Y := (Y_t)_{t\in[0,T]}$ and $Z_t := \nabla u_t(W_t)$.
As a consequence of Theorem 3.2.2, we have the following result :

Theorem 3.2.3. *Under the hypothesis of the preceding theorem, we have the following results :*
(i) One has the following stochastic representation for u^2, $\mathbb{P} \otimes \mathbb{P}^m$-a.e., for any $0 \leqslant t \leqslant T$,

$$\begin{aligned} u_t^2\left(W_t\right) - \Phi^2\big(W_T\big) = 2\int_t^T \big[u_s f_s(W_s) - \frac{1}{2}|\nabla u_s|^2(W_s) - \langle \nabla u_s, g_s\rangle(W_s) + \frac{1}{2}|h_s|^2(W_s)\,\big]\, ds \\ + \int_t^T \big(u_r g_r\big)(W_r) * dW_r - 2\sum_i \int_t^T \big(u_r \partial_i u_r\big)\left(W_r\right) dW_r^i + 2\int_t^T \big(u_r h_r\big)\left(W_r\right) \cdot \overleftarrow{dB}_r. \end{aligned} \quad (3.2.7)$$

(ii) With the notation introduced in (3.2.6), one can write the relation (3.2.7) as

$$\begin{aligned} |Y_t|^2 + \int_t^T |Z_r|^2 dr = |Y_T|^2 + 2\int_t^T Y_r f_r(W_r)dr - 2\int_t^T \langle Z_r, g_r(W_r)\rangle\, dr + \int_t^T Y_r g_r(W_r) * dW_r \\ - 2\sum_i \int_t^T Y_r Z_{i,r} dW_r^i + 2\int_t^T Y_r h_r(W_r) \cdot \overleftarrow{dB}_r + \int_t^T |h_r|^2(W_r)dr. \end{aligned} \quad (3.2.8)$$

(iii) One has the estimate

$$\mathbb{E}^m\, \mathbb{E}\left[\left(\sup_{t\leqslant s \leqslant T}|Y_s|^2\right) + \int_t^T |Z_s|^2\, ds\right] \leqslant c\left\{\|\phi\|_2^2 + \int_t^T (\|f_s\|_2^2 + \|g_s\|_2^2 + \|h_s\|_2^2\,)\, ds\right\}, \quad (3.2.9)$$

for each $t \in [0,T]$.

We note that in the deterministic case, it was proven in [68] that the solution of a quasilinear equation has a quasicontinuous version. Here we have the same property for the solution of an SPDE as stated in Proposition 1, [54] :

Theorem 3.2.4. *Under the hypothesis of Theorem 3.2.2, there exists a function* $\bar{u} : [0,T]\times\Omega\times\mathbb{R}^d \longrightarrow \mathbb{R}$ *which is a quasicontinuous version of* u*, in the sense that for each* $\varepsilon > 0$*, there exits a predictable random set* $D^\varepsilon \subset [0,T]\times\Omega\times\mathbb{R}^d$ *such that* $\mathbb{P}$*-a.s. the section* D^ε_ω *is open and* $\bar{u}(\cdot,\omega,\cdot)$ *is continuous on its complement* $(D^\varepsilon_\omega)^c$ *and*

$$\mathbb{P}\otimes\mathbb{P}^m\ \big((\omega,\omega')\,\big|\,\exists t\in[0,T]\ s.t.\ (t,\omega,W_t(\omega'))\in D^\varepsilon\,\big) \leqslant \varepsilon.$$

In particular the process $\big(\overline{u}_t(W_t)\big)_{t\in[0,T]}$ *has continuous trajectories,* $\mathbb{P}\otimes\mathbb{P}^m$*-a.s.*

3.2.3 Numerical Scheme for decoupled Forward-BDSDE

In order to approximate the solution of the SPDE (3.1.1), we introduce the following discretized version of the assiciated BDSDE (3.2.5). Let $\pi = \{t_0 = 0 < t_1 < ... < t_N = T\}$ be an equidistant partition of the time interval $[0,T]$ i.e. $\Delta_N = \frac{T}{N}$ and $t_n = n\Delta_N$, $0\leqslant n\leqslant N$. Throughout the rest, we will use the notations $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ and $\Delta B_n = B_{t_{n+1}} - B_{t_n}$, for $n = 1,...,N$.
Quite naturally, the solution (Y,Z) of (3.2.5) is approximated by (Y^N, Z^N) defined by :

$$Y^N_{t_N} = \Phi(W^N_T),\, Z^N_{t_N} = 0, \tag{3.2.10}$$

and for $0\leqslant n\leqslant N-1$,

$$\begin{aligned} Y^N_{t_n} &= E_{t_n}[Y^N_{t_{n+1}}] + \Delta_N E_{t_n}[f(t_n,\Theta^N_n)] + \frac{1}{2}E_{t_n}[g(t_{n+1},\Theta^N_{n+1})\Delta W_n] \\ &+ E_{t_n}[h(t_{n+1},\Theta^N_{n+1})\Delta B_n], \end{aligned} \tag{3.2.11}$$

$$\begin{aligned} \Delta_N Z^N_{t_n} &= E_{t_n}\Bigg[Y^N_{t_{n+1}}\Delta W^*_n + h(t_{n+1},\Theta^N_{n+1})\Delta B_n\Delta W^*_n \\ &+ \frac{1}{2}\big\{g(t_n,\Theta^N_n) + g(t_{n+1},\Theta^N_{n+1})\big\}\Delta W_n\Delta W^*_n\Bigg], \end{aligned} \tag{3.2.12}$$

where

$$\Theta^N_n \;:=\; (W^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}), \quad \forall\, n = 0,\dots,N.$$

$*$ denotes the transposition operator and E_{t_n} denotes the conditional expectation over the σ-algebra $\mathcal{F}_{t_n} := \mathcal{F}^0_{t_n}\vee\mathcal{F}^B_{t_n,T}$.
We define also for all $n = 0,..,N-1$, $(Y^N,Z^N)_{t_n\,\leqslant\, s<t_{n+1}}$ as the solution of the following BDSDE :

$$\begin{cases} dY^N_s = -f(t_n,\theta^N_n)ds - \frac{1}{2}\{g(t_n,\Theta^N_n)dW_s + g(t_{n+1},\Theta^N_{n+1})\overleftarrow{dW_s}\} - h(t_{n+1},\Theta^N_{n+1})\overleftarrow{dB_s} + Z^N_s dW_s, \\ Y^N_{t_{n+1}} \text{ is given by our numerical scheme.} \end{cases} \tag{3.2.13}$$

This is the continuous approximation of the solution of the BDSDE (3.2.5). The superscript (t,x) indicates the dependence of the solution (W,Y,Z) on the initial date (t,x). To alleviate notations, we omit the dependence in (t,x) of $(Y^{t,x},Z^{t,x})$ and $(Y^{N,t,x},Z^{N,t,x})$ when the context is clear.
Notations : For a real matrix A, $\|A\|$ is the Frobenius norm defined by $\|A\| = (\sum_{i,j} a^2_{i,j})^{1/2}$.
For a vector x, $|x|$ stands for its Euclidean norme defined by $|x| = (\sum_i |x_i|^2)^{1/2}$.
In the next computations, the constant C denotes a generic constant that may change from line to line.

3.3 Time discretization error

In this section, we study the time discretization error of the Euler numerical scheme (3.2.11)-(3.2.12) of our BDSDEs. First, We give in the following an upper bound for the time discretization error. Then we give a regularity result which allows us to derive the rate of convergence for our numerical scheme.

3.3.1 Upper bound for the time discretization error

First, we need the following Lemma, which is a generalization of the Itô formula for BDSDEs given in [63] :

Lemma 3.3.1. *Let $\zeta \in \mathbb{S}^2_1([0,T])$, $\vartheta \in \mathbb{H}^2_1([0,T])$, $\gamma \in \mathbb{H}^2_{d_1([0,T])}$ and ϕ, ψ and $\delta \in \mathbb{H}^2_{d([0,T])}$ such that :*

$$\zeta_t = \zeta_0 + \int_0^t \vartheta_s ds + \int_0^t \gamma_s \overleftarrow{dB_s} + \int_0^t \phi_s dW_s + \int_0^t \psi_s dW_s + \int_0^t \delta_s \overleftarrow{dW_s}. \tag{3.3.1}$$

Then

$$\begin{aligned} |\zeta_t|^2 - |\zeta_0|^2 &= 2\int_0^t (\zeta_s, \vartheta_s)ds + 2\int_0^t (\zeta_s, \gamma_s \overleftarrow{dB_s}) + 2\int_0^t (\zeta_s, \phi_s dW_s) \\ &+ 2\int_0^t (\zeta_s, \psi_s dW_s) + 2\int_0^t (\zeta_s, \delta_s \overleftarrow{dW_s}) + 2\int_0^t (\phi_s, \psi_s)ds \\ &+ \int_0^t |\phi_s|^2 ds + \int_0^t |\psi_s|^2 ds - \int_0^t |\gamma_s|^2 ds - \int_0^t |\delta_s|^2 ds. \end{aligned}$$

Proof. Following [63], we write

$$\begin{aligned} |\zeta_{t_{i+1}} - \zeta_{t_i}|^2 &= (\zeta_{t_{i+1}} - \zeta_{t_i}, \zeta_{t_{i+1}} - \zeta_{t_i}) \\ &= |\zeta_{t_{i+1}}|^2 - |\zeta_{t_i}|^2 - 2(\zeta_{t_{i+1}} - \zeta_{t_i}, \zeta_{t_i}). \end{aligned}$$

$$\begin{aligned} 2(\zeta_{t_{i+1}} - \zeta_{t_i}, \zeta_{t_i}) &= 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \vartheta_s)ds + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \gamma_s \overleftarrow{dB_s}) + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \phi_s dW_s) \\ &+ 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \psi_s dW_s) + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \delta_s \overleftarrow{dW_s}) \\ &= 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \vartheta_s)ds + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_{i+1}}, \gamma_s \overleftarrow{dB_s}) + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \phi_s dW_s) \\ &+ 2\int_{t_i}^{t_{i+1}} (\zeta_{t_i}, \psi_s dW_s) + 2\int_{t_i}^{t_{i+1}} (\zeta_{t_{i+1}}, \delta_s \overleftarrow{dW_s}) - 2\int_{t_i}^{t_{i+1}} (\zeta_{t_{i+1}} - \zeta_{t_i}, \gamma_s \overleftarrow{dB_s}) \\ &- 2\int_{t_i}^{t_{i+1}} (\zeta_{t_{i+1}} - \zeta_{t_i}, \delta_s \overleftarrow{dW_s}). \end{aligned}$$

On the other hand

$$\begin{aligned} |\zeta_{t_{i+1}} - \zeta_{t_i}|^2 &= |\int_{t_i}^{t_{i+1}} \vartheta_s ds|^2 + |\int_{t_i}^{t_{i+1}} \gamma_s dB_s|^2 + |\int_{t_i}^{t_{i+1}} \phi_s dW_s|^2 \\ &+ |\int_{t_i}^{t_{i+1}} \psi_s dW_s|^2 + |\int_{t_i}^{t_{i+1}} \delta_s dW_s|^2 + 2(\int_{t_i}^{t_{i+1}} \vartheta_s ds, \int_{t_i}^{t_{i+1}} \gamma_s \overleftarrow{dB_s}) \\ &+ 2(\int_{t_i}^{t_{i+1}} \vartheta_s ds, \int_{t_i}^{t_{i+1}} \phi_s dW_s) + 2(\int_{t_i}^{t_{i+1}} \vartheta_s ds, \int_{t_i}^{t_{i+1}} \psi_s dW_s) \\ &+ 2(\int_{t_i}^{t_{i+1}} \vartheta_s ds, \int_{t_i}^{t_{i+1}} \delta_s \overleftarrow{dW_s}) + 2(\int_{t_i}^{t_{i+1}} \gamma_s \overleftarrow{dB_s}, \int_{t_i}^{t_{i+1}} \phi_s dW_s) \\ &+ 2(\int_{t_i}^{t_{i+1}} \gamma_s \overleftarrow{dB_s}, \int_{t_i}^{t_{i+1}} \psi_s dW_s) + 2(\int_{t_i}^{t_{i+1}} \gamma_s \overleftarrow{dB_s}, \int_{t_i}^{t_{i+1}} \delta_s \overleftarrow{dW_s}) \\ &+ 2(\int_{t_i}^{t_{i+1}} \phi_s dW_s, \int_{t_i}^{t_{i+1}} \psi_s dW_s) + 2(\int_{t_i}^{t_{i+1}} \phi_s dW_s, \int_{t_i}^{t_{i+1}} \delta_s \overleftarrow{dW_s}) \\ &+ 2(\int_{t_i}^{t_{i+1}} \psi_s dW_s, \int_{t_i}^{t_{i+1}} \delta_s \overleftarrow{dW_s}). \end{aligned}$$

Using that

$$\begin{aligned}
-2\int_{t_i}^{t_{i+1}}(\zeta_{t_{i+1}}-\zeta_{t_i},\gamma_s\overleftarrow{dB_s}) &= -2(\int_{t_i}^{t_{i+1}}\vartheta_s ds,\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})-2(\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s},\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})\\
&- 2(\int_{t_i}^{t_{i+1}}\phi_s dW_s,\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})-2(\int_{t_i}^{t_{i+1}}\psi_s dW_s,\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})\\
&- 2(\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s},\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})
\end{aligned}$$

and

$$\begin{aligned}
-2\int_{t_i}^{t_{i+1}}(\zeta_{t_{i+1}}-\zeta_{t_i},\delta_s\overleftarrow{dW_s}) &= -2(\int_{t_i}^{t_{i+1}}\vartheta_s ds,\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s})-2(\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s},\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s})\\
&- 2(\int_{t_i}^{t_{i+1}}\phi_s dW_s,\int_{t_i}^{t_{i+1}i+1}\delta_s\overleftarrow{dW_s})-2(\int_{t_i}^{t_{i+1}}\psi_s dW_s,\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s})\\
&- 2(\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s},\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s}),
\end{aligned}$$

we get

$$\begin{aligned}
|\zeta_{t_{i+1}}|^2-|\zeta_{t_i}|^2 &= 2\int_{t_i}^{t_{i+1}}(\zeta_{t_i},\vartheta_s)ds+2\int_{t_i}^{t_{i+1}}(\zeta_{t_{i+1}},\gamma_s\overleftarrow{dB_s})+2\int_{t_i}^{t_{i+1}}(\zeta_{t_i},\psi_s dW_s)\\
&+ \int_{t_i}^{t_{i+1}}(\zeta_{t_i},\psi_s dW_s)+2(\int_{t_i}^{t_{i+1}}\zeta_{t_{i+1}},\delta_s\overleftarrow{dW_s})\\
&+ |\int_{t_i}^{t_{i+1}}\vartheta_s ds|^2+|\int_{t_i}^{t_{i+1}}\phi_s dW_s|^2+|\int_{t_i}^{t_{i+1}}\psi_s dW_s|^2\\
&+ 2(\int_{t_i}^{t_{i+1}}\phi_s dW_s,\int_{t_i}^{t_{i+1}}\psi_s dW_s)+2(\int_{t_i}^{t_{i+1}}\vartheta_s ds,\int_{t_i}^{t_{i+1}}\phi_s dW_s)\\
&+ 2(\int_{t_i}^{t_{i+1}}\vartheta_s ds,\int_{t_i}^{t_{i+1}}\psi_s dW_s)-2(\int_{t_i}^{t_{i+1}}\delta_s\overleftarrow{dW_s},\int_{t_i}^{t_{i+1}}\gamma_s\overleftarrow{dB_s})\\
&- |\int_{t_i}^{t_{i+1}}\gamma_s dB_s|^2-\int_{t_i}^{t_{i+1}}|\delta_s dW_s|^2.
\end{aligned}$$

Finally, we can write

$$\begin{aligned}
|\zeta_{t_{i+1}}|^2-|\zeta_{t_i}|^2 &= 2\int_{t_i}^{t_{i+1}}(\zeta_{t_i},\vartheta_s)ds+2(\int_{t_i}^{t_{i+1}}\zeta_{t_{i+1}},\gamma_s\overleftarrow{dB_s})+2(\int_{t_i}^{t_{i+1}}\zeta_{t_i},\psi_s dW_s)\\
&+ (\int_{t_i}^{t_{i+1}}\zeta_{t_i},\psi_s dW_s)+2(\int_{t_i}^{t_{i+1}}\zeta_{t_{i+1}},\delta_s\overleftarrow{dW_s})\\
&+ |\int_{t_i}^{t_{i+1}}\phi_s dW_s|^2+|\int_{t_i}^{t_{i+1}}\psi_s dW_s|^2\\
&+ 2(\int_{t_i}^{t_{i+1}}\phi_s dW_s,\int_{t_i}^{t_{i+1}}\psi_s dW_s)-|\int_{t_i}^{t_{i+1}}\gamma_s dB_s|^2+|\int_{t_i}^{t_{i+1}}\delta_s dW_s|^2+\rho_i,
\end{aligned}$$

where $\sum_{i=0}^{N-1}\rho_i\longrightarrow 0$ in probability, as $\sup_i|t_{i+1}-t_i|\longrightarrow 0$. □

The next theorem states an upper bound result regarding the time discretization error. First, let Us define the process $\bar{Z}$ by

$$\left\{\begin{array}{l}\bar{Z}_t=\frac{1}{h}E_{t_n}[\int_{t_n}^{t_{n+1}}Z_s ds],\forall t\in[t_n,t_{n+1}),\forall n\in\{0,\ldots,N-1\},\\ \bar{Z}_{t_N}=0.\end{array}\right. \tag{3.3.2}$$

Theorem 3.3.1. *Assume that Assumptions* ***(H)*** *and* ***(HD2')*** *hold, define the error*

$$\begin{aligned} Error_N(Y,Z) &:= \sup_{0\leqslant t\leqslant T} E^m E[|Y_t - Y_t^N|^2] \\ &+ E^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds], \end{aligned} \tag{3.3.3}$$

where Y^N and Z^N are given by (3.2.13). Then

$$\begin{aligned} Error_N(Y,Z) &\leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds\}. \end{aligned} \tag{3.3.4}$$

Proof. We first define the following quantities :

$$\theta_s := (X_s, Y_s, Z_s), \forall s \in [t_n, t_{n+1}), \forall n = 0, \dots, N-1.$$

$\forall s \in [t_n, t_{n+1}), \forall n = 0, \dots, N-1$:

$$\begin{cases} \delta Y_t^N := Y_t - Y_t^N,\ \delta Z_t^N := Z_t - Z_t^N, \\ \delta f_t := f(t,\theta_t) - f(t_n, \Theta_n^N), \\ \delta g_{1,t} := g(t,\theta_t) - g(t_n, \Theta_n^N), \\ \delta g_{2,t} := g(t,\theta_t) - g(t_{n+1}, \Theta_{n+1}^N), \\ \delta h_t := h(t,\theta_t) - h(t_{n+1}, \Theta_{n+1}^N). \end{cases} \tag{3.3.5}$$

From the definition of δY^N, we have

$$\delta Y_s^N = -\delta f_s dt - \delta h_s \overleftarrow{dB_s} - \frac{1}{2}\delta g_{1,s} dW_s - \frac{1}{2}\delta g_{2,s}\overleftarrow{dW_s} + \delta Z_s^N dW_s, \tag{3.3.6}$$

Using the generalized Ito Formula given in Lemma 3.3.1 , we have $\forall t \in [t_n, t_{n+1})$

$$\begin{aligned} |\delta Y_t^N|^2 - |\delta Y_{t_{n+1}}^N|^2 &= 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta f_s)ds + 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta h_s \overleftarrow{dB_s}) + \int_t^{t_{n+1}} (\delta Y_s^N, \delta g_{1,s} dW_s) \\ &+ \int_t^{t_{n+1}} (\delta Y_s^N, \delta g_{2,s}\overleftarrow{dW_s}) - 2\int_t^{t_{n+1}} (\delta Y_s^N, \delta Z_s^N dW_s) + \int_t^{t_{n+1}} (\delta Z_s^N, \delta g_{1,s})ds \\ &+ \int_t^{t_{n+1}} |\delta h_s|^2 ds + \frac{1}{4}\int_t^{t_{n+1}} |\delta g_{2,s}|^2 ds - \int_t^{t_{n+1}} |\delta Z_s^N|^2 ds - \frac{1}{4}\int_t^{t_{n+1}} |\delta g_{1,s}|^2 ds. \end{aligned}$$

Then, tacking the expectation, we get

$$\begin{aligned} A_t^n &:= E^m E|\delta Y_t^N|^2 + \int_t^{t_{n+1}} E^m E|\delta Z_s^N|^2 ds - E^m E|\delta Y_{t_{n+1}}^N|^2 \\ &= 2\int_t^{t_{n+1}} E^m E(\delta Y_s^N, \delta f_s)ds + \int_t^{t_{n+1}} E^m E(\delta Z_s^N, \delta g_{1,s})ds + \int_t^{t_{n+1}} E^m E|\delta h_s|^2 ds \\ &+ \frac{1}{4}\int_t^{t_{n+1}} E^m E|\delta g_{2,s}|^2 ds - \frac{1}{4}\int_t^{t_{n+1}} E^m E|\delta g_{1,s}|^2 ds. \end{aligned} \tag{3.3.7}$$

After that, we have

$$\begin{aligned} &2\int_t^{t_{n+1}} E^m E(\delta Z_s^N, \delta g_{1,s})ds \leqslant 2\int_t^{t_{n+1}} E^m E(|\delta Z_s^N|, |\delta g_{1,s}|)ds \\ &\leqslant 2\int_t^{t_{n+1}} E^m E(|\delta Z_s^N|, C|W_s - W_{t_n}^N| + C|Y_s - Y_{t_n}^N| + \alpha|Z_s - Z_{t_n}^N|)ds \\ &\leqslant 2C\int_t^{t_{n+1}} E^m E(|\delta Z_s^N|, |W_s - W_{t_n}^N|)ds + 2C\int_t^{t_{n+1}} E^m E(|\delta Z_s^N|, |Y_s - Y_{t_n}^N|)ds \\ &+ 2\alpha\int_{t_n}^{t_{n+1}} E^m E(|\delta Z_s^N|, |Z_s - \bar{Z}_{t_n}|)ds + 2\alpha\int_{t_n}^{t_{n+1}} E^m E(|\delta Z_s^N|, |\bar{Z}_{t_n} - Z_{t_n}^N|)ds \\ &\leqslant (2\alpha + C\varepsilon(Z))\int_{t_n}^{t_{n+1}} E^m E|\delta Z_s^N|^2 ds + \frac{C}{\varepsilon(Z)}\int_t^{t_{n+1}} E^m E|W_s - W_{t_n}^N|^2 ds + \frac{C}{\varepsilon(Z)}\int_t^{t_{n+1}} E^m E|Y_s - Y_{t_n}^N|^2 ds \\ &+ \frac{\alpha}{\varepsilon(Z)}\int_{t_n}^{t_{n+1}} E^m E|Z_s - \bar{Z}_{t_n}|^2 ds, \end{aligned}$$

where $\varepsilon(Z)$ is a positive constante which will be specified later.
Plugging Y_{t_n}, we get :

$$2\int_t^{t_{n+1}} E^m E(\delta Z_s^N, \delta g_{1,s})ds \leqslant \frac{C}{\varepsilon(Z)}\int_t^{t_{n+1}} E^m E|W_s - W_{t_n}^N|^2 ds + \frac{C}{\varepsilon(Z)}\int_t^{t_{n+1}} E^m E|Y_s - Y_{t_n}|^2 ds$$
$$+ \frac{C}{\varepsilon(Z)}\Delta_N E^m E|\delta Y_{t_n}^N|^2 + (2\alpha + C\varepsilon(Z))\int_{t_n}^{t_{n+1}} E^m E|\delta Z_s^N|^2 ds + \frac{C}{\varepsilon(Z)}\int_t^{t_{n+1}} E^m E|Z_s^N - Z_{t_n}^N|^2 ds \tag{3.3.8}$$

From the assumption **(H)**, we have

$$\int_t^{t_{n+1}} E^m E[|\delta h_s|^2]ds \leqslant C\int_t^{t_{n+1}} E^m E[|W_s - W_{t_{n+1}}^N|^2]ds$$
$$+ C\int_t^{t_{n+1}} [|Y_s - Y_{t_{n+1}}^N|^2]ds + (1+\varepsilon)\beta^2 E^m E\big[\int_t^{t_{n+1}} |Z_s - Z_{t_{n+1}}^N|^2 ds\big]. \tag{3.3.9}$$

where ε is a positive constant.
We plug $\bar{Z}_{t_{n+1}}$ in the following and we use the Young's inequality, with a positive constant ε_1 (to be specified later),

$$\begin{aligned} E^m E\big[\int_t^{t_{n+1}} |Z_s - Z_{t_{n+1}}^N|^2 ds\big] &\leqslant (1+\frac{1}{\varepsilon_1})E^m E\big[\int_t^{t_{n+1}} |Z_s - \bar{Z}_{t_{n+1}}|^2 ds\big] \\ &+ (1+\varepsilon_1)E^m E\big[\int_{t_n}^{t_{n+1}} |\bar{Z}_{t_{n+1}} - Z_{t_{n+1}}^N|^2 ds\big] \end{aligned}$$

Using the definition of $\bar{Z}_{t_{n+1}}$, we get

$$\begin{aligned} E^m E\big[\int_t^{t_{n+1}} |Z_s - Z_{t_{n+1}}^N|^2 ds\big] &\leqslant (1+\frac{1}{\varepsilon_1})E^m E\big[\int_t^{t_{n+1}} |Z_s - \bar{Z}_{t_{n+1}}|^2 ds\big] \\ &+ (1+\varepsilon_1)E^m E\big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\big]. \end{aligned}$$

Now we plug the last inequality in (3.3.9) then insert $Y_{t_{n+1}}$ to get

$$\int_t^{t_{n+1}} E^m E[|\delta h_s|^2]ds \leqslant C\int_t^{t_{n+1}} E^m E[|W_s - W_{t_{n+1}}^N|^2] + C\int_t^{t_{n+1}} E^m E[|Y_s - Y_{t_{n+1}}|^2]ds$$
$$+ C\Delta_N E^m E[|\delta Y_{t_{n+1}}^N|^2] + (1+\varepsilon)(1+\varepsilon_1)\beta^2 E^m E\big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\big] \tag{3.3.10}$$
$$+ (1+\varepsilon)(1+\frac{1}{\varepsilon_1})\beta^2\int_t^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds.$$

The same arguments give us

$$\int_t^{t_{n+1}} E^m E[|\delta g_{2,s}|^2]ds \leqslant C\int_t^{t_{n+1}} E^m E[|W_s - W_{t_{n+1}}^N|^2]ds + C\int_t^{t_{n+1}} E^m E[|Y_s - Y_{t_{n+1}}|^2]ds$$
$$+ C\Delta_N E^m E[|\delta Y_{t_{n+1}}^N|^2] + (1+\varepsilon_3)(1+\varepsilon_4)\alpha^2 E^m E\big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\big] \tag{3.3.11}$$
$$+ (1+\varepsilon_3)(1+\frac{1}{\varepsilon_4})\alpha^2\int_t^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds$$

and

$$\int_t^{t_{n+1}} E^m E[|\delta g_{1,s}|^2]ds \leqslant C'\int_t^{t_{n+1}} E^m E[|W_s - W_{t_n}^N|^2]ds + C'\int_t^{t_{n+1}} E^m E[|Y_s - Y_{t_n}|^2]ds$$
$$+ C'\Delta_N E^m E[|\delta Y_{t_n}^N|^2] + (1+\varepsilon_5)\alpha^2 E^m E\big[\int_{t_n}^{t_{n+1}} |\delta Z_s^N|^2 ds\big] \tag{3.3.12}$$
$$+ (1+\varepsilon_5)\alpha^2\int_{t_n}^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_n}|^2]ds.$$

We remind that

$$E^mE[|W_s - W_{t_n}^N|^2] \leqslant C\Delta_N \text{ and } E^mE[|W_s - W_{t_{n+1}}^N|^2] \leqslant C\Delta_N \tag{3.3.13}$$

and

$$\begin{aligned} E^mE[|Y_s - Y_{t_{n+1}}^N|^2] &\leqslant C\{E^mE[|Y_s - Y_{t_{n+1}}|^2] + E^mE[|Y_{t_{n+1}} - Y_{t_{n+1}}^N|^2]\} \\ &\leqslant C\{\Delta_N + E^mE[|\delta Y_{t_{n+1}}^N|^2]\}, \end{aligned} \tag{3.3.14}$$

where C is a positive constant.
Plugging inequalities (3.3.11), (3.3.10) and (3.3.8) in (3.3.7) then using estimations (3.3.13) and (3.3.14), we obtain

$$\begin{aligned} A_t^n &\leqslant 2\int_t^{t_{n+1}} E^mE(\delta Y_s^N, \delta f_s)ds + \int_t^{t_{n+1}} E^mE(\delta Z_s^N, \delta g_{1,s})ds + \int_t^{t_{n+1}} E^mE|\delta h_s|^2 ds \\ &+ \frac{1}{4}\int_t^{t_{n+1}} E^mE|\delta g_{2,s}|^2 ds - \frac{1}{4}\int_t^{t_{n+1}} E^mE|\delta g_{1,s}|^2 ds \\ &\leqslant 2\int_t^{t_{n+1}} E^mE(\delta Y_s^N, \delta f_s)ds + (\alpha + C\varepsilon(Z))\int_{t_n}^{t_{n+1}} E^mE|\delta Z_s^N|^2 ds + C\Delta_N(1+|x|^2) \\ &+ C\Delta_N E^mE|\delta Y_{t_n}^N|^2 + \Big\{(1+\varepsilon)(1+\varepsilon_1)\beta^2 + (1+\varepsilon_3)(1+\varepsilon_4)\frac{\alpha^2}{4}\Big\}E\Big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\Big] \\ &+ C\int_{t_n}^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_n}|^2]ds + C\int_t^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds + C\Delta_N E^mE|\delta Y_{t_{n+1}}^N|^2 \end{aligned} \tag{3.3.15}$$

where C is a generic positive constant depending on $\varepsilon(Z)$ and independent from x.
We set $I_\alpha^\varepsilon := (\alpha + C\varepsilon(Z))$ and $I_\beta^\varepsilon := \Big\{(1+\varepsilon)(1+\varepsilon_1)\beta^2 + (1+\varepsilon_3)(1+\varepsilon_4)\frac{\alpha^2}{4}\Big\}$. Since $(2I_\alpha^\varepsilon + I_\beta^\varepsilon) \in]0,1[$, we can use the inequality $2ab \leqslant \dfrac{1-(2I_\alpha^\varepsilon + I_\beta^\varepsilon)}{4C}$ in the last estimation to get,

$$\begin{aligned} A_t^n &\leqslant \frac{4C}{1-(2I_\alpha^\varepsilon + I_\beta^\varepsilon)}\int_t^{t_{n+1}} E^mE|\delta Y_s^N|^2 ds + \frac{1-(2I_\alpha^\varepsilon + I_\beta^\varepsilon)}{4C}\int_t^{t_{n+1}} E^mE|\delta f_s|^2 ds \\ &+ I_\alpha^\varepsilon \int_{t_n}^{t_{n+1}} E^mE|\delta Z_s^N|^2 ds + C\Delta_N(1+|x|^2) + C\Delta_N E^mE|\delta Y_{t_{n+1}}^N|^2 + C\Delta_N E^mE|\delta Y_{t_n}^N|^2 \\ &+ I_\beta^\varepsilon E^mE\Big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\Big] + C\int_{t_n}^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_n}|^2]ds + C\int_t^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds. \end{aligned} \tag{3.3.16}$$

From the hypothesis (**H**), we have

$$\begin{aligned} &\int_t^{t_{n+1}} E^mE[|\delta f_s|^2]ds \leqslant C\int_t^{t_{n+1}} E^mE[|W_s - W_{t_n}^N|^2]ds + C\int_t^{t_{n+1}} E^mE[|Y_s - Y_{t_n}^N|^2]ds \\ &+ C\int_t^{t_n} E^mE[|Z_s - Z_{t_n}^N|^2]ds \\ &\leqslant C\int_t^{t_{n+1}} E^mE[|W_s - W_{t_n}^N|^2]ds + C\int_t^{t_{n+1}} E^mE[|Y_s - Y_{t_{n+1}}|^2]ds + C\Delta_N E^mE[|\delta Y_{t_{n+1}}^N|^2] \\ &+ 2C\int_{t_n}^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_n}|^2]ds + 2C\int_t^{t_{n+1}} E^mE[|\bar{Z}_{t_n} - Z_{t_n}^N|^2]ds \\ &\leqslant C\Delta_N + C\Delta_N E^mE[|\delta Y_{t_n}^N|^2] + 2C\int_{t_n}^{t_{n+1}} E^mE[|\delta Z_s^N|^2]ds \\ &+ 2C\int_t^{t_{n+1}} E^mE[|Z_s - \bar{Z}_{t_n}|^2]ds \end{aligned} \tag{3.3.17}$$

Plugging the last inequality in (3.3.16), we get

$$\begin{aligned} A_t^n &\leqslant C\int_t^{t_{n+1}} E^m E|\delta Y_s^N|^2 ds + \frac{1-I_\beta^\varepsilon}{2}\int_{t_n}^{t_{n+1}} E^m E|\delta Z_s^N|^2 ds \\ &+ C\Delta_N(1+|x|^2) + C\Delta_N E^m E|\delta Y_{t_{n+1}}^N|^2 + C\Delta_N E^m E|\delta Y_{t_n}^N|^2 + I_\beta^\varepsilon E^m E\Big[\int_{t_{n+1}}^{t_{n+2}} |\delta Z_s^N|^2 ds\Big] \\ &+ C\int_{t_n}^{t_{n+1}} E[|Z_s - \bar{Z}_{t_n}|^2]ds + C\int_t^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds \end{aligned} \tag{3.3.18}$$

From (3.3.7) and (3.3.18), we have

$$\begin{aligned} E^m E[|\delta Y_t^N|^2] &\leqslant A_t^n + E^m E[|\delta Y_{t_{n+1}}^N|^2] \\ &= E^m E[|\delta Y_t^N|^2] + \int_t^{t_{n+1}} E^m E[\|\delta Z_s^N\|^2] ds \\ &\leqslant C\int_t^{t_{n+1}} E^m E[|\delta Y_s^N|^2] ds + B_n, \ \forall t \in [t_n, t_{n+1}), \end{aligned} \tag{3.3.19}$$

where we set for all $n = 0, \ldots, N-2$:

$$\begin{aligned} B_n \quad := \quad & E^m E[|\delta Y_{t_{n+1}}^N|^2] + C\Delta_N E^m E[|\delta Y_{t_{n+1}}^N|^2] + C\Delta_N E^m E[|\delta Y_{t_n}^N|^2] + C\Delta_N^2(1+|x|^2) \\ + \quad & C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds + \frac{1-I_\beta^\varepsilon}{2}\int_{t_n}^{t_{n+1}} E^m E[||\delta Z_s^N||^2] ds \\ + \quad & C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds + I_\beta^\varepsilon \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z_s^N\|^2] ds. \end{aligned}$$

Using Gronwall Lemma, we have

$$E^m E[|\delta Y_t^N|^2] \leqslant B_n e^{C\Delta_N}, \ \forall t \in [t_n, t_{n+1}). \tag{3.3.20}$$

From inequalities (3.3.20) and (3.3.19), we get for Δ_N small enough

$$\begin{aligned} E^m E[|\delta Y_t^N|^2] + \int_t^{t_{n+1}} E^m E[\|\delta Z_s^N\|^2] ds &\leqslant (1 + C\Delta_N e^{C\Delta_N}) B_n \\ &\leqslant (1 + C\Delta_N) B_n, \ \forall t \in [t_n, t_{n+1}). \end{aligned} \tag{3.3.21}$$

By taking $t = t_n$ in the last inequality, we obtain

$$\begin{aligned} & E^m E[|\delta Y_{t_n}^N|^2] + \int_{t_n}^{t_{n+1}} E^m E[\|\delta Z_s^N\|^2] ds \leqslant (1 + C\Delta_N)\Big\{E^m E[|\delta Y_{t_{n+1}}^N|^2] + C\Delta_N E^m E[|\delta Y_{t_n}^N|^2] \\ + \quad & C\Delta_N^2 + C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds + \frac{1-I_\beta^\varepsilon}{2}\int_{t_n}^{t_{n+1}} E^m E[||\delta Z_s^N||^2] ds \\ + \quad & C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds + I_\beta^\varepsilon \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z_s^N\|^2] ds\Big\}. \end{aligned}$$

Then

$$\begin{aligned} & (1 - C\Delta_N) E^m E[|\delta Y_{t_n}^N|^2] + [1 - (1 + C\Delta_N)\frac{1-I_\beta^\varepsilon}{2}] \int_{t_n}^{t_{n+1}} E^m E[\|\delta Z_s^N\|^2] ds \\ \leqslant \quad & (1 + C\Delta_N)\Big\{E^m E[|\delta Y_{t_{n+1}}^N|^2] + C\Delta_N^2 + C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds \\ + \quad & C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds + I_\beta^\varepsilon \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z_s^N\|^2] ds\Big\}. \end{aligned}$$

For Δ_N small enough, we get

$$\begin{aligned} & E^m E[|\delta Y_{t_n}^N|^2] + \frac{1+I_\beta^\varepsilon}{2}\int_{t_n}^{t_{n+1}} E^m E[\|\delta Z_s^N\|^2] ds \leqslant (1 + C\Delta_N)\Big\{E^m E[|\delta Y_{t_{n+1}}^N|^2] \\ & + I_\beta^\varepsilon \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z_s^N\|^2] ds + C\Delta_N^2 + C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds \\ & + C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds\Big\}. \end{aligned} \tag{3.3.22}$$

Iterating the last inequality, we obtain for all $n = 0, \ldots, N-1$

$$\begin{aligned}
& E^m E[|\delta Y^N_{t_n}|^2] + \frac{1+I^\varepsilon_\beta}{2} \int_{t_n}^{t_{n+1}} E^m E[\|\delta Z^N_s\|^2] ds \leqslant (1+C\Delta_N)^{N-1} \Big\{ E^m E[|\delta Y^N_T|^2] + C\Delta_N \\
+ \quad & C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds + C \sum_{n=0}^{N-2} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds \\
+ \quad & (\beta^2 + \frac{\alpha^2}{4}) \int_{t_{N-1}}^{t_N} E^m E[||Z_s||^2] ds \Big\}.
\end{aligned}$$

Using the Lipschitz condition on Φ, we get

$$\begin{aligned}
& E^m E[|\delta Y^N_{t_n}|^2] + \frac{1+I^\varepsilon_\beta}{2} \int_{t_n}^{t_{n+1}} E^m E[\|\delta Z^N_s\|^2] ds \leqslant (1+C\Delta_N)^{N-1} \Big\{ C\Delta_N \\
& + C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds + C \sum_{n=0}^{N-2} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds \\
& + (\beta^2 + \frac{\alpha^2}{4}) \int_{t_{N-1}}^{t_N} E^m E[||Z_s||^2] ds \Big\}.
\end{aligned} \tag{3.3.23}$$

Now we sum up inequality (3.3.22) over n, we get

$$\begin{aligned}
& \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_n}|^2] + \frac{1+I^\varepsilon_\beta}{2} \int_0^T E^m E[\|\delta Z^N_s\|^2] ds \leqslant (1+C\Delta_N) \Big\{ \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_{n+1}}|^2] \\
+ \quad & \sum_{n=0}^{N-1} C\Delta_N^2 + C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E^m E[\|Z_s - \bar{Z}_{t_n}\|^2] ds + C \sum_{n=0}^{N-2} \int_{t_n}^{t_{n+1}} E^m E[\|Z_s - \bar{Z}_{t_{n+1}}\|^2] ds \\
+ \quad & (\beta^2 + \frac{\alpha^2}{4}) E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds \Big] \Big\} + (1+C\Delta_N) I^\varepsilon_\beta \sum_{n=0}^{N-2} \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z^N_s\|^2] ds.
\end{aligned}$$

Using that $N\Delta_N = T$ and $\sum_{n=0}^{N-2} \int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z^N_s\|^2] ds = \int_{t_1}^{T} E^m E[\|\delta Z^N_s\|^2] ds$, we get

$$\begin{aligned}
& \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_n}|^2] + \frac{1+I^\varepsilon_\beta}{2} \int_0^T E^m E[\|\delta Z^N_s\|^2] ds \leqslant (1+C\Delta_N) \Big\{ \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_{n+1}}|^2] \\
+ \quad & C\Delta_N + C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds + C \sum_{n=0}^{N-2} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds \\
+ \quad & (\beta^2 + \frac{\alpha^2}{4}) E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds \Big] \Big\} + (1+C\Delta_N) I^\varepsilon_\beta \int_0^T E^m E[\|\delta Z^N_s\|^2] ds.
\end{aligned}$$

Then

$$\begin{aligned}
& \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_n}|^2] + \Big[\frac{1+I^\varepsilon_\beta}{2} - (1+C\Delta_N) I^\varepsilon_\beta \Big] \int_0^T E^m E[\|\delta Z^N_s\|^2] ds \\
\leqslant \quad & (1+C\Delta_N) \Big\{ \sum_{n=0}^{N-1} E^m E[|\delta Y^N_{t_{n+1}}|^2] + C\Delta_N + C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2] ds \\
+ \quad & C \sum_{n=0}^{N-2} \int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2] ds + (\beta^2 + \frac{\alpha^2}{4}) E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds \Big] \Big\}.
\end{aligned}$$

We obtain for Δ_N small enough

$$\begin{aligned}&\sum_{n=0}^{N-1} E^m E[|\delta Y_{t_n}^N|^2] + \frac{1-I_\beta^\varepsilon}{2}\int_0^T E^m E[\|\delta Z_s^N\|^2]ds \leqslant (1+C\Delta_N)\Big\{\sum_{n=0}^{N-1} E^m E[|\delta Y_{t_{n+1}}^N|^2]\\ + \quad & C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds\\ + \quad & (\beta^2 + \frac{\alpha^2}{4})E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big]\Big\}.\end{aligned}$$

Hence, we get

$$\begin{aligned}&\frac{1-I_\beta^\varepsilon}{2}\int_0^T E^m E[\|\delta Z_s^N\|^2]ds \leqslant (1+C\Delta_N)E^m E[|\delta Y_T^N|^2] + [(1+C\Delta_N)-1]\sum_{n=1}^{N-1} E^m E[|\delta Y_{t_n}^N|^2]\\ - \quad & E^m E[|\delta Y_{t_0}^N|^2] + C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds\\ + \quad & C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + (\beta^2 + \frac{\alpha^2}{4})E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big].\end{aligned}$$

Using the lipschitz on Φ, we get

$$\begin{aligned}&\frac{1-I_\beta^\varepsilon}{2}\int_0^T E^m E[\|\delta Z_s^N\|^2]ds \leqslant C\Delta_N + C\Delta_N\sum_{n=1}^{N-1} E^m E[|\delta Y_{t_n}^N|^2] + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds\\ &+ C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + (\beta^2 + \frac{\alpha^2}{4})E^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big].\end{aligned} \tag{3.3.24}$$

Summing up (3.3.23) over n, we have

$$\begin{aligned}&\Delta_N\sum_{n=0}^{N-1} E^m E[|\delta Y_{t_n}^N|^2 \leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds\\ + \quad & C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + C\int_{t_{N-1}}^{t_N} E^m E[||Z_s||^2]ds.\end{aligned}$$

Plugging the last inequality in (3.3.24), we obtain

$$\begin{aligned}&\frac{1-I_\beta^\varepsilon}{2}\int_0^T E^m E[\|\delta Z_s^N\|^2]ds \leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds\\ &+ C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE^m E\Big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\Big].\end{aligned} \tag{3.3.25}$$

Now, turning Back to equation (3.3.21), we have for all $n = 0, \dots, N-2$

$$\begin{aligned}E^m E[|\delta Y_t^N|^2] \quad &\leqslant \quad (1+C\Delta_N)B_n\\ &\leqslant \quad (1+C\Delta_N)\Big\{E^m E[|\delta Y_{t_{n+1}}^N|^2] + I_\beta^\varepsilon\int_{t_{n+1}}^{t_{n+2}} E^m E[\|\delta Z_s^N\|^2]ds\\ &+ \quad C\Delta_N E^m E[|\delta Y_{t_n}^N|^2] + (\frac{1-I_\beta^\varepsilon}{2})\int_{t_n}^{t_{n+1}} E^m E[||\delta Z_s^N||^2]ds\\ &+ \quad C\Delta_N^2 + C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds\\ &+ \quad C\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds\Big\}, \forall t\in[t_n, t_{n+1}).\end{aligned}$$

Using inequality (3.3.23), we get

$$\begin{aligned} E^m E[|\delta Y_t^N|^2] &\leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ &+ C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE^m E\big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\big]. \end{aligned}$$

Then by taking the supremum over t in the last inequality, we obtain

$$\begin{aligned} \sup_{0\leqslant t\leqslant T} E^m E[|\delta Y_t^N|^2] \leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE^m E\big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\big]. \end{aligned} \tag{3.3.26}$$

Inequalities (3.3.26) and (3.3.25) give together

$$\begin{aligned} \sup_{0\leqslant t\leqslant T} E^m E[|\delta Y_t^N|^2] + \int_0^T E^m E[\|\delta Z_s^N\|^2]ds \leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds \\ + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE^m E\big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\big]. \end{aligned} \tag{3.3.27}$$

Plugging $\bar{Z}_{t_n}$, we deduce that

$$\begin{aligned} E^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds] \leqslant CE^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||Z_s - \bar{Z}_{t_n}||^2 ds] \\ + CE^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||\delta Z_s^N||^2 ds]. \end{aligned} \tag{3.3.28}$$

Using the last inequality in (3.3.27), we get

$$\begin{aligned} \sup_{0\leqslant t\leqslant T} E^m E[|\delta Y_t^N|^2] + E^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds] \leqslant C\Delta_N \\ + \quad C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-2}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds + CE^m E\big[\int_{t_{N-1}}^{t_N} ||Z_s||^2 ds\big] \end{aligned}$$

which can be written, if we set $\bar{Z}_{t_N} := 0$

$$\begin{aligned} \sup_{0\leqslant t\leqslant T} E^m E[|\delta Y_t^N|^2] + E^m E[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds] \leqslant C\Delta_N \\ + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_n}||^2]ds + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[||Z_s - \bar{Z}_{t_{n+1}}||^2]ds. \end{aligned} \tag{3.3.29}$$

$\square$

3.3.2 Zhang L^2-Regularity and rate of convergence

In this subsection, we extend to our framework the result concerning the Zhang L^2-regularity of the martingale integrand Z proved by Zhang [70] for standard FBSDEs. This result is very important to derive the rate of convergence of our numerical scheme. So, we use the stochastic representation (3.2.5). Thus, the solution of the SPDE (3.1.1) is given by the couple (Y, Z), solution of the following FBDSDE :

$$Y_t - Y_s = \int_s^t Z_r dW_r - \int_s^t f_r(W_r)\, dr - \frac{1}{2}\int_s^t g_r * dW_r - \int_s^t h_r(W_r)\cdot \overleftarrow{dB}_r. \tag{3.3.30}$$

Hence, we will focus in the following on the FBDSDEs of the following form :

$$\begin{aligned}Y_t = \Phi(T) + \int_t^T f_r\left(r, W_r, Y_r, Z_r\right) dr + \frac{1}{2}\int_t^T g(r, W_r, Y_r, Z_r) dW_r + \frac{1}{2}\int_t^T g(r, W_r, Y_r, Z_r)\overleftarrow{dW_r} \\ + \int_t^T h\left(r, W_r, Y_r, Z_r\right) \cdot \overleftarrow{dB}_r - \int_t^T Z_r dW_r.\end{aligned} \tag{3.3.31}$$

We start by giving two important lemmas which are crucial to prove the main regularity result. The first lemma is classical and concerns the L^2-estimates and stability results for the FBDSDE's solution. We omit its proof which is based on standard computations for BSDEs.

Lemma 3.3.2. *Assume that Assumptions* **(H)** *and* **(HD2')** *hold. Let* $\Theta := (W, Y, Z)$ *denote the solution of the FBDSDE (3.3.31) Then we have the following*
(i)There exists a constant C_2*, depending on* T, K, α *and* β *such that*

$$\begin{aligned}E^m E\Big[\sup_{t \leqslant s \leqslant T}|Y_s|^2 + \int_t^T \|Z_s\|^2 ds\Big] \leqslant C_2 E^m E\Big\{|\Phi(W_T)|^2 + \int_t^T |f(s,0,0)|^2 ds \\ + \int_t^T |h(s,0,0,0)|^2 ds + \int_t^T |g(s,0,0,0)|^2 ds\Big\}\end{aligned} \tag{3.3.32}$$

and

$$\begin{aligned}E^m E\Big[|Y_s - Y_t|^2\Big] \leqslant C_2 E^m E\Big\{|\Phi(W_T)|^2 + \sup_{0 \leqslant r \leqslant T}|f(r,0,0,0)|^2 \\ + \sup_{0 \leqslant r \leqslant T}|h(r,0,0,0)|^2 + \sup_{0 \leqslant r \leqslant T}|g(r,0,0,0)|^2 ds\Big\}|s-t| \\ + E^m E\Big[\int_t^s \|Z_u\|^2 du\Big].\end{aligned} \tag{3.3.33}$$

(ii)Let $\Theta^\varepsilon := (W, Y^\varepsilon, Z^\varepsilon)$ *denote the solution of the perturbed FBDSDE* (3.3.31) *with coefficients replaced by* $f^\varepsilon, g^\varepsilon$ *and* h^ε *and* Φ^ε *as a terminal value. Assume that* $f^\varepsilon, g^\varepsilon, h^\varepsilon$ *and* Φ^ε *satisfy Assumptions* (**H**) *and* (**HD2′**) *and that for fixed (x,y,z) in* $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$,

$$\begin{aligned}\lim_{\varepsilon \longrightarrow 0} E^m E\Big\{|\Phi^\varepsilon(W_T) - \Phi(W_T)|^2 + \int_t^T |h^\varepsilon(s,x,y,z) - h(s,x,y,z))|^2 ds \\ + \int_t^T |g^\varepsilon(s,x,y,z) - g(s,x,y,z))|^2 ds + \int_t^T |f^\varepsilon(s,x,y,z) - f(s,x,y,z)|^2 ds\Big\} = 0.\end{aligned}$$

Then we have

$$\lim_{\varepsilon \longrightarrow 0} E^m E\Big\{\sup_{t \leqslant s \leqslant T}|Y_s^\varepsilon - Y_s|^2 + \int_t^T |Z_s^\varepsilon - Z_s|^2 ds\Big\} = 0. \tag{3.3.34}$$

The second Lemma gives estimates which are needed for the upper bound result (3.3.4) on our error and for the Zhang L^2–regularity result given in Theorem 3.3.2.

Lemma 3.3.3. *Assume that Assumptions* **(H)** *and* **(HD2')** *hold. Then, there exists a constant* $C_2 > 0$ *depending only on* T, K, α *and* β *such that*

$$\Big(E^m E[\|Z_s^{t,0}\|^2]\Big)^{\frac{1}{2}} \leqslant C_2 \ a.e. s \in [t,T]. \tag{3.3.35}$$

In addition, there exist a positive constant C *independent from* Δ_N *the time step of our uniform time-grid such that*

$$\max_{0 \leqslant n \leqslant N-1}\Big\{\sup_{t_n \leqslant s \leqslant t_{n+1}} E^m E[|Y_s^{t,0} - Y_{t_n}^{t,0}|^2] + \sup_{t_n \leqslant s \leqslant t_{n+1}} E^m E[|Y_s^{t,0} - Y_{t_{n+1}}^{t,0}|^2]\Big\} \leqslant C\Delta_N. \tag{3.3.36}$$

Proof. See Appendix. ◻

Now, we can state the main regularity result of this subsection

Theorem 3.3.2. *Under Assumptions* (**H**) *and* (**HD2′**), *we have the following estimation*

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} \Big\{\|Z_s - Z_{t_n}\|^2 + \|Z_s - Z_{t_{n+1}}\|^2\Big\} ds\Big] \leqslant C\Delta_N. \tag{3.3.37}$$

Proof. Let $\pi' : 0 = t_0, \dots, t_N = T$ denote a uniform fixed partition of $[0,T]$ with time step Δ_N and $\pi : 0 = s_0, \dots, s_m = T$ any partition finer than π'. We will prove the theorem for the fixed partition π'. Without loss of generality, we assume that $s_{l_i} = t_i$ for $i = 1, \dots, N$. Let Φ^π, f^π, h^π and $g^\pi \in C_b^1$ smooth molifiers of Φ, f, h and g, such that all the derivatives are bounded by K. We denote by $\Theta^\pi = (W, Y^\pi, Z^\pi)$ the solution of the following FBDSDE :

$$\begin{aligned} Y_s^\pi &= \Phi(W_T^\pi) + \int_s^T f^\pi(u, \Theta^\pi) ds + \int_s^T h^\pi(u, \Theta^\pi) \overleftarrow{dB_u} \\ &+ \frac{1}{2}\int_s^T g^\pi(u, \Theta^\pi) dW_u + \frac{1}{2}\int_s^T g^\pi(u, \Theta^\pi) \overleftarrow{dW_u} - \int_s^T Z_u^\pi dW_u. \end{aligned} \tag{3.3.38}$$

By the stability result (3.3.34), we have

$$\lim_{|\pi| \longrightarrow 0} E^m E\Big\{\sup_{t \leqslant s \leqslant T} |Y_s^\pi - Y_s|^2 + \int_t^T |Z_s^\pi - Z_s|^2 ds\Big\} = 0. \tag{3.3.39}$$

By (3.3.39), there exists a subsequence denoted again by π such that $\lim_{|\pi| \longrightarrow 0} E^m E|Z_s^\pi - Z_s|^2 = 0$ for a.e. s $\in [t, T]$. Let Us note that, for $s \in [t_n, t_{n+1})$, we have

$$E^m E|Z_s - Z_{t_n}|^2 + E^m E|Z_s - Z_{t_{n+1}}|^2 \leqslant C E^m E\Big\{|Z_s - Z_s^\pi|^2 + |Z_s^\pi - Z_{t_n}^\pi|^2 + |Z_s^\pi - Z_{t_{n+1}}^\pi|^2\Big\}. \tag{3.3.40}$$

By (3.3.39), proving the theorem remains to estimate $E^m E|Z_s^\pi - Z_{t_n}^\pi|^2$ and $E^m E|Z_s^\pi - Z_{t_{n+1}}^\pi|^2$ for $s \in [t_n, t_{n+1})$. To this end, we denote by $(I_d, \nabla^\pi Y)$ the solution of the linear equation (3.5.1) with coefficients replaced by Φ^π, f^π, h^π and g^π.
Using the representaion result (3.5.3) for (Z^π), we have

$$Z_s^\pi - Z_{s'}^\pi = \nabla Y_s^\pi - \nabla Y_{s'}^\pi, s, s' \in [t_n, t_{n+1}). \tag{3.3.41}$$

We conclude by the estimation (3.3.36) that

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} |Z_s^\pi - Z_{t_n}^\pi|^2 ds\Big] \leqslant C\Delta_N.$$

By the same arguments used on $E^m E|Z_s^\pi - Z_{t_{n+1}}^\pi|^2$ (taking $s' = t_{n+1}$), we conclude that

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} |Z_s^\pi - Z_{t_{n+1}}^\pi|^2 ds\Big] \leqslant C\Delta_N.$$

This gives

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s^\pi - Z_{t_n}^\pi|^2 + |Z_s^\pi - Z_{t_{n+1}}^\pi|^2\Big\} ds\Big] \leqslant C\Delta_N. \tag{3.3.42}$$

Using (3.3.40), we obtain

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s - Z_{t_n}|^2 + |Z_s - Z_{t_{n+1}}|^2\Big\} ds\Big] \leqslant C E^m E\Big[\int_t^T |Z_s - Z_s^\pi|^2 ds\Big] + C\Delta_N. \tag{3.3.43}$$

Recalling (3.3.39) and letting $|\pi| \longrightarrow 0$, we finally get

$$\sum_{n=0}^{N-1} E^m E\Big[\int_{t_n}^{t_{n+1}} \Big\{|Z_s - Z_{t_n}|^2 + |Z_s - Z_{t_{n+1}}|^2\Big\} ds\Big] \leqslant C\Delta_N. \tag{3.3.44}$$

□

Now, we are able to give the rate of convergence for our numerical scheme Under Assumptions (**H**) and (**HD2′**), we have

$$Error_N(Y,Z) := \sup_{0\leqslant s\leqslant T} E^m E[|Y_s - Y_s^N|^2] + \sum_{n=0}^{N-1} E^m E[\int_{t_n}^{t_{n+1}} ||Z_s - Z_{t_n}^N||^2 ds] \leqslant C\Delta_N.$$

Proof. First we recall that under (**H**) and (**HD2′**), we have by (3.3.4)

$$\begin{aligned} Error_N(Y,Z) &\leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_n}|^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds. \end{aligned} \tag{3.3.45}$$

Then, as the conditional expectation minimizes the conditional mean square error, we have

$$\int_{t_n}^{t_{n+1}} E^m E|Z_s - \bar{Z}_{t_n}|^2 ds \quad \leqslant \quad \int_{t_n}^{t_{n+1}} E^m E|Z_s - Z_{t_n}|^2 ds.$$

On the other hand, plugging $Z_{t_{n+1}}$ in the following, we get

$$\int_{t_n}^{t_{n+1}} E^m E[|Z_s - \bar{Z}_{t_{n+1}}|^2]ds \leqslant C\int_{t_n}^{t_{n+1}} E^m E[|Z_s - Z_{t_{n+1}}|^2]ds + C\Delta_N E^m E[|Z_{t_{n+1}} - \bar{Z}_{t_{n+1}}|^2]ds$$

By the defintion of $\bar{Z}_{t_{n+1}}$, Jensen's inequality and Cauchy-Schwarz inequality, we have for all $n = 0,\ldots,N-2$

$$\begin{aligned} \Delta_N E^m E[|Z_{t_{n+1}} - \bar{Z}_{t_{n+1}}|^2]ds &= C\frac{1}{\Delta_N}E^m E\Big\{\Big|E_{t_{n+1}}\Big[\int_{t_{n+1}}^{t_{n+2}}\{Z_{t_{n+1}} - Z_s\}ds\Big]\Big|^2\Big\} \\ &\leqslant C\frac{1}{\Delta_N}E^m E\Big\{\Big|\Big[\int_{t_{n+1}}^{t_{n+2}}\{Z_{t_{n+1}} - Z_s\}ds\Big]\Big|^2\Big\} \\ &\leqslant CE^m E\int_{t_{n+1}}^{t_{n+2}} |Z_{t_{n+1}} - Z_s|^2 ds. \end{aligned}$$

Hence, the inequality (3.3.45) becomes

$$\begin{aligned} Error_N(Y,Z) &\leqslant C\Delta_N + C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[|Z_s - Z_{t_n}|^2]ds \\ &+ C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} E^m E[|Z_s - Z_{t_{n+1}}|^2]ds. \end{aligned}$$

We conclude the rate of convergence by using Theorem 3.3.2. □

3.4 Implementation and numerical tests

In the following, we test statically the convergence of our method. In this part, we are interested in implementing our numerical scheme. Our aim is only to test statically its convergence. Further analysis of the convergence of the used method and of the error bounds will be accomplished in a future work.

3.4.1 Notations and algorithm

We use a path-dependent algorithm, for every fixed path of the brownian motion B, we approximate by a regression method the solution of the associated PDE. Then, we replace the conditional expectations which appear in (3.4.1) and (3.4.2) by $\mathbb{L}^2(\Omega,\mathcal{P})$ projections on the function basis approximating $\mathbb{L}^2(\Omega,\mathcal{F}_{t_n})$. We compute $Z_{t_n}^N$ and $Y_{t_n}^N$ in a implicit way. Actually, we proceed as in [34], except that in our case the solutions $Y_{t_n}^N$ and $Z_{t_n}^N$ are measurable functions of $(W_{t_n}^N, (\Delta B_i)_{n\leqslant i\leqslant N-1})$. So, each solution given by our algorithm depends on the fixed path of B.

3.4.1.1 Numerical scheme

For each fixed path of B, the solution of (3.2.11)-(3.2.12) is approximated by (Y^N, Z^N) defined by the following algorithm, given in the multidimensional case.
For $0 \leqslant n \leqslant N-1$:

$$Y^N_{t_n} = E_{t_n}\Big[Y^N_{t_{n+1}} + \Delta_N f(W^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}) + \sum_{j_1=1}^{d^1} h_{j_1}(W^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Delta B_{n,j_1}\Big] + \frac{1}{2}E_{t_n}\Big[\sum_{j=1}^{d} g_j(W^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\}\Delta W_{n,j}\Big] \tag{3.4.1}$$

and $\forall j \in \{1, \dots, d\}$

$$\Delta_N Z^N_{t_n,j} = E_{t_n}\Big[Y^N_{t_{n+1}}\Delta W_{n,j} + \sum_{j_1=1}^{l} g_{j_1}(W^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Delta B_{n,j_1}\Delta W_{n,j}\Big] + \frac{1}{2}E_{t_n}\Big[\Big\{g_j(W^N_{t_n}, Y^N_{t_n}, Z^N_{t_n}) + g_j(W^N_{t_{n+1}}, Y^N_{t_{n+1}}, Z^N_{t_{n+1}})\Big\}(\Delta W_{n,j})^2\Big]. \tag{3.4.2}$$

We stress that at each discretization time, the solution of the algorithm depends on the fixed path of the brownian motion B.

3.4.1.2 Vector spaces of functions

At every t_n, we select $k(d+1)$ deterministic functions bases $(p_{i,n}(.))_{1 \leqslant i \leqslant k(d+1)}$ and we look for approximations of $Y^N_{t_n}$ and $Z^N_{t_n}$, which will be denoted respectively by y^N_n and z^N_n, in the vector spaces spanned respectively by the basis $(p_{j_1,n}(.))_{1 \leqslant j_1 \leqslant k}$ and the basis $(p_{j_1,j_2,n}(.))_{1 \leqslant j_1 \leqslant k, 1 \leqslant j_2 \leqslant d}$. Each basis $p_{i,n}(.)$ is considered as a vector of functions of dimension $L_{i,n}$. In other words, $P_{i,n}(.) = \{\alpha.p_{i,n}(.), \alpha \in \mathbb{R}^{L_{i,n}}\}$.
As an example, we cite the hypercube basis (**HC**) used in [34]. In this case, $p_{i,n}(.)$ does not depend nor on i neither on n and its dimension is simply denoted by L. A domain $D \subset \mathbb{R}^d$ centered on $W_0 = 0$, that is $D = \prod_{i=1}^d(-a, +a]$, can be partitionned on small hypercubes of edge δ. Then, $D = \bigcup_{i_1,\dots,i_d} D_{i_1,\dots,i_d}$ where $D_{i_1,\dots,i_d} = (-a + i_1\delta, -a + i_1\delta] \times \dots \times (-a + i_d\delta, -a + i_d\delta]$. Finally we define $p_{i,n}(.)$ as the indicator functions of this set of hypercubes.

3.4.1.3 Monte Carlo simulations

To compute the projection coefficients α, we will use M independent Monte Carlo simulations of $W_{t_n}^N$ and ΔW_n which will be respectively denoted by $W^{N,m}_{t_n}$ and $\Delta W^m_n, m = 1, \dots, M$.

3.4.1.4 Description of the algorithm

We will perform I Picard iterations both on (3.4.2) and(3.4.1) :
→ Initialization :
For $n = N$, take $(y^{N,m,I}_N) = (\Phi(W^{N,m}_{t_N}))$ and $(z^{N,m,I}_N) = 0$.
→ Iterations : For $n = N-1, \dots, 0$:
• For $i = 0$, we set $\alpha^{M,0}_{0,n} = 0$ and $\alpha^{M,0}_{j,n} = 0$, for all $j \in \{1, \dots, d\}$.
• For $i = 1, \dots, I$:
We approximate (3.4.2) by computing for all $j \in \{1, \dots, d\}$

$$\alpha^{M,i}_{j,n} = \arg\inf_{\alpha} \frac{1}{M}\sum_{m=1}^{M}\Big|\Big\{y^{N,M,I}_{n+1}(W^{N,m}_{t_{n+1}}) + \sum_{j_1=1}^{l} h_{j_1}\Big(W^{N,m}_{t_{n+1}}, y^{N,M,I}_{n+1}(W^{N,m}_{t_{n+1}}), z^{N,M,I}_{n+1}(W^{N,m}_{t_{n+1}})\Big)\Delta B_{n,j_1}\Big\}\frac{\Delta W^m_{n,j}}{\Delta_N}$$
$$+ \frac{1}{2}\Big\{g_j(W^{N,m}_{t_n}, y^{N,M,i-1}_n(W^{N,m}_{t_n}), z^{N,M,,i-1}_n(W^{N,m}_{t_n})) + g_j(W^{N,m}_{t_{n+1}}, y^{N,M,I}_{n+1}(W^{N,m}_{t_{n+1}}), z^{N,M,I}_{n+1}(W^{N,m}_{t_{n+1}}))\Big\}\frac{(\Delta W_{n,j})^2}{\Delta_N}$$
$$- \alpha.p^m_{j,n}\Big|^2.$$

Then, we approximate (3.4.1) by calculating $\alpha_{0,n}^{M,i}$ as the minimizer of :

$$\begin{aligned} & \frac{1}{M}\sum_{m=1}^{M}\Big|y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})+hf\Big(W_{t_n}^{N,m},y_n^{N,M,i-1}(W_{t_n}^{N,m}),z_n^{N,M,i}(W_{t_n}^{N,m})\Big)\\ + & \frac{1}{2}\sum_{j_1=1}^{d^1}g_{j_1}\Big(W_{t_{n+1}}^{N,m},y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}),z_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\Big)\Delta B_{n,j_1}\\ + & \sum_{j=1}^{d}h_j\Big(W_{t_{n+1}}^{N,m},y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}),z_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\Big)\Delta W_{n,j}-\alpha p_k^m\Big|^2. \end{aligned}$$

Finally, we set $y_n^{N,M,I}(.) = (\alpha_{0,n}^{M,I}.p_n(.))$ and $z_{n,j}^{N,M,I}(.) = (\alpha_{j,n}^{M,I}.p_{j,n}(.))$. $j \in \{1,\ldots,d\}$.

3.4.2 One-dimensional case (Case when $d = d^l = 1$)

3.4.2.1 Function bases

We use the basis (**HC**) defined above. So we set :

$$d_1 = \min_{n,m} W_{t_n}^m, \quad d_2 = \max_{n,m} W_{t_n}^m \text{ and } L = \frac{d_2 - d_1}{\delta}$$

where δ is the edge of the hypercubes $(D_j)_{1\leqslant j\leqslant L}$ defined by $D_j = \Big[d+(j-1)\delta, d+j\delta\Big), \forall j$. We take at each time t_n

$$1_{D_j}(W_{t_n}^{N,m}) = 1_{[d+(j-1)\delta,d+j\delta)}(W_{t_n}^{N,m}), j = 1,\ldots,L$$

and

$$(p_{i,n}^m(.)) = \Big\{\sqrt{\frac{M}{card(D_j)}}1_{D_j}(W_{t_n}^{N,m}), 1\leqslant j\leqslant L\Big\}, i = 0,1.$$

$Card(D_j)$ denotes the number of simulations of $W_{t_n}^N$ which are in our cube D_j.
This system is orthonormal with respect to the empirical scalar product defined by

$$< \psi_1, \psi_2 >_{n,M} := \frac{1}{M}\sum_{m=1}^{M}\psi_1(W_{t_n}^{N,m})\psi_2(W_{t_n}^{N,m}).$$

In this case, the solutions of our least squares problems are given by :

$$\begin{aligned} \alpha_{1,n}^{M,i} &= \frac{1}{M}\sum_{m=1}^{M}p_{1,n}(W_{t_n}^{N,m})\Big\{\Big[y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\\ &+ \Delta_N\Big(W_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}), z_{n+1}^{N,M,}(W_{t_{n+1}}^{N,m})\Big)\Delta B_n\Big]\frac{\Delta W_n^m}{\Delta_N}\\ &+ \frac{1}{2}\Big[g\Big(W_{t_n}^{N,m}, y_n^{N,M,i-1}(W_{t_n}^{N,m}), z_n^{N,M,i-1}(W_{t_n}^{N,m})\Big)\\ &+ g\Big(W_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}), z_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\Big)\Big]\frac{(\Delta W_n^m)^2}{\Delta_N}\Big\} \end{aligned}$$

$$\begin{aligned} \alpha_{0,n}^{M,i} &= \frac{1}{M}\sum_{m=1}^{M}p_{0,n}(W_{t_n}^{N,m})\Big\{y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\\ &+ \Delta_N f\Big(W_{t_n}^{N,m}, y_n^{N,M,i-1}(W_{t_n}^{N,m}), z_n^{N,M,i}(W_{t_n}^{N,m})\Big)\\ &+ h\Big(W_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}), z_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\Big)\Delta B_n\\ &+ \frac{1}{2}g\Big(W_{t_{n+1}}^{N,m}, y_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m}), z_{n+1}^{N,M,I}(W_{t_{n+1}}^{N,m})\Big)\Delta B_n^m\Big\}. \end{aligned}$$

We note that for each value of M, N and δ, we launch the algorithm 50 times and we denote by $(Y_{0,m'}^{0,0,N,M,I})_{1 \leqslant m' \leqslant 50}$ the set of collected values. Then we calculate the empirical mean $\overline{Y}_0^{0,0,N,M,I}$ and the empirical standard deviation $\sigma^{N,M,I}$defined by :

$$\overline{Y}_0^{0,0,N,M,I} = \frac{1}{50}\sum_{m'=1}^{50} Y_{0,m'}^{0,0,N,M,I} \text{ and } \sigma^{N,M,I} = \sqrt{\frac{1}{49}\sum_{m'=1}^{50} |Y_{0,m'}^{0,0,N,M,I} - \overline{Y}_0^{0,0,N,M,I}|^2}. \tag{3.4.3}$$

3.4.2.2 Case when f, h and g are independant of y and z

Let Us denote in this case the solution of the linear equation (3.2.5) by Y_{linear}. Then, Y_{linear} admits a representation in terms of a solution of a classical BDSDE. Indeed, we can define $\widetilde{Y}_t := Y_{t,linear} + \frac{1}{2}\int_0^t g_r * dW_r$, $\forall t \in [0,T]$. Thus, $(\widetilde{Y}_t, Z_t)_{0 \leqslant t \leqslant T}$ is the solution of the following BDSDE :

$$\begin{cases} \widetilde{Y}_t = \widetilde{Y}_T + \int_t^T f s ds + \int_t^T h_s \overleftarrow{dB_s} - Z_s dW_s, \\ \widetilde{Y}_T = \Phi(W_T) + \frac{1}{2}\int_0^T g_s * dW_s. \end{cases} \tag{3.4.4}$$

The algorithm for resolving the latter equation is given by :

$$\begin{cases} \widetilde{Y}_{t_N}^N = \Phi(W_T^N) + \frac{1}{2}\sum_0^{N-1} g_{t_l}\Delta W_l, \\ \widetilde{Y}_{t_n}^N = E_{t_n}[\widetilde{Y}_{t_{n+1}}^N + \Delta_N f_{t_n} + h_{t_{n+1}}\Delta B_n], 0 \leqslant n \leqslant N-1. \end{cases} \tag{3.4.5}$$

For each n, $Y_{t_n,linear}^N$ is given by $Y_{t_n,linear}^N = \widetilde{Y}_{t_n}^N - \frac{1}{2}\sum_0^{n-1} g_{t_l}\Delta W_l$ and finally we get $Y_{0,linear}^N = \widetilde{Y}_0^N$. We note that in the linear case above, we resolve a BDSDE with linear coefficients (independent from Y and Z). We can now implement numerically our algorithm by taking :

$$\Big\{ \quad \Phi(x) = \tfrac{10}{1+|x|}, \ f(t,x,y,Z) = a_0, \ , h(t,x,y,z) = \beta_0, \ g(t,x,y,z) = \alpha_0 t$$

and we set a_0, β_0 and α_0 are fixed constants.
The following table gives the approximated solution $\overline{Y}_{0,linear}^{0,x,N,M}$ $(\sigma^{N,M})$ of the previous BDSDE-Algorithm for different values of M and N.
For $a_0 = 0.5$, $\beta_0 = 0.5$, $\alpha_0 = 0.2$ $\delta = 0.2$

	N=10	N=20	N=30
M=128	7.310(0.12)	7.491(0.13)	7.519(0.12)
M=512	7.292(0.06)	7.476(0.05)	7.498(0.05)
M=2048	7.283(0.02)	7.438(0.03)	7.504(0.07)
M=8192	7.282(0.01)	7.475(0.01)	7.504(0.01)

After that, we implement our algorithm in the general case (the approximated solution is denoted $\overline{Y}_{0,div}^{0,x,N,M}$ $(\sigma^{N,M})$), but taking the same linear coefficients token above, and we compare the non linear algorithm and the linear one in this case.
For $a_0 = 0.5$, $\beta_0 = 0.5$, $\alpha_0 = 0.2$ $\delta = 0.2$

	N=10	N=20	N=30
M=128	7.310(0.12)	7.491(0.13)	7.518(0.12)
M=512	7.292(0.06)	7.476(0.05)	7.498(0.05)
M=2048	7.283(0.02)	7.473(0.03)	7.504(0.02)
M=8192	7.282(0.01)	7.475(0.01)	7.504(0.01)

We note that the two algorithms give exactly the same results.

3.4.2.3 Comparison of numerical approximations of the solutions of the FBDSDE with divergence term, the FBDSDE and the FBSDE in the general case

Now we take

$$\begin{cases} \Phi(x) = \frac{10}{1+|x|}, \\ f(t,x,y,z) = -\theta z - ry + (y - \frac{z}{\sigma})^{-}(R-r) \\ h(t,x,y,z) = -5x + 0.5y + \beta z \\ g(t,x,y,z) = -2x + 0.5y + \alpha z \end{cases}$$

and we set $\theta = (\mu - r)/\sigma$, $\mu = 0,05$, $\sigma = 0,2$, $r = 0,01$, $R = 0,06$ and $T = 0,25$. Then we take $\beta = 0.5$, $\alpha = 0.2$ $\delta = 0.2$
The approximated solution of the BDSDE with divergence term is denoted by $\overline{Y}^{0,x,N,M,I}_{0,div}$, the one of the BDSDE is denoted by $\overline{Y}^{0,x,N,M,I}_{0,BDSDE}$ and obtained by taking $g = 0$ and the one for the standard BSDE is denoted by $\overline{Y}^{0,x,N,M,I}_{0,BSDE}$ and obtained by taking $g = h = 0$.

Algorithm with 20 Picard Iterations for Y and Z at the same time
For $N = 10$

M	$\overline{Y}^{0,x,N,M}_{0,BSDE}(\sigma^{N,M})$	$\overline{Y}^{0,x,N,M,I}_{0,BDSDE}(\sigma^{N,M,I})$	$\overline{Y}^{0,x,N,M,I}_{0,div}(\sigma^{N,M,I})$
128	6.965(0.31)	5.426(0.859)	3.967(0.903)
512	7.318(0.13)	5.259(0.39)	3.890(0.37)
2048	7.403(0.07)	5.165(0.16)	3.836(0.16)
8192	7.440(0.02)	5.142(0.08)	3.829(0.08)

For $N = 20$

M	$\overline{Y}^{0,x,N,M}_{0,BSDE}(\sigma^{N,M})$	$\overline{Y}^{0,x,N,M,I}_{0,BDSDE}(\sigma^{N,M,I})$	$\overline{Y}^{0,x,N,M,I}_{0,div}(\sigma^{N,M,I})$
128	6.31(0.56)	7.066(1.33)	5.697(1.32)
512	7.117(0.25)	6.929(0.59)	5.614(0.55)
2048	7.287(0.13)	6.820(0.29)	5.555(0.27)
8192	7.375(0.09)	6.785(0.12)	5.546(0.12)

For $N = 30$

M	$\overline{Y}^{0,x,N,M}_{0,BSDE}(\sigma^{N,M})$	$\overline{Y}^{0,x,N,M,I}_{0,BDSDE}(\sigma^{N,M,I})$	$\overline{Y}^{0,x,N,M,I}_{0,div}(\sigma^{N,M,I})$
128	6.208(0.78)	7.605(1.73)	6.540(1.81)
512	6.688(0.42)	7.281(0.73)	6.143(1.17)
2048	7.145(0.36)	7.141(0.35)	6.176(0.34)
8192	7.381(0.02)	7.055(0.19)	6.09(0.19)

Graphical comparison between BDSDEs with divergence term, the BDSDEs and standard BSDEs

Here we take the same coefficients as in the previous paragraph for the Quasilinear BSDE, we take $g = 0$ for the BDSDE and we take $h = 0$ for the standard BSDE. Then we variate N, M and δ, by taking these quantities as follows : Let $j \in \mathbb{N}$, we take $\alpha_M = 3.4$, $\beta_\delta = 1$, $N = 2(\sqrt{2})^{(j-1)}$, $M = 2(\sqrt{2})^{\alpha_M(j-1)}$ and $\delta = 1.2(\sqrt{2})^{(j-1)(\beta_\delta+1)/2}$. Then, we draw the map of each solution at $t = 0$ with respect to j.

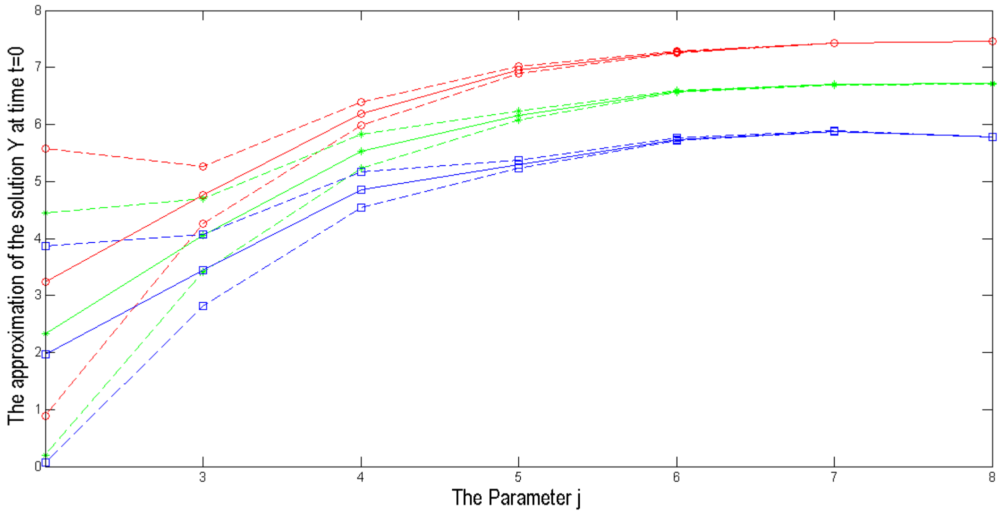

FIGURE 3.1 – Comparison of the BSDE's solution, the BDSDE's one and the solution of the BDSDE with divergence term : The solution of the BSDE is in red with circle markers, the solution of the BDSDE is in green with star markers and the one for the BDSDE with divergence term is in blue with square markers.

3.5 Appendix

3.5.1 Proof of Lemma 3.3.3

Fist, we consider the case when b, σ, f, g, h and $\Phi \in C_b^1$ and satisfying assumptions **(H)** and **(HD2')**. Let $\nabla\Theta^{t,0} := (I_d, \nabla Y^{t,0}, \nabla Z^{t,0})$ be the solution of the following equation

$$\begin{aligned}
\nabla Y_s^{t,0} &= \nabla\Phi(W_T^{t,0}) + \int_s^T \Big(\nabla_x f(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0}) \\
&+ \nabla_y f(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Y_r^{t,0} + \nabla_z f(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Z_r^{t,0}\Big) dr \\
&+ \int_s^T \Big(\nabla_x h(r, W_r^{t,0}, Y_r^{t,x}, Z_r^{t,x}) + \nabla_y h(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Y_r^{t,0} \\
&+ \nabla_z h(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Z_r^{t,0}\Big) \overleftarrow{dB_r} + \frac{1}{2}\int_s^T \Big(\nabla_x g(r, W_r^{t,0}, Y_r^{t,x}, Z_r^{t,x}) \\
&+ \nabla_y g(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Y_r^{t,0} + \nabla_z g(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Z_r^{t,0}\Big) dW_r \\
&+ \frac{1}{2}\int_s^T \Big(\nabla_x g(r, W_r^{t,0}, Y_r^{t,x}, Z_r^{t,x}) + \nabla_y g(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Y_r^{t,0} \\
&+ \nabla_z g(r, W_r^{t,0}, Y_r^{t,0}, Z_r^{t,0})\nabla Z_r^{t,0}\Big) \overleftarrow{dW_r} - \int_s^T \nabla Z_r^{t,0} dW_r.
\end{aligned} \tag{3.5.1}$$

Since $\nabla Y^{t,0}$ is the solution of the linear FBDSDE (3.5.1). Using estimation (3.3.32), we get

$$E^m E[\sup_{0 \leqslant t \leqslant T} |\nabla Y_s^{t,0}|^2] \leqslant C_2. \tag{3.5.2}$$

Now, we use the following representation result

$$Z_s^{t,0} = \nabla Y_s^{t,0}, s \in [t, T], \tag{3.5.3}$$

which gives

$$E^m \|Z_s^{t,0}\|_2 \leqslant E^m \|\nabla Y_s^{t,0}\|_2 \leqslant C_2. \tag{3.5.4}$$

Now the aim is to generalize the previous estimation to Lipschitz coefficients case. So let Φ, f, h and g satisfying Assumptions (**H**) and (**HD2′**) and let Φ^k, f^k, h^k and g^k smooth molifiers of these functions. Denoting $Z^{t,0,k}$ the solution of the F-BDSDE associated to the regular coefficients, we deduce from (3.5.4) that $E^m\|Z_s^{k,t,0}\|_2 \leqslant C_2$. Using the stability result (3.3.34) , we get

$$\lim_{k\longrightarrow+\infty} E^m E\Big[\int_t^T |Z_s^{k,t,0} - Z_s^{t,0}|^2\Big] ds = 0. \tag{3.5.5}$$

We deduce that for a.e. $s \in [t,T]$, there exist a subsequence of $(Z^{k,t,0})_k$ such that $\lim_{k\longrightarrow+\infty} Z_s^{k,t,0} = Z_s^{t,0}$ in probability. By the Fatou's Lemma, we get $E\|Z_s^{t,0}\|_2 \leqslant C_2$. Inserting the latter inequality in estimation (3.3.33), we get the estimation (3.3.36).

□

Chapitre 4

Empirical Regression method for BDSDEs

4.1 Introduction

In this work, we analyse the regression error arising from an algorithm approximating the solution of a discrete time BDSDE. We use the least-squares regression approach developed by Gobet, Lemor and Warin [34] and more recently by Gobet and Turkedjiev [35]. The tools for the regression error analysis, arising from discrete BSDE's approximation, was developed recently in [35] in a very general context. These tools will allow us to analyse the regression error in the doubly stochastic framework. The BDSDE of our interest is of the following form

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr + \int_s^T h(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\overleftarrow{dB}_r - \int_s^T Z_r^{t,x} dW_r, \quad (4.1.1)$$

where $(X_s^{t,x})_{t \leqslant s \leqslant T}$ is a diffusion process starting from x at time t driven by the finite dimensional brownian motion $(W_t)_{t \geqslant 0}$. The differential term with $\overleftarrow{dB}_t$ refers to the backward stochastic integral with respect to a l-dimensional brownian motion on $\big(\Omega, \mathcal{F}, \mathbb{P}, (B_t)_{t \geqslant 0}\big)$.

First, let Us recall the principle of the least squares algorithm as presented in the work of Gobet, Lemor and Warin [34]. The discrete BSDE arising from the time discretization of the standard BSDE, given by (4.1.1) for $h = 0$, is the following One step forward Dynamic Programming (ODP for short) equation : $Y_{t_N} = \Phi(X_{t_N})$ and for all $i \in \{N-1, \dots, 0\}$

$$\begin{aligned} Y_{t_i} &= E_{t_i}[Y_{t_{i+1}} + \Delta_i f(t_i, X_{t_i}, Y_{t_{i+1}}, Z_{t_i})], \\ \Delta_i Z_{t_i} &= E_{t_i}\Big[Y_{t_{i+1}} \Delta W_i^*\Big], \end{aligned} \quad (4.1.2)$$

where $t_i \in \pi := \{t_0, \dots, t_N\}$ and π is a discrete time grid of the time interval $[0,T]$, $\Delta W_i := W_{t_{i+1}} - W_{t_i}$, E_{t_i} denotes the conditional expectation given $\mathcal{F}_{t_i}$ and * denotes the transposition operator.
Since $(X_{t_i})_i$ is a Markov chain, there exist deterministic measurable functions $y_i(.)$ and $z_i(.)$, but unknown, such that $Y_{t_i} = y_i(X_{t_i})$ and $Z_{t_i} = z_i(X_{t_i})$. The functions $y_i(.)$ and $z_i(.)$ are solutions of least squares problems in $L_2(\Omega, \mathcal{F}_{t_i})$ and can be approximated on a finite dimensional subspace. The approximations are computed by empirical least-squares regression using Monte Carlo simulations of the process X.
Recently, Gobet and Turkedjiev [35] studied a discrete BSDE in the form of a Multi step forward Dynamic Programming (MDP for short) equation given by

$$\begin{aligned} Y_{t_i} &= E_{t_i}[\Phi_{t_{i_N}} + \sum_{k=i}^{N-1} \Delta_i f(t_k, X_{t_k}, Y_{t_{k+1}}, Z_{t_k})], \\ \Delta_i Z_{t_i} &= E_{t_i}\Big[\{\Phi_{t_{i_N}} + \sum_{k=i}^{N-1} \Delta_i f(t_k, X_{t_k}, Y_{t_{k+1}}, Z_{t_k})\} \Delta W_i^*\Big]. \end{aligned} \quad (4.1.3)$$

Using the tower property of conditional expectations, we note that the ODP and the MDP coincide.
In their recent work [35], Gobet and Turkedjiev proved that the MDP leads to better error estimates. Indeed, the quadratic error is the average of local error terms rather than the sum, as in the ODP's case. In this sense, the MDP gives better estimates.

For BDSDEs, the ODP scheme was studied by Aboura [1]. He proposed an algorithm based on the empirical least-squares regression approach to resolve numerically BDSDEs, following [34]. He considered the solution of the BDSDE at time t_i as a measurable deterministic function of $\Big(X_{t_i}, (B_{t_{k+1}} - B_{t_k})_{i \leqslant k \leqslant N}\Big)$. His approach imply a high-dimensionality problem, since he is dealing with a dimension of $d + l \times N$ and the parameter N goes to infinity. In our case, we study the MDP scheme for BDSDEs and we proceed to a conditional analysis of the error, given the trajectories of the Brownian motion B. Thus, we reduce the dimension of the regression from $d + l \times N$, to d. In this work, in order to simplify the computations, we will treat the one dimensional case for the process Y and the Brownian motion B.

4.2 Preliminaries and notations

4.2.1 Forward Backward Doubly Stochastic Differential Equation

Let $\{W_s, 0 \leqslant s \leqslant T\}$ and $\{B_s, 0 \leqslant s \leqslant T\}$ be two mutually independent standard Brownian motion processes, with values respectively in $\mathbb{R}^d$ and $\mathbb{R}$ where $T > 0$ is a fixed horizon time, on the probability space $(\Omega, \mathcal{F}, P)$.
We shall work on the product space $\Omega := \Omega_W \times \Omega_B$, where Ω_W is the set of continuous functions from $[0,T]$ into $\mathbb{R}^d$ and Ω_B is the set of continuous functions from $[0,T]$ into $\mathbb{R}$.
We fix $t \in [0,T]$. For each $s \in [t,T]$, we define

$$\mathcal{F}_s^t := \mathcal{F}_{t,s}^W \vee \mathcal{F}_{s,T}^B$$

where $\mathcal{F}_{t,s}^W = \sigma\{W_r - W_t, t \leqslant r \leqslant s\}$, and $\mathcal{F}_{s,T}^B = \sigma\{B_r - B_s, s \leqslant r \leqslant T\}$. We take $\mathcal{F}^W = \mathcal{F}_{0,T}^W$, $\mathcal{F}^B = \mathcal{F}_{0,T}^B$ and $\mathcal{F} = \mathcal{F}^W \vee \mathcal{F}^B$.
We define the probability measures P_W on $(\Omega_W, \mathcal{F}^W)$ and P_B on $(\Omega_B, \mathcal{F}^B)$. We then define the probability measure $P := P_W \otimes P_B$ on $(\Omega, \mathcal{F}^W \times \mathcal{F}^B)$. Without loss of generality, we assume that $\mathcal{F}^W$ and $\mathcal{F}^B$ are complete.
Note that the collection $\{\mathcal{F}_s^t, s \in [t,T]\}$ is neither increasing nor decreasing, and it does not constitute a filtration.
For all $(t,x) \in [0,T] \times \mathbb{R}^d$, let $(X_s^{t,x})_s$ be the unique strong solution of the following stochastic differential equation :

$$dX_s^{t,x} = b(X_s^{t,x})ds + \sigma(X_s^{t,x})dW_s, \quad s \in [t,T], \qquad X_s^{t,x} = x, \quad 0 \leqslant s \leqslant t, \tag{4.2.1}$$

where b and σ are two functions on $\mathbb{R}^d$ with values respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$. We will omit the dependence of the forward process X in the initial condition if it starts at time $t = 0$.
We consider the following BDSDE : For all $t \leqslant s \leqslant T$,

$$\begin{cases} dY_s^{t,x} &= -f(Y_s^{t,x})ds - h(Y_s^{t,x})\overleftarrow{dB_s} + Z_s^{t,x}dW_s, \\ Y_T^{t,x} &= \Phi(X_T^{t,x}), \end{cases} \tag{4.2.2}$$

where f and h are two $\mathbb{R}$-valued functions on $\mathbb{R}$ and Φ is a an $\mathbb{R}$-valued function on $\mathbb{R}^d$.
The following assumptions will be needed in our work :

Assumption (H1) There exist a positive constant K such that

$$|b(x) - b(x')| + \|\sigma(x) - \sigma(x')\| \leqslant K|x - x'|, \forall x, x' \in \mathbb{R}^d.$$

Assumption (H2) There exist non-negative constants $C_f, C_h, C_\xi, L_f, L_h$ and L_ξ such that

- **(i)** $\forall y_1, y_2 \in \mathbb{R}^k, |f(y_1) - f(y_2)| \leqslant L_f|y_1 - y_2|$,
- **(ii)** $\forall y_1, y_2 \in \mathbb{R}^k, |h(y_1) - h(y_2)| \leqslant L_h|y_1 - y_2|$,
- **(iii)** $|f(0)|$ and $|h(0)|$ are bounded by C_f and C_h respectively,
- **(iv)** The function Φ is a measurable function on $\mathbb{R}^d$ bounded by C_ξ ,
- **(v)** $\forall x_1, x_2 \in \mathbb{R}^d, |\Phi(x_1) - \Phi(x_2)| \leqslant L_\xi|x_1 - x_2|$.

4.2.2 Numerical Scheme for decoupled Forward-BDSDE

In order to approximate the solution of the F-BDSDE (4.2.1)-(4.2.2), we introduce the following discretized version. Let

$$\pi : t_0 = 0 < t_1 < \ldots < t_N = T, \tag{4.2.3}$$

be a partition of the time interval $[0, T]$ with time step $\Delta_i := t_{i+1} - t_i$, $0 \leqslant i \leqslant N-1$. Throughout the rest, we will use the notations $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ and $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, for $i = 1, \ldots, i$.
The forward component X will be approximated by the classical Euler scheme :

$$\begin{cases} X^N_{t_0} &= X_{t_0}, \\ X^N_{t_i} &= X^N_{t_{i-1}} + b(X^N_{t_{n-1}})(t_i - t_{i-1}) + \sigma(X^N_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}), \text{ for } i = 1, \ldots, N. \end{cases} \tag{4.2.4}$$

It is known that as N goes to infinity, one has $\sup_{0 \leqslant i \leqslant N} E|X_{t_i} - X^N_{t_i}|^2 \to 0$.
The solution Y of (4.2.2) is approximated by (Y^N) defined by the following **Multi step-forward Dynamic Programming (MDP)** equation : For $i = N-1, \ldots, 0$, we set

$$Y^N_{t_i} \quad = \quad E_{t_i}\Big[\Phi(X^N_T) + \sum_{k=i}^{N-1} \Delta_k f(Y^N_{t_{k+1}}) + h(Y^N_{t_{k+1}})\Delta B_k\Big] \tag{4.2.5}$$

We define the σ-algebra $\mathcal{G}_t := \mathcal{F}^W_{0,t} \vee \mathcal{F}^B_{0,T}$ and we note by $E_{t_i}[.]$ the conditional expectation over $\mathcal{G}_{t_i}$.
We note that in the case we are treating, when the drivers are independent from the variable z, we do not have to approximate the control process Z, since we don't need it to approximate Y.

4.3 A Priori estimates

In this section, we establish a priori estimates on discrete BDSDEs. These estimates will be needed later for the regression analysis. The discrete BDSDEs that we will study in this section are of the following form :
For $j = 1, 2$, we set $Y_{j,N} = \xi_j$ and for all $i = N-1, \ldots, 0$

$$Y_{j,i} = E_{t_i}[Y_{j,i+1} + \Delta_i f_{j,i}(Y_{j,i+1}) + h_{j,i+1}(Y_{j,i+1})\Delta B_i], \tag{4.3.1}$$

where for all $i \in \{0, \ldots, N-1\}$,

$y \mapsto f_{1,i}(y)$ and $y \mapsto f_{2,i}(y)$ are given real-valued and $\mathcal{G}_{t_i} \otimes \mathcal{B}(\mathbb{R})$-measurable functions on $\mathbb{R}$
and
$y \mapsto h_{1,i+1}(y)$ and $y \mapsto h_{2,i+1}(y)$ are given real-valued and $\mathcal{G}_{t_{i+1}} \otimes \mathcal{B}(\mathbb{R})$-measurable functions on $\mathbb{R}$.

We set $\delta Y_i := Y_{1,i} - Y_{i,2}$, $\delta Y^h_{i+1} := \delta Y_{i+1} + \{h_{1,i+1}(Y_{1,i+1}) - h_{2,i+1}(Y_{2,i+1})\}\Delta B_i$, $\delta f_i := f_{1,i}(Y_{1,i}) - f_{2,i}(Y_{1,i})$ and $\delta h_{i+1} := h_{1,i+1}(Y_{1,i}) - h_{2,i+1}(Y_{1,i})$. Then we state the following lemma, which gives a local estimate on the solutions of two discrete BDSDEs :

Lemma 4.3.1. *We assume that for $j \in \{\mathbf{1}, 2\}$, ξ_j belongs to $\mathrm{L}_2(\mathcal{F}^0_T)$ and that for $i \in \{0, \ldots, N-1\}$, $f_{1,i}(Y_{1,i+1})$ and $h_{1,i+1}(Y_{1,i+1})$ belong to $\mathrm{L}_2(\mathcal{F}^0_T)$. In addition, we assume that $f_{2,i}$ and $h_{2,i+1}$ are Lipschitz continuous with finite non-negative Lipschitz constants $L_{f_{2,i}}$ and $L_{h_{2,i+1}}$. Finally, for each $i \in \{0, \ldots, N-1\}$, we assume that γ^B_i is a positive constant and that we can choose $\gamma_i > 0$ such that $4(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}} \leqslant 1$. Then, we have*

$$\begin{aligned} |\delta Y_i|^2 \quad &\leqslant \quad \Big\{(1 + \gamma_i \Delta_i)(1 + C^Y_{\Delta B_i}|\Delta B_i|) + \frac{\Delta_i}{2}\Big\}E_{t_i}[|\delta Y_{i+1}|^2] + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2] \\ &+ \quad 2(1 + \gamma_i \Delta_i)(|\Delta B_i| + \frac{1}{\gamma^B_i})|\Delta B_i| E_{t_i}[|\delta h_{i+1}|^2], \end{aligned} \tag{4.3.2}$$

where

$$C^Y_{\Delta B_i} := \gamma^B_i + 2(|\Delta B_i| + \frac{1}{\gamma^B_i})L^2_{h_{2,i+1}}. \tag{4.3.3}$$

Proof. From (4.3.1), we have

$$\delta Y_i = E_{t_i}[\delta Y^h_{i+1} + \Delta_i\{f_{1,i}(Y_{1,i+1}) - f_{2,i}(Y_{2,i+1})\}].$$

Using the Young inequality with the positive constant γ_i, we get

$$\begin{aligned}|\delta Y_i|^2 &\leqslant (1+\gamma_i\Delta_i)|E_{t_i}[\delta Y^h_{i+1}]|^2 \\ &+ (1+\frac{1}{\gamma_i\Delta_i})|E_{t_i}[\Delta_i\{f_{1,i}(Y_{1,i+1}) - f_{2,i}(Y_{2,i+1})\}]|^2.\end{aligned}$$

Plugging $f_2(Y_{1,i+1})$, we get

$$\begin{aligned}|\delta Y_i|^2 &\leqslant (1+\gamma_i\Delta_i)|E_{t_i}[\delta Y^h_{i+1}]|^2 + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2] \\ &+ 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i L^2_{f_{2,i}} E_{t_i}[|\delta Y_{i+1}|^2]. \end{aligned} \tag{4.3.4}$$

On the other hand, we have

$$L_{h_{2,i+1}}\delta Y^h_{i+1} := \delta Y_{i+1} + \{h_{1,i+1}(Y_{1,i+1}) - h_{2,i+1}(Y_{2,i+1})\}\Delta B_i.$$

Plugging $h_2(Y_{1,i+1})$ and using the Young inequality with γ_i^B, we get

$$\begin{aligned}|\delta Y^h_{i+1}|^2 &\leqslant (1+\gamma_i^B|\Delta B_i|)|\delta Y_{i+1}|^2 + 2(|\Delta B_i| + \frac{1}{\gamma_i^B})L^2_{h_{2,i+1}}|\Delta B_i||\delta Y_{i+1}|^2 \\ &+ 2(|\Delta B_i| + \frac{1}{\gamma_i^B})|\Delta B_i||\delta h_{i+1}|^2. \end{aligned} \tag{4.3.5}$$

Inserting the estimation (4.3.5) in (4.3.4), we obtain

$$\begin{aligned}|\delta Y_i|^2 &\leqslant \Big\{(1+\gamma_i\Delta_i)[(1+\gamma_i^B|\Delta B_i|) + 2(|\Delta B_i| + \frac{1}{\gamma_i^B})L^2_{h_{2,i+1}}|\Delta B_i|] \\ &+ 2(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}}\Delta_i\Big\}E_{t_i}[|\delta Y_{i+1}|^2] + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2] \\ &+ 2(1+\gamma_i\Delta_i)(|\Delta B_i| + \frac{1}{\gamma_i^B})|\Delta B_i|E_{t_i}[|\delta h_{i+1}|^2 \\ &\leqslant \Big\{(1+\gamma_i\Delta_i)\Big(1 + [\gamma_i^B + 2(|\Delta B_i| + \frac{1}{\gamma_i^B})L^2_{h_{2,i+1}}]|\Delta B_i|\Big) \\ &+ 2(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}}\Delta_i\Big\}E_{t_i}[|\delta Y_{i+1}|^2] + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2] \\ &+ 2(1+\gamma_i\Delta_i)(|\Delta B_i| + \frac{1}{\gamma_i^B})|\Delta B_i|E_{t_i}[|\delta h_{i+1}|^2].\end{aligned}$$

The assumption on γ_i insures that $2(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}} \leqslant \frac{1}{2}$. Then, we get

$$\begin{aligned}|\delta Y_i|^2 \leqslant \Big\{(1+\gamma_i\Delta_i)\Big(1 + [\gamma_i^B + 2(|\Delta B_i| + \frac{1}{\gamma_i^B})L^2_{h_{2,i+1}}]|\Delta B_i|\Big) + \frac{\Delta_i}{2}\Big\}E_{t_i}[|\delta Y_{i+1}|^2] \\ + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2] + 2(1+\gamma_i\Delta_i)(|\Delta B_i| + \frac{1}{\gamma_i^B})|\Delta B_i|E_{t_i}[|\delta h_{i+1}|^2].\end{aligned}$$

□

The following proposition gives a global stability result on the solutions of two discrete BDSDEs :

Proposition 4.3.1. *We assume that for $j \in \{\mathbf{1}, 2\}$, ξ_j belongs to $\mathrm{L}_2(\mathcal{F}^0_T)$ and that for $i \in \{0, \dots, N-1\}$, $f_{1,i}(Y_{1,i+1})$ and $h_{1,i+1}(Y_{1,i+1})$ belong to $\mathrm{L}_2(\mathcal{F}^0_T)$. In addition, we assume that $f_{2,i}$ and $h_{2,i+1}$ are Lipschitz continuous with finite non-negative Lipschitz constants $L_{f_{2,i}}$ and $L_{h_{2,i+1}}$. Finally, for each $i \in \{0, \dots, N-1\}$,* *we assume that γ_i^B is a positive constant and that we can choose $\gamma_i > 0$ such that $4(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}} \leqslant 1$. Then, the following estimation holds a.s. for all i :*

$$\Gamma_i\Gamma_i^B|\delta Y_{t_i}|^2 \leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B E_{t_i}[|\delta\xi|^2] + \sum_{k=i}^{N-1} 2(\Delta_k + \frac{1}{\gamma_k})\Delta_k E_{t_i}[|\delta f_k|^2]\Gamma_k\Gamma_k^B$$
$$+ \sum_{k=i}^{N-1} 2(1+\gamma_k\Delta_k)(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| E_{t_i}[|\delta h_k|^2]\Gamma_k\Gamma_k^B\Big\}, \tag{4.3.6}$$

where we set for all $k \in \{1,\ldots,N\}$:

$$\Gamma_k := \Pi_{j=0}^{k-1}(1+\gamma_j\Delta_j), \Gamma_0 := 1, \tag{4.3.7}$$
$$\Gamma_k^B := \Pi_{j=0}^{k-1}(1+C^Y_{\Delta B_j}|\Delta B_j|), \Gamma_0^B := 1 \tag{4.3.8}$$
$$\text{and} \qquad C^Y_{\Delta B_i} := \gamma_i^B + 2(|\Delta B_i| + \frac{1}{\gamma_i^B})L^2_{h_{2,i+1}}. \tag{4.3.9}$$

Proof. We set $\lambda_i := \{(1+\gamma_{i-1}\Delta_{i-1})(1+C^Y_{\Delta B_{i-1}}\Delta_j|\Delta B_{i-1}|) + \frac{\Delta_{i-1}}{2}\}\lambda_{i-1}$, for $i \in \{1,\ldots,N\}$ and $\lambda_0 := 1$. We multiply the two sides of (4.3.2) by λ_i, we get

$$|\delta Y_i|^2\lambda_i \leqslant E_{t_i}[|\delta Y_{i+1}|^2]\lambda_{i+1} + 2(\Delta_i + \frac{1}{\gamma_i})\Delta_i E_{t_i}[|\delta f_i|^2]\lambda_i$$
$$+ 2(1+\gamma_i\Delta_i)(|\Delta B_i| + \frac{1}{\gamma_i^B})|\Delta B_i| E_{t_i}[|\delta h_{i+1}|^2]\lambda_i.$$

Now, we sum up the previous inequallity from i to $N-1$ after applying $E_{t_i}[.]$, we obtain

$$|\delta Y_i|^2\lambda_i \leqslant E_{t_i}[|\delta\xi|^2]\lambda_N + \sum_{k=i}^{N-1} 2(\Delta_k + \frac{1}{\gamma_k})\Delta_k E_{t_i}[|\delta f_k|^2]\lambda_k$$
$$+ \sum_{k=i}^{N-1} 2(1+\gamma_k\Delta_k)(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| E_{t_i}[|\delta h_k|^2]\lambda_k.$$

Using the inequality $(1+x+h) \leqslant (1+x)e^h, \forall x \in \mathbb{R}$ and h small enough, we obtain

$$\Gamma_i\Gamma_i^B \leqslant \lambda_i = exp\Big\{\sum_{k=0}^{i-1} \ln\{(1+\gamma_k\Delta_k)(1+C^Y_{\Delta B_k}|\Delta B_k|) + \frac{\Delta_k}{2}\}\Big\} \leqslant e^{\frac{T}{2}}\Gamma_i\Gamma_i^B,$$

The last inequality gives

$$|\delta Y_i|^2\Gamma_i\Gamma_i^B \leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B E_{t_i}[|\delta\xi|^2] + \sum_{k=i}^{N-1} 2(\Delta_k + \frac{1}{\gamma_k})\Delta_k E_{t_i}[|\delta f_k|^2]\Gamma_k\Gamma_k^B$$
$$+ \sum_{k=i}^{N-1} 2(1+\gamma_k\Delta_k)(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| E_{t_i}[|\delta h_k|^2]\Gamma_k\Gamma_k^B\Big\}.$$

□

As an application of the global a priori estimates given in Proposition (4.3.1), we give the following result, which is an a.s. upper bound to the solution of the discrete BDSDE (4.2.5).

Proposition 4.3.2. *Under Assumptions **(H1)** and **(H2)**, the solution of the discrete BDSDE (4.3.1) has an a.s. upper bound given in the following : For all $i \in \{0,\ldots,N\}$,*

$$|Y^N_{t_i}| \leqslant C_y^B := \exp\Big\{(\frac{1}{4} + 4L_f^2)T + L_h^2\sum_{j=0}^{N-1}|\Delta B_j|^2\Big\}\exp\Big\{(\frac{1}{2} + L_h^2)\sum_{j=0}^{N-1}|\Delta B_j|\Big\}\Big\{C_\xi^2 + \frac{C_f^2 T}{4L_f^2}$$
$$+ 2C_h^2\sum_{j=0}^{N-1}|\Delta B_j|\Big\}^{\frac{1}{2}} \ a.s. \tag{4.3.10}$$

Proof. In order to obtain the a.s. upper bound result, we apply Proposition (4.3.1) with two constants $\gamma_i^B > 0$ and $\gamma_i > 0$ such that $4(\Delta_i + \frac{1}{\gamma_i})L^2_{f_{2,i}} \leqslant 1$. In one hand, we set $Y_{1,i} := 0$, $\xi_1 = 0$, and $h_{1,i}(y) := 0$. On the other hand, we set $Y_{2,i} := Y^N_{t_i}$, $\xi_2 = \xi$, $f_{2,i}(y) := f(y)$ and $h_{2,i+1}(y) := h(y)$. This gives

$$\begin{aligned}
|Y^N_{t_i}|^2\Gamma_i\Gamma_i^B &\leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B E_{t_i}[|\xi|^2] + \sum_{k=i}^{N-1} 2(\Delta_k + \frac{1}{\gamma_k})\Delta_k E_{t_i}[|f(0)|^2]\Gamma_k\Gamma_k^B \\
&+ \sum_{k=i}^{N-1} 2(1+\gamma_k\Delta_k)(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| E_{t_i}[|h(0)|^2]\Gamma_k\Gamma_k^B\Big\} \\
&= e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B E_{t_i}[|\xi|^2] + 2\sum_{k=i}^{N-1} \frac{1}{\gamma_k}\Delta_k E_{t_i}[|f(0)|^2]\Gamma_{k+1}\Gamma_k^B \\
&+ 2\sum_{k=i}^{N-1}(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| E_{t_i}[|h(0)|^2]\Gamma_{k+1}\Gamma_k^B\Big\}.
\end{aligned}$$

Using Assumption **(H2)**, we get

$$\begin{aligned}
|Y^N_{t_i}|^2\Gamma_i\Gamma_i^B &\leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B C_\xi^2 + \sum_{k=i}^{N-1} \frac{1}{\gamma_k}\Delta_k C_f^2\Gamma_{k+1}\Gamma_k^B \\
&+ 2\sum_{k=i}^{N-1}(|\Delta B_k| + \frac{1}{\gamma_k^B})|\Delta B_k| C_h^2\Gamma_{k+1}\Gamma_k^B\Big\}.
\end{aligned}$$

We take $\gamma_k^B := 1$ and $\gamma_k := 8L_f^2$, so that the condition $4(\Delta_i + \frac{1}{\gamma_i})L_f^2 \leqslant 1$ is satisfied for N enough large. Indeed, we can choose N enough large to get $4\Delta_i L_f^2 \leqslant \frac{1}{2}$, which implies that, when we take $\gamma_k := 8L_f^2$ for all k, we have

$$4(\Delta_i + \frac{1}{\gamma_i})L_f^2 \leqslant \frac{1}{2} + \frac{4L_f^2}{\gamma_i} = 1.$$

Then, the inequality (4.3.11) becomes

$$\begin{aligned}
|Y^N_{t_i}|^2\Gamma_i\Gamma_i^B &\leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B C_\xi^2 + \frac{C_f^2}{4L_f^2}\sum_{k=i}^{N-1} \Delta_k\Gamma_{k+1}\Gamma_k^B \\
&+ 2C_h^2\sum_{k=i}^{N-1}(|\Delta B_k| + 1)|\Delta B_k|\Gamma_{k+1}\Gamma_k^B\Big\}.
\end{aligned}$$

Since $|\Delta B_k| + 1 \leqslant C_{\Delta B_k}$, we have $\Gamma_k^B \leqslant (|\Delta B_k| + 1)\Gamma_k^B \leqslant \Gamma_{k+1}^B$. Thus, the last inequality becomes

$$\begin{aligned}
|Y^N_{t_i}|^2\Gamma_i\Gamma_i^B &\leqslant e^{\frac{T}{2}}\Big\{\Gamma_N\Gamma_N^B C_\xi^2 + \frac{C_f^2}{4L_f^2}\sum_{k=i}^{N-1} \Delta_k\Gamma_{k+1}\Gamma_{k+1}^B \\
&+ 2C_h^2\sum_{k=i}^{N-1}|\Delta B_k|\Gamma_{k+1}\Gamma_{k+1}^B\Big\}. \qquad (4.3.11)
\end{aligned}$$

Now, giving the a.s. upper bound for $Y^N_{t_i}$ remains to give an a.s. upper bound for $\Gamma_k\Gamma_k^B$ for all k in $\{0,\dots,N\}$. This can be easily done for our explicit choice $\gamma_k^B := 1$ and $\gamma_k := 8L_f^2$ for all k in $\{0,\dots,N-1\}$. Indeed, for all k in $\{0,\dots,N\}$, we have

$$\Gamma_k := \Pi_{j=0}^{k-1}(1 + 8L_f^2\Delta_j) \leqslant e^{\sum_{j=0}^{N-1} 8L_f^2\Delta_j} = e^{8L_f^2 T}. \qquad (4.3.12)$$

In the same manner, we have also

$$\begin{aligned}
\Gamma_k^B &:= \Pi_{j=0}^{k-1}(1 + C_{\Delta B_j}|\Delta B_j|) \leqslant \exp\Big\{\sum_{j=0}^{N-1} C_{\Delta B_j}|\Delta B_j|\Big\} \\
&= \exp\Big\{\sum_{j=0}^{N-1}\Big(1 + 2(|\Delta B_i| + 1)L^2_{h_{2,i+1}}\Big)|\Delta B_j|\Big\} \\
&= \exp\Big\{(1 + 2L_h^2)\sum_{j=0}^{N-1}|\Delta B_i|\Big\}\exp\Big\{2L_h^2\sum_{j=0}^{N-1}|\Delta B_i|^2\Big\}. \qquad (4.3.13)
\end{aligned}$$

Plugging the estimations (4.3.12) and (4.3.13) in the inequality (4.3.11), we obtain the following a.s. upper bound

$$
\begin{aligned}
|Y_{t_i}^N|^2 &\leqslant \exp\Big\{(\frac{1}{2}+8L_f^2)T+2L_h^2\sum_{j=0}^{N-1}|\Delta B_i|^2\Big\}\exp\Big\{(1+2L_h^2)\sum_{j=0}^{N-1}|\Delta B_i|\Big\}\Big\{C_\xi^2+\frac{C_f^2T}{4L_f^2}\\
&+ 2C_h^2\sum_{k=0}^{N-1}|\Delta B_k|\Big\}.
\end{aligned}
$$

□

4.4 Monte Carlo regression scheme

The purpose of this section is the approximation of conditional expectations involved in (4.2.5) using linear least squares regressions.

4.4.1 Preliminaries

Let $\Delta B := \{\Delta B_i, i=0,\dots,N-1\}$ be the sequence of N random variables generated by the discretized Brownian motion B. For the computations of the conditional expectations and in order to alleviate notations, $E_{\Delta B}[.]$ will denote the conditional expectation given the sigma-algebra $\sigma(\Delta B)$. Using the Markov property of (X^N) with respect to the filtration $(\mathcal{G}_t)_{0\,\leqslant\, t\,\leqslant\, T}$, we deduce that there exists a sequence of random real-valued functions $(y_i(\Delta B,.))_{0\,\leqslant\, i\,\leqslant\, N-1}$ such that, $\forall i=0,\dots,N-1$, $y_i(\Delta B,.)$ is $\mathcal{G}_{t_i}$-measurable and satisfies a.s.

$$
Y_{t_i}^N = y_i(\Delta B, X_{t_i}^N). \tag{4.4.1}
$$

In the following, we recall the definition of the Least-squares regression as stated in [35] :

Definition 4.4.1. *Let $l_1,l_2\geqslant 1$. We consider the two probability spaces $(\Omega_B,\mathcal{F}_{0,T}^B,P_B)$ and $(\mathbb{R}^{l_1},\mathcal{B}(\mathbb{R}^{l_1}),\nu)$. S is a $\mathcal{F}_{0,T}^B\otimes\mathcal{B}(\mathbb{R}^{l_1})$-measurable $\mathbb{R}^{l_2}$-valued function such that $S(w,.)\in L_2(\mathcal{B}(\mathbb{R}^{l_1}))$ for P_B-a.e. $w\in\Omega_B$ and $\mathcal{K}$ is the linear vector subspace of $L_2(\mathcal{B}(\mathbb{R}_1^l))$ spanned by deterministic $\mathbb{R}^{l_2}$-valued functions $\{p_k, k\geqslant 1\}$. The least squares approximation of S in the space $\mathcal{K}$ with respect to ν is a $P_B\times\nu$-a.e. unique and $\mathcal{F}_{0,T}^B\otimes\mathcal{B}(\mathbb{R}^{l_1})$-measurable function S^* given by :*

$$
S^*(w,.) = \operatorname*{arginf}_{\phi\in\mathcal{K}} \int |\phi(x)-S(w,x)|^2\nu(dx). \tag{4.4.2}
$$

Then, we say that S^* solves $OLS(S,\mathcal{K},\nu)$.
In the same manner, let $\nu_M(\mathcal{X}) := \frac{1}{M}\sum_{m=1}^M \delta_{\mathcal{X}^{(m)}}$ be a discrete probability measure on $(\mathbb{R}^{l_1},\mathcal{B}(\mathbb{R}^{l_1}))$, where δ_x is the Dirac measure on x and $\mathcal{X}^1,\dots,\mathcal{X}^M:\Omega_B\longrightarrow\mathbb{R}^{l_1}$ are i.i.d. random variables. For an $\mathcal{F}_{0,T}^B\otimes\mathcal{B}(\mathbb{R}^{l_1})$-measurable $\mathbb{R}^{l_2}$-valued function S such that $|S(w,\mathcal{X}^{(m)}(w))|<\infty$ for all m and $P_B-a.e. w\in\Omega_B$, the least squares approximation of S in the space $\mathcal{K}$ with respect to ν_M is the P_B-a.e. unique, $\mathcal{F}_{0,T}^B\otimes\mathcal{B}(\mathbb{R}^{l_1})$-measurable function S^* given by

$$
S^*(w,.) = \operatorname*{arginf}_{\phi\in\mathcal{K}} \frac{1}{M}\sum_{m=1}^M |\phi(\mathcal{X}^{(m)})-S(w,\mathcal{X}^{(m)})|^2. \tag{4.4.3}
$$

Due to (4.4.1), the MDP equation (4.2.5) is interpreted in terms of Definition (4.4.1) as follows :
For all $i\in\{0,\dots,N-1\}$, $y_i(\Delta B,.)$ is the measurable function given by :

$$
y_i(\Delta B,.) \text{ is the solution of } OLS\Big(\mathcal{Y}_i(\Delta B,.),\mathcal{K}_i,\nu_i\Big), \tag{4.4.4}
$$

where $\nu_i := P\circ(X_i,\dots,X_N)^{-1}$, K_i is any dense subset in the R^{l_2}-valued functions to $(L_2(\mathcal{B}(\mathbb{R}_2^l)), P\circ(X_i)^{-1})$ and $\forall \underline{x} := (x_0,\dots,x_N)\ \in (R^d)^{N+1}$,

$$
\mathcal{Y}_i(\Delta B,\underline{x}) := \Phi(x_N) + \sum_{k=i}^{N-1}\Delta_k f(y_{k+1}(\Delta B,x_{k+1})) + h(y_{k+1}(\Delta B,x_{k+1}))\Delta B_k. \tag{4.4.5}
$$

4.4.2 Algorithm Notations and algorithm

The solution $y_i(\Delta B,.)$ of (4.4.4) will be approximated in a finite dimensional functional linear space, as defined in the following :

Definition 4.4.2. ***Finite dimensional approximation spaces***
For each i in $\{0,\dots,N-1\}$, the finite dimensional approximation space, denoted by $\mathcal{K}_{Y,i}$, is given by :

$$\mathcal{K}_{Y,i} := \{p^j_{\mathcal{K}_{Y,i}}, j := 1,\dots, card(\mathcal{K}_{Y,i})\} \tag{4.4.6}$$

where for all j, $p^j_{\mathcal{K}_{Y,i}} : \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfies the condition $E[p^j_{\mathcal{K}_{Y,i}}|^2] < +\infty$.

The best approximation error of $y_i(\Delta B,.)$ on the linear space $\mathcal{K}_{Y,i}$ is given by

$$\tau^{\Delta B,Y}_{1,i} := E_{\Delta B}\Big[\inf_{\phi\in\mathcal{K}_{Y,i}}\Big|y_i(\Delta B_i,.)-\phi\Big|^2\Big] \tag{4.4.7}$$

The approximation error (4.4.7) involves explicit computations using the law of $X_i,\dots,X_N$. We avoid that by regressing using the empirical measure, instead of the law. The empirical measure is defined in the following :

Definition 4.4.3. ***Simulations and empirical measure***
Let $M_i^{\Delta B}$ denote the number of Monte Carlo simulations used for the regression at time t_i. This number will be denoted by M_i to alleviate notations. At each discretization time t_i, we sample M_i independent copies of X, that we denote by $\mathcal{C}_i$. i.e.

$$\mathcal{C}_i := \{X_j^{(i,m)}, 0 \leqslant j \leqslant N, 1 \leqslant m \leqslant M_i\}.$$

$\mathcal{C}_i$ is the cloud of M_i simulations used for computations at time t_i, for all $i \in \{1,\dots,N-1\}$. In addition, we assume that the clouds $\{\mathcal{C}_i, 0,\dots,N-1\})$ are sampled independently. The random variables described below are supported by a probability space $(\Omega^{(M)},\mathcal{F}^{(M)},P^{(M)})$ equipped with the empirical probability measure associated to the cloud $\mathcal{C}_i$:

$$\nu_{i,M} := \frac{1}{M_i}\sum_{m=1}^{M_i}\delta_{(X_i^{(i,m)},\dots,X_N^{(i,m)})}. \tag{4.4.8}$$

Then, the full probability space used to analyse the algorithm is $(\bar\Omega,\bar{\mathcal{F}},\bar P) = (\Omega,\mathcal{G},P)\otimes(\Omega^{(M)},\mathcal{F}^{(M)},P^{(M)})$.

· We define the truncation operator $[.]_i$ by : $[y]_i = -C_y^B \vee y \wedge C_y^B$, where C_y^B is the bound computed in Proposition (4.3.2).

Definition 4.4.4. ***Least-Squares MDP (LSMDP) Algorithm***
We define $y_i^{(M)}(\Delta B,.)$, for all i, by the following algorithm :

$$y_N^{(M)}(.) = \Phi(.) \tag{4.4.9}$$

and for $i = N-1,\dots,0$, we set

$$\psi_i^{(M)}(\Delta B,.) \text{ as the solution of } OLS\Big(\mathcal{Y}_i^{(M)}(\Delta B,.),\mathcal{K}_{Y,i},\nu_{i,M}\Big) \tag{4.4.10}$$

where $\forall \underline{x} \in (R^d)^{N+1}$,

$$\begin{aligned}\mathcal{Y}_i^{(M)}(\Delta B,\underline{x}) &:= \Phi(x_N) + \sum_{k=i}^{N-1}\Delta_k f(y_{k+1}^{(M)}(\Delta B,x_{k+1}))\\ &+ h(y_{k+1}^{(M)}(\Delta B,x_{k+1}))\Delta B_k.\end{aligned} \tag{4.4.11}$$

Afterthat, we set

$$y_i^{(M)}(\Delta B,.) := \Big[\psi_i^{(M)}(\Delta B,.)\Big]_i. \tag{4.4.12}$$

4.4.3 Error on the regression scheme

In this part, we will give Non asymptotic estimates for our algorithm. We will need the following lemma

Lemma 4.4.1. *We define* $\tau_{1,i,M}^{\Delta B,Y} := E_{\Delta B}\Big[\inf_{\phi\in\mathcal{K}_{Y,i}}\Big|y_i(\Delta B_i,.)-\phi\Big|^2_{\nu_{i,M}}\Big]$. *This is the squared approximation error of* y_i *in the linear space* $\mathcal{K}_{Y,i}$ *with respect to the empirical measure* $\nu_{i,M}$. *Then, we have :*
For all $i\in\{0,\dots,N-1\}$,

$$\tau_{1,i,M}^{\Delta B,Y} \leqslant \tau_{1,i}^{\Delta B,Y}. \tag{4.4.13}$$

We will make use of the following Proposition, based on the Proposition (4.5.1).

Proposition 4.4.1. *For all* $i\in\{0,\dots,N-1\}$, *we have*

$$\begin{aligned} E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2\Big] &\leqslant 2E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2_{\nu_{i,M}}\Big] \\ &+ (C_y^B)^2\frac{2028(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}. \end{aligned} \tag{4.4.14}$$

Proof. For all $i\in\{0,\dots,N-1\}$, one can write

$$\begin{aligned} &E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2\Big] \leqslant 2E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2_{\nu_{i,M}}\Big] \\ &+ \Big(E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2\Big]-2E_{\Delta B}\Big[\Big|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)\Big|^2_{\nu_{i,M}}\Big]\Big)_+. \end{aligned}$$

Then, we apply the Proposition (4.5.1). □

Now, we are able to give the estimates on our conditional error. We set

$$\eta_i^{Y,B} := \sqrt{E_{\Delta B}\Big[|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)|^2_{\nu_{i,M}}\Big]} = \Big\|\,|y_i(\Delta B,.)-y_i^{(M)}(\Delta B,.)|^2_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}$$

as the the conditional error with respect to the empirical measure $\nu_{i,M}$.

The following theorem gives the conditional regression error of the algorithm (4.4.9)-(4.4.12) for approximating solutions of (4.2.5).

Theorem 4.4.1. *We assume that* ***Assumptions*** *(H1-H2) hold. Then, we have*

$$\eta_i^{Y,B} \leqslant \delta_i + \exp\{L_fT+L_h\sum_{k=i}^{N-1}|\Delta B_k|\}\Big\{\sum_{k=i}^{N-2}(L_f\Delta_k+L_h|\Delta B_k|)\delta_{k+1}\Big\}, \tag{4.4.15}$$

where for all k in $\{0,\dots,N-1\}$

$$\begin{aligned} \delta_k &:= \Big(\tau_{1,k,M}^{\Delta B,Y}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,k})}{M_k}\Big)^{\frac{1}{2}}\sigma_{\mathcal{Y}_k}(\Delta B) \\ &+ \sqrt{1014}C_y^B\sum_{j=k}^{N-2}(L_f\Delta_j+L_h|\Delta B_j|)\sqrt{\frac{(card(\mathcal{K}_{Y,j+1})+1)\log(3M_{j+1})}{M_{j+1}}}, \\ \text{and}\quad C_i &:= \sqrt{2}(L_f\Delta_i+L_h|\Delta B_i|). \end{aligned} \tag{4.4.16}$$

$$\begin{aligned} &\sigma_{\mathcal{Y}_k}(\Delta B) := C_\xi + T(L_fC_y^B+C_f)+(L_hC_y^B+C_h)\sum_{j=k}^{N-1}|\Delta B_j| \\ \text{and}\quad &(C_y^B)^2 := \exp\Big\{(\frac{1}{2}+8L_f^2)T+2L_h^2\sum_{j=0}^{N-1}|\Delta B_j|^2\Big\}\exp\Big\{(1+2L_h^2)\sum_{j=0}^{N-1}|\Delta B_j|\Big\}\Big\{C_\xi^2+\frac{C_f^2T}{4L_f^2} \\ &+ 2C_h^2\sum_{j=0}^{N-1}|\Delta B_j|\Big\}. \end{aligned} \tag{4.4.17}$$

This estimation can be written also

$$\eta_i^{Y,B} \leqslant \Big(\tau_{1,i,M}^{\Delta B,Y}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,i})}{M_i}\Big)^{\frac{1}{2}}\sigma_{\mathcal{Y}_i}(\Delta B) + \sqrt{1014}C_y^B\sum_{k=i}^{N-2}(L_f\Delta_i + L_h|\Delta B_i|)\Big(\frac{(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}\Big)^{\frac{1}{2}}$$
$$+ \quad \sqrt{2}\exp\{\sqrt{2}L_fT + \sqrt{2}L_h\sum_{k=i}^{N-1}|\Delta B_k|\}\sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\Big\{\Big(\tau_{1,k+1,M}^{\Delta B,Y}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,k+1})}{M_{k+1}}\Big)^{\frac{1}{2}}\sigma_{\mathcal{Y}_{k+1}}(\Delta B)\Big\}$$
$$+ \quad (2028)^{\frac{1}{2}}C_y^B\exp\{\sqrt{2}(L_fT+L_h\sum_{k=i}^{N-1}|\Delta B_k|)\}\{L_fT+L_h\sum_{k=i}^{N-1}|\Delta B_k|\}\sum_{k=i}^{N-2}(L_f\Delta_k+L_h|\Delta B_k|)\Big(\frac{(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}\Big)^{\frac{1}{2}}.$$

Remark 4.4.1. *Theorem (4.4.1) gives non asymptotic error estimates for our algorithm, since the constants of the upper bound for the error depend on the time discretization number N and the fixed path $(\Delta B_k)_{0\leqslant k\leqslant N-1}$.*

Remark 4.4.2. *For a fixed time discretization number N and a fixed path $(\Delta B_k)_{0\leqslant k\leqslant N-1}$, we obtain the convergence of our algorithm by taking $(M_k)_{0\leqslant k\leqslant N-1}$ and $(card(\mathcal{K}_{Y,k}))_{0\leqslant k\leqslant N-1}$ enough large.*

Remark 4.4.3. *In the case of BDSDEs and unlike the standard BSDEs case , the we don't have good a.s. estimates for the process Y because of the backward martingale term. Obtaining such estimates is an open question until now.*

Now we will give the proof of Theorem (4.4.1), the main result of this section. We will introduce these notations :

Notations for the proof of Theorem (4.4.1)

· First, we define the following $\sigma - algebras$:

$\mathcal{G}_i^* := \sigma\Big(\Delta B, \mathcal{C}^{i+1}, \ldots, \mathcal{C}^{N-1}\Big)$ and $\mathcal{G}_i^{i,1:M} := \sigma\Big(\mathcal{G}_i^*, X_i^{(i,m):1\leqslant m\leqslant M}\Big)$ for all $i = 0, \ldots, N-1$.

·After that, we define also $\psi_i(\Delta B,.)$ as the solution of $OLS\Big(\mathcal{Y}_i(\Delta B,.), \mathcal{K}_i, \nu_{i,M}\Big)$.

Proof. of Theorem 4.4.1

Since the truncation operator $\Big[.\Big]_i$ is 1-Lipschitz,

$$\begin{aligned}\eta_i^{Y,B} &= \Big\|\Big|y_i(\Delta B,.) - y_i^{(M)}(\Delta B,.)\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}\\ &= \Big\|\Big|\Big[y_i(\Delta B,.\Big]_i - \Big[\psi_i^{(M)}(\Delta B,.)\Big]_i\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}\end{aligned}$$

Then, plugging $E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.)\Big]$ and $E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i^{(M)}(\Delta B,.)\Big]$ and using the triangle inequality, we get

$$\begin{aligned}\eta_i^{Y,B} &\leqslant \Big\|\Big|y_i(\Delta B,.) - E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big]\Big|_{L^2(\Delta B)} + \Big\|\Big|E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.) - \psi_i^{(M)}(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}\\ &+ \Big\|\Big|E_{\mathcal{G}_i^{i,1:M}}[\psi_i^{(M)}(\Delta B,.)] - \psi_i^{(M)}(\Delta B,.)\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}.\end{aligned} \tag{4.4.18}$$

We deal with each term in the previous inequality separately.

- Term $\Big\|\Big|y_i(\Delta B,.) - E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}$ in (4.4.18) :

We note that for all m in $\{1,\ldots,M_i\}$, $E_{\mathcal{G}_i^{i,1:M}}\Big[\mathcal{Y}_i(\Delta B, X_i^{(i,m)})\Big] = y_i(\Delta B, X_i^{(i,m)})$. It follows from Proposition (4.5.2)-(iii) that $E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.)\Big]$ solves $OLS\Big(y_i(\Delta B_i,.), \mathcal{K}_{Y,i}, \nu_{i,M}\Big)$. Then, we have

$$\begin{aligned}\Big\|\Big|y_i(\Delta B,.) - E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)} &= \Big\|\Big|y_i(\Delta B,.) - OLS\Big(y_i(\Delta B_i,.), \mathcal{K}_{Y,i}, \nu_{i,M}\Big)\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}\\ &= \sqrt{\tau_{1,i,M}^{\Delta B,Y}}\end{aligned} \tag{4.4.19}$$

where $\tau_{1,i,M}^{\Delta B,Y} := E_{\Delta B}\Big[\inf_{\phi\in\mathcal{K}_{Y,i}}\Big|y_i(\Delta B_i,.) - \phi\Big|_{\nu_{i,M}}^2\Big]$ is the approximation error with respect to the empirical norm.

- Term $\Big\|\Big|E_{\mathcal{G}_i^{i,1:M}}[\psi_i^{(M)}(\Delta B,.)] - \psi_i^{(M)}(\Delta B,.)\Big|_{\nu_{i,M}}\Big]\Big\|_{L^2(\Delta B)}$ in (4.4.18) :

It can be controlled as follows :

The terms $\mathcal{Y}_i^{(M)}(\Delta B,.)$ are computed only using the clouds $\{\mathcal{C}_k, k \geqslant i+1\}$. Thus, we obtain by Proposition (4.5.2)-(iv)

$$E_{\Delta B}\Big[\Big|\psi_i^{(M)}(\Delta B,.) - E_{\mathcal{G}_i^{i,1:M}}[\psi_i^{(M)}(\Delta B,.)]\Big|^2_{\nu_{i,M}}\Big] \quad \leqslant \quad \frac{card(\mathcal{K}_{Y,i})}{M_i}\sigma^2_{\mathcal{Y}_i}(\Delta B), \qquad (4.4.20)$$

where $\sigma^2_{\mathcal{Y}_i}(\Delta B)$ is a $\sigma(\Delta B)$-measurable random variable bounding the conditional variance $Var(\mathcal{Y}_i^{(M)}(\Delta B, X\Big|X_i = x_i)$ uniformly in x_i.

- Term $\Big\|\Big|E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.) - \psi_i^{(M)}(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)}$ in (4.4.18) :

We set $\mathcal{E}^M_{Y,i}(\Delta B, x) := E\Big[\mathcal{Y}_i(\Delta B, X) - \mathcal{Y}_i^{(M)}(\Delta B, X)|X_i = x_i, \mathcal{G}_i^*\Big]$. As $\mathcal{Y}_i(\Delta B, X) - \mathcal{Y}_i^{(M)}(\Delta B, X)$ are computed only with the clouds $\{\mathcal{C}_k, k \geqslant i+1\}$, we have for all m

$$E_{\mathcal{G}_i^{i,1:M}}\Big[\mathcal{Y}_i(\Delta B, X^{(i,m)}) - \mathcal{Y}_i^{(M)}(\Delta B, X^{(i,m)})\Big] \quad = \quad \mathcal{E}^M_{Y,i}(\Delta B, X_i^{(i,m)}).$$

Thus, by Proposition (4.5.2)-(i-iii), $E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.) - \psi_i^{(M)}(\Delta B,.)\Big]$ solves $OLS\Big(\mathcal{E}^M_{Y,i}(\Delta B,.), \mathcal{K}_{Y,i}, \nu_{i,M}\Big)$. Using Proposition (4.5.2)-(ii) in [35] (the property on Norm stability of the OLS operator), we get

$$\Big\|\Big|E_{\mathcal{G}_i^{i,1:M}}\Big[\psi_i(\Delta B,.) - \psi_i^{(M)}(\Delta B,.)\Big]\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)} \quad \leqslant \quad \Big\|\Big|\mathcal{E}^M_{Y,i}(\Delta B,.)\Big|_{\nu_{i,M}}\Big\|_{L^2(\Delta B)} = \Big\|\mathcal{E}^M_{Y,i}(\Delta B, X_i)\Big\|_{L^2(\Delta B)}.$$

Using the triangle inequality, we get

$$\Big\|\mathcal{E}^M_{Y,i}(\Delta B, X_i)\Big\|_{L^2(\Delta B)} \quad \leqslant \quad \sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\Big\|y_{k+1}(\Delta B, X_{k+1}) - y_{k+1}^{(M)}(\Delta B, X_{k+1})\Big\|_{L^2(\Delta B)}$$

Hence, by plugging the last estimation and the estimations (4.4.19) and (4.4.20) in the inequality (4.4.18), we have

$$\begin{aligned}\eta_i^{Y,B} \quad &\leqslant \quad \Big(\tau^{\Delta B,Y}_{1,i,M}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,i})}{M_i}\Big)^{\frac{1}{2}}\sigma_{\mathcal{Y}_i}(\Delta B) \\ &+ \quad \sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\Big\|y_{k+1}(\Delta B, X_{k+1}) - y_{k+1}^{(M)}(\Delta B, X_{k+1})\Big\|_{L^2(\Delta B)}. \qquad (4.4.21)\end{aligned}$$

The Proposition (4.4.1) allows us to link the term $\Big\|y_{k+1}(\Delta B, X_{k+1}) - y_{k+1}^{(M)}(\Delta B, X_{k+1})\Big\|_{L^2(\Delta B)}$ to the term $\eta_{k+1}^{Y,B}$ as follows :

$$\Big\|y_{k+1}(\Delta B, X_{k+1}) - y_{k+1}^{(M)}(\Delta B, X_{k+1})\Big\|_{L^2(\Delta B)} \leqslant \sqrt{2}\eta_{k+1}^{Y,B} + C_y^B\sqrt{\frac{2028(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}}.$$

The inequality (4.4.21) becomes

$$\begin{aligned}\eta_i^{Y,B} \quad &\leqslant \quad + \sum_{k=i}^{N-2}\sqrt{2}(L_f\Delta_i + L_h|\Delta B_i|)\eta_{k+1}^{Y,B} + \Big(\tau^{\Delta B,Y}_{1,i,M}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,i})}{M_i}\Big)^{\frac{1}{2}}\sigma_{\mathcal{Y}_i}(\Delta B) \\ &+ \quad \sqrt{1014}C_y^B\sum_{k=i}^{N-2}(L_f\Delta_i + L_h|\Delta B_i|)\sqrt{\frac{(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}}.\end{aligned}$$

Thus, one can write

$$\eta_i^{Y,B} \quad \leqslant \quad \delta_i + \sum_{k=i}^{N-2}C_k\eta_{k+1}^{Y,B} \qquad (4.4.22)$$

where we set for all $i \in \{0, \dots, N-1\}$

$$\begin{aligned}\delta_i &:= \left(\tau_{1,i,M}^{\Delta B,Y}\right)^{\frac{1}{2}} + \left(\frac{card(\mathcal{K}_{Y,i})}{M_i}\right)^{\frac{1}{2}} \sigma_{\mathcal{Y}_i}(\Delta B) \\ &+ \sqrt{1014} C_y^B \sum_{k=i}^{N-2} (L_f \Delta_i + L_h |\Delta B_i|) \sqrt{\frac{(card(\mathcal{K}_{Y,k+1}) + 1) \log(3M_{k+1})}{M_{k+1}}}, \quad (4.4.23)\end{aligned}$$

and $$C_i := \sqrt{2}(L_f \Delta_i + L_h |\Delta B_i|). \quad (4.4.24)$$

We note that in the estimation (4.4.22), the error $\eta_i^{Y,B}$ is bounded by a local error term δ_i and the sum of the errors $\eta_k^{Y,B}$ arising before the step i. So we have to iterate this inequality to obtain an estimation on $\eta_i^{Y,B}$.

Writing the estimation (4.4.22) for the discrete time index i, we have

$$\begin{aligned}\eta_i^{Y,B} &\leqslant \delta_i + \sum_{k=i}^{N-1} C_k \eta_{k+1}^{Y,B}, \\ &= \delta_i + C_i \eta_{i+1}^{Y,B} + C_{i+1} \eta_{i+2}^{Y,B} + \sum_{k=i+3}^{N-1} C_k \eta_{k+1}^{Y,B} \quad (4.4.25)\end{aligned}$$

First, we would like to bound the term $C_i \eta_{i+1}^{Y,B}$ in the right side of the previous inequality, then $\eta_{i+2}^{Y,B}$ and so on, until $\eta_{N-1}^{Y,B}$. For that, we write the inequality (4.4.22) for the discrete time $i+1$ and we multiply by C_i. This gives

$$\begin{aligned}C_i \eta_{i+1}^{Y,B} &\leqslant C_i \delta_{i+1} + C_i \sum_{k=i+1}^{N-1} C_k \eta_{k+1}^{Y,B} \\ &= C_i \delta_{i+1} + C_i C_{i+1} \eta_{i+2}^{Y,B} + C_i \sum_{k=i+2}^{N-1} C_k \eta_{k+1}^{Y,B}\end{aligned}$$

Plugging the last inequality in (4.4.25)

$$\eta_i^{Y,B} \leqslant \delta_i + C_i \delta_{i+1} + (1 + C_i) C_{i+1} \eta_{i+2}^{Y,B} + (1 + C_i) \sum_{k=i+2}^{N-1} C_k \eta_{k+1}^{Y,B}. \quad (4.4.26)$$

Now, we would like to bound the term $(1 + C_i) C_{i+1} \eta_{i+2}^{Y,B}$ in the previous estimation. So, we wright the inequality (4.4.22) for the discrete time $i+2$ and multiply by $(1 + C_i) C_{i+1}$. This gives

$$(1 + C_i) C_{i+1} \eta_{i+2}^{Y,B} \leqslant (1 + C_i) C_{i+1} \delta_{i+2} + (1 + C_i) C_{i+1} \sum_{k=i+2}^{N-1} C_k \eta_{k+1}^{Y,B},$$

Plugging the last inequality in (4.4.26), we obtain

$$\eta_i^{Y,B} \leqslant \delta_i + C_i \delta_{i+1} + (1 + C_i) C_{i+1} \delta_{i+2} + (1 + C_i)(1 + C_{i+1}) \sum_{k=i+2}^{N-1} C_k \eta_{k+1}^{Y,B}. \quad (4.4.27)$$

We see that the coefficient appearing in (4.4.27) for each error term $\eta_{k+1}^{Y,B}$, for $k \geqslant i$, is $\Gamma_{i,k}^C C_k$ where $\Gamma_{i,k}^C$ is defined by

$$\begin{aligned}&\Gamma_{i,k}^C := \Pi_{j=i}^{k-1} (1 + C_j), \forall k \in \{i+1, \dots, N-1\} \\ \text{and} \quad &\Gamma_{i,i}^C := 1.\end{aligned}$$

Thus, the error terms from i to $N-1$ are bounded as follows :

$$\begin{aligned}
&\eta_i^{Y,B} \leqslant \delta_i + \Gamma^C_{i,i} C_i \eta^{Y,B}_{i+1} + \Gamma^C_{i,i+1} C_{i+1} \eta^{Y,B}_{i+2} + \sum_{k=i+2}^{N-1} C_k \eta^{Y,B}_{k+1} \\
&\Gamma^C_{i,i} C_i \eta^{Y,B}_{i+1} \leqslant \Gamma^C_{i,i} C_i \delta_{i+1} + \Gamma^C_{i,i} C_i C_{i+1} \eta^{Y,B}_{i+2} + \Gamma^C_{i,i} C_i \sum_{k=i+2}^{N-1} C_k \eta^{Y,B}_{k+1} \\
&\Gamma^C_{i,i+1} C_{i+1} \eta^{Y,B}_{i+2} \leqslant \Gamma^C_{i,i+1} C_{i+1} \delta_{i+2} + \sum_{k=i+2}^{N-1} \Gamma^C_{i,i+1} C_{i+1} C_k \eta^{Y,B}_{k+1} \\
&. \\
&. \\
&. \\
&\Gamma^C_{i,N-3} C_{N-3} \eta^{Y,B}_{N-2} \leqslant \Gamma^C_{i,N-3} C_{N-3} \delta_{N-2} + \Gamma^C_{i,N-3} C_{N-3} C_{N-2} \eta^{Y,B}_{N-1} \\
&\Gamma^C_{i,N-2} C_{N-2} \eta^{Y,B}_{N-1} \leqslant \Gamma^C_{i,N-2} C_{N-2} \delta_{N-1}.
\end{aligned}$$

Summing up the last inequalities from i to $N-1$, we obtain

$$\eta_i^{Y,B} \quad \leqslant \quad \delta_i + \sum_{k=i}^{N-2} \Gamma^C_{i,k} C_k \delta_{k+1}. \tag{4.4.28}$$

Now, we have to bound $\Gamma^C_{i,k}$ in the estimation (4.4.28). This is easily done as follows

$$\begin{aligned}
\Gamma^C_{i,k} \quad &\leqslant \quad \Gamma^C_{i,N} = \Pi_{j=i}^{N-1}(1+\sqrt{2}L_f\Delta_j + \sqrt{2}L_h|\Delta B_j|) \\
&\leqslant \quad \exp\{\sum_{k=i}^{N-1}(\sqrt{2}L_f\Delta_k + \sqrt{2}L_h|\Delta B_k|)\} \\
&= \quad \exp\{\sqrt{2}L_f T + \sqrt{2}L_h \sum_{k=i}^{N-1}|\Delta B_k|\}.
\end{aligned}$$

Finally we obtain

$$\begin{aligned}
&\eta_i^{Y,B} \leqslant \delta_i + \sqrt{2}\exp\{\sqrt{2}L_f T + 2L_h \sum_{k=i}^{N-1}|\Delta B_k|\} \sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\delta_{k+1} \\
\leqslant \quad &\Big(\tau^{\Delta B,Y}_{1,i,M}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,i})}{M_i}\Big)^{\frac{1}{2}} \sigma_{\mathcal{Y}_i}(\Delta B) + \sqrt{1014}C^B_y \sum_{k=i}^{N-2}(L_f\Delta_i + L_h|\Delta B_i|)\Big(\frac{(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}\Big)^{\frac{1}{2}} \\
+ \quad &\sqrt{2}\exp\{\sqrt{2}L_f T + \sqrt{2}L_h \sum_{k=i}^{N-1}|\Delta B_k|\} \sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\Big\{\Big(\tau^{\Delta B,Y}_{1,k+1,M}\Big)^{\frac{1}{2}} + \Big(\frac{card(\mathcal{K}_{Y,k+1})}{M_{k+1}}\Big)^{\frac{1}{2}} \sigma_{\mathcal{Y}_{k+1}}(\Delta B)\Big\} \\
+ \quad &(2028)^{\frac{1}{2}} C^B_y \exp\{\sqrt{2}(L_f T + L_h \sum_{k=i}^{N-1}|\Delta B_k|)\}\{L_f T + L_h \sum_{k=i}^{N-1}|\Delta B_k|\} \sum_{k=i}^{N-2}(L_f\Delta_k + L_h|\Delta B_k|)\Big(\frac{(card(\mathcal{K}_{Y,k+1})+1)\log(3M_{k+1})}{M_{k+1}}\Big)^{\frac{1}{2}}.
\end{aligned}$$

□

4.5 Appendix

From [35], we state an upper bound result, for a sample deviation, uniformly on the function spaces.

Proposition 4.5.1. *For finite $B > 0$, let $\mathrm{G} := \{\psi(\tau_B\phi(.)) - \vartheta(.) : \phi(.) \in \mathcal{K}\}$, where $\tau_B := -B \vee \phi \wedge B$, $\vartheta : \mathbb{R} \to [0, +\infty)$ is Lipschitz continuous with $\psi(0) = 0$ and Lipschitz constant L_ψ, $\vartheta : \mathbb{R}^d \to [-B, B]$, and $\mathcal{K}$ is a finite K-dimensional vector space of functions. Then, for $\mathcal{X}^{(1)}, \ldots, \mathcal{X}^{(M)}$ i.i.d.random variables distributed as $\mathcal{X}$, we have*

$$E_{\Delta B}\Big[\sup_{g\in \mathrm{G}}\Big(\int_{\mathbb{R}} g(x)P\circ\mathcal{X}^{-1}(dx) - \frac{2}{M}\sum_{m=1}^{M} g(\mathcal{X}^{(M)}\Big)_+\Big] \leqslant \frac{507(card(\mathcal{K})+1)\log(3M)}{M}. \tag{4.5.1}$$

After that, we state a proposition, given always in [35] , which gives the properties of the least-squares operator.

Proposition 4.5.2. *Assume that* $\mathcal{K} := span\{p^j_{\mathcal{K}}, j := 1, \dots, card(\mathcal{K})\}$*. Let* S^* *solve* $OLS(S, \mathcal{K}, \nu)$ *(repectively* $OLS(S, \mathcal{K}, \nu_M)$*), according to (4.4.2) (repectively (4.4.3)). The following properties are satisfied :*
(i) linearity : the mapping $S \to S^*$ *is linear.*
(ii) Norm stability property : $\|S^*\|_{L_2(\mathcal{B}(\mathbb{R}^{l_1}),\mu)} \leqslant \|S\|_{L_2(\mathcal{B}(\mathbb{R}^{l_1}),\mu)}$*, where* $\mu = \nu$ *(respectively* $\mu = \nu_M$*).*
(iii) conditional expectation solution : in the case of the discrete probability measure ν_M*, assume additionally that the* $sub-\sigma-algebra$ $\mathcal{Q}$ *is such that* $\{p^j_{\mathcal{K}}(\mathcal{X}^{(m)}), m := 1, \dots, M\}$ *is* $\mathcal{Q}-$*measurable, for every* $j \in \{1, \dots, card(\mathcal{K})\}$*. Setting* $S_{\mathcal{Q}}(\mathcal{X}^{(m)}) = E\Big[S(\mathcal{X}^{(m)})\Big|\mathcal{Q}\Big]$ *for each* $m \in \{1, \dots, M\}$*, then* $E\Big[S^*\Big|\mathcal{Q}\Big]$ *solves* $OLS(S_{\mathcal{Q}}, \mathcal{K}, \nu_M)$*.*
(iv) bounded conditional variance : in the case of the discrete probability measure ν_M*, suppose that* $S(w, x)$ *is* $G \otimes \mathcal{B}(\mathbb{R}^{l_1})-$ *measurable, for a* $sub-\sigma-algebra$ G*, independent of* $\sigma(\mathcal{X}^{(1)}, \dots, \mathcal{X}^{(M)})$*, there exists a Borel measurable function* $h : \mathbb{R}^{l_1} \to \varepsilon$*, for some Euclidean space* ε*, such that the random variables* $\{p^j_{\mathcal{K}}(\mathcal{X}^{(m)}), m =, \dots, M, j = 1, \dots, M\}$ *are* $\mathcal{H} := \sigma(\mathcal{X}^{(m)}, m = 1 \dots, M)-$*measurable, and there exists a finite constant* $\sigma^2 \geqslant 0$ *that uniformly bounds the conditional variances*

$$E\Big[|S(\mathcal{X}^{(m)}) - E[S(\mathcal{X}^{(m)})|G \vee \mathcal{H}]|^2\Big|G \vee \mathcal{H}\Big] \leqslant \sigma^2 P_B - a.s. \text{ and for all } m \in \{0, \dots, M\}.$$

Then

$$E\Big[\|S^*(.) - E[S^*(.)|G \vee \mathcal{H}]\|^2_{L_2(\mathcal{B}(\mathbb{R}^{l_1}),\nu_M)}\Big|G \vee \mathcal{H}\Big] \leqslant \sigma^2 \frac{card(\mathcal{K})}{M}.$$

Chapitre 5

Hastings-Metropolis algorithm on Markov chains for small probability estimation

5.1 Introduction

The study of neutronics began in the 40s, when nuclear energy was on the verge of being used both for setting up nuclear devices like bombs and for civil purposes like the production of energy. Neutronics is the study of neutron population in fissile media that can be modeled using the linear Boltzmann equation, also known as the transport equation. More precisely, it can be subdivided in two different sub-domains. On the one hand, criticality studies aim at understanding the neutron population dynamics due to the branching process that mimics fission reaction (see for instance [71] for a recent survey on branching processes in neutronics). On the other hand, when neutrons are propagated through media where fission reactions do not occur, or can safely be neglected, their transport can be modeled by simple exponential flights [72] : indeed, between each collisions, neutrons travel along straight path distributed exponentially.

Among this last category, shielding studies allow to size shielding structures so as to protect humans from ionizing particles, and imply, by definition, the attenuation of initial neutron flux typically by several decades. For instance, the vessel structure of a nuclear reactor core attenuates the thermal neutron flux inside the code by a factor roughly equal to 10^{13}. Many different national nuclear authorities require shielding studies of nuclear systems before giving their agreement for the design of these systems. Examples are reactor cores, but also devices for nuclear medicine (proton-therapy, gamma-therapy, etcâ). The study of those nuclear systems is complicated by 3-dimensional effects due to the geometry and by non-trivial energetic spectrum that can hardly be modeled.

Since Monte-Carlo transport codes (like MCNP [50], Geant4 [2], Tripoli-4 [29]) require very few hypotheses, they are often used for shielding studies. Nevertheless, those studies represent a long-standing numerical challenge for Monte-Carlo codes in the sense that they schematically require to evaluate the proportion of neutrons that "pass" through the shielding system. This proportion is, by construction, very small. Hence a shielding study by Monte-Carlo code requires to evaluate a small probability, which is the motivation of the present paper.

There is a fair amount of literature on classical techniques for reducing the variance in these small-probability estimation problems for Monte-Carlo codes. Those techniques often rely on a zero-variance scheme [43, 42, 12] adapted to the Boltzmann equation, allied with âweight-watchingâ techniques [13]. The particular forms that this scheme takes when concretely developed in various transport codes range from the use of weight windows [17, 42, 43, 50], like in MCNP, to the use of the exponential transform [14, 29] like in Tripoli-4. Nowadays, all those techniques have proven to be often limited in view of full-filling the requirements made by national nuclear authorities for the precise measurements of radiation, which standards are progressively strengthen. Thus, new variance reduction techniques have been recently proposed in the literature (see for instance [30] for the use of neural networks for evaluating the importance function).

This paper deals with the application of the recent interacting-particle method developed in [36], for small probability estimation, to a neutronic shielding's Monte Carlo code. The method proposed in [36] have interesting theoretical properties and is particularly efficient in practical cases. Nevertheless, its application to shielding studies with Monte-Carlo codes is not straightforward. Monte-Carlo codes consist in sampling the trajectory of a neutron which can, depending on the complexity of the physical modeling, be the realization of a discrete-time branching process or stochastic process. Indeed, since a neutron travels along straight paths between collisions, there is no loss of information in considering only the characteristics of the collisions (dates, positions, energies, subparticle creations) as random.

In view of simplifying the matter, the subparticle-creation phenomena are not taken into account, neither is the energy dependence. We consider here the simplified but realistic case of a monokinetic particle (constant speed, no offsprings). Then a particle trajectory is just a set of successive collisions and constitutes a Markov chain. Furthermore, with probability one, the particle is absorbed after a finite number of collisions. The small probability we are interested in is the probability that a particle "pass" through a shielding system and reach a domain of interest before absorption.

The method proposed in [36] relies on the Hastings-Metropolis algorithm [55, 41] for practical implementation. This algorithm is clearly a textbook method when applied to probability distributions on the Euclidean space. Nevertheless, we have discussed below that small probability estimation problems in neutronic codes involve Markov Chains instead of random vectors. Thus, it is not automatic to apply the method of [36] to these kind of problems. Our contribution is two-fold. First, we show how the Hastings-Metropolis algorithm can be extended to the case of Markov chains that are absorbed after finite time. Second, we apply the resulting method to an academic one-dimensional case and to a two-dimensional case, representing a monokinetic-particle model in a simplified but realistic shielding system. The smaller the probability to estimate is, the more the interacting-particle method clearly outperforms a simple-Monte Carlo method.

The manuscript is organized as follows. In section 5.2 we give a reminder of the interacting-particle method [36], and highlight the need of the Hastings-Metropolis algorithm. In section 5.3, we prove the validity and the convergence of the Hastings-Metropolis algorithm applied to Markov Chains absorbed after a finite time. Then in section 5.4, we present the one and two-dimensional cases. We give the actual equations for the small probability estimation method. At last in section 5.5, we present numerical results in the one and two-dimensional cases.

5.2 Reminder on the interacting-particle method for small probability estimation

We present the interacting-particle method [36] and highlight its need of the Hastings-Metropolis (HM) algorithm for practical application.

We consider a probability space $(\Omega, \mathcal{F}, P)$, and a measurable space $(S, \mathcal{S}, Q)$. We consider a random variable X from $(\Omega, \mathcal{F}, P)$ to $(S, \mathcal{S}, Q)$. We assume that we are able to sample realizations of X.

We consider an objective function $\Phi : S \to \mathbb{R}$, for which we only assume a continuous cumulative distribution function F. The interacting-particle method aims at estimating the probability of the event $\Phi(X) \geqslant l$, for a given level $l \in \mathbb{R}$. We denote this probability p.

The method can be presented in two steps. First, we assume that an ideal, or a theoretical, method can be implemented exactly. In this case, the finite-sample distribution of the corresponding estimator of the probability p is known exactly, so that exact finite-sample confidence intervals are available. Furthermore, the limit, for large number of sampling from X, of the probability estimation error, has attractive properties as shown in [36]. The ideal method is presented in section 5.2.1.

Nevertheless, this ideal method can not be implemented exactly for a large range of practical problems. Thus, it is proposed in [36] to approximate the ideal method by using a HM algorithm. This is presented in subsection 5.2.2.

5.2.1 The theoretical version of the interacting-particle method

In this section 5.2.1, we assume that we are able to sample realizations of X, conditionally to the event $\Phi(X) \geqslant t$, for any $t \in \mathbb{R}$. This is a strong assumption, which is why the corresponding method that we present is called the ideal method.

The ideal algorithm for estimating p is then parameterized by a number of particle N and is as follows.

Algorithm 5.2.1

– Generate an *iid* sample $(X_1, .., X_N)$, from the distribution of X, and initialize $m = 1$, $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ and $X_1^1 = X_1, ..., X_N^1 = X_N$.

- While $L_m \leqslant l$ do
  - For $i = 1, ..., N$
    - Set $X_i^{m+1} = X_i^m$ if $\Phi(X_i^m) > L_m$, and else $X_i^{m+1} = X^*$, where X^* follows the distribution of X conditionally to $\Phi(X) \geqslant L_m$, and is independent of any other random variables involved in the algorithm.
  - Set $m = m + 1$.
  - Set $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$.
- The estimate of the probability p is $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

For each finite N, the ideal estimator $\hat{p}_{ipm}$ obtained from algorithm 5.2.1 has an explicit distribution that is detailed in [36]. In this paper, we just consider two properties of $\hat{p}_{ipm}$. First, the estimator is unbiased : $\mathbb{E}(\hat{p}_{ipm}) = p$. Second, asymptotic 95% confidence intervals, for N large, are of the form

$$I_{\hat{p}_{ipm}} = \left[\hat{p}_{ipm}\exp\left(-1.96\sqrt{\frac{-\log\hat{p}_{ipm}}{N}}\right), \hat{p}_{ipm}\exp\left(1.96\sqrt{\frac{-\log\hat{p}_{ipm}}{N}}\right)\right]. \quad (5.2.1)$$

Finally, we notice that the event $p \in I_{\hat{p}_{ipm}}$ is asymptotically equivalent (for N large) to the event $\hat{p}_{ipm} \in I_p$, with I_P as in (5.2.1), with $\hat{p}_{ipm}$ replaced by p. We will use this property in section 5.5.

5.2.2 Practical implementation of the interacting-particle method with the Hastings-Metropolis algorithm

For practical implementation of algorithm 5.2.1, the only problem we have to solve is the conditional sampling, with the distribution of X, conditionally to $\Phi(X) \geqslant t$, for any $t \in \mathbb{R}$.

An application of the HM algorithm is proposed in [36]. For this, the following is assumed

- The distribution of X has a probability distribution function (pdf) f with respect to $(S, \mathcal{S}, Q)$. For any $x \in S$ we can compute $f(x)$.
- We dispose of a transition kernel on $(S, \mathcal{S}, Q)$ with conditional pdf $\kappa(x, y)$ (pdf of y conditionally to x). We are able to sample from $\kappa(x, .)$ for any $x \in S$ and we can compute $\kappa(x, y)$ for any $x, y \in S$.

Let $t \in \mathbb{R}$ and $x \in S$ so that $\Phi(x) \geqslant t$. Then, the following algorithm enables to, starting from x, sample approximately with the distribution of X, conditionally to $\Phi(X) \geqslant t$. The algorithm is parameterized by a number of iterations $T \in \mathbb{N}^*$.

Algorithm 5.2.2.1

- Let $X = x$.
- For $i = 1, ..., T$
  - Independently from any other random variable, generate X^* following the $\kappa(X, .)$ distribution.
  - If $\Phi(X^*) \geqslant t$
    - Let $r = \frac{f(X^*)\kappa(X^*,X)}{f(X)\kappa(X,X^*)}$.
    - With probability $min(r, 1)$, let $X = X^*$.
- Return X.

The random variable returned by algorithm 5.2.2.1 is denoted $X_T(x)$.

For consistency, we now give the actual interacting-particle method, involving algorithm 5.2.2.1. This method is parameterized by the number of particles N and the number of HM iterations T.

Algorithm 5.2.2.2

- Generate an *iid* sample $(X_1, .., X_N)$ from the distribution of X and initialize $m = 1$, $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ and $X_1^1 = X_1, ..., X_N^1 = X_N$.
- While $L_m \leqslant l$ do
  - For $i = 1, ..., N$
    - Set $X_i^{m+1} = X_i^m$ if $\Phi(X_i^m) > L_m$, and else pick at random an integer J among the integers $1 \leqslant j \leqslant N$ so that $\Phi(X_j^m) > L_m$. Then, let $X_i^{m+1} = X_T(X_J^m)$, with the notation of algorithm 5.2.2.1.
  - Set $m = m + 1$.
  - Set $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$.
- The estimate of the probability p is $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

The estimator $\hat{p}_{ipm}$ of algorithm 5.2.2.2 is the practical estimator that we will study in the numerical results of section 5.5.

In [36], it is shown that, when the space S is a subset of $\mathbb{R}^d$, under mild assumptions, the distribution of the estimator of algorithm 5.2.2.2 converges, as $T \to +\infty$, to the distribution of the ideal estimator of algorithm 5.2.1. For this reason, we call the estimator of algorithm 5.2.1 the estimator corresponding to the case $T = +\infty$. We also call the confidence intervals (5.2.1) the confidence intervals of the case $T = +\infty$.

Nevertheless, the space S we are interested in is a space of sequences that are killed after a finite time. Thus, it is not straightforward that the convergence, as $T \to +\infty$, discussed above, hold in our case. Furthermore, even the notion of pdf on this space of sequences has to be defined.

This is the object of the next section 5.3, that defines the notion of pdf, on a space of sequences that are killed after a finite time, and that gives a convergence result for the HM algorithm. The definition of the pdf is also restated in section 5.4, so that sections 5.2 and 5.4 are self-sufficient for the implementation of the small-probability estimation method in the one and two-dimensional cases.

5.3 An extension of Hastings-Metropolis algorithm to Markov chain sampling

5.3.1 Introducing Markov Chains Killed Out of a Domain

The dynamic of the collisions is described by a Markov chain $(X_n)_{n\geqslant 0}$ with values in $\mathbb{R}^d$ and a probability transition function q. The study of the detection probability for the collisions occurs only in a restricted area. We decide to censor the model and redefine it to obtain accurate theoretical results.

Let D be an open bounded subset of $\mathbb{R}^d$ and ∂D be its $\mathcal{C}^\infty$ boundary. D constitutes the domain of interest. We modify the transition function $(X_n)_{n\geqslant 0}$ as follows

$$k(x,dy) = (q(x,y)1_D(y)\,dy + p(x,D^C)\delta_\Delta(dy))1_D(x) + \delta_\Delta(dy)1_\Delta(x)$$

where Δ is a cemetery point and

$$p(x,D^C) = \int_{D^C} q(x,y)\,dy.$$

This representation describes the following dynamic :

- while $(X_n)_{n\geqslant 0}$ is inside D, it behaves with a transition kernel that could push it outside D. We use the already define q which reflects the collision dynamic.
- when $(X_n)_{n\geqslant 0}$ hits D^C, it is killed and send to the cemetery point Δ where it stays. This way we censor the neutrons and keep only the collisions inside D.

We call this model Markov Chain Killed Out of a Domain, MCKOD for short.

5.3.2 Formulation of the Hastings-Metropolis algorithm

For self-sufficiency of section 5.3, we start by giving a formulation of the Hastings-Metropolis Algorithm [55, 41] for sampling a distribution γ that admits a density versus a measure Π. The main idea is to define a Markov chain $(Y_n)_{n\geqslant 0}$ that converges in distribution to γ. To construct $(Y_n)_{n\geqslant 0}$, we use an instrumental Markov chain $(Z_n)_{n\geqslant 0}$ and an acceptation-rejection function r. We denote by κ the probability transition function of $(Z_n)_{n\geqslant 0}$ and Γ the transition kernel of $(Y_n)_{n\geqslant 0}$. We assume that κ and Γ admit a density versus the same measure Π. Step by step, the algorithm is :

- Introduce a starting point x and use it to sample a transition y of $(Z_n)_{n\geqslant 0}$.
- Accept or reject this transition using r.
- If the sample is accepted redo the procedure with y , else redo with the starting point.

With enough repetitions of this procedure, the distribution of y is approximately γ. Following [41] and [67], a formula of Γ is

$$\Gamma(u,dv) = \overline{\kappa}(u,v)\Pi(dv) + \overline{r}(u)\delta_u(dv)$$

where

$$\overline{\kappa}(u,v) = \begin{cases} \kappa(u,v)r(u,v), & \text{if } x \neq y, \\ 0, & \text{if } x = y, \end{cases}$$

and

$$\overline{r}(u) = 1 - \int p(u,v)\Pi(dv).$$

For the acceptation-rejection function, following [67], we choose

$$r(u,v) = \begin{cases} \min\left\{\dfrac{\gamma(v)\kappa(v,u)}{\gamma(u)\kappa(u,v)}, 1\right\}, & \text{if } \gamma(u)\kappa(u,v) > 0 \\ 1, & \text{if } \gamma(u)\kappa(u,v) = 0 \end{cases}.$$

This is the case presented in algorithm 5.2.2.1. It ensures the reversibility condition required in [67] to prove that Π is invariant for Γ. In those books, a state space endowed with a countably generated σ-algebra and a good topology is a basic condition for most of the definitions and results. A priori, our MCKOD do not stick with that. In addition, we need to introduce a measure Π. Thus, section 5.3.3 is devoted to solve these two points. In addition, according to [67], the kernel Γ is irreducible only if the instrumental kernel κ is irreducible. Then, section 5.3.4 defines a family of instrumental and prove their irreducibility. Finally section 5.3.5 gives the proof using the results in [56], [60] and [67].

5.3.3 State space, distribution and density for MCKOD

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $q(x,y)$ a probability transition kernel on $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ with density w.r.t. the Lebesgue measure. Let $D \subset \mathbb{R}^d$ be an open bounded subset with $\mathcal{C}^\infty$ boundary ∂D. We suppose that $0 \notin D$ and is the cemetery point. We set $D_0 = D \cup \{0\}$. We introduce the space

$$c_0 = \{(u_n)_{n \geqslant 0} \in D_0^{\mathbb{N}} : \exists n_0 \in \mathbb{N}, \forall n \geqslant n_0, u_n = 0\},$$

equipped with the distance

$$d_\infty(u,v) = \max_{n \geqslant 0} d(u_n - v_n)_{\mathbb{R}^d}$$

and denote by $\mathcal{B}(c_0)$ the Borelian σ-algebra. We want to see the Markov chain $(X_n)_{n \geqslant 0}$ as a random variable :

$$\begin{array}{rcl} X : (\Omega, \mathcal{F}, P) & \mapsto & (c_0, \mathcal{B}(c_0)) \\ \omega & \mapsto & (X_n(\omega))_{n \geqslant 0}. \end{array}$$

A tricky point is the measurability with $\mathcal{B}(c_0)$. But, we have the measurability with

$$\overline{\mathcal{F}} = \bigvee_{i=0}^{+\infty} \mathcal{B}(D_0).$$

Thus, the following result give it :

Proposition 5.3.1. *Let $\overline{\mathcal{F}}_{c_0}$ be the restriction of $\overline{\mathcal{F}}$ to the subspace c_0. Then,*

$$\overline{\mathcal{F}}_{c_o} = \mathcal{B}(c_0)$$

Proof. Let p_n be the projection from c_0 in D_0 who associates u_n to u. This application is Lipschitz. In fact, let u and v be in c_0, we have $d(u_n - v_n) \leqslant d(u - v)$. Consequently, every projection is measurable and $\overline{\mathcal{F}}_{c_o} \subset \mathcal{B}(c_0)$.

Let show the inclusion $\mathcal{B}(c_0) \subset \overline{\mathcal{F}}_{c_0}$. We know that (c_0, d_∞) is separable. We denote by S a dense subset. $\mathcal{B}(c_0)$ is generated by the ball of radius $\rho \in \mathbb{Q} \cap D_0$ and center point $u \in S$. Thus, it is enough to show that the ball $B(\rho, u)$ is in $\overline{\mathcal{F}}_{c_0}$. So, we write

$$B(\rho, u) = \bigcap_{n=0}^{+\infty} \{v \in c_0, d(u_n - v_n) \leqslant \rho\}$$

and because each member of this intersection is in $\overline{\mathcal{F}}_{c_0}$, we have the desired result. ◻

Let Us define

$$A_n = \{u \in c_0 : u_k \in D, \forall k < n \text{ and } u_k = 0, \forall k \geqslant n\}.$$

The family $(A_n)_{n \geqslant 0}$ is a partition of the space c_0. We introduce the family of projections $(\pi_n)_{n \geqslant 0}$, with :

$$\begin{aligned} \pi_n : A_n &\mapsto D^n \\ u &\mapsto (u_1, \cdots, u_n), \end{aligned}$$

and the measure on c_0

$$\Pi(dx) = \sum_{n=1}^{+\infty} \lambda_n(\pi_n^{-1}(dx)) 1_{A_n}(x),$$

where λ_n is the Lebesgue measure on D^n. We have the following result :

Proposition 5.3.2. *The distribution γ of a MCKOD $(X_n)_{n \geqslant 0}$ starting from x is absolutely continuous versus Π. The density is given by :*

$$\sum_{n=1}^{+\infty} \int_{D^C} \int_{D^C} q_D^n(x, x_n) q(x_n, x_{n+1})\, dx_n dx_{n+1}\, 1_{A_n}(x)$$

where

$$q_D^n(x, x_n) = q(x, x_1) \cdots q(x_{n-1}, x_n)\, 1_D(x_1) \cdots 1_D(x_n).$$

Proof. We fix $n > 0$ and we restrain Π to A_n. Suppose that $A \in \mathcal{B}(c_0)$ with $\Pi(A) = 0$, we have

$$\begin{aligned} \gamma(A \cap A_n) &= P((X_1, \cdots, X_n) \in p_n(A \cap A_n), X_{n+1} \in D^C) \\ &= P((X_1, \cdots, X_n) \in p_n(A \cap A_n)) \times p(X_n, D^C) \\ &= \int_{\pi_n(A \cap A_n)} q_D^n(x, x_n)\, dx_n \times p(X_n, D^C) \end{aligned}$$

$q^n(x, y)$ is absolutely continuous versus λ_n and $\lambda_n(A \cap A_n) = \Pi_{|A_n}(A) = 0$. Using the classical formula :

$$P(X \in A) = \sum_{n=0}^{+\infty} P(X \in (A \cap A_n)) = 0,$$

we conclude that $P(X \in A) = 0$. For the density, the result comes from the expression

$$\begin{aligned} P(T = n+1) &= \int_{\pi_n(A \cap A_n)} q_D^n(x, x_n)\, dx_n \times p(X_n, 0) \\ &= \int_{D^C} \int_{D^C} q_D^n(x, x_n) q(x_n, x_{n+1})\, dx_n dx_{n+1} \\ &= \int_{D^C} \int_{D^C} \int q(x, x_1) \cdots q(x_{n-1}, x_n)\, 1_D(x_1) \\ &\cdots 1_D(x_n)\, q(x_n, x_{n+1})\, dx_1 \cdots dx_n\, dx_{n+1}, \end{aligned}$$

where T is the first time X hits D^C. Thus, the mass is one. □

5.3.4 Some Π-irreducible instrumental kernels on c_0

We introduce and study the Markov chain on c_0 starting from w with kernel :

$$\kappa(u, dv) = \sum_{k=1}^{+\infty} \Theta(u, A_k) \nu_k(u, dv) 1_{A_k}(v)$$

where we suppose

- For each $u \in c_0$, the sum of the $(\Theta(u, A_k))_{k \geqslant 0}$ is one.
- $(\nu_k(u, dv))_{k \geqslant 1}$ is a family of probability transition kernels on D^n.

This statements insures that κ is a probability transition kernel on c_0. We describe the behavior of the chain :

– from a sequence u, we change the number of non-null points using the family $(\Theta(u, A_k))_{k \geqslant 0}$.
– for each non-null point, we change its position using a probability transition kernel from $(\nu_k(u, dv))_{k \geqslant 1}$.

This kind of Markov chains will serve as the so-called instrumental chains for the Hastings-Metropolis algorithm. We give the following definition for the irreducibility of a chain :

Definition 5.3.1. *Let G be a topological space, $\mathcal{G}$ a σ-algebra on G, m a probability measure and μ a probability transition kernel. We say that $A \in \mathcal{G}$ is attainable from $x \in G$ if :*

$$\mu^n(x, A) > 0 \text{ for some } n \geqslant 1.$$

We say that the set $B \in \mathcal{G}$ is m-communicating if :

$$\forall x \in B, \forall A \in \mathcal{G} \text{ such that } A \subset B,\ m(A) > 0, A \text{ is attainable from } x.$$

In addition, if G is m-communicating, the chain is m-irreducible

The γ-irreducibility of the instrumental kernel is required for the irreducibility of the Metropolis chain and the convergence of the algorithm. Thus, the following results are crucial and permit us to have usable kernels.

Proposition 5.3.3. *If κ is a probability transition kernel satisfying*
– *For every $u \in c_0$ and $k \geqslant 0$, $\Theta(u, A_k) > 0$.*
– *for each $k \geqslant 1$, $\nu_k(u, dv)$ is absolutely continuous w.r.t. the Lebesgue measure in D^n.*
– *for each $k \geqslant 1$, $\nu_k(u, dv)$ is irreducible for the Lebesgue measure in D^n.*

then
– *κ is absolutely continuous versus Π.*
– *κ is Π-irreducible.*

Proof. The absolute continuity is induced by the definition and the third hypothesis. The proof is totally analog to 5.3.2. Thus, we only show the Π-irreducibility. Let $A \in \mathcal{B}(c_0)$ be a Π-positive subset and $u \in c_0$ a sequence. We want to prove that :

$$\kappa^n(u, A) > 0, \text{ for some } n \geqslant 1.$$

This result naturally holds if we show for all $k \geqslant 0$ and for all $A \subset A_k$ that A is attainable from $u \in c_0$. We fix $k \geqslant 0$ and choose $A \subset A_k$. From the definition of κ, we have

$$\kappa(u, A) = \int_A \Theta(u, A_k)\nu_k(u, dw)\, dw.$$

Thus, we only have to prove that

$$\int_A \nu_k(u, dw)\, dw > 0.$$

The absolute continuity and the irreducibility of the $(\nu_k(u, dv))_{k \geqslant 0}$ induce that

$$\text{for every } k \geqslant 0 \text{ and } \lambda_n\text{-positive set A, } \nu_k^m(u, A) > 0 \text{ for } m = 1.$$

Indeed, suppose the opposite, for each $m \geqslant 1$ we have

$$\nu_k^m(u, A) = \int_A \int_D \cdots \int_D \nu_k(u, v_1) \cdots \nu_k(u, v_m)\, dv_1 \cdots dv_m = 0$$

and we have a conflict with the irreducibility. Finally,

$$\int_A \nu_k(u, y) dy > 0$$

and the result is proved. ◻

Corollary 5.3.1. *With the same hypothesis as in 5.3.3, κ is γ-irreducible.*

Proof. A result of [60] says that : if a kernel κ is Π-irreducible and there is a measure γ which is absolutely continuous versus Π, then κ is γ-irreducible. ◻

5.3.5 Convergence of the extended Hastings-Metropolis Algorithm

Before proving the convergence of the algorithm, we give an example of $(\Theta(u, A_k))_{k \geqslant 0}$ and $(\nu_k(u, dv))_{k \geqslant 1}$ for which the result hold. G denotes the geometric distribution adapted to $\mathbb{N}$. For each $u \in c_0$, we choose it for $\Theta(u, A_k) = P(G = k)$. Let g denote the density of the uniform distribution inside D. We take

$$\nu_k(u, v) = \prod_{i=1}^{k} g(v)$$

and one could see that it satisfies the hypothesis of the result below.

Proposition 5.3.4. *If κ is probability transition kernel satisfying satisfying the same hypothesis as in 5.3.3 and*

- *for every $k \geqslant 0$,*
$$\inf_{u \in c_0} \Theta(u, A_k) > 0.$$
- *for each $k \geqslant 1$,*
$$\inf_{u \in D} \nu_k(u, dv) > 0.$$

Then, the Hastings-Metropolis kernel Γ converge to γ.

Proof. In order to prove the convergence, we follow [67]. Consequently, we have to show that Γ is γ-irreducible and $\gamma\{\overline{r}(u) > 0\} > 0$. We start with the γ-irreducibility. Let $A \in \mathcal{B}(c_0)$ be a γ-positive subset and $u \in c_0$ a sequence. We want

$$\Gamma^n(u, A) > 0, \text{ for some } n \geqslant 1.$$

We use the same approach as in the proof of proposition 5.3.3. For fixed $k \geqslant 0$ and $A \subset A_k$, we study

$$\overline{\kappa}(u, A) = \int_A \kappa(u, w) r(u, w) \Pi(dw).$$

Since the second term in the expression of Γ is positive, the γ-irreducibility of $\overline{\kappa}$ is fairly enough. We recall that

$$r(u, v) = \begin{cases} \min\left\{ \dfrac{\gamma(w)\kappa(w, u)}{\gamma(u)\kappa(u, w)}, 1 \right\}, & \text{if } \gamma(u)\kappa(u, w) > 0 \\ 1, & \text{if } \gamma(u)\kappa(u, w) = 0 \end{cases}.$$

If $\gamma(u)\Theta(u, A_k)\nu_k(u, w) = 0$ on A γ-almost-surely, then the proof is trivial with the γ-irreducibility of κ. This is the same if

$$\frac{\gamma(w)\kappa(w, u)}{\gamma(u)\kappa(u, w)} > 1$$

on A γ-almost-surely. Thus, we have to check that

$$\overline{\kappa}(u, A) = \int_A \kappa(u, w) \cdot \frac{\gamma(w)\kappa(w, u)}{\gamma(u)\kappa(u, w)} \Pi(dw) > 0.$$

Suppose that $u \in A_l$, then

$$\begin{aligned} \overline{\kappa}(u, A) &= \int_A \frac{\gamma_{A_k}(w)\Theta(w, A_l)\nu_l(w, u)}{\gamma(u)} \, dw_1 \cdots dw_k \\ &= \int_A \frac{\gamma_{A_k}(w)\Theta(w, A_l)\nu_l(w, u)}{\gamma_{A_l}(u)} \, dw_1 \cdots dw_k. \end{aligned}$$

Because $\gamma_{A_k}(u)$ is strictly positive on A γ-almost-surely, otherwise we are back to the previous case. With the two lower-bound hypotheses, the problem is reduced to

$$\int_A \gamma_{A_k}(w) \, dw_1 \cdots dw_k > 0,$$

which is trivially true. For the aperiodicity, the probability to stay inside a A_k is positive. If we take $A \subset A_k$ such that $\gamma_{A_k}(A)$, the probability to reach it is positive, since v_k is γ_{A_k}-irreducible (see hypotheses in 5.3.3 and corollary 5.3.1). □

5.4 Practical implementation in dimension one and two

In section 5.4 we present the one and two-dimensional cases, for which the results of the interacting-particle method of section 5.2 are presented in section 5.5.

The interacting-particle method 5.2.2.2 necessitates, as we have seen, to evaluate pdf on Markov-chain trajectories with finite number of non-absorbed points. These pdf have been defined in section 5.3. They are redefined in definition 5.4.1 and proposition 5.4.7 so that section 5.4 is self-sufficient.

The actual values of these pdf, for the one and two-dimensional cases, are also given in section 5.4.

5.4.1 Some general vocabulary and notation

Throughout section 5.4, we consider a monokinetic particle (a particle with constant speed and yielding no subparticle birth) evolving in $\mathbb{R}^d$, with $d = 1, 2$. This particle is created at the source point $s \in \mathbb{R}^d$, that is the birth of the particle takes place at s.

The trajectory of the monokinetic particle is characterized by its collision points, which constitute a Markov chain. The sequence of collision points is written $(X_n)_{n\in\mathbb{N}^*}$. The birth of the particle takes place at s, that is $X_0 = s$. After its birth, the monokinetic particle travels along straight lines, with random distances and directions, between its collision points, until it is absorbed. The absorption happens almost surely after a finite number of collisions.

The distribution of the Markov chain of the birth and collision points $(X_n)_{n\in\mathbb{N}}$ is characterized by, first, a function $P_a(t) : \mathbb{R}^d \to [0, 1]$, so that $P_a(t)$ is the probability of absorption for a collision taking place at t. Second, the distribution is characterized by the pdf $q(t_1, t_2) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$, so that $q(t_1, .) : \mathbb{R}^d \to \mathbb{R}^+$ is a pdf.

The behavior of the monokinetic particle is then as follows. At a collision point X_n, for which the monokinetic particle has not been absorbed yet, one and only one of the two following events can randomly happen. First, the particle can be absorbed during the collision. In this case, we define this absorption by the equality $X_m = \Delta$ for any $m > n$. The point $\Delta \in \mathbb{R}^d$ symbolizes that the monokinetic particle has been absorbed. It is called the cemetery point, similarly to section 5.3. The choice of the cemetery point Δ is of course arbitrary, as long as it is distinguished from the source point, that is $\Delta \neq s$. If the monokinetic particle is not absorbed during the collision, it is scattered. In this case, the position of the next collision point X_{n+1} has the $q(X_n, .)$ pdf.

Thus, eventually, the sequence $(X_n)_{n\in\mathbb{N}}$ of collisions points of the monokinetic particle is a Markov chain that has the property that, almost surely, there exists $m \in \mathbb{N}$ so that $X_n = \Delta$ for $n \geqslant m$. This is the type of Markov Chains that are covered in section 5.3. We say that the monokinetic particle is active at time n, or at X_n, or before collision n, if $X_n \neq \Delta$.

Finally, when the Markov Chain of a monokinetic particle $(X_i)_{i\in\mathbb{N}}$ has been sampled, and is equal to $(x_i)_{i\in\mathbb{N}}$ we call the sequence $(x_i)_{i\in\mathbb{N}^*}$ of its collision points a trajectory of the monokinetic-particle. (There is no loss of information in that a trajectory does not store the deterministic source point $x_0 = s$.)

Remark 5.4.1. *Strictly speaking, it is possible that a collision n takes place at position Δ, that is $X_n = \Delta$, though the monokinetic particle has not been absorbed before X_n. Nevertheless, in this section 5.4, we only consider transition kernels $q(., .)$ that are absolutely continuous with respect to the Lebesgue measure. Thus, almost-surely, the cemetery point Δ identifies without ambiguity that the monokinetic-particle has been absorbed.*

5.4.2 Description of the one-dimensional case and expression of the probability density functions

In this section 5.4.2, we consider an academic Monte-Carlo problem similar to the shielding studies, but for which a monokinetic particle evolves in a one-dimensional space.

5.4.2.1 A one-dimensional random walk

We consider that the monokinetic particle evolves in $\mathbb{R}$. As described in section 5.4.1, the birth of the particle takes place at 0, that is $X_0 = s = 0$. In the one-dimensional model we define, when the monokinetic particle has not been absorbed, the signed-distance traveled between two collisions is a Gaussian variable. That is, if $X_n \neq \Delta$, we have $X_{n+1} = X_n + \varepsilon_{n+1}$, where the $(\varepsilon_i)_{i\in\mathbb{N}^*}$ are *iid* and follow a $\mathcal{N}(0, \sigma^2_{mk})$ distribution.

Finally, at each collision point X_n, the probability of absorption is one if $X_n \leqslant L_{inf}$ or $X_n \geqslant L_{sup}$, where $L_{inf} < 0 < L_{sup}$. If a collision point is $L_{inf} < X_n < L_{sup}$, the monokinetic particle is absorbed with probability $0 \leqslant P_a < 0$, and is scattered with probability $1 - P_a$.

The event of interest is here that the monokinetic particle reaches the domain $(-\infty, L_{inf}]$. When using the interacting-particle method of section 5.2, this event is traduced by the event $\Phi(x) \geqslant 0$, with $\Phi(x) = L_{inf} - \inf_{i\in\mathbb{N}^*; x_i \neq \Delta} x_i$. Notice that, almost-surely, the infimum is over a finite number of points.

The one-dimensional case presented here reproduces some key features of the shielding studies by Monte Carlo code described in sections 5.1 and 5.4.1. Indeed, the monokinetic particle travels a random distance, toward a random direction (positive or negative), between two collision points, and random absorption is considered. The particle is absorbed after a number of collisions that is random and almost-surely finite. By setting P_a sufficiently large, and L_{inf} sufficiently away from 0, we will see that we can tackle problems with arbitrary-small probabilities. Thus, in section 5.5, we will consider a probability small enough so that the interacting-particle method of section 5.2 outperforms a simple-Monte Carlo method.

Finally, notice that an important feature of the two-dimensional case of section 5.4.3, that is not reproduced by the one-dimensional case, is the presence of different media, and the medium-crossing phenomena.

5.4.2.2 Expression of the probability density function of a trajectory

We now give the expression of the pdf (with respect to the setting of definition 5.4.1 and proposition 5.4.7) of a trajectory obtained from the one-dimensional model above. We let $(x_i)_{i\in\mathbb{N}^*}$ be the sequence of collision points (the trajectory) of a monokinetic particle. We let $\mathcal{D} = (L_{inf}, L_{sup})$. We denote $\phi_{m,\sigma^2}(t)$ the pdf at t of the one-dimensional Gaussian distribution with mean m and variance σ^2.

Similarly to section 5.3, we define $A_n = \{(x_i)_{i\in\mathbb{N}^*}, x_j \neq \Delta, \forall 1 \leqslant j \leqslant n-1, x_k = \Delta, \forall k \geqslant n\}$, that is the set of trajectories that are absorbed at collision point $n-1$ (so that they are in the absorbed state from collision point n and onward).

Proposition 5.4.1. *The pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 5.4.1 and proposition 5.4.7, of a trajectory $(x_n)_{n\in\mathbb{N}^*}$, sampled from the procedure of section 5.4.2.1, is $f(x) = \sum_{n\in\mathbb{N}^*} \mathbf{1}_{A_{n+1}}(x) f_n(x)$, with*

$$f_n(x) = \prod_{i=1}^{n-1} \left(\mathbf{1}_{x_i\in\mathcal{D}} \phi_{x_{i-1},\sigma^2_{mk}}(x_i)(1-P_a) \right) \phi_{x_{n-1},\sigma^2_{mk}}(x_n) \left(\mathbf{1}_{x_n\notin\mathcal{D}} + P_a \mathbf{1}_{x_n\in\mathcal{D}} \right),$$

where $x_0 = 0$ by convention.

Several comments can be made on proposition 5.4.1.

The pdf of proposition 5.4.1 has to be evaluated for each trajectory, either sampled from its initial distribution, or from a perturbation method in the HM algorithm 5.2.2.1. The perturbation methods are presented below in sections 5.4.2.3, 5.4.2.4, 5.4.2.5 and 5.4.2.6.

There is a mild difference between proposition 5.3.2 and proposition 5.4.1. In proposition 5.3.2, when the monokinetic particle goes outside the domain of interest D, and consequently is absorbed, the collision point in the exterior of D is not stored in the trajectory of the monokinetic particle. Indeed, the exact value of this point is not needed to assess if the monokinetic has reached the domain $(-\infty, L_{inf}]$. However, by not storing this point, the pdf of proposition 5.3.2 necessitates to know the probability that, starting from a birth or scattering point in the domain D, the next collision point lies outside D. This probability is not known explicitly in the framework of section 5.4.3, and *a fortiori* in shielding studies involving more complex Monte Carlo codes. To avoid evaluating this probability numerically each time a pdf of a trajectory is computed, we store the collision points outside the domain D. Notice that this, inevitably, add some variance in the HM method, because we use a source of randomness (the exact collision point at which the monokinetic particle leaves D) that does not impact the event of interest.

The evaluation of a pdf like that of proposition 5.4.1 is an intrusive operation on a Monte Carlo code. Indeed, it necessitates to know all the random-quantity sampling that are done when this code samples a monokinetic-particle trajectory. Thus, the Monte Carlo code is not used as a black box.

Nevertheless, the computational cost of the pdf evaluation is of the same order as the computational cost of a trajectory sampling, and the same kind of operations are involved. Namely, both tasks require a loop which length is the number of collisions made by the monokinetic-particle before its absorption. Furthermore, for each random quantity that is sampled for a trajectory sampling, the pdf evaluation requires to compute the corresponding pdf. For example, in the case of proposition 5.4.1, when a trajectory sampling requires to sample n Gaussian variables and n or $n-1$ Bernoulli variables, the trajectory-pdf evaluation requires to compute the corresponding Gaussian pdf and Bernoulli probabilities.

The discussion above holds similarly for the two-dimensional case of section 5.4.3.

5.4.2.3 Description of the trajectory perturbation method when $P_a = 0$

For clarity of exposition, we present first the perturbation method when $P_a = 0$. In this case, the monokinetic particle is a random walk on $\mathbb{R}$, that is absorbed once it goes outside $\mathcal{D}$.

The perturbation method is parameterized by $\sigma^2_{hm} > 0$. Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Then, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$. This inhomogeneous Markov chain is so that $Y_0 = 0$. Then, if $i \leqslant n-1$, and if the perturbed monokinetic particle is still in $\mathcal{D}$ at collision point i, we have $Y_{i+1} = Y_i + \varepsilon_{i+1}$, where the $(\varepsilon_i)_{1\leqslant i\leqslant n}$ are independent and where ε_i follows a $\mathcal{N}(x_i - x_{i-1}, \sigma^2_{hm})$ distribution.

Similarly to the initial sampling, the perturbed monokinetic particle is absorbed at the first collision point outside $\mathcal{D}$. If the collision point Y_n of the perturbed monokinetic particle is in $\mathcal{D}$ (contrary to x_n for the initial trajectory), the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial monokinetic particle would be sampled if its collision point n was Y_n.

This conditional sampling method for perturbed trajectories is intrusive : it necessitates to change the stochastic dynamic of the monokinetic particle. Nevertheless, the new dynamic is here chosen as to have the same cost as the unconditional sampling, and to require the same type of computations. This is similar to the discussion following proposition 5.4.1.

5.4.2.4 Expression of the probability density function of a perturbed trajectory when $P_a = 0$

We now give the expression of the conditional pdf (with respect to the setting of proposition 5.3.2) of a trajectory obtained from the one-dimensional perturbation method above.

Proposition 5.4.2. *Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. The conditional pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 5.4.1 and proposition 5.4.7, of a trajectory $(y_n)_{n\in\mathbb{N}^*}$ sampled from the procedure of section 5.4.2.3, is $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ where, if $m \leqslant n$*

$$f_{n,m}(x,y) = \prod_{i=1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma^2_{hm}}(y_i) \right) \mathbf{1}_{y_m\notin\mathcal{D}} \phi_{y_{m-1}+(x_m-x_{m-1}),\sigma^2_{hm}}(y_m),$$

and if $m > n$,

$$\begin{aligned} f_{n,m}(x,y) \quad &= \quad \prod_{i=1}^{n} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma^2_{hm}}(y_i) \right) \\ &\quad \times \prod_{i=n+1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1},\sigma^2_{mk}}(y_i) \right) \\ &\quad \mathbf{1}_{y_m\notin\mathcal{D}} \phi_{y_{m-1},\sigma^2_{mk}}(y_m), \end{aligned}$$

where $y_0 = 0$ by convention.

Similarly to the discussion following 5.4.1, the computation of the conditional pdf of a perturbed trajectory has the same computational cost as the sampling of this perturbed trajectory.

5.4.2.5 Description of the trajectory perturbation method when $P_a > 0$

Let us now consider the general case where $P_a > 0$.

The perturbation method is parameterized by $\sigma_{hm}^2 > 0$ and $0 < P_c < 1$. Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n.

As when $P_a = 0$, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$, so that $Y_0 = 0$. As when $P_a = 0$, we modify the increments of the initial trajectory, and, if the perturbed trajectory outsurvives the initial one, we generate the sequel with the initial distribution. Specifically to this case $P_a > 0$, we perturb the absorption/non-absorption sampling by changing the initial values with probability P_c.

More precisely, for $i \leqslant n-1$, and if the perturbed monokinetic particle has not been absorbed before collision point i, it is absorbed with probability $max(P_c, \mathbf{1}_{Y_i \notin D})$. If it is scattered instead, we have $Y_{i+1} = Y_i + \varepsilon_{i+1}$, where the $(\varepsilon_i)_{1 \leqslant i \leqslant n}$ are independent and where ε_i follows a $\mathcal{N}(x_i - x_{i-1}, \sigma_{hm}^2)$ distribution. If the perturbed monokinetic particle has not been absorbed before collision point n, then it is absorbed if $Y_n \notin \mathcal{D}$. If $Y_n \in \mathcal{D}$, the perturbed monokinetic particle is absorbed with probability p, where $p = 1 - P_c$ if $x_n \in \mathcal{D}$ and $p = P_a$ if $x_n \notin \mathcal{D}$.

As when $P_a = 0$, if the perturbed monokinetic particle has not been absorbed before collision point Y_n, the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial particle would be sampled if its collision point n was Y_n.

The idea is that, by selecting the difference between P_c and $\min(P_a, 1-P_a)$, the closeness between the perturbed and initial trajectories can be specified, from the point of view of the absorption/non-absorption events.

5.4.2.6 Expression of the probability density function of a perturbed trajectory when $P_a > 0$

Proposition 5.4.3. *Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Let $y_0 = x_0 = 0$ by convention. The conditional pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 5.4.1 and proposition 5.4.7, of a trajectory $(y_n)_{n\in\mathbb{N}^*}$ sampled from the procedure of section 5.4.2.5, is $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ where, if $m \leqslant n-1$,*

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c) \right) \\ &\quad \phi_{y_{m-1}+(x_m-x_{m-1}),\sigma_{hm}^2}(y_m) \left(\mathbf{1}_{y_m\in\mathcal{D}} P_c + \mathbf{1}_{y_m\notin\mathcal{D}} \right), \end{aligned}$$

if $m = n$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c) \right) \\ &\quad \phi_{y_{n-1}+(x_n-x_{n-1}),\sigma_{hm}^2}(y_n) \left(\mathbf{1}_{y_n\notin\mathcal{D}} + \mathbf{1}_{y_n\in\mathcal{D}} \mathbf{1}_{x_n\in\mathcal{D}}(1-P_c) + \mathbf{1}_{y_n\in\mathcal{D}} \mathbf{1}_{x_n\notin\mathcal{D}} P_a \right), \end{aligned}$$

and if $m \geqslant n+1$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c) \right) \\ &\quad \mathbf{1}_{y_n\in\mathcal{D}} \phi_{y_{n-1}+(x_n-x_{n-1}),\sigma_{hm}^2}(y_n) \left(\mathbf{1}_{x_n\in\mathcal{D}} P_c + \mathbf{1}_{x_n\notin\mathcal{D}}(1-P_a) \right), \\ &\quad \prod_{i=n+1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1},\sigma_{mk}^2}(y_i)(1-P_a) \right) \\ &\quad \phi_{y_{m-1},\sigma_{mk}^2}(y_m) \left(\mathbf{1}_{y_m\notin\mathcal{D}} + \mathbf{1}_{y_m\in\mathcal{D}} P_a \right). \end{aligned}$$

5.4.3 Description of the two-dimensional case and expression of the probability density functions

5.4.3.1 Description of the neutron transport problem

The monokinetic particle evolves in $\mathbb{R}^2$, and its birth takes place at the source point $s = (-p_{s,x}, 0)$, with $p_{s,x} > 0$. The domain of interest is a box $[-\frac{L}{2}, \frac{L}{2}]^2$ with $p_{s,x} < \frac{L}{2}$, in which there is an obstacle sphere with center 0 and radius l, with $l < \frac{L}{2}$. The obstacle sphere is the set $\{x \in \mathbb{R}^2 | |x| \leqslant l\}$, with $|x|$ the Euclidean norm of $x \in \mathbb{R}^2$.

The box is composed of two media. The obstacle sphere is composed of "poison" and the rest of the box is composed of "water". The exterior of the box is also composed of "water". Nevertheless, if the monokinetic-particle reaches this exterior, it is considered to have gone too far away, and subsequently is is absorbed at the first collision point in the exterior of the box.

We consider a detector, in the box, that is a sphere with center $(p_{det,x}, 0)$ and radius l_d. The detector is the set $\{x \in \mathbb{R}^2 | |x - (p_{det,x}, 0)| \leqslant l_d\}$. The detector is in the exterior of the obstacle sphere, that is $l < p_{det,x} - l_d$. The event of interest is that the monokinetic particle makes a collision in the detector, before being absorbed. The cemetery point Δ is, consequently, just chosen to be different from s and in the exterior of the detector.

Let $(x_i)_{i \in \mathbb{N}^*}$ be a trajectory of the monokinetic particle. When using the interacting-particle method of section 5.2, the event of interest is traduced by the event $\Phi(x) \geqslant 0$, with $\Phi(x) = l_d - \inf_{i \in \mathbb{N}^*; x_i \neq \Delta} |x_i - (p_{det,x}, 0)^t|$.

We now discuss the probabilities of absorption. If a collision takes place outside the box, then we have discussed that the monokinetic particle has left the domain of interest. With respect to the probability of reaching the detector, this is equivalent to writing that the probability of absorption outside the box is one. This is what we do in the sequel. Then, the probability of absorption in the box, but outside the obstacle sphere, is written $P_{a,w}$ and the probability of absorption in the obstacle sphere is written $P_{a,p}$, with $P_{a,w} \leqslant P_{a,p}$.

Finally, let us discuss the distribution of the jumps between collision points. Following the neutron-transport models, after a scattering, or birth, at X_n, of the monokinetic-particle, the direction toward which the monokinetic particle travels has isotropic distribution. This direction is here denoted u, with u an unit two-dimensional vector. Then, the sampling of the distance to the next collision point X_{n+1} is as follows. First, the distance τ is sampled from an exponential distribution with rate λ_w, if X_n is in the medium "water", or $\lambda_p > \lambda_w$ if X_n is in the medium "poison". Then, two case are possible. First, if the sampled distance is so that the monokinetic particle stays in the same medium while it travels this distance, then the next collision point is $X_{n+1} = X_n + \tau u$. Second, if between X_n and $X_n + \tau u$, there is a change of medium, then the monokinetic particle is virtually stopped at the first medium-change point between X_n and $X_n + \tau u$. At this point, the travel direction remains the same, but the remaining distance to travel is resampled, from the exponential distribution with the rate corresponding to the new medium. These resampling are iterated each time a sampled distance causes a medium-change. The new collision point X_{n+1} is the point reached by the first sampled distance that does not cause a medium change. Notice that, in this precise setting with two media, the maximum number of distance sampling between two collision points is three. This can happen in the case where the collision point X_n is in the box but not in the obstacle sphere, where the sampled direction points toward the obstacle sphere, and where toward this direction, the monokinetic particle enters and leaves the obstacle sphere.

The actual pdf, corresponding to the medium-change process described, of a collision point X_{n+1}, conditionally to a collision point X_n, is given in proposition 5.4.4.

Notice that the setting described does constitute a model for a shielding system in neutron transport. The source point corresponds to a neutron production area. This neutron production area is separated from a sensible area, modeled by the detector. The shielding system is constituted first by the obstacle sphere, which is placed between the source and the detector and has the largest probability of absorption $P_{a,p}$. Second, the standard "water" medium also constitutes a milder protection, because it also has a probability of absorption $P_{a,w}$.

We are interested in evaluating the number of monokinetic particles that reach the detector. Since the number of monokinetic-particles produced at the source point is approximately known, the problem is to evaluate the probability that a monokinetic-particle produced at the source reach the detector.

5.4.3.2 Expression of the probability density function of a trajectory

We first set some notations for the two-dimensional problem presented in section 5.4.3.1. We write B as the box $[-\frac{L}{2},\frac{L}{2}]^2$. We write B_{ext} as the exterior of the box, $B_{ext}=\mathbb{R}^2\backslash B$. The obstacle sphere is denoted S, with $S=\left\{x\in\mathbb{R}^2 | |x|\leqslant l\right\}$.

We write $|x|$ as the Euclidean norm of $x\in\mathbb{R}^2$. We write $[v,w]$ as the segment between two points $v,w\in\mathbb{R}^2$.

Consider two points $v,w\in\mathbb{R}^2$ so that v is strictly in the interior of S ($|v|<l$) and w is strictly in the exterior of S ($|v|>l$). Then $c_S(v,w)$ is defined as the unique point in the boundary of S that belongs to $[v,w]$.

Similarly, for $v,w\in\mathbb{R}^2\backslash S$ and when $[v,w]$ has a non-empty intersection with S, we denote by $c_{S_1}(v,w)$ and $c_{S_2}(v,w)$ the two intersection points between $[v,w]$ and the boundary of S. The indexes 1 and 2 are so that $|v-c_{S,1}(v,w)|\leqslant|v-c_{S,2}(v,w)|$.

For $v,w\in\mathbb{R}^2\backslash S$, we let $I_S(v,w)$ be equal to 1 if $[v,w]$ has a non-empty intersection with S, and 0 otherwise.

The computation of $c_s(v,w)$, $I_S(v,w)$, $c_{S,1}(v,w)$, and $c_{S,2}(v,w)$ are equally needed for a monokinetic-particle simulation, and for the computation of the corresponding pdf of proposition 5.4.5. The four quantities can be computed explicitly.

We now give the pdf of the collision point X_{n+1}, conditionally to a scattering or a birth point X_n.

Proposition 5.4.4. *Consider a scattering, or birth, point $x_n\in B$. Then, the pdf of the collision point X_{n+1}, conditionally to x_n, is denoted $q(x_n,x_{n+1})$ and is given by, if $x_n\in B\backslash S$*

$$\begin{aligned}
q(x_n,x_{n+1}) &= \frac{1}{2\pi|x_n-x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S}(1-I_S(x_n,x_{n+1}))\lambda_w e^{-\lambda_w|x_n-x_{n+1}|}\\
&\quad+\frac{1}{2\pi|x_n-x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S}I_S(x_n,x_{n+1})e^{-\lambda_w|x_n-c_{S_1}(x_n,x_{n+1})|}\\
&\quad\ e^{-\lambda_p|c_{S_1}(x_n,x_{n+1})-c_{S_2}(x_n,x_{n+1})|}\lambda_w e^{-\lambda_w|c_{S_2}(x_n,x_{n+1})-x_{n+1}|}\\
&\quad+\frac{1}{2\pi|x_n-x_{n+1}|}\mathbf{1}_{x_{n+1}\in S}e^{-\lambda_w|x_n-c_S(x_n,x_{n+1})|}\lambda_p e^{-\lambda_p|c_S(x_n,x_{n+1})-x_{n+1}|} \qquad (5.4.1)
\end{aligned}$$

and, if $x_n\in S$,

$$\begin{aligned}
q(x_n,x_{n+1}) &= \frac{1}{2\pi|x_n-x_{n+1}|}\mathbf{1}_{x_{n+1}\in S}\lambda_p e^{-\lambda_p|x_n-x_{n+1}|} \qquad (5.4.2)\\
&\quad+\frac{1}{2\pi|x_n-x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S}e^{-\lambda_p|x_n-c_S(x_n,x_{n+1})|}\lambda_w e^{-\lambda_w|c_S(x_n,x_{n+1})-x_{n+1}|}.
\end{aligned}$$

Proof. The proposition is obtained by using the properties of the exponential distribution, the definitions of $c_s(x_n,x_{n+1})$, $I_S(x_n,x_{n+1})$, $c_{S,1}(x_n,x_{n+1})$, and $c_{S,2}(x_n,x_{n+1})$ and a two-dimensional polar change of variables. The proof is straightforward but burdensome. □

Using proposition 5.4.4 above, we now give the pdf of the monokinetic-particle trajectories obtained from the sampling procedure of section 5.4.3.1.

Proposition 5.4.5. *The pdf, with respect to $(c_0,\mathcal{S},\Pi)$ of definition 5.4.1 and proposition 5.4.7, of a trajectory $(x_n)_{n\in\mathbb{N}^*}$, sampled from the procedure of section 5.4.3.1, is $f(x)=\sum_{n\in\mathbb{N}^*}\mathbf{1}_{A_{n+1}}(x)f_n(x)$, with*

$$f_n(x)=\prod_{i=1}^{n-1}\left(q(x_{i-1},x_i)\left[\mathbf{1}_{x_i\in B\backslash S}(1-P_{a,w})+\mathbf{1}_{x_i\in S}(1-P_{a,p})\right]\right)q(x_{n-1},x_n)\left[\mathbf{1}_{x_n\notin\mathcal{B}}+\mathbf{1}_{x_n\in B\backslash S}P_{a,w}+\mathbf{1}_{x_n\in S}P_{a,p}\right],$$

where $x_0=0$ by convention, and with $q(x_{i-1},x_i)$ and $q(x_{n-1},x_n)$ as in proposition 5.4.4.

5.4.3.3 Description of the trajectory perturbation method

The perturbation method is parameterized by $\sigma_{hm}^2>0$, $0<P_{c,w}<1$ and $0<P_{c,p}<1$.

Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n.

As in section 5.4.2, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$, so that $Y_0=0$. We modify independently the collision points of the initial trajectory, and, if the perturbed trajectory outsurvive the initial one, we generate the sequel with the initial distribution. Similarly to section 5.4.2.5, we perturb the absorption/non-absorption sampling by changing the initial values with probabilities $P_{c,w}$ and $P_{c,p}$, if the initial and perturbed collision points are both in $B\backslash S$ or both in S. If this is not the case, we sample the absorption/non-absorption for the perturbed monokinetic-particle with the initial probabilities $P_{a,w}$ and $P_{a,p}$.

More precisely, for $i \leqslant n-1$, and if the perturbed monokinetic particle has not been absorbed before collision point Y_i, it is absorbed at collision point Y_i with probability $P(x_i, Y_i)$ with

$$P(x_i, Y_i) = \begin{cases} 1 & \text{if } Y_i \in \mathbb{R}^2\backslash B \\ P_{a,w} & \text{if } Y_i \in B\backslash S \text{ and } x_i \in S \\ P_{a,p} & \text{if } Y_i \in S \text{ and } x_i \in B\backslash S \\ P_{c,w} & \text{if } Y_i \in B\backslash S \text{ and } x_i \in B\backslash S \\ P_{c,p} & \text{if } Y_i \in S \text{ and } x_i \in S \end{cases}. \quad (5.4.3)$$

Similarly to the one-dimensional case, by taking $P_{a,w}$ smaller than $\min(P_{a,w}, 1-P_{a,w})$, and $P_{a,p}$ smaller than $\min(P_{a,p}, 1-P_{a,p})$, we can modify rather mildly the initial trajectories.

If the perturbed monokinetic particle is not absorbed at collision point Y_i, its next collision point is $Y_{i+1} = x_{i+1}+\varepsilon_{i+1}$, where the $(\varepsilon_i)_{1 \leqslant i \leqslant n}$ are independent and where ε_i follows a $\mathcal{N}(0, \sigma_{hm}^2 I_2)$ distribution, where I_2 is the 2×2 identity matrix. If the perturbed monokinetic particle has not been absorbed before collision point Y_n, then it is absorbed with probability $P(x_n, Y_n)$ given by

$$P(x_n, Y_n) = \begin{cases} 1 & \text{if } Y_n \in \mathbb{R}^2\backslash B \\ P_{a,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in S \\ P_{a,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in \mathbb{R}^2\backslash B \\ P_{a,p} & \text{if } Y_n \in S \text{ and } x_n \in B\backslash S \\ P_{a,p} & \text{if } Y_n \in S \text{ and } x_n \in \mathbb{R}^2\backslash B \\ 1-P_{c,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in B\backslash S \\ 1-P_{c,p} & \text{if } Y_n \in S \text{ and } x_n \in S \end{cases}. \quad (5.4.4)$$

As in section 5.4.2, if the perturbed monokinetic particle has not been absorbed before collision point Y_n, the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial particle would be sampled if its collision point n was Y_n.

5.4.3.4 Expression of the probability density function of a perturbed trajectory

Proposition 5.4.6. *Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Let $y_0 = x_0 = 0$ by convention. The conditional pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 5.4.1 and proposition 5.4.7, of a trajectory $(y_n)_{n\in\mathbb{N}^*}$ sampled from the procedure of section 5.4.3.3, is $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ where the $f_{n,m}$ are given by the following. If $m \leqslant n-1$,*

$$f_{n,m}(x,y) = \prod_{i=1}^{m-1} \left(\phi_{x_i, \sigma_{hm}^2 I_2}(y_i) \left[1 - P(x_i, y_i)\right] \right) \phi_{x_m, \sigma_{hm}^2 I_2}(y_m) P(x_m, y_m),$$

with $P(x_i, y_i)$ and $P(x_m, y_m)$ as in (5.4.3). If $m = n$,

$$f_{n,m}(x,y) = \prod_{i=1}^{n-1} \left(\phi_{x_i, \sigma_{hm}^2 I_2}(y_i) \left[1 - P(x_i, y_i)\right] \right) \phi_{x_n, \sigma_{hm}^2 I_2}(y_n) P(x_n, y_n),$$

with $P(x_i, y_i)$ *as in* (5.4.3) *and* $P(x_n, y_n)$ *as in* (5.4.4). *If* $m \geqslant n+1$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\phi_{x_i,\sigma_{hm}^2 I_2}(y_i) \left[1 - P(x_i,y_i)\right]\right) \\ &\quad \phi_{x_n,\sigma_{hm}^2 I_2}(y_n) \left[1 - P(x_n,y_n)\right] \\ &\quad \prod_{i=n+1}^{m-1} \left(q(y_{i-1},y_i) \left[\mathbf{1}_{y_i \in B \setminus S}(1 - P_{a,w}) + \mathbf{1}_{y_i \in S}(1 - P_{a,p})\right]\right) \\ &\quad q(y_{m-1},y_m) \left[\mathbf{1}_{y_m \notin B} + \mathbf{1}_{y_m \in B \setminus S} P_{a,w} + \mathbf{1}_{y_m \in S} P_{a,p}\right], \end{aligned}$$

with $P(x_i, y_i)$ *as in* (5.4.3), $P(x_n, y_n)$ *as in* (5.4.4) *and* $q(y_{i-1}, y_i)$ *and* $q(y_{m-1}, y_m)$ *as in proposition 5.4.4.*

5.4.4 Proofs for section 5.4

For the proofs for section 5.4, we first define the space of the monokinetic-particle trajectories, and the corresponding σ-algebra and measure, in definition 5.4.1 and proposition 5.4.7. These definitions are similar to those given in section 5.3. They are stated here so that section 5.4 is self-sufficient.

Definition 5.4.1. *Define*

$$c_0 = \{(u_n)_{n \geqslant 1} \in \left(\mathbb{R}^d\right)^{\mathbb{N}^*} : \exists n_0 \in \mathbb{N}^*, \forall n \geqslant n_0, u_n = 0\}.$$

We define $\mathcal{S}$ *as the smallest sigma-algebra on* c_0 *containing the sets* $\{x | x_1 \in A_1, ..., x_n \in A_n\}$, *for* $n \in \mathbb{N}^*$ *and* $A_i \in \mathcal{B}\left(\mathbb{R}^d\right)$, *where* $\mathcal{B}\left(\mathbb{R}^d\right)$ *is the Borel sigma-algebra on* $\mathbb{R}^d$.

Proposition 5.4.7. *There exists a unique measure* Π *on* $(c_0, \mathcal{S})$ *that verifies the following relation, for any* $E_n = \{x | x \in A_{n+1}, (x_1, ..., x_n) \in A_1 \times ... \times A_n\}$, *with* $A_1, ..., A_n \in \mathcal{B}\left(\mathbb{R}^d\right)$ *and* $n \in \mathbb{N}^*$.

$$\Pi(E_n) = \lambda(A_1 \times ... \times A_n), \tag{5.4.5}$$

with λ *the Lebesgue measure.*

Proof. The proof is carried out in the same way as in section 5.3. Alternatively, the Carathéodory extension theorem can be used. □

We now give the following general proposition 5.4.8, for the expression of pdf for inhomogeneous Markov chains that are absorbed in finite-time.

Proposition 5.4.8. *Consider a sequence of measurable applications* $a_n : \mathbb{R}^d \to [0,1]$, $n \in \mathbb{N}$, *with* $a_0 = 0$. *Consider a sequence* $(q_n)_{n \in \mathbb{N}^*}$ *of conditional pdf, that is to say* $\forall n, y_{n-1}$, $p_n(y_{n-1}, y_n)$ *is a pdf on* $\mathbb{R}^d$ *with respect to* y_n.

Consider a probability space $(\Omega, \mathcal{F}, P)$. *Consider a Markov Chain on* $(\Omega, \mathcal{F}, P)$, $(Y_n)_{n \in \mathbb{N}}$, *so that first* $Y_0 = y_o$ *a.s, when* y_0 *in a non-zero constant of* $\mathbb{R}^d$. *Second,* Y_n *has the non-homogeneous transition kernel defined by*

$$k(y_{n-1}, dy_n) = \mathbf{1}_{y_{n-1}=0}\delta_0(dy_n) + \mathbf{1}_{y_{n-1}\neq 0}\left\{a_n(y_{n-1})\delta_0(dy_n) + \left[1 - a_n(y_{n-1})\right] q_n(y_{n-1}, y_n) dy_n\right\}. \tag{5.4.6}$$

Assume finally that, almost surely, the Markov Chain Y_n *reaches* 0 *after a finite time. Then, the application* $\omega \to (Y_i(\omega))_{i \in \mathbb{N}^*}$ *is a random variable on* $(c_0, \mathcal{S}, \Pi)$ *(see definition 5.4.1 and proposition 5.4.7), with probability density function, for* $y = (y_i)_{i \in \mathbb{N}^*}$, $f(y) = \sum_{n=1}^{+\infty} \mathbf{1}_{A_{n+1}}(y) f_n(y)$, *with*

$$f_n(y) = \prod_{i=1}^{n} \left[(1 - a_{i-1}(y_{i-1})) q_i(y_{i-1}, y_i)\right] a_n(y_n),$$

where s_0 *is the constant value of* Y_0 *by convention.*

Proof. Proposition 5.4.8 is proved in the same way as proposition 5.3.2. □

Let us now comment proposition 5.4.8.

Similarly to proposition 5.3.2, the value 0 corresponds to the cemetery point, symbolizing the absorption of the Markov chain. Thus, the dynamic (5.4.6) is that, at each time n, if the Markov chain with value Y_n is absorbed, it stays absorbed. If it is not absorbed, then is is absorbed with probability $a_n(Y_n)$, thus a probability both depending on position and time. This is consistent with the sampling of perturbed trajectories in sections 5.4.2.5 and 5.4.3.3, where these probabilities of absorption depend on the initial trajectories, and thus depend on time and position. If the Markov chain is not absorbed at Y_n, then its next value Y_{n+1} has, conditionally to Y_n, the pdf $q_{n+1}(Y_n, y_{n+1})$. This conditional distribution depends on time, as is the case in the perturbation procedures of sections 5.4.2.3, 5.4.2.5 and 5.4.3.3.

Hence, proposition 5.4.8 can be applied to calculate the pdf of initial and perturbed trajectories in propositions 5.4.1, 5.4.2, 5.4.3, 5.4.5 and 5.4.6.

Proof. [Proof of proposition 5.4.1] We apply proposition 5.4.8 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_a + \mathbf{1}_{y_i \notin D}$$

and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i).$$

□

Proof. [Proof of proposition 5.4.2] We denote $x = (x_i)_{i \in \mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 5.4.8 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \notin D},$$

for $i \geqslant 1$,

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1} + x_i - x_{i-1}, \sigma^2_{hm}}(y_i),$$

for $1 \leqslant i \leqslant n$ and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i),$$

for $i \geqslant n+1$. □

Proof. [Proof of proposition 5.4.3] We denote $x = (x_i)_{i \in \mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 5.4.8 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_c + \mathbf{1}_{y_i \notin D},$$

for $1 \leqslant i \leqslant n-1$,

$$a_n(y_n) = \mathbf{1}_{y_i \in D} \left(\mathbf{1}_{x_i \in D}(1 - P_c) + \mathbf{1}_{x_i \notin D} P_a \right) + \mathbf{1}_{y_i \notin D},$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_a + \mathbf{1}_{y_i \notin D},$$

for $i \geqslant n+1$,

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1} + x_i - x_{i-1}, \sigma^2_{hm}}(y_i),$$

for $1 \leqslant i \leqslant n$ and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i),$$

for $i \geqslant n+1$. □

Proof. [Proof of proposition 5.4.5] We apply proposition 5.4.8 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in S} P_{a,p} + \mathbf{1}_{y_i \in B \setminus S} P_{a,w} + \mathbf{1}_{y_i \in \mathbb{R}^2 \setminus B}$$

and

$$q_i(y_{i-1}, y_i) = q(y_{i-1}, y_i),$$

with $q(y_{i-1}, y_i)$ as in proposition 5.4.4. □

Proof. [Proof of proposition 5.4.6] We denote $x = (x_i)_{i\in\mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 5.4.8 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = P(x_i, y_i),$$

for $1 \leqslant i \leqslant n-1$ and with $P(x_i, y_i)$ as in (5.4.3),

$$a_n(y_n) = P(x_n, y_n),$$

with $P(x_n, y_n)$ as in (5.4.4),

$$a_i(y_i) = \mathbf{1}_{y_i\in S} P_{a,p} + \mathbf{1}_{y_i\in B\setminus S} P_{a,w} + \mathbf{1}_{y_i\in\mathbb{R}^2\setminus B},$$

for $i \geqslant n+1$

$$q_i(y_{i-1}, y_i) = \phi_{x_i, \sigma_{hm}^2 I_2}(y_i),$$

for $1 \leqslant i \leqslant n$ and

$$q_i(y_{i-1}, y_i) = q(y_{i-1}, y_i),$$

for $i \geqslant n+1$ and with $q(y_{i-1}, y_i)$ as in proposition 5.4.4. ◻

5.5 Numerical results in dimension one and two

In this section 5.5, we present numerical results for the interacting-particle method of section 5.2, in the one and two-dimensional cases of section 5.4. We follow a double objective. First we aim at investigating to what extent the ideal results of the interacting-particle method hold (in term of bias and of theoretical confidence intervals). Second, we want to confirm that, when the objective probability is small, the method outperforms a simple-Monte Carlo method.

The simple-Monte Carlo method is parameterized by a number of Monte Carlo sample n_{mc}. It consists in generating n_{mc} independent trajectories $x_1, ..., x_{n_{mc}}$ and in estimating p by the empirical proportion of these trajectories that verify the small-probability event. We denote by $\hat{p}_{mc}$ the simple-Monte Carlo estimator of p.

5.5.1 Numerical results in dimension one

5.5.1.1 Features of the interacting-particle method

We first present a simple one-dimensional setting, with no-absorption ($P_a = 0$). We set for the domain $L_{inf} = -10$, $L_{sup} = 1$, and for the variance of the increments $\sigma_{mk}^2 = 1$. As a result, the probability p to estimate is not small. It is easily estimated to be $p = 0.13$ by the simple-Monte Carlo method.

For the perturbation method, we set $\sigma_{hm}^2 = 0.1^2$. This choice may not be optimal, but it is reasonable and can be considered as typical for the implementation of the interacting-particle method in this one-dimensional case.

The results we obtain for 100 independent estimations for the interacting-particle method are regrouped in figure 5.1. We have used $N = 200$ particles and $T = 300$ and $T = 30$ iterations in the HM algorithm 5.2.2.1. Let us first interpret the results for $T = 300$ iterations. In this case, we observe that the estimator is empirically non-biased. Furthermore, we also plot the theoretical 95% confidence intervals for the ideal estimator with $T = +\infty$, that are approximately (for N large) $I_p = \left[p\exp\left(-1.96\sqrt{\left(\frac{-\log p}{N}\right)}\right), p\exp\left(1.96\sqrt{\left(\frac{-\log p}{N}\right)}\right)\right]$. We also recall from the discussion after (5.2.1) that the events $\hat{p}_{ipm} \in I_p$ and $p \in I_{\hat{p}_{ipm}}$ are approximately equivalent when N is large. Hence the coverage probability of I_p for $\hat{p}_{ipm}$ is approximately the probability that $I_{\hat{p}_{ipm}}$ contains p, which is the practical quantity of interest. We see on figure 5.1 that I_p approximately matches the empirical distribution of the estimator $\hat{p}_{ipm}$. The overall conclusion of this case $T = 300$ is that there is a good agreement between theory and practice. This emphasizes the validity of using the interacting-particle method of algorithm 5.2.2.2, involving the HM algorithm, in a space that is not a subset of $\mathbb{R}^d$.

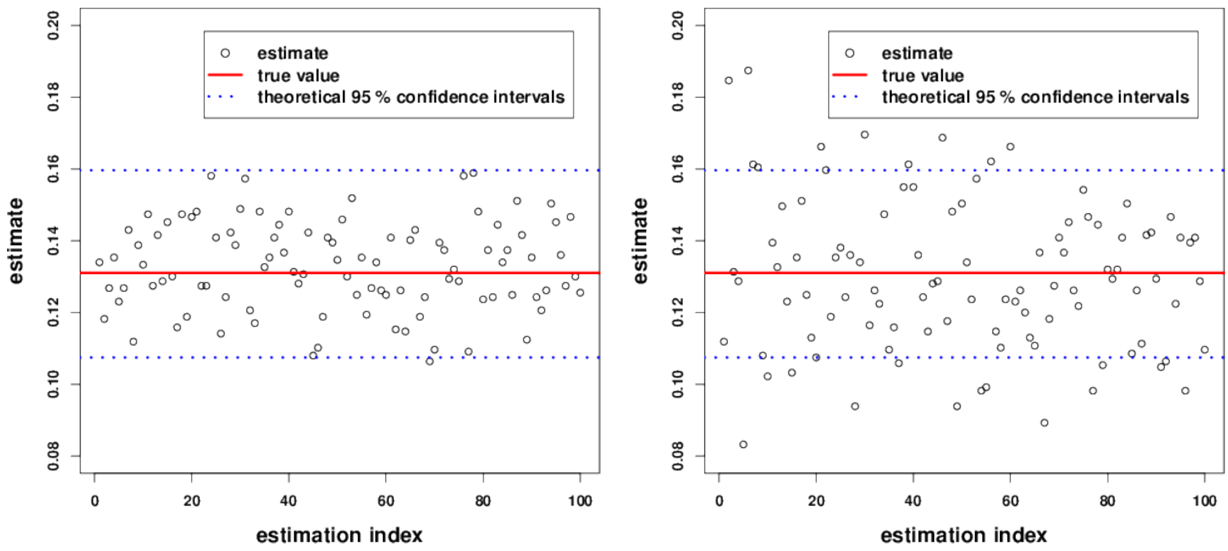

FIGURE 5.1 – One-dimensional case. Plot of 100 independently estimated probabilities with the interacting-particle method 5.2.2.2, for number of particles $N = 200$, and number of iterations in the HM algorithm 5.2.2.1 $T = 300$ (left) and $T = 30$ (right). We also plot the theoretical 95% confidence intervals (5.2.1) of the case $T = +\infty$. The true probability $p = 0.13$ is evaluated quasi-exactly by a simple-Monte Carlo method. In both cases, the interacting-particle estimator is empirically unbiased. For $T = 300$, the theoretical confidence interval, obtained in the case $T = +\infty$ is adapted to the practical estimator. For $T = 30$ however, the estimator has more variance that the ideal estimator $T = +\infty$ has.

In figure 5.1, we also consider the case $T = 30$. The estimator is still empirically unbiased. However, its empirical variance is larger, so that the theoretical 95% confidence interval I_p is non negligibly too thin. This can be interpreted, because when T is small, a new particle at a given conditional sampling step of algorithm 5.2.2.2 is not independent of the $N - 1$ particles that have been kept. Thus, one can argue that, at each step of algorithm 5.2.2.2, the overall set of N particles has more interdependence, so that eventually the estimator has more variance. Nevertheless, on the other hand, an estimation with $T = 30$ is 10 times less time-consuming that an estimation with $T = 300$. We further discuss this trade-off problem in section 5.5.3.

Finally, for this case of a probability that is not small, we have used simple Monte Carlo as a mean to estimate it quasi-exactly. We have found that the interacting-particle method 5.2.2.2 requires more computation time than the Monte Carlo method, for reaching the same accuracy. We do not elaborate on this fact, since we especially expect the interacting-particle-method to be competitive for estimating a small probability. This is the object of section 5.5.1.2. For this case of a probability that is not small, we have just investigated the features of the interacting-particle method.

5.5.1.2 Comparison with simple Monte Carlo in a small-probability case

We now consider a case with possible absorption of the monokinetic particle. Thus we set $P_a = 0.45$. We keep the same values $\sigma_{mk}^2 = 1$ and $L_{sup} = 1$ as in section 5.5.1.1, but we set $L_{inf} = -15$. As a result of these parameters for the monokinetic-particle transition kernel, the probability of interest is small. In fact, we have not estimated it with negligible uncertainty. With a simple-Monte Carlo estimation of sample size 10^9, the probability estimate is $\hat{p}_{vlmc} = 6.6 \times 10^{-8}$. We call this estimate the very large Monte Carlo (VLMC) estimate. Given that the number of successes in this estimate is 66, which is not very large, we are reluctant to use the Central Limit Theorem approximation for computing 95% confidence intervals. Instead, we use the Clopper-Pearson interval [18], for which the actual coverage probability is always larger than 95%. This 95% confidence interval is there equal to $[5.1 \times 10^{-8}, 8.4 \times 10^{-8}]$. This uncertainty is small enough for the conclusions we will draw from this case. Finally, notice that this very large Monte Carlo estimate is not a benchmark for the interacting-particle method, because it is much more time consuming.

For the interacting-particle method, we set $N = 200$ particles, and for the HM algorithm, we set $T = 300$ iterations. We use $\sigma_{hm}^2 = 0.1^2$ and $P_c = 0.2$ for the perturbation method. We still denote $\hat{p}_{ipm}$ the obtained estimator for p. We consider a third estimator, that we denote p_{mc} and that consists in the simple-Monte Carlo estimator with sample size 5×10^6. This sample size is appropriate to compare the efficiency of the interacting-particle and Monte Carlo method, as we will show below.

The first criterion for comparing the two estimators $\hat{p}_{ipm}$ and $\hat{p}_{mc}$ is their computation time. We have two possible ways to make this comparison. First, we can evaluate the complexities of the two methods. The Monte Carlo method requires to perform 5×10^6 monokinetic-particle simulations. For each proposed perturbation, the interacting-particle method requires to sample one perturbed trajectory, and to compute its unconditional and conditional pdf. This has to be done approximately $T \times \frac{\log \hat{p}_{vlmc}}{\log(1-\frac{1}{N})} \approx 10^6$ times. Thus, from this point of view, the costs of the two methods have the same orders of magnitude. We can not give a more precise comparison, since the trajectories sampled by the two methods do not necessarily have the same length in the mean sense. Furthermore, it is not obvious to compare the computational cost of an initial sampling, with the costs of a conditional sampling and pdf computations.

Hence, we will just compare the computational costs of the two methods by considering their actual computational times, for the implementation we have used. Averaged over all the estimations, the time for the interacting-particle method is 58% of the time for the Monte Carlo method. Hence, we confirm that the computational costs are of the same order of magnitude, the comparison being nevertheless beneficial to the interacting-particle method.

We now compare the accuracy of the two methods for estimating the true probability p. On figure 5.2, we plot the results of 100 independent estimations for $\hat{p}_{ipm}$ and 50 independent estimations for $\hat{p}_{mc}$. It appears clearly that the interacting-particle method is more precise in this small probability case. Especially, consider the empirical Root Mean Square Error criterion, for n independent estimates $\hat{p}^1, ..., \hat{p}^n$, for any estimator $\hat{p}$ of p : $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^n (p - \hat{p}^i)^2}$. Regardless of the value of p in the very large Monte Carlo 95% confidence interval $[5.1 \times 10^{-8}, 8.4 \times 10^{-8}]$, the RMSE is smaller for $\hat{p}_{ipm}$ than for $\hat{p}_{mc}$. If we assume $p = \hat{p}_{vlmc}$, then the RMSE is 10^{-7} for $\hat{p}_{mc}$ and 2×10^{-8} for $\hat{p}_{ipm}$.

A comparison ratio for $\hat{p}_{ipm}$ and $\hat{p}_{mc}$, taking into account both computational time and estimation accuracy (in line with the efficiency in [40]), is the quality ratio defined by $\frac{\sqrt{TIME_{mc}} \times RMSE_{mc}}{\sqrt{TIME_{ipm}} \times RMSE_{ipm}}$, where the four notations $TIME_{mc}$, $TIME_{ipm}$, $RMSE_{mc}$ and $RMSE_{ipm}$ are self-explanatory. This ratio is 6.7 here. This is interpreted as : if the two estimation methods were set as to require the same computational time, then the interacting-particle method would be 6.7 times as accurate (in term of RMSE) as the Monte Carlo method.

Notice that, if we had done the comparison from the point of view of the relative estimation errors, instead of the absolute errors, it would have been even more beneficial to the interacting-particle method. Indeed, assuming again $p = \hat{p}_{vlmv} = 6.6 \times 10^{-8}$ for discussion, the interacting-particle method does a maximum relative error of 250%. On the other hand, the Monte Carlo estimator takes only 3 different values in figure 5.2. When it takes value $\frac{2}{5 \times 10^6}$ it does a relative error of 600%, when it takes value $\frac{1}{5 \times 10^6}$ it does a relative error of 300% and when it takes value 0 it does an infinite relative error. Alternatively, we can also say that, in the majority of the cases, the Monte Carlo estimator does not see any realization of the rare event, so that it can provide only an overly-conservative upper-bound for p.

5.5.2 Numerical results in dimension two

5.5.2.1 Features of the interacting-particle method

We now present the numerical results for the two-dimensional case. We set the absorption probability in the background $P_{a,w} = 0.2$, the absorption probability in the obstacle sphere $P_{a,p} = 0.5$, the dimension of the box $[-\frac{L}{2}, -\frac{L}{2}]$ $L = 10$, the radius of the obstacle sphere $l = 2$, the radius of the detector $l_d = 0.5$. The positions of the detector and the source are given by $p_{det,x} = p_{s,x} = 3$. We set the rate of collisions in the water medium $\lambda_w = 0.2$ and in the poison medium $\lambda_p = 2$. As a result, the probability is $p = 2 \times 10^{-4}$ and is evaluated quasi-exactly by a Monte Carlo sampling, similarly to section 5.5.1.1.

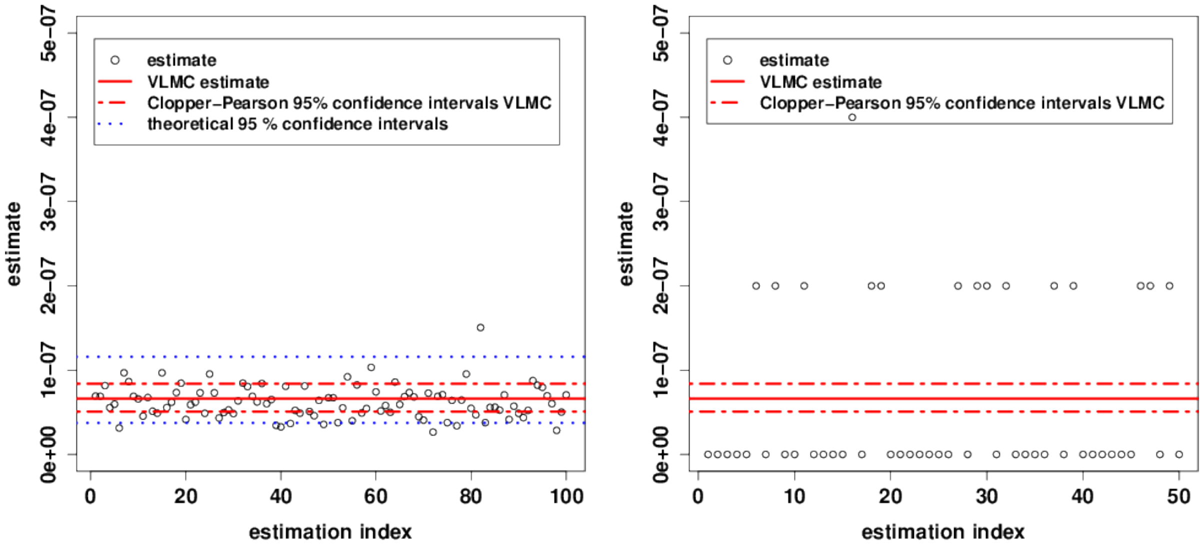

FIGURE 5.2 – One-dimensional case. Plot of 100 independently estimated probabilities with the interacting-particle method 5.2.2.2 for $N = 200$ and $T = 300$ (left) and 50 independently estimated probabilities with the Monte Carlo method with sample size 5×10^6 (right). We plot a very large Monte Carlo estimate of the true probability $\hat{p}_{vlmc} = 6.6 \times 10^{-8}$, together with the associated Clopper-Pearson 95% confidence intervals. For the interacting-particle method, we also plot the theoretical 95% confidence intervals of the case $T = +\infty$, assuming the true probability is the VLMC estimate. The uncertainty on the VLMC estimate of the true value p of the probability is small enough for our conclusions to hold. These conclusions are that the interacting-particle method outperforms the Monte Carlo method (with sample size 5×10^6), both in term of computation time and of accuracy.

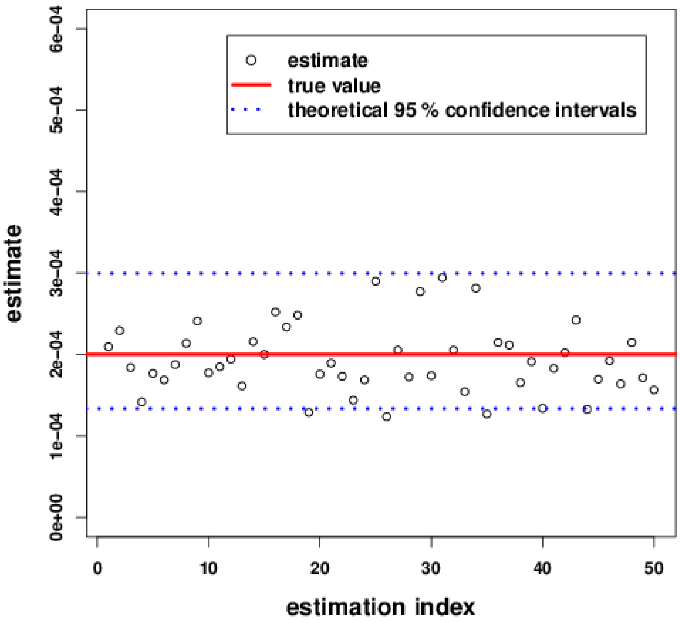

FIGURE 5.3 – Two-dimensional-case. Plot of 50 independently estimated probabilities with the interacting-particle method 5.2.2.2, for number of particles $N = 200$ and number of iterations in the HM algorithm $T = 300$. We also plot the theoretical 95% confidence intervals (5.2.1) of the case $T = +\infty$. The true probability $p = 0.2 \times 10^{-4}$ is evaluated quasi-exactly by a simple-Monte Carlo method. The interacting-particle estimator is empirically unbiased and the 95% theoretical confidence interval, obtained in the case $T = +\infty$, is adapted to the practical estimator.

This value is not very small, so that we do not compare the interacting-particle method with the Monte Carlo method. We just aim at showing that the interacting-particle method is valid in this two-dimensional setting. Indeed, this setting has many features that are representative of shielding studies with Monte Carlo codes. Namely, the setting involves absorption, the presence of two media with different collision rates and the presence of medium-border crossing phenomena.

For the HM perturbations of algorithm 5.2.2.1, we set the collision-point perturbation variance $\sigma_{hm}^2 = 0.5^2$, the probability of changing the absorption/non absorption in the obstacle sphere $P_{c,p} = 0.1$ and in the rest of the box $P_{c,w} = 0.05$. As in section 5.5.1.1, these settings are reasonable, but are not tuned as to yield an optimal performance of the interacting-particle method.

In figure 5.3, we present the results for 50 independent estimations with the interacting-particle method. Empirically, the estimator is unbiased and the theoretical 95% confidence intervals are valid. This is the same conclusion as in section 5.5.1.1, and is again a validation of the HM algorithm in the space of the monokinetic-particle trajectories.

5.5.2.2 Comparison with simple Monte Carlo in a small-probability case

We now consider the case of a small probability. For this, we set the absorption probability in the obstacle sphere $P_{a,p} = 0.7$ and in the rest of the box $P_{a,w} = 0.5$, the dimension of the box $[-\frac{L}{2}, -\frac{L}{2}]$ $L = 10$, the radius of the obstacle sphere $l = 2.5$, the radius of the detector $l_d = 0.5$. The positions of the detector and the source are given by $p_{det,x} = p_{s,x} = 3$. We set the rate of collisions in the water medium $\lambda_w = 2$ and in the poison medium $\lambda_p = 3$. In essence, the obstacle sphere is larger than in subsection 5.5.2.1, the absorption probabilities are larger, and the collision rates are larger, thus yielding all the more frequent absorption.

The probability p is estimated by very large Monte Carlo with sample size 1.25×10^9. The estimate is $\frac{22}{1.25 \times 10^9} \approx 1.76 \times 10^{-8}$. Similarly to section 5.5.1.2, the Clopper-Pearson 95% confidence interval for the probability is $[10^{-8}, 2.5 \times 10^{-8}]$. It is also small enough for validating the discussion that follows.

We compare the estimators $\hat{p}_{ipm}$, with $N = 200$ particles and $T = 300$ iterations in the HM algorithm, and the estimator $\hat{p}_{mc}$ with sample size 5×10^6. We have found that the computation time for the estimator

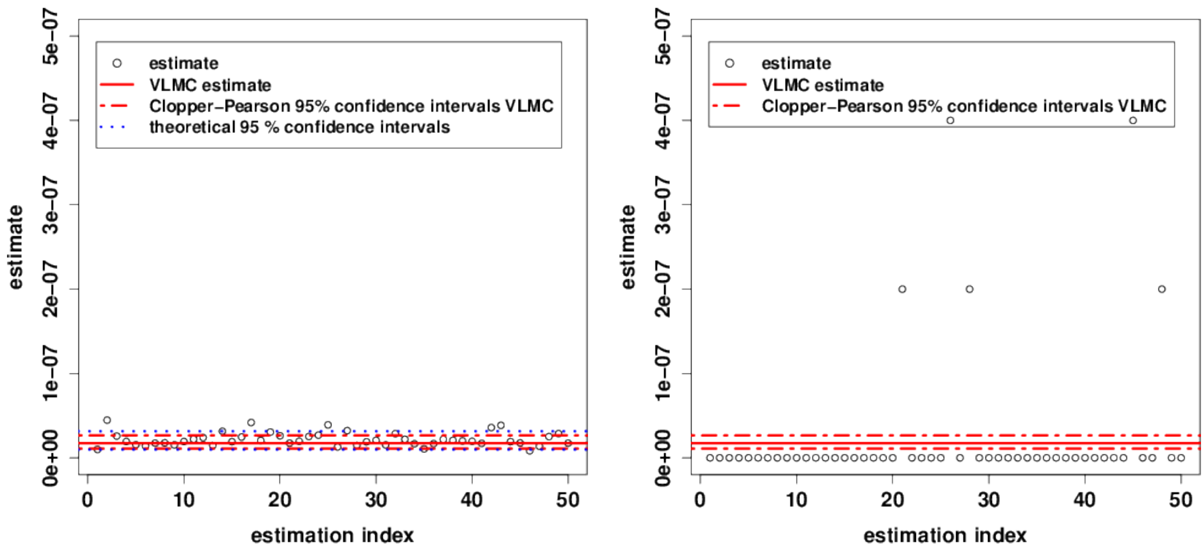

FIGURE 5.4 – Two-dimensional-case. Plot of 50 independently estimated probabilities with the interacting-particle method of algorithm 5.2.2.2 for $N = 200$ and $T = 300$ (left) and with the simple-Monte Carlo method with sample size 5×10^6 (right). We plot a very large Monte Carlo estimate of the true probability $\hat{p}_{vlmc} = 1.76 \times 10^{-8}$, together with the associated Clopper-Pearson 95% confidence intervals. For the interacting-particle method, we also plot the theoretical 95% confidence intervals of the case $T = +\infty$, assuming the true probability is the VLMC estimate. The uncertainty on the VLMC estimate of the true value p of the probability is small enough for our conclusions to hold. These conclusions are that the interacting-particle method outperforms the Monte Carlo method (with sample size 5×10^6), both in term of computation time and of accuracy.

$\hat{p}_{ipm}$ is, on average, 88% of that of the estimator $\hat{p}_{mc}$.

Now, concerning estimation accuracy, the results are presented in figure 5.4. The interacting-particle method outperforms the simple-Monte Carlo method, to a greater extent that in figure 5.2. As a confirmation, the quality ratio $\frac{\sqrt{TIME_{mc}} \times RMSE_{mc}}{\sqrt{TIME_{ipm}} \times RMSE_{ipm}}$ is 10.5, against 6.7 in figure 5.2.

5.5.3 Discussion on the numerical results

We now discuss some conclusions on the numerical results of section 5.5. First, in two cases with a probability that is not small (figures 5.1 and 5.3), the interacting-particle method is empirically unbiased. The theoretical confidence intervals $T = +\infty$ are in agreement with the empirical distribution for finite T, provided that T is large enough. For the two cases of small probabilities (figures 5.2 and 5.4), we do not state conclusions on this question, in one sense or another, because we do not know the probability with negligible uncertainty.

However, for figures 5.2 and 5.4, the uncertainty on the probability is small enough to compare the performances of the simple-Monte Carlo and interacting-particle methods. The conclusion of this comparison is strongly unilateral, and is that, for a small probability, the interacting-particle method is preferable over a simple-Monte Carlo sampling.

We have not carried out numerical test for extremely small probabilities (say, under 10^{-10}). The reason for that is that we would not have an estimate of these probabilities similar to the very large Monte Carlo estimate $\hat{p}_{vlmc}$. That it to say an estimate that comes with confidence intervals with guaranteed coverage probability. Nevertheless, a simple-Monte Carlo method, with computational time similar to that of the interacting-particle method, would most likely never see the rare event, and thus only provide an overly conservative upper bound. Thus, the comparison would be even more in favor of the interacting-particle method than for figures 5.2 and 5.4.

In figure 5.1, we have mentioned the trade-off problem between the number of particles N and the

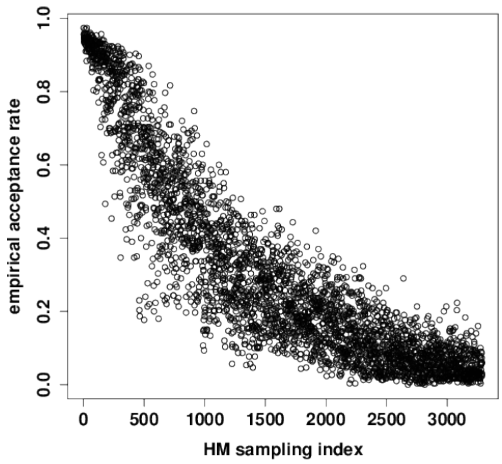

FIGURE 5.5 – Same setting as in figure 5.2 for the interacting-particle method. We consider one estimation $\hat{p}_{ipm}$ of the interacting-particle method. We plot the empirical acceptance rate, over the $T = 300$ proposed perturbations, for each of the $\frac{\log \hat{p}_{ipm}}{\log\left(1-\frac{1}{N}\right)}$ HM samplings of algorithm 5.2.2.1. The acceptance rate decreases considerably when one gets closer to the rare-event.

number of HM iterations T. The average complexity of the interacting-particle method is proportional to the product NT. Naturally, increasing N improves the accuracy of the interacting-particle method. Especially, the variance is proportional to N when N is large, in the ideal case $T = +\infty$. We have seen in figure 5.2 that increasing T also reduces the variance, which is well interpreted. It is however quite difficult to quantify the dependence between T and the variance of the estimator. We think that the question of this trade-off between N and T would benefit from further investigation.

In our experiments, we have not optimized the choice of the perturbation method. This would naturally bring a potential additional benefit for the interacting-particle method. Perhaps less natural is the prospect of allowing the perturbation method to vary with the progression of the algorithm. For example, one could use a perturbation method that propose perturbed trajectories that are closer to the initial ones, when these trajectories are close to the rare event. The results we now present in figure 5.5 support this idea. In figure 5.5, we plot the acceptance rate in the HM algorithm 5.2.2.1 (by acceptance we mean that both the pdf ratio and the objective function conditions are fulfilled), as a function of the progression in the interacting-particle method. This acceptance rate is decreasing, and is small when the interacting-particle method is in the rare-event state. Notice that this was not the case in the experiments conducted in [36].

An other potential tuning of the interacting-particle method is the choice of the objective function Φ, for which the event "the monokinetic particle makes a collision in the detector" is equivalent to the event that Φ, evaluated on the trajectory of the monokinetic particle, exceeds a threshold. We have used as a function Φ the (opposite of the) minimum, over the collision points of the trajectory, of the Euclidean distance to the center of the detector. This choice could be improved. One natural possibility is to replace the Euclidean distance by the optical distance. That is to say the distance traveled in each medium would be weighted by the collision rate in the medium. For some neutron-transport problem, it is also possible to use more specific objective functions, by finding approximations of the importance function, see e.g. [30].

Conclusion

We have considered the adaptation of the interacting-particle method [36] to a small-probability estimation problem, motivated by shielding studies in neutron transport. The adaptation is not straightforward, because shielding studies involve working on probability distributions on a set of trajectories that are killed after a finite time.

The contribution brought by the paper it two-fold. First, it has been shown that probability density functions can be defined on this set. This enables to use the Hastings-Metropolis algorithm, which is necessary to implement the method [36] in practice. A convergence result has also been shown for the Hastings-Metropolis algorithm in this setting.

The second contribution of the paper is to give the actual probability density function equations, for implementing the interacting-particle method in an academic one-dimensional problem, and a simplified but realistic two-dimensional problem. In both cases, the method is shown to be valid. Furthermore, the method outperforms a simple-Monte Carlo estimator, for estimating a small probability.

Prospects are possible for both contributions. First, the proof of the convergence of the Hastings-Metropolis could be extended under more general assumptions. Second, several possibilities for practical improvement of the interacting-particle method are presented in section 5.5.3.

Bibliographie

[1] Aboura, O.(2011) : A Regression Monte-Carlo method for Backward Doubly Stochastic Differential Equations. hal-00607274, version 1. (Cited on pages 6, 12, 13, 20, 62 and 87.)

[2] Agostinelli, S. & al, *Geant4 : a simulation toolkit*, Nuclear instruments and methods in physics research section A 506, 250-303, 2003. (Cited on pages 14 and 101.)

[3] Aman, A.(2011) : A numerical scheme for backward doubly stochastic differential equations. *Arxiv :1011.6170v2*. Bernouilli, to appear. (Cited on pages 6, 20 and 62.)

[4] Bachoc, F., Bachouch, A., Lenotre, L. (2014) : Hastings-Metropolis algorithm on Markov chains for small-probability estimation. *Submitted to ESAIM Proceedings*. (Cited on page 13.)

[5] Bachouch A., Ben Lasmer M. A., Matoussi A., Mnif M. (2013) : Numerical scheme for semilinear Stochastic PDEs via Backward Doubly Stochastic Differential Equations. *Arxiv : 1302.0440 v3. In revision*. (Cited on page 3.)

[6] Bachouch, A., Gobet, E., Matoussi, A. (2014) : Empirical Regression Method for Backward Doubly Stochastic Differential Equations. *In progress*. (Cited on page 12.)

[7] Bachouch A., Matoussi, A. (2014) : Numerical Computations for quasilinear Stochastic PDEs. *Preprint*. (Cited on page 10.)

[8] Bally, V. (1997). Approximation scheme for solutions of BSDE. *Pitman Research Notes in Mathematics*, 364 :177–191. (Cited on pages 6, 20 and 62.)

[9] Bally V., Matoussi A. (1999) : Weak solutions for SPDEs and Backward Doubly Stochastic Differential Equations. *J. Theoretical Probability*, Vol.14,No.1, 125-164. (Cited on pages 1, 5, 6, 8, 19, 39, 40 and 41.)

[10] Barles G., Lesigne E. (1997) : SDE, BSDE and PDE. *Pitman Research Notes in Mathematics Series*, 364, 47-80 ; (Cited on pages 8, 39 and 41.)

[11] Bismut, J.M. (1973). Théorie Probabiliste du ContrÃ´le des Diffusions, Mem. Amer. Math. Soc. 176, Providence, Rhone Island. *Pitman Research Notes in Mathematics*, 364 :177–191. (Cited on page 1.)

[12] Booth, T.E., *Comments on Monte Carlo probability of initiation estimates for neutron fission chains*, Nucl. Sci. Eng. 166, 175-178, 2010. (Cited on pages 14 and 101.)

[13] Booth, T.E., *Common misconceptions in Monte Carlo particle transport*, Applied Radiation and Isotopes 70, 1042-1051, 2012. (Cited on pages 14 and 101.)

[14] Both, J.P. & AL,, *Automated importance generation and biasing techniques for Monte-Carlo shielding techniques by the Tripoli-3-code*, Progress in Nuclear Energy 24, 273-281, 1990. (Cited on page 101.)

[15] Bouchard B., Elie R.(2008) : Discrete-time approximation of decoupled forward-backward SDEs with jumps. *Stochastic Processes and Their applications*, 118(1), 53-57. (Cited on pages 6, 20 and 62.)

[16] Bouchard B., Touzi N. (2004) : Discrete time approximation and Monte-Carlo Simulation of Backward Stochastic differential equations. *Stochastic Processes and Their applications*, 111, 175-206. (Cited on pages 1, 6, 20 and 62.)

[17] Burn, K. W., *Complete optimization of space/energy cell importance with the dsa cell importance model*, Ann. Nucl. Energy 19(2), 65-98, 1992. (Cited on pages 14 and 101.)

[18] Clopper, C. and Pearson, E. S., *The use of confidence or fiducial limits illustrated in the case of the binomial*, Biometrika 26(4), 404â413, 1934. (Cited on page 119.)

[19] Crisan D.(2006) : Particle Approximations for a Class of Stochastic Partial Differential Equations. *Appl Math Optim*, 54, 293-314. (Cited on pages 7 and 20.)

[20] Crisan D., Manolarakis K. and Touzi N.(2010) : On the Monte Carlo simulation of BSDEs : An improvement on the Malliavin weights. *Stochastic Processes and Their applications*, to appear. (Cited on pages 6, 20 and 62.)

[21] Crisan D., Manolarakis K.(2010) : Second order discretization of Backward SDEs. *Arixv :1012.5650v1.* Annals of the Applied Probability, to appear. (Cited on page 6.)

[22] Crisan D., Manolarakis K.(2012) : Solving Backward Stochastic Differential Equations using the Cubature Method. *SIAM Journal on Mathematical Finance*, 534-571. (Cited on page 6.)

[23] Del Moral, P. & Garnier, J., *Genealogical particle analysis of rare events* Ann. Appli. Probab. 15, 2496-2534, 2005. (Not cited.)

[24] Denis, L. : Solutions of stochastic partial differential equations considered as Dirichlet processes. *Bernoulli J.of Probability* **10**, 783-827 (2004).2093611 (Not cited.)

[25] Denis L., Matoussi A. and Stoica, I. L. : L^p estimates for the uniform norm of solutions of quasilinear SPDE's. *Probability Theory Related Fileds* **133**, 437-463 (2005).2197109 (Cited on page 62.)

[26] Denis L., Matoussi A. and Stoïca L. : Maximum Principle and Comparison Theorem for Quasilinear Stochastic PDE's. *Electronic Journal of Probability*, 14, p. 500-530 (2009). (Cited on page 62.)

[27] L. Denis, A. Matoussi, L. Stoica. Moser iteration applied to parabolic SPDEâs : first approach. Quaderni di Matematica, Stochastic Partial Differential Equations and Applications Vol. 25, pp 99-125 (2010). (Cited on page 62.)

[28] Denis L. and Stoica, I. L. : A general analytical result for non-linear s.p.d.e.'s and applications, *Electronic Journal of Probability* **9**, 674-709 (2004).2480551 (Cited on pages 62, 64, 65 and 66.)

[29] Diop, C. & al, *Tripoli-4 : a 3D continuous-energy Monte Carlo transport code*, Trans. Am. Nuc. Soc. 95, 661, 2006 (Cited on pages 14 and 101.)

[30] Dumonteil, E., *On a New Variance Reduction Technique : Neural Network Biasing - A Study of Two Test Cases with the Monte Carlo Code Tripoli4*, Nuclear Technology 168(3), 793-798, 2009. (Cited on pages 14, 101 and 124.)

[31] El Karoui N., Peng S.,Quenez M.C.(1997) : Backward Stochastic Differential Equations in finance. Mathematical Finance, Vol.7, No. 1, 1-71. (Cited on pages 2, 23, 33, 34, 56 and 59.)

[32] Gerencsér M., Gyongy I. (2013) : Finite element schemes for stochastic partial differential equations in sobolev spaces. Arxiv : 1308.4614v1. (Cited on pages 6 and 7.)

[33] Gobet, E., Lemor, J.P, Warin X.(2005) : A regression-based Monte Carlo method to solve backward stochastic differntiel equations. *The Annals of Applied Probability*, vol.15 , no. 3, 2172-2202. (Cited on pages 6, 12, 20 and 62.)

[34] Gobet, E., Lemor, J.P, Warin X.(2006) : Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli* 12 , no. 5, 889-916. (Cited on pages 1, 6, 12, 20, 44, 49, 62, 79, 80 and 87.)

[35] Gobet, E., Turkedjiev, P.(2013) : Linear Regression MDP Sheme For Discrete Backward Stochastic Differential Equation. *Mathematics Of Computation.* (Cited on pages 1, 12, 13, 87, 93, 97, 99 and 100.)

[36] Guyader, A., Hengartner, N., Matzner-Lober,E., *Simulation and Estimation of Extreme Quantiles and Extreme Probabilities*, Applied Mathematics & Optimization 64(2), 171-196, 2011. (Cited on pages 1, 13, 14, 15, 16, 17, 101, 102, 103, 104, 124 and 125.)

[37] Gyongy I., Krylov N. (2010) : Accelerated finite difference schemes for linear stochastic partial differential equations in the whole space. *SIAM J.Math.Anal.*,42, 2275-2296. (Cited on pages 6, 9, 19, 20, 21, 39, 42 and 43.)

[38] Gyongy I., Nualart D. (1995) : Implicit scheme for quasi-linear parabolic partial differential equations perturbed by space-time white noise. *Stochastic Processes and Their applications* 58(1995), 57-72. (Cited on pages 6, 20 and 39.)

[39] Gyongy I. (1995) : Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by a space-time white noise I. *Potential Anal.* 11 (1999), 1-37. (Cited on pages 6, 20 and 39.)

[40] Hammersley, J. and Handscomb, D., *Monte Carlo methods*, Methuen, London, 1965. (Cited on page 120.)

[41] Hastings, W. K., *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika 57, 97-109, 1970. (Cited on pages 14, 102 and 104.)

[42] Hoogenboom, E., *Zero-variance Monte Carlo schemes revisited*, Nucl. Sci. Eng. 160, 1-22, 2008. (Cited on pages 14 and 101.)

[43] Kahn, H., *Applications of Monte Carlo*, RM-1237-AEC, Rand Corporation, 1956. (Cited on pages 14 and 101.)

[44] Jentzen A., Kloeden P. (2010) : Taylor Expansions of solutions of stochastic partial differential equation with additive noise. *The Annals of Probability*, 2010. Vol.38, No.2,532-569. (Cited on pages 7, 19, 20 and 39.)

[45] Kloeden P.E., Platen,E. (1992) : Numerical Solution of Stochastic Differential Equations. Springer, Berlin. (Cited on page 28.)

[46] Kunita H. : Stochastic differential equations and stochastic flows of diffeomorphisms. *Ecole d'étd́e Probabilitd́e Saint-Flour XII -1982* (PL Hennequin, ed), Lecture Notes in Mathematics, vol.1097, Springer-Verlag, Berlin Heidelberg New York Tokyo, 1984, 143-303. (Cited on page 23.)

[47] Kunita, H. (1994) : Stochastic Flows Acting on Schwartz Distributions. *Journal of Theoretical Probability*, **7**, 247-278. (Cited on page 19.)

[48] Kunita, H. (1994) : Generalized Solutions of a Stochastic Partial Differential Equation. *Journal of Theoretical Probability*, **7**, 279-308. (Cited on page 19.)

[49] Kunita, H. (1990) : Stochastic Flows and Stochastic Differential Equations. *Cambridge University Press*. (Cited on pages 19 and 23.)

[50] Los Alamos National Laboratory, *Monte Carlo Code Group*, https ://mcnp.lanl.gov/. (Cited on pages 14 and 101.)

[51] Lototsky S., Mikulevicius R., Rozovskii B. L. (1997) : Nonlinear filtering revisited : a spectral approach. *SIAM J. Control Optim.*, 35 (2) : 435-461. (Cited on pages 7 and 20.)

[52] Lototsky S. (2004) : Optimal filtering of stochastic parabolic equations. *in recent developments stochastic analysis and related topics*, pages 330-353. World Sci. Publ., Hackensack, NJ. (Cited on page 20.)

[53] Matoussi A. and Scheutzow, M. (2002) : Semilinear Stochastic PDE's with nonlinear noise and Backward Doubly SDE's. *J. of Theoret. Probab.* **15**, 1-39. (Cited on pages 19 and 39.)

[54] Matoussi A., Stoica L. : The Obstacle Problem for Quasilinear Stochastic PDE's. *Annals of Probability*, Vol. 38, N.3, 1143-1179 (2010). (Cited on pages 1, 11, 62, 63 and 66.)

[55] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E., *Equations of state calculations by fast computing machines*, J. Chem. Phys. 21, 1087-32, 1953. (Cited on pages 14, 102 and 104.)

[56] Meyn, S.P. & Tweedie, R. L., *Markov chains and stochastic stability*, Cambridge university press, 2009. (Cited on pages 14 and 105.)

[57] Milstein G.N. and Tretyakov M.V.(2009) : Solving Parabolic Stochastic Partial Differential Equations via Averaging Over Characteristics. *Mathematics of computation* **78**, 2075-2106. (Cited on pages 7 and 20.)

[58] Nualart D. (2006) : The Malliavin calculus and related topics. *Springer-Verlag Berlin Heidelberg*. (Cited on page 33.)

[59] Nualart, D. and E. Pardoux (1988) : Stochastic calculus with anticipating integrands. *Prob. Theory Relat. Fields* **78**, 535-581. (Cited on pages 9, 21 and 23.)

[60] Nummelin, E., *General irreducible Markov chains and non-negative operators*, Cambridge University Press, 1984. (Cited on pages 105 and 107.)

[61] Pardoux, E., Peng, S (1990). Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, 14 : 55–61.

[62] Pardoux E., Peng S. (1992) : Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. *Lect. Notes Inf. Sci.*, vol. 176, 200-217. (Cited on pages 1 and 2.)

[63] Pardoux E., Peng S. (1994) : Backward doubly stochastic differential equations and systems of quasilinear SPDE's. *Probab. Theory Relat. Fields* 98, 209-227. (Cited on pages 33 and 34.)

[64] Pardoux E., Pradeilles F., Rao Z. (1997) : Probabilistic interpretation of a system of semi-linear parabolic partial differential equations. *Annales de l'I.H.P.* 33(4), 467-490. (Cited on pages 3, 4, 9, 19, 21, 23, 26, 33, 34, 37, 43, 52, 61 and 68.)

[65] Picard J. (1984) : Approximation of nonlinear filtering problems and order of convergence. Lecture Notes in *Control and Inf. Sci.* 61, Springer (1984) 219-236.. (Cited on page 33.)

[66] Prevot, C., Rochner, M. (2007). A concise course on stochastic partial differential equations. *Lecture Notes in Mathematics* 1905, Springer, Berlin. (Cited on page 20.)

[67] TIERNEY, L., *Markov chains for exploring posterior distributions*, The annals of statistics 22, 1702-1762, 1994. (Cited on page 19.)
(Cited on pages 104, 105 and 108.)

[68] Stoica, L. : A probabilistic interpretation of the divergence and BSDE's. *Stochastic Process and their Applications* **103**, 31-55 (2003). 1947959 (Cited on pages 10, 61, 63, 64 and 66.)

[69] Walsh J. B. (2005) : Finite Elements Methods for Parabolic Stochastic PDE's. *Potential Analysis* (2005) 23, 1-43. (Cited on pages 7, 20 and 39.)

[70] Zhang J. (2004) : A numerical scheme for BSDE's. *The Annals of Applied Probability*, Vol. 14 ; No. 1, 459-488 (Cited on pages 1, 6, 11, 20, 36, 62 and 76.)

[71] ZOIA, A., & AL, *Branching exponential flights : traveled lengths and collision statistics*, J. Phys. A : Math. Theor. 45, 425002, 2012. (Cited on pages 14 and 101.)

[72] ZOIA, A., & AL, *Collision densities and mean residence times for d-dimensional exponential flights*, Physical Review E83, 041137, 2011. (Cited on pages 14 and 101.)

Thèse de Doctorat

Achref BACHOUCH

Numerical computations for Backward Doubly Stochastic Differential Equations and Non-linear Stochastic PDEs.

Résumé

L'objectif de cette thèse est l'étude d'un schéma numérique pour les équations différentielles doublement stochastiques rétrogrades (EDDSR). On propose une méthode numérique qui nous permet d'attaquer une large gamme d'équations aux dérivées partielles stochastiques (EDPS) non-linéaires. Ceci est possible par le biais de leur représentation probabiliste en termes d'EDDSRs. Dans la dernière partie, nous étudions une nouvelle méthode des particules dans le cadre des études de protection en neutroniques.

Dans le chapitre 2, on propose un schéma numérique pour les EDDSRs. Nous étudions l'erreur de discrétisation en temps et nous donnons la vitesse de convergence associée. Ensuite, nous déduisons un schéma numérique pour les solutions faibles d'EDPS semilinéaires et nous déduisons la vitesse de convergence en temps.

Dans le chapitre 3, nous étendons notre méthode aux équations différentielles doublement stochastiques rétrogrades généralisées (EDDSRG). Nous étudions l'erreur de discrétisation en temps et donnons la vitesse de convergence associée. Nous déduisons un schéma numérique pour les EDPS quasilinéaires associées aux EDDSRG, en donnant la vitesse de convergence.

Dans le chapitre 4, on propose un schéma pour l'approximation par projections et simulations de Monte-Carlo des EDDSRs discrètes. On étudie l'erreur de régression dans un cas particulier mais très instructif. On procède à une analyse conditionnelle de l'erreur sachant les trajectoires de ce bruit extérieur. On obtient des bornes supérieures presque sûres non asymptotiques mais explicites de l'erreur de régression conditionnelle, qui assurent la convergence de notre schéma.

Dans le chapitre 5, on étudie un problème d'estimation de probabilités faibles dans le cadre des études de protection en neutroniques. On adapte une méthode récente d'estimation de faibles probabilités par un système de particules en interaction, se basant sur l'algorithme de Hastings-Metropolis et qui est proposée initialement pour les variables aléatoires, au cas des chaînes de Markov. On montre la convergence de notre algorithme. L'implémentation de la méthode est donnée en détails dans le cas unidimensionnel ainsi que dans le cas bidimensionnel.

Mots clés: Equations différentielles doublement stochastiques rétrogrades, Equations aux dérivées partielles stochastiques nonlinéaires, Projections, Simulations de Monte-Carlo, régression , Système de particules en intéraction , Algorithme de Hastings-Metropolis, Chaînes dee Markov.

Abstract

The purpose of this thesis is to study a numerical method for backward doubly stochastic differential equations (BDSDEs in short). Our numerical method allows us to tackle a large class of nonlinear stochastic partial differential equations (SPDEs in short), thanks to their probabilistic interpretation in terms of BDSDEs. In the last part, we study a new particle method in the context of shielding studies.

In chapter 2, we propose a numerical scheme for BDSDEs. We study the error arising from the time discretization and we give the associated rate of convergence. Then, we deduce a numerical scheme for the weak solutions of the associated semilinear SPDEs and we deduce the rate of convergence in time.

In chapter 3, we extend our method to generalized backward doubly stochastic differential equations (GBDSDEs in short). We study the time discretization error and we give the associated rate of convergence. Then, we deduce a numerical scheme for the quasilinear SPDEs associated to the GBDSDEs and we deduce the rate of convergence.

In chapter 4, we propose a scheme based on projections and Monte-Carlo simulations to approximate solutions of discrete BDSDEs. We study the regression error in a particular but very instructive case. We proceed to a conditional analysis of the error given the trajectories of the exterior noise. We obtain non asymptotic but explicit almost sure upper bounds for the regression conditional error.

In chapter 5, we study a problem of small probability estimation in the context of shielding studies in neutron transport. We adapt a recent interacting particle method for small probabilities estimation, based on Hastings-Metropolis algorithm and given initially for random variables, to the case of Markov chains. We show the convergence of our algorithm. Then, the practical implementation is given in details in the one and two-dimensional cases.

Key Words: Backward Doubly Stochastic Differential Equations, Semilinear Stochastic PDEs, Generalized Backward Doubly Stochastic Differential Equations, Quasilinear Stochastic PDEs, Projections, Monte-Carlo simulations, regression, Interacting particle systems, Hastings-Metropolis algorithm, Markov chains.

L'Université Nantes Angers Le Mans