



HAL
open science

Quelques contributions aux schémas numériques pour les équations différentielles stochastiques rétrogrades

Celine Labart

► **To cite this version:**

Celine Labart. Quelques contributions aux schémas numériques pour les équations différentielles stochastiques rétrogrades. Mathématiques [math]. Université Savoie Mont Blanc, 2024. tel-04678206

HAL Id: tel-04678206

<https://theses.hal.science/tel-04678206>

Submitted on 26 Aug 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Université Savoie Mont Blanc

Habilitation à diriger des recherches

présentée par

Céline Labart

Spécialité: Informatique et Mathématiques appliquées

**Some contributions to numerical schemes for
Backward Stochastic Differential Equations**

Soutenue le 26 juin 2024 devant le jury composé de

M. Christian BENDER	Rapporteur
Mme Mireille BOSSY	Rapporteure
M. Philippe BRIAND	Examinateur
M. Jean-François CHASSAGNEUX	Examinateur
Mme Adeline LECLERCQ-SAMSON	Examinatrice
Mme Agnès SULEM	Examinatrice
M. Nizar TOUZI	Rapporteur

A Pierre.

Remerciements

Mes premiers remerciements vont à Christian Bender, Mireille Bossy et Nizar Touzi pour avoir rapporté mon manuscrit. Merci à eux pour le temps et l'énergie qu'ils y ont consacrés. Je remercie aussi chaleureusement Philippe Briand, Jean-François Chassagneux, Adeline Leclercq-Samson et Agnès Sulem d'avoir fait partie de mon jury.

Je remercie aussi mes co-auteurs, et plus particulièrement Philippe Briand, Hannah et Stefan Geiss, pour tous ces échanges fructueux qui nous ont permis de publier des articles ensemble. J'espère que l'avenir nous permettra de nous retrouver encore face à un problème qui nous tient à coeur.

La rédaction de ce manuscrit est en grande partie dûe à Anne Siegel, DAS à l'INS2I, qui a su trouver les mots lors de l'une de ses interventions en tant que chargée de mission parité. Je lui en suis très reconnaissante.

Ce travail est l'aboutissement de nombreuses années de travail passées au Laboratoire de Mathématiques de l'université Savoie Mont Blanc. Je tiens à remercier tous ses membres qui contribuent à en faire un cadre de travail agréable. Merci à tous mes collègues avec qui je partage l'amour de la montagne, de la cuisine, et le plaisir d'enseigner. Merci à l'équipe administrative pour leur soutien dans l'organisation de manifestations scientifiques, leur accompagnement au quotidien dans la gestion du master et leur bonne humeur.

Pour conclure, merci infiniment à Jérôme et Thomas qui ont tout fait pour me changer les idées ces quatre derniers mois (je me serai tout de même passée de l'épisode coqueluche!). Merci aussi à toi Pierre pour tous ces kilomètres parcourus ensemble, pour ton soutien et ta gentillesse. Tu n'as pas fini de me manquer.

Contents

Publications	8
1 Introduction	12
1.1 Preliminaries	12
1.1.1 Backward stochastic differential equations	13
1.1.2 Markovian BSDEs	14
1.1.3 Applications of BSDE	15
1.2 Numerical simulation of BSDEs	16
1.2.1 Random walk approximation	16
1.2.2 First backward Euler method and computation of the conditional expectations	17
1.2.3 Other numerical methods to solve standard BSDEs	19
1.2.4 The case of non standard BSDEs	20
1.2.5 Contributions of this manuscript	21
1.3 Plan	22
2 Random walk	23
2.1 Strong error approximation	23
2.1.1 Introduction	23
2.1.2 Skorokhod embedding	24
2.1.3 Some properties of solutions to BSDEs and associated PDEs	25
2.1.4 Case of Markovian BSDE with Hölder terminal condition	28
2.1.5 Case of a decoupled FBSDE	30
2.2 Donsker type result	36
2.2.1 Introduction	36
2.2.2 Notations	37
2.2.3 Scaled random walk and Wasserstein distance	41
2.2.4 Regularity results on $u, U^n, \nabla u$ and Δ^n	43
2.2.5 Main results	44
2.3 Case of Doubly reflected BSDEs with RCLL obstacles	46
2.3.1 General framework	48
2.3.2 Discrete time Approximation	50
2.3.3 Penalized scheme	52
2.3.4 Reflected scheme	56
3 Monte Carlo method and Wiener Chaos expansion	59
3.1 Introduction	59
3.1.1 Definitions and Notations	60

3.2	Standard case	61
3.2.1	Wiener Chaos Expansion	62
3.2.2	Chaos decomposition approximations	64
3.2.3	Approximation steps of the BSDE	66
3.2.4	Convergence results	68
3.2.5	Numerical Examples	70
3.3	Case with jumps	73
3.3.1	Preliminaries	75
3.3.2	Wiener chaos expansion	75
3.3.3	Chaos decomposition approximations	78
3.3.4	Approximation steps of the BSDE	80
3.3.5	Convergence results	82
3.3.6	Numerical examples	83
3.4	McKean-Vlasov Case	84
3.4.1	Preliminaries	86
3.4.2	Approximation steps of the BSDE	86
3.4.3	Convergence results	90
3.4.4	Numerical examples	93

Publications

Published papers

- [L-1] Ph. Briand, Ch. Geiss, S. Geiss and C. Labart. Donsker-Type Theorem for BSDEs: Rate of Convergence. *Bernoulli*, 27(2), 2021, <https://hal.archives-ouvertes.fr/hal-02165750>.
- [L-21] A. Abdalazeez, I. Didenkulova, D. Dutykh and C. Labart. Extreme Inundation Statistics on a Composite Beach. *Water*, 12(6), 2020, https://www.researchgate.net/publication/342408944_Extreme_Inundation_Statistics_on_a_Composite_Beach, submitted.
- [L-3] Ch. Geiss, C. Labart and A. Luoto. Mean square rate of convergence for random walk approximation of forward-backward SDEs. *Advances in Applied Probability*, 52(3), 2020, <https://hal.archives-ouvertes.fr/hal-01838449>.
- [L-4] Ph. Briand, A. Ghannoum and C. Labart. Mean Reflected Stochastic Differential Equations with Jumps. *Advances in Applied Probability*, 52(2):523-562, 2020, <https://hal.archives-ouvertes.fr/hal-01742164>.
- [L-5] Ph. Briand, P-E. Chaudru de Raynal, A. Guillin and C. Labart. Particle systems and Numerical Schemes for Mean Reflected Stochastic Differential Equations. *Annals of Applied Probability*, 30(4):1884-1909, 2020, <https://arxiv.org/abs/1612.06886v3>.
- [L-6] Ch. Geiss, C. Labart and A. Luoto. Random walk approximation of BSDEs with Holder continuous terminal condition. *Bernoulli*, 26(1):159–190, 2020, <https://hal.archives-ouvertes.fr/hal-01818668>.
- [L-7] Ch. Geiss and C. Labart. Simulation of BSDEs with jumps by Wiener Chaos Expansion. *Stochastic Processes and their Applications*, 126:2123-2162, 2016, <https://hal.archives-ouvertes.fr/hal-01118172>.
- [L-8] R. Dumitrescu and C. Labart. Numerical approximation of doubly reflected BSDEs with jumps and RCLL obstacles. *Journal of Mathematical Analysis and Applications*, 442(1):206–243; 2016, <http://hal.archives-ouvertes.fr/hal-01006131>.
- [L-9] R. Dumitrescu and C. Labart. Reflected scheme for doubly reflected BSDEs with jumps and RCLL obstacles. *Journal of computational and Applied Mathematics*, 296:827–839, 2016, <https://hal.archives-ouvertes.fr/hal-01114996>.

- [L-10] Ph. Briand and C. Labart. Simulation of BSDEs by Wiener Chaos Expansion, *Annals of Applied Probability*, 24(3):1129–1171, 2014, <http://hal.archives-ouvertes.fr/hal-00688523>.
- [L-11] A. Alfonsi, C. Labart, and J. Lelong. Stochastic local intensity loss models with interacting particle systems. *Math. Finance*, 26(2):366–394, <http://hal.archives-ouvertes.fr/hal-00786239>, 2016.
- [L-12] C. Labart and J. Lelong. A Parallel Algorithm for solving BSDEs. *Monte Carlo Methods Appl.*, 9(1), 2013, <http://hal.archives-ouvertes.fr/hal-00680652>.
- [L-13] D. Dutykh, C. Labart, and D. Mitsotakis. Long wave runup on random beaches. *Physical Review Letter*, 107(13), 2011, <http://hal.archives-ouvertes.fr/hal-00606627>.
- [L-14] E. Gobet and C. Labart. Solving BSDE with adaptive control variate. *SIAM Journal on Numerical Analysis*, 48(1):257–277, 2010, <http://hal.archives-ouvertes.fr/hal-00373350>.
- [L-15] C. Labart and J. Lelong. Pricing Double Parisian options using Laplace transforms. *Int. Journal of Theo. and Appl. Finance*, 12(1):19–44, 2009.
- [L-16] C. Labart and J. Lelong. Pricing Parisian options using Laplace transforms. *Bankers, Markets Investors*, 99:29–43, March-April 2009.
- [L-17] E. Gobet and C. Labart. Sharp estimates for the convergence of the density of the Euler scheme in small time. *Elect. Comm. in Probab.*, 1:352–363, 2008, <http://hal.archives-ouvertes.fr/hal-00281365>.
- [L-18] E. Gobet and C. Labart. Error expansion for the discretization of Backward Stochastic Differential Equations. *Stochastic Processes and their Applications*, 117:803–829, 2007, <http://hal.archives-ouvertes.fr/hal-00019463>.

Book chapters, conference proceedings

- [L-19] J-O. Lachaud, D. Coeurjolly, C. Labart, P. Romon and B. Thibert. *Lightweight Curvature Estimation on Point Clouds with Randomized Corrected Curvature Measures*. Computer Graphics Forum (Proceedings of Symposium on Geometry Processing), 2023.
- [L-20] C. Labart. *Parisian*. Encyclopedia of Quantitative Finance, Wiley, 2010.
- [L-21] A. Abdalazeez, I. Didenkulova, D. Dutykh and C. Labart. Run-up of narrow and wide-banded irregular waves on a beach. In *European Geosciences Union General Assembly 2020*, Vienna, Austria, 2020, https://www.researchgate.net/publication/338949437_Run-up_of_narrow_and_wide-banded_irregular_waves_on_a_beach.
- [L-22] E. Gobet and C. Labart. A sequential Monte Carlo algorithm for solving BSDE. In *6th International Congress on Industrial Applied Mathematics (ICIAM07)*, Zurich, Switzerland, 2007, <http://hal.archives-ouvertes.fr/hal-00393602/>.

Phd Thesis

- [L-23] C. Labart. *EDSR: analyse de discrétisation et résolution par méthodes de Monte Carlo adaptatives; Perturbation de domaines pour les options américaines*. PhD thesis, Ecole Polytechnique, <http://tel.archives-ouvertes.fr/tel-00199861/fr/>, 2007.

Submitted papers

- [L-24] K. Firdova, C. Labart and A. Martel. Distance Rank Score: Unsupervised filter method for feature selection on imbalanced dataset. Technical report, <https://arxiv.org/abs/2305.19804>, submitted.
- [L-25] C. Acary-Robert, Ph. Briand, A. Ghannoum and C. Labart. Simulation of McKean-Vlasov BSDEs by Wiener Chaos Expansion. Technical report, <https://hal.archives-ouvertes.fr/hal-01976770>, submitted.
- [L-26] Ph. Briand, E. Idée and C. Labart. An asymptotical method to estimate the parameters of a deteriorating system under condition-based maintenance. <http://hal.archives-ouvertes.fr/hal-00843566>.

Co-supervised PhD theses

During the last few years, I have had the great opportunity to co-supervise two PhD theses with Philippe Briand and Laurent Vuillon.

- Abir Ghannoum. Supervised with Philippe Briand. *Mean reflected SDEs with jumps*. 2016-2019.
- Katarina Firdova. Supervised with Laurent Vuillon. CIFRE contract with Optimistik. *Time series anomaly detection using machine learning*. 2020-2023.

Chapter 1

Introduction

This document presents a survey of my research work on numerical schemes for backward stochastic differential equations (BSDEs) since I defended my PhD in 2007. This thesis was supervised by Emmanuel Gobet at CMAP (Ecole Polytechnique) and dealt with the study of some numerical methods for BSDEs and American options. Initially recruited at LPSM (Sorbonne University) and then at LAMA (Savoie Mont Blanc University), I have continued to work on numerical methods for BSDEs. I have also worked on numerical methods for mean reflected SDEs. It is within this context that I co-supervised a PhD thesis with Philippe Briand on Mean reflected SDEs with jumps from 2016 to 2019. I also had the opportunity to co-supervise a PhD thesis on times series anomaly detection using machine learning from 2020 to 2023. In order to present a coherent manuscript, despite having published around twenty articles, I have chosen to focus on my articles related to numerical methods for BSDEs. As such, it is centered around 8 specific articles written with several co-authors. Before presenting my contributions, I briefly introduce BSDEs and present a survey on the different numerical methods introduced so far to approximate solutions to BSDEs.

1.1 Preliminaries

The aim of this section is to briefly review the basic theory of BSDEs (Section 1.1.1) and forward BSDEs (Section 1.1.2). In particular, we give the standard sufficient conditions that ensure existence and uniqueness of their solutions. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a given filtered probability space which satisfies the usual conditions of completeness and right continuity. On this space, we define a d -dimensional standard Brownian motion $(W_t)_{t \in [0, T]}$ whose natural filtration, augmented by the class of \mathbb{P} -null sets of \mathcal{F}_T , is $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$. T denotes a strictly positive constant which represents a fixed terminal time. We define the following probability spaces and sets.

- \mathcal{P} : the σ -field of predictable sets in $\Omega \times [0, T]$
- $L_p(\mathcal{F}_T) := L_p(\Omega, \mathcal{F}_T, \mathbb{P})$, $p \in \mathbb{N}^*$, the space of all \mathcal{F}_T -measurable random variables (r.v. in the following) satisfying $\|X\|_p^p := \mathbb{E}(|X|^p) < \infty$.
- $\mathbb{E}_t(X)$ denotes $\mathbb{E}(X|\mathcal{F}_t)$ for any X in $L_1(\mathcal{F}_T)$.
- $S_T^p(\mathbb{R}^d)$, $p \in \mathbb{N}$, $p \geq 2$, the space of all càdlàg predictable processes $\phi : \Omega \times [0, T] \mapsto \mathbb{R}^d$ such that $\|\phi\|_{S_T^p}^p = \mathbb{E}(\sup_{t \in [0, T]} |\phi_t|^p) < \infty$.
- $H_T^p(\mathbb{R}^d)$, $p \in \mathbb{N}$, $p \geq 2$, the space of all predictable processes $\phi : \Omega \times [0, T] \mapsto \mathbb{R}^d$ such that $\|\phi\|_{H_T^p}^p = \mathbb{E} \int_0^T |\phi_t|^p dt < \infty$.

- $L_2([0, T]; \mathbb{R}^d)$, the space of all square integrable functions on $[0, T]$ taking values in \mathbb{R}^d .
- $C^{k,l}$, the set of continuously differentiable functions $\phi : (t, x) \in [0, T] \times \mathbb{R}^d$ with continuous derivatives w.r.t. t (resp. w.r.t. x) up to order k (resp. up to order l).
- $C_b^{k,l}$, the set of continuously differentiable bounded functions $\phi : (t, x) \in [0, T] \times \mathbb{R}^d$ with continuous and uniformly bounded derivatives w.r.t. t (resp. w.r.t. x) up to order k (resp. up to order l).
- $C^{k+\alpha}$, $\alpha \in]0, 1]$ is the set of C^k functions whose k -th derivative is Hölder continuous of order α .
- $C_b^{k+\alpha}$, $\alpha \in]0, 1]$ is the set of C^k bounded functions whose k -th derivative is Hölder continuous of order α with bounded derivatives.
- \mathcal{W}_p , $p \geq 1$, stands for the L_p -Wasserstein distance: let X and X' be two r.v. and let us denote by μ the law of X and by μ' the law of X' . With the usual abuse of notation, the Wasserstein distance is defined by

$$\mathcal{W}_p(\mu, \mu') = \mathcal{W}_p(X, X') := \inf\{\mathbb{E}(|Y - Y'|^p)^{\frac{1}{p}} : \text{law}(Y) = \mu, \text{law}(Y') = \mu'\}.$$

The infimum is taken over all the couplings μ and μ' . Notice that if X and X' are random variables of order p with values in \mathbb{R}^d , then by definition we have

$$\mathcal{W}_p(X, X') \leq \left[\mathbb{E}|X - X'|^p \right]^{1/p}. \quad (1.1)$$

- $\mathcal{P}_2(\mathbb{R}^d)$ is the set of probability measures with a finite second-order moment, endowed with the L_2 -Wasserstein distance.

1.1.1 Backward stochastic differential equations

Backward Stochastic Differential Equations (BSDEs) were first introduced by Jean-Michel Bismut in 1973 ([Bis73]). They were generalized by Etienne Pardoux and Shige Peng in 1990 ([PP90]) to an equation of the following form

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \quad (1.2)$$

where $(Y, Z)_{t \in [0, T]}$ takes values in $\mathbb{R}^m \times \mathbb{R}^{m \times d}$. The function $f : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m$ is called the generator or driver. The r.v. ζ is a \mathcal{F}_T -measurable random variable taking values in \mathbb{R}^m . The pair (f, ζ) are called the parameters of the BSDE. Solving such an equation consists in finding a pair of processes $(Y_t, Z_t)_{0 \leq t \leq T}$ adapted to the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ and satisfying (1.2).

Definition 1.1.1 A solution to the BSDE (1.2) is a pair of processes (Y, Z) taking values in $\mathbb{R}^m \times \mathbb{R}^{m \times d}$ such that Y is an adapted and continuous process, Z is a predictable process such that $\int_0^T Z_s^2 ds < +\infty$ \mathbb{P} -a.s. and (Y, Z) satisfy

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T.$$

The following theorem identifies the standard sufficient conditions imposed to ensure the existence and uniqueness of a solution to the BSDE (1.2) in $S_T^2(\mathbb{R}^m) \times H_T^2(\mathbb{R}^{m \times d})$.

Hypothesis 1.1 We assume

- $f : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{m \times d})$ -measurable
- the generator f is uniformly Lipschitz continuous in (y, z) : there exists a constant L_f such that for all $y_1, y_2 \in \mathbb{R}^m$ and $z_1, z_2 \in \mathbb{R}^{m \times d}$

$$\|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)\| \leq L_f (\|y_1 - y_2\| + \|z_1 - z_2\|) dt \otimes d\mathbb{P} - a.e.$$

- the terminal condition ζ and the generator f satisfy $\mathbb{E}(|\zeta|^2 + \int_0^T |f(s, 0, 0)|^2 ds) < +\infty$.

Theorem 1.1.2 ([PP90]) Under Hypothesis 1.1, there exists a unique solution $(Y, Z) \in \mathcal{S}_T^2(\mathbb{R}^m) \times \mathcal{H}_T^2(\mathbb{R}^{m \times d})$ which solves (1.2).

The proof of this result is based on Picard's iterations. A fixed point argument can also be used to prove the result (see [EKKP⁺97]).

Before introducing Markovian BSDEs, we present some intuition regarding the solution (Y, Z) to (1.2). If we consider the simple case $f = 0$ in (1.2) we obtain the following BSDE

$$\begin{cases} dY_t = Z_t dW_t, & t \in [0, T] \\ Y_T = \zeta \end{cases} \quad (1.3)$$

If $\zeta \in L_2(\mathcal{F}_T)$, f and ζ satisfy Hypothesis 1.1 then Theorem 1.1.2 ensures that (1.3) admits a unique solution in $\mathcal{S}_T^2(\mathbb{R}^m) \times \mathcal{H}_T^2(\mathbb{R}^{m \times d})$. Then Y satisfies

$$Y_t = \mathbb{E}(Y_T | \mathcal{F}_t) = \mathbb{E}(\zeta | \mathcal{F}_t)$$

and thus Y is an $(\mathcal{F}_t)_{t \in [0, T]}$ -martingale. From (1.3) we can also write

$$Y_t = \mathbb{E}(\zeta | \mathcal{F}_t) = Y_0 + \int_0^t Z_s dW_s = \mathbb{E}(\zeta | \mathcal{F}_0) + \int_0^t Z_s dW_s = \mathbb{E}(\zeta) + \int_0^t Z_s dW_s.$$

Since Y is a martingale, the martingale representation theorem ensures that there exists a unique process $\Gamma \in \mathcal{H}_T^2(\mathbb{R}^{m \times d})$ satisfying $Y_t = Y_0 + \int_0^t \Gamma_s dW_s$ for all $t \in [0, T]$. Hence, we obtain

$$\zeta = Y_T = Y_0 + \int_0^T \Gamma_s dW_s = Y_0 + \int_0^t \Gamma_s dW_s + \int_t^T \Gamma_s dW_s = Y_t + \int_t^T \Gamma_s dW_s,$$

which shows that (Y, Γ) solves (1.3). By the uniqueness of the solution to the BSDE we get $Z = \Gamma$. We see in this simple case that an adapted solution to the BSDE can only be a pair of processes (Y, Z) such that Z is the component needed to ensure that Y is adapted.

1.1.2 Markovian BSDEs

Markovian BSDEs (called FBSDEs in the following) are made of a BSDE coupled with a forward Stochastic Differential Equation (SDE) in the sense that the driver may depend on the solution to the SDE and the terminal condition of the BSDE is a function of the terminal value of the SDE:

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \quad (1.4)$$

where $f : [0, T] \times \mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^q \rightarrow \mathbb{R}^m$ are deterministic functions, $(Y, Z)_{t \in [0, T]}$ takes values in $\mathbb{R}^m \times \mathbb{R}^{m \times d}$ and X is the \mathbb{R}^q -valued diffusion process solving the following standard SDE

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) \cdot dW_s, \quad 0 \leq t \leq T, \quad (1.5)$$

where $b : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ and $\sigma : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^{q \times d}$ are given functions and $x \in \mathbb{R}^q$. We recall that under the following hypothesis, there exists a unique strong solution to the SDE.

Hypothesis 1.2 We assume

- b and σ are Lipschitz functions in space
- there exists a constant $M > 0$ such that $\|b(t, 0)\| + \|\sigma(t, 0)\| \leq M$ for all $t \in [0, T]$.

FBSDEs widely appear in the literature, especially in regards to numerical methods, since FBSDEs can be linked to partial differential equations (PDE). This equivalence is exploited in many numerical methods. In many cases the dimension of Y is set to one (i.e. $m = 1$), particularly when dealing with deterministic PDEs. Let us consider the following parabolic PDE

$$\begin{cases} \partial_t u(t, x) + \mathcal{L}u(t, x) + f(t, x, u(t, x), (\partial_x u \sigma)(t, x)) = 0, \\ u(T, x) = g(x), \end{cases} \quad (1.6)$$

where \mathcal{L} is defined by

$$\mathcal{L}_{(t,x)} u(t, x) = \frac{1}{2} \sum_{i,j} [\sigma \sigma^*]_{ij}(t, x) \partial_{x_i x_j}^2 u(t, x) + \sum_i b_i(t, x) \partial_{x_i} u(t, x). \quad (1.7)$$

If this PDE has a sufficiently smooth solution u , Pardoux and Peng proved in [PP92, Theorem 3.1] that, by applying Itô's formula, $(u(t, X_t), \partial_x u \sigma(t, X_t))_{0 \leq t \leq T}$ is a solution to the BSDE (1.4). This generalizes the Feynman-Kac formula to semi-linear parabolic PDEs. If $X^{t,x}$ denotes the solution to (1.5) starting from x at time t , then we have $(Y_s^{t,x}, Z_s^{t,x}) = (u(s, X_s^{t,x}), \partial_x u \sigma(s, X_s^{t,x}))$ for all s in $[t, T]$. $Y_t^{t,x}$ is deterministic and equals $u(t, x)$. Conversely, Pardoux and Peng proved in [PP92] that when b, σ, f and g are Lipschitz w.r.t. (x, y, z) uniformly in t , the FBSDE provides the unique viscosity solution to the semilinear parabolic PDE (1.6). We also refer to [EKPQ97b, Proposition 4.3 and Theorem 4.2] for a proof of these results.

1.1.3 Applications of BSDE

BSDEs find extensive applications across various domains within the social and natural sciences. Notably, the challenge of pricing and hedging a European option can be articulated in terms of a BSDE. The seminal work by El Karoui, Peng, and Quenez ([EKPQ97b]) illustrates this concept, wherein BSDEs accommodate portfolio constraints in pricing problems (see also [CK96], [BEM18], [BCS98], [CKS98]).

These equations have been instrumental in valuing numerous financial derivatives within both complete and incomplete or constrained markets, encompassing European and American options (see [EKKP⁺97], [EKPQ97a]). BSDEs can also provide necessary and sufficient conditions for optimality in optimal control problems such as utility maximization control problems with constraints and risk-sensitive control problems (see [HHI⁺14], [EKH03], [HIM05], [REK00]).

Because of the intricate nature of BSDEs, obtaining analytical solutions for these equations is seldom feasible. Consequently, resorting to numerical methods becomes imperative to approximate the exact solution. To address this challenge, various numerical techniques have been developed, and research in this domain remains highly active. In the following sections, we present a chronological survey of the different methods to solve BSDEs (Sections 1.2.1, 1.2.2). In Section 1.2.3, we present other numerical methods to solve standard BSDEs. Since our contributions also deal with non-standard BSDEs (reflected BSDEs, BSDEs with jumps, Mc-Kean Vlasov BSDEs) we dedicate Section 1.2.4 to the presentation of different existing algorithms to solve these types of BSDEs. Our overview emphasizes methods close to our contributions and is by no means exhaustive. We refer to [CKSY23] for a complete survey on numerical methods for BSDEs. Finally Section 1.2.5 presents our contributions to the numerical methods to solve standard and non-standard BSDEs.

1.2 Numerical simulation of BSDEs

1.2.1 Random walk approximation

The first results on the numerical approximation of (1.2) go back to the end of the 90's. They were dealing with the case of a generator f independent of z (see [Che97] and [CMM99]). The authors proposed a temporal and spatial discretization approach for the BSDE, a methodology that bears some resemblance to the dynamic programming equation, which was introduced a few years later (see (1.11)). The case of a generator dependent of z was first treated in [Bal97], where the author introduced a random discretization. We also refer to the work of [MPSMT02]. In [BDM01], the authors generalized the scheme proposed in [Che97] to the case of f depending on z and proved the weak convergence of their scheme. The idea was to approximate the Brownian motion by a random walk, i.e. by considering an i.i.d. Rademacher symmetric sequence $\{\xi_k\}_{1 \leq k \leq n}$ and

$$W_t^n := \sqrt{h} \sum_{k=1}^{\lceil t/h \rceil} \xi_k^n, \quad 0 \leq t \leq T$$

Then, they considered the following discretization of (1.2), where f does not depend on time

$$y_k = y_{k+1} + hf(y_k, z_k) - \sqrt{h}z_k\xi_{k+1}, \quad k = n-1, \dots, 0, \quad y_n = \zeta^n \quad (1.8)$$

and ζ^n is a square integrable random variable measurable w.r.t. \mathcal{G}_n , where $\mathcal{G}_k = \sigma(\xi_1, \dots, \xi_k)$. A solution is a process $\{y_k, z_k\}_{0 \leq k \leq n-1}$ adapted w.r.t. \mathcal{G}_k . To solve (1.8), first choose

$$z_k = \frac{1}{\sqrt{h}} \mathbb{E}(y_{k+1}\xi_{k+1} | \mathcal{G}_k)$$

and take y_k as the solution to (1.8): if f is a Lipschitz function and if h is sufficiently small, a fix point argument ensures the existence and uniqueness of the solution. Moreover y_k is \mathcal{G}_k -measurable since it is \mathcal{G}_{k+1} -measurable and orthogonal to ξ_{k+1} . We get the following fully implementable scheme, since in this case the computation of the conditional expectation can be done easily (y_{k+1} is a function of $(\xi_1, \dots, \xi_{k+1})$)

$$\begin{cases} y_n = \zeta^n \\ z_k = \frac{1}{\sqrt{h}} \mathbb{E}(y_{k+1}\xi_{k+1} | \mathcal{G}_k) \\ y_k = \mathbb{E}(y_{k+1} | \mathcal{G}_k) + hf(y_k, z_k) \end{cases} \quad (1.9)$$

(1.9) ensues from (1.8) by taking the conditional expectation w.r.t. \mathcal{G}_k .

Remark 1.2.1 The equation (1.9) is implicit in y_k . If f is a Lipschitz function and if the time step of the grid is sufficiently small, a fix point argument ensures the existence and uniqueness of the solution. We can replace y_k by y_{k+1} in f to make (1.9) explicit.

By defining the following two continuous time processes

$$Y_t^n = y_{[t/h]}, \quad Z_t^n = z_{[t/h]} \quad \forall t \in [0, T]$$

the authors prove the convergence of (Y^n, Z^n) to (Y, Z) in probability. In addition, [AKH00], [CMM99] and [CMS01] analysed the convergence of the discretized filtration. Using these results, [BDM02] proved the convergence of a sequence of solutions to a BSDE driven by a martingale to the exact solution (1.2). It was generalized to BSDEs with random terminal time in [Tol06]. [PX11] presented an algorithm based on random walk and proved its convergence in the case of a one dimensional Brownian motion. The rate of the convergence of this method was left as an open problem. To deal with this question, an approach based on the dynamic programming equation was introduced by Bouchard and Touzi in [BT04] and Zhang in [Zha04].

1.2.2 First backward Euler method and computation of the conditional expectations

1.2.2.1 Dynamic programming equation

The idea of the dynamic programming equation introduced in [BT04] and [Zha04] is the following. Instead of approximating W by a random walk, we can approximate the solution to (1.2) in a backward way by setting $Y_{t_n}^n = \zeta$ and $(Y_{t_k}^n, Z_{t_k}^n)$ is a couple of adapted processes such that for all k in $\{0, \dots, n-1\}$

$$Y_{t_k}^n = Y_{t_{k+1}}^n + hf(t_k, Y_{t_k}^n, Z_{t_k}^n) - \int_{t_k}^{t_{k+1}} \tilde{Z}_s^n dW_s, \quad \forall k \in 0, \dots, n-1 \quad (1.10)$$

where $(\tilde{Z}_s^n)_{s \in [0, T]}$ is a process ensuing from the martingale representation theorem. Since $Z_{t_k}^n$ has to be \mathcal{F}_{t_k} -measurable, we take the best approximation \mathcal{F}_{t_k} -measurable of \tilde{Z}^n in the L^2 sense. We get

$$Z_{t_k}^n = \frac{1}{h} \mathbb{E} \left(\int_{t_k}^{t_{k+1}} \tilde{Z}_s^n ds \middle| \mathcal{F}_{t_k} \right) = \frac{1}{h} \mathbb{E}(Y_{t_{k+1}}^n (W_{t_{k+1}} - W_{t_k}) \middle| \mathcal{F}_{t_k})$$

Taking conditional expectation in (1.10) leads to the following scheme

$$\begin{cases} Y_{t_n}^n = \zeta \\ Z_{t_k}^n = \frac{1}{h} \mathbb{E}(Y_{t_{k+1}}^n (W_{t_{k+1}} - W_{t_k}) \middle| \mathcal{F}_{t_k}) \\ Y_{t_k}^n = \mathbb{E}(Y_{t_{k+1}}^n \middle| \mathcal{F}_{t_k}) + hf(t_k, Y_{t_k}^n, Z_{t_k}^n). \end{cases} \quad (1.11)$$

Unlike (1.9), (1.11) is not directly implementable, the computation of the conditional expectation is the main issue. To be fully implementable, this algorithm requires to have a good approximation of its associated conditional expectation. To address this issue, Bouchard and Touzi in [BT04] and Zhang in [Zha04] dealt with the Markovian case, i.e. they assumed that $\zeta = g(X_T)$ where X is a solution to a stochastic differential equation (see (1.4) and (1.5)).

X is approximated by its Euler scheme on the regular grid $0 = t_0 < t_1 < \dots < t_n = T$ with n time steps. We set $h := \frac{T}{n}$. A time discretization of X is

$$\begin{cases} X_0^n = x, \\ X_t^n = X_{t_k}^n + b(t_k, X_{t_k}^n)(t - t_k) + \sigma(t_k, X_{t_k}^n)(W_t - W_{t_k}), \quad \forall t \in]t_k, t_{k+1}] \end{cases} \quad (1.12)$$

This scheme enables to simulate X^n easily at each time step of the grid. If b and σ satisfy Hypothesis 1.2, we get a rate of convergence in L_p for $X^n - X$:

$$\forall p \geq 1, \left(\mathbb{E} \left(\sup_{t \leq T} |X_t^n - X_t|^p \right) \right)^{\frac{1}{p}} \leq \frac{C}{\sqrt{n}}.$$

[BT04] and [Zha04] proved that the convergence in L_2 of the backward Euler scheme (1.11) is of order $\frac{1}{2}$ under standard Lipschitz assumptions and with Lipschitz terminal condition

$$\sup_{t \in [0, T]} \mathbb{E} (|Y_t - Y_t^n|^2) + \mathbb{E} \left(\int_0^T \|Z_t - Z_t^n\|^2 dt \right) \leq \frac{C}{n}.$$

The scheme has a convergence rate of 1 if the forward scheme can be simulated perfectly on the grid or if it is approximated by a higher order scheme and if the coefficients are sufficiently smooth (see [L-18]). The convergence may also depend on the regularity of the terminal condition through the L_2 -regularity of Z (see [GGG12]). The numerical stability analysis of (1.11) was studied in [CR15].

1.2.2.2 Computation of conditional expectations

In case of Markovian BSDEs, the computation of the conditional expectation can be well approximated. Various methods have been developed:

- Least-squares regression methods were first introduced in [LGW05] and [GLW06], [BS12] and [Zha17]. Additionally, we point to [GT16], where the authors devise a numerical scheme intended for solving the multi-step forward dynamic programming equation derived from discretizing FBSDEs. Their approach involved computing the resulting sequence of conditional expectations by employing empirical least-squares regressions. Furthermore, in [GLSTV16], the authors introduced an alternative algorithm based on least-squares Monte Carlo, specifically designed to facilitate extensive parallelization of computations on highly multicore processors like GPUs.
- Quantization technics were introduced in [BP03a], [BP03b] and [BPJ01] to solve multidimensional optimal stopping problems and then applied to develop discretization schemes for reflected BSDEs (see [PS18]). Some improvements can be found in [CGG23], [ENP21] and in [PS18].
- Malliavin calculus is used in two different ways to compute conditional expectations: by using integration by parts (see [BT04], [BET04], [FLLL01] and [CMT10]) or by introducing Malliavin weights in the Z representation (see [GT16] inspired by the representation given in [MZ02]).
- Cubature methods (see [CM12], [CM14], [NS19] and [CGT20]). These methods construct explicit finitely supported measures, along a partition of $[0, T]$, that approximate the law of the forward diffusion.

Remark 1.2.2 The random walk approximation can also be seen as a method to compute conditional expectations. It is however more general than the methods mentioned above since it also includes the non Markovian case.

1.2.3 Other numerical methods to solve standard BSDEs

- High order schemes generalizing (1.11) have also been proposed. First introduced in [ZCP06], the θ -scheme is proved to have a rate of convergence of order 2 when the driver does not depend on Z (see [ZWP09]). This result is generalized to the case of a driver depending on the control process by introducing more parameters (see [ZLZ12]). To achieve higher orders of convergence, a multi-step scheme was developed in [ZZJ10]. The error analysis of linear multi-step schemes was done in [Cha14] in case where the driver depends on Z . Runge-Kutta schemes for BSDEs are studied in [CC14]. We refer to [YZZ20] to get a general framework where the stability, consistency and convergence of different discretization schemes for FBSDEs are studied (backward Euler method, θ -scheme and different types of multi-step schemes).
- Forward methods solve the BSDE in a forward way and thus usually avoid the computation of conditional expectations. We list some of them

- Picard iteration methods construct an approximation of the solution to (1.2) in an iterative way: it incorporates in the driver the solution computed at step $k - 1$ to compute the solution at step k : we set $(Y^0, Z^0) = (0, 0)$ and (Y^k, Z^k) is the solution to the following BSDE

$$Y_t^k = \zeta + \int_t^T f(s, Y_s^{k-1}, Z_s^{k-1}) ds - \int_t^T Z_s^k dW_s.$$

These methods were introduced in [BD07]. This work was extended in [BM10] and in [Mos10] where the authors used a variance reduction technique based on importance sampling. In [L-14], we proposed to solve the equivalent PDE by using Picard iterations and a control variate technique. A parallel version of this algorithm was proposed in [L-12].

- Branching diffusion methods are used to solve non linear PDEs and corresponding BSDEs (see [HL12] and [HLTT14]). The concept of branching diffusion system can also be employed for constructing backward numerical methods where the driver depends on Z (see [BTW17]) or not (see [BTZW17]).
- Asymptotic expansion consists in expanding a nonlinear BSDE into a sequence of linear BSDEs (see [TY15]). The case of non smooth drivers was studied in [GP15].
- Multi-level Picard methods consist in three steps (see [EHJ20]): reformulate the PDE as a stochastic fixed point problem, approximate the unique fixed point by Picard iteration and then approximate the iterations by multilevel Monte Carlo methods. The idea was initiated in [EHJK19] in which Picard iterations and multilevel Monte-Carlo methods are used to solve high dimensional nonlinear PDEs coming from physics and finance. In [HKN22], Picard iterations for backward stochastic differential equations with globally Lipschitz continuous nonlinearity were shown to converge. An extension of multilevel Picard approximation was applied to general forward diffusion in [HN22].
- Machine learning techniques, particularly deep learning methods, have emerged as a highly dynamic and actively researched category within the realm of numerical methods for BSDEs. These methods demonstrate remarkable capabilities in tackling high-dimensional nonlinear BSDEs and their associated nonlinear PDEs. This is accomplished through the use of neural network approximations, enabling the estimation of nested high-dimensional expectations and gradients.

While the overall complexity of deep learning methods remains challenging to comprehensively analyze presently, a handful of theoretical outcomes documented in the literature illustrate that the complexity tends to increase at most polynomially concerning the PDE dimension and inversely proportional to the approximation accuracy (see [BHJK23] and [HJKN23] among others).

- The Deep BSDE method, initially proposed in [EHJ17] and subsequently discussed in [HJE18], revolves around utilizing the solution’s gradient concerning the control process Z as the policy function. This gradient is approximated using a neural network, akin to the approach employed in standard deep reinforcement learning methodologies. In the work outlined in [HL20], an error analysis was undertaken for deep BSDE methods. This analysis specifically involves deriving a posteriori estimates for coupled FBSDEs, establishing a connection between the quadratic terminal loss and the approximation error inherent in the numerical solution to the FBSDE.
- Deep backward dynamic programming was proposed in [HPW19] and improved in [GPW22]. Unlike deep BSDE methods, deep backward dynamic programming is constructed based on the backward resolution technique employing an implicit backward Euler scheme. This approach incorporates machine learning techniques to estimate both the solution and its gradient. This estimation is achieved by minimizing a loss function at each time step. Subsequently, these localized problems are recursively addressed using a stochastic gradient algorithm, progressing backward in time.
- The Deep Splitting method, detailed in [BBC⁺21], draws upon the Feynman-Kac representation of the solution for a nonlinear parabolic PDE. This method involves creating a recursive approximation of the solution derived from the Feynman-Kac formula, systematically progressing backward in time. To achieve this, the recursion is approximated using a continuous function across the support of the marginal distribution.

1.2.4 The case of non standard BSDEs

To address problems in stochastic control, finance and partial differential equations, different extensions of (1.2) have to be considered. We list some of them in the following.

- Reflected BSDEs (RBSDEs) were introduced in [EKKP⁺97]. Indeed, these equations are used in numerous practical scenarios such as optimal stopping problems, pricing of American options, stochastic optimal control, and differential games. Their solutions are influenced by a stochastic process known as the obstacle, against which they are reflected. A variety of methods have been employed to tackle and solve these equations: quantization technique (see [BP03a], [BP03b], [PS18]), random walk (see [MSMT11], [PX11]) and least-squares Monte-Carlo regression with Malliavin approach (see [BT04]). The stability analysis was done in [MZ05] and [BC08]. Other types of RBSDEs have also been studied as discrete RBSDEs, RBSDEs with multiple obstacles and RBSDEs with jumps. The use of the penalization method provides efficient numerical schemes (see [MPX08]). It has been applied to RBSDEs with oblique reflections (see [CÉK12] and [HT10]), doubly RBSDEs (see [Xu11]) and double RBSDE with jumps (see [L-9]).
- Fully-coupled BSDEs are FBSDEs where the coefficients of the forward SDE can depend on

the backward components (Y, Z) :

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s \end{cases}$$

for $t \in [0, T]$. Numerically, the first scheme was proposed in the 1990s in [MPY94]. A implementable numerical method was developed in [DMP96]. We also refer to [BZ08] (Picard iterations), multi-step schemes (see [LSZ20] and [Zfz14]) and deep learning methods (see [HL20], [JPPZ20] and [JPPZ22])

- BSDEs with generators of quadratic growth (i.e. such that $|f(t, x, y, z)| \leq C(1 + \|y\| + \|z\|^2)$) play an important role in mathematical finance. They were introduced in [Kob00]. In contrast to standard BSDEs, dealing with the quadratic driver is not trivial. In the development of numerical methods this issue is addressed by truncating the quadratic driver. Time-discretization schemes are developed in [Ric11] and in [CR16].
- BSDEs with jumps contain jumps in their backward component

$$-dY_t = f(t, Y_t, Z_t, U_t)dt - Z_t dW_t - U_t dN_t$$

with a suitable jump process $(N_t)_{t \in [0, T]}$ and a control process $(U_t)_{t \in [0, T]}$. The presence of jumps increases the complexity of numerical implementation. Discretization schemes were developed based on Malliavin calculus (see [BE08]). A numerical algorithm was developed in [Aaz13] to approximate the solution to a decoupled FBSDE driven by a pure Lévy process. An explicit prediction-correction scheme was developed for solving decoupled FBSDEs with jumps in [FYZ18]. In [LMT13], Brownian and Poissonian components were independently approximated by random walks. An asymptotic expansion was proposed in [FT19]. We also refer to [GPP22] for a method based on deep learning.

- McKean-Vlasov BSDEs are BSDEs where the laws of Y and Z are involved in the driver ($[\theta]$ is the notation for the law of a random variable θ)

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s, [Y_s], [Z_s]) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T. \quad (1.13)$$

They were introduced in [BDLP09] and [BLP09]. The study of numerical methods for McKean-Vlasov BSDEs was done by using tree method in [CCD19], first-order and Crank-Nicolson schemes in [ZM23], higher order discretization method for decoupled cases in [CdRGT15] and deep learning (see [CCSJ21], [FZ20], [GMW22] and [HHL23]).

1.2.5 Contributions of this manuscript

In this document, we explore two distinct approaches for approximating BSDEs. The initial approach relies on Random Walk Approximation (as detailed in Section 1.2.1), while the second approach involves a forward method that integrates Picard iterations (we refer to Section 1.2.3) along with Wiener chaos expansion.

Our contribution regarding random walk approximation encompasses two main aspects. The first pertains to establishing a convergence rate for the random walk approximation, achieved in collaboration with Philippe Briand, Christel Geiss, Stefan Geiss, and Antti Luoto. This convergence rate was established in L_2 for a Forward-Backward Stochastic Differential Equation (FBSDE) whose

diffusion is a Brownian motion and with a Hölder terminal condition (we refer to [L-6]). Additionally, in the case of a general FBSDE, albeit with stronger assumptions on f and g , another convergence rate was determined in [L-3]. Subsequently, a Donsker-type result was proved in [L-1] for a FBSDE whose diffusion is a Brownian motion and with a Hölder terminal condition. The second contribution focusing on random walk approximation involves Doubly Reflected Backward Stochastic Differential Equations (RBSDEs) with jumps and Right-Continuous with Left Limits (RCLL) obstacles. This work was conducted in collaboration with Roxana Dumitrescu. In [L-8] and [L-9], two distinct schemes have been proposed to approximate RBSDEs with jumps and RCLL obstacles. The first scheme is based on penalization, while the second relies on reflection. Convergence of the numerical scheme has been rigorously established for both approaches.

Our collaborative contribution, alongside Céline Acary-Robert, Philippe Briand, Christel Geiss, and Abir Ghannoum, centers on employing Picard iterations and Wiener chaos expansion to devise a novel method for approximating standard BSDEs. This forward method involves initially applying Picard iterations, followed by Wiener chaos expansion to approximate conditional expectations. Unlike most methods outlined in Section 1.2, the primary advantage of this approach is its applicability to non-Markovian BSDEs. Additionally, Wiener chaos expansion provides an approximation of Z without incurring any additional costs. The convergence of this algorithm was rigorously established in [L-10]. Subsequently, this method was extended to encompass BSDEs with jumps in [L-7] and to the realm of McKean-Vlasov BSDEs as detailed in [L-25].

1.3 Plan

This manuscript is divided into two chapters. Chapter 2 deals with random walk approximation and Chapter 3 deals with the method based on Picard iterations and Wiener chaos expansion.

Chapter 2 is divided into three sections. Section 2.1 presents the strong error results on the rate of convergence of a random walk approximation, in the case of a FBSDE driven by a Brownian motion with an Hölder terminal condition (Section 2.1.4, based on [L-6]) and in the case of a FBSDE driven by a diffusion with more regular coefficients (Section 2.1.5, based on [L-3]). Section 2.2 presents a Donsker type result in the case of a FBSDE driven by a Brownian motion with an Hölder terminal condition ([L-1]). Section 2.3 presents two schemes to approximate doubly reflected BSDEs with jumps and RCLL obstacles and prove their convergence ([L-8], [L-9]).

Chapter 3 is divided into four sections. The first one (Section 3.1) is a general introduction giving the common notations and the intuition of the method. Section 3.2 presents the algorithm and its convergence in the standard case, ensuing from [L-10]. Section 3.3 deals with the case with jumps ([L-7]) and Section 3.4 deals with the case of a McKean-Vlasov BSDE ([L-25]).

Chapter 2

Random walk

This Chapter deals with the study of a numerical scheme based on random walk to approximate BSDEs as introduced in Section 1.2.1. Section 2.1 presents the strong error results on the rate of convergence of a random walk approximation, in the case of a FBSDE driven by a Brownian motion with an Hölder terminal condition (see Section 2.1.4, based on [L-6]) and in the case of a FBSDE driven by a diffusion with more regular coefficients (see Section 2.1.5, based on [L-3]). Section 2.2 presents a Donsker type result in the case of a FBSDE driven by a Brownian motion with an Hölder terminal condition (see [L-1]). Section 2.3 presents two schemes to approximate doubly reflected BSDEs with jumps and RCLL obstacles and prove their convergence (see [L-8] and [L-9]).

2.1 Strong error approximation

2.1.1 Introduction

In this section we present results [L-6] and [L-3] written with Christel Geiss and Antti Luoto. We are concerned with the discretization of solutions to BSDEs of the following form

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (2.1)$$

where X is either equals to W (i.e. a standard Brownian motion), f is Lipschitz continuous and g is a Hölder continuous function (see Section 2.1.4) or X is a diffusion and f and g are more regular functions (see Section 2.1.5). In one of the first studies on this topic, in the case $X = W$ and the generator f may depend on z as well, Philippe Briand, Bernard Delyon and Jean Mémin in [BDM01] proposed an approximation based on Donsker's theorem. They showed that the solution (Y, Z) to the previous BSDE can be approximated by the solution (Y^n, Z^n) to the following discretized BSDE

$$Y_{t_k}^n = g(W_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_m}^n \xi_{m+1}, \quad 0 \leq k \leq n-1 \quad (2.2)$$

where W^n is the scaled random walk

$$W_t^n = \sqrt{h} \sum_{k=1}^{\lfloor t/h \rfloor} \xi_k, \quad 0 \leq t \leq T, \quad (2.3)$$

and where $\lfloor x \rfloor := \max\{r \in \mathbb{Z} : r \leq x\}$ for any real number x , $h := \frac{T}{n}$ and $(\xi_k)_{k \geq 1}$ is an i.i.d. sequence of Rademacher random variables. In this section we denote $\mathcal{G}_k := \sigma(\xi_i : 1 \leq i \leq k)$

and $\mathcal{G}_0 = \{\emptyset, \Omega\}$. The associated discrete-time random walk $(W_{t_k}^n)_{k=0}^n$ is $(\mathcal{G}_k)_{k=0}^n$ -adapted. They proved that, in the case of a terminal condition of type $G(W)$ (where $G(W)$ is only required to be a square integrable random variable), (Y^n, Z^n) converges to (Y, Z) . However, the question of the rate of convergence was left open. In Section 2.1.4 we deal with the Markovian case, meaning that $G(W) = g(W_T)$. The case $G(X) = g(X_T)$ is studied in Section 2.1.5. The question on the rate of convergence seems to be tractable, in particular due to the PDE structure behind. Indeed, in the diffusion case, (Y, Z) is related to the semilinear equation

$$\begin{cases} \partial_t u(t, x) + \frac{\sigma^2(t, x)}{2} \Delta u(t, x) + b(t, x) \nabla_x u(t, x) + f(t, x, u(t, x), \sigma(t, x) \nabla_x u(t, x)) = 0, \\ u(T, x) = g(x), \quad x \in \mathbb{R} \end{cases} \quad t \in [0, T], x \in \mathbb{R}, \quad (2.4)$$

where, under certain regularity conditions, we can choose

$$Y_s = u(s, X_s) \quad \text{and} \quad Z_s = \nabla_x u(s, X_s).$$

The convergence of this scheme was already proved in [BDM02, Proposition 13] for a general terminal condition and a generator that is Lipschitz in its spatial coordinates but without any rate of convergence. Even though the link with PDEs was pointed out in [BDM01], the rate of convergence of the approximation of BSDEs given by scaled random walks was completely open. In [L-6] we give a first answer to this question in the case $X = W$ (see Section 2.1.4). We show that the error between (Y^n, Z^n) and (Y, Z) is of order $n^{-\varepsilon/4}$ when g is assumed to be ε -Hölder continuous and $f(t, \cdot)$ is Lipschitz continuous. In [L-3] we answer to this question in the general case of a diffusion (see Section 2.1.5). Compared to the case $X = W$, we need rather strong conditions on the smoothness and boundedness on f and g . We show that the error between (Y^n, Z^n) and (Y, Z) is of order $n^{-\frac{1}{4} \wedge \frac{\alpha}{2}}$ provided that g'' is locally α -Hölder.

One of the main arguments of Section 2.1 consists in constructing the random walk from the Brownian motion W using the Skorohod embedding (see [Wal03]) together with generalizations of the pioneering work of Jin Ma and Jianfeng Zhang [MZ02] on representation theorems for BSDEs. This approach allows to work with convergence in the L_2 -sense even if the problem naturally arises in the weak sense. The drawback is that the rate of convergence $n^{-\varepsilon/4}$ obtained in Section 2.1.4 is not optimal as one can expect $n^{-\varepsilon/2}$. This will be discussed in Section 2.2. Before presenting our results, we present Skorokhod embedding in Section 2.1.2 and some properties of solutions to BSDEs and associated PDEs in Section 2.1.3.

2.1.2 Skorokhod embedding

The rate of the L_2 -convergence of (Y^n, Z^n) to (Y, Z) will rely on the fact that the random walk W^n can be constructed from the Brownian motion W by Skorokhod embedding. This enables to work on the same probability space. Let $\tau_0 := 0$ and define

$$\tau_k := \inf\{t > \tau_{k-1} : |W_t - W_{\tau_{k-1}}| = \sqrt{h}\}, \quad k \geq 1.$$

Then $(W_{\tau_k} - W_{\tau_{k-1}})_{k=1}^\infty$ is a sequence of i.i.d. random variables with

$$\mathbb{P}(W_{\tau_k} - W_{\tau_{k-1}} = \pm\sqrt{h}) = \frac{1}{2},$$

which means that $\sqrt{h}\xi_k \stackrel{d}{=} W_{\tau_k} - W_{\tau_{k-1}}$. We use this random walk for our approximation, i.e. we require

$$W_t^n = \sum_{k=1}^{\lfloor t/h \rfloor} (W_{\tau_k} - W_{\tau_{k-1}}), \quad 0 \leq t \leq T. \quad (2.5)$$

Properties satisfied by τ_k and W_{τ_k} are stated in the following Lemma. We denote by \mathbb{E}_{τ_k} the conditional expectation w.r.t. \mathcal{F}_{τ_k} .

Lemma 2.1.1 *For all $0 \leq k \leq m \leq n$ and $p > 0$, it holds for $h = \frac{T}{n}$ that*

1. $\mathbb{E}\tau_k = kh$,
2. $\mathbb{E}|\tau_1|^p \leq C(p)h^p$,
3. $\mathbb{E}|W_{\tau_m} - W_{\tau_k}|^2 = t_m - t_k$,
4. $\mathbb{E}|W_{\tau_k} - W_{t_k}|^{2p} \leq C(p)\mathbb{E}|\tau_k - t_k|^p \leq C(p)(t_k h)^{\frac{p}{2}}$.

Remark 2.1.2 Note that for $p > 0$ there exists a $C(p) > 0$ such that for all $k = 1, \dots, n$ it holds

$$\frac{1}{C(p)}(t_k h)^{\frac{1}{4}} \leq (\mathbb{E}|W_{\tau_k} - W_{t_k}|^p)^{\frac{1}{p}} \leq C(p)(t_k h)^{\frac{1}{4}}.$$

The upper estimate is given in Lemma 2.1.1. For $p \in [4, \infty)$ the lower estimate follows from [AKU21, Proposition 5.3]. We get the lower estimate for $p \in (0, 4)$ by choosing $0 < \theta < 1$ and $0 < p < p_1$ such that $\frac{1}{4} = \frac{1-\theta}{p} + \frac{\theta}{p_1}$. Then it holds by the log-convexity of L_p norms (see, for example [Tao10, Lemma 1.11.5]) that

$$\|W_{\tau_k} - W_{t_k}\|_p^{1-\theta} \geq \frac{\|W_{\tau_k} - W_{t_k}\|_4}{\|W_{\tau_k} - W_{t_k}\|_{p_1}^\theta} \geq \frac{C(4)^{-1}(t_k h)^{\frac{1}{4}}}{\left(C(p_1)(t_k h)^{\frac{1}{4}}\right)^\theta} \geq \left(C(p)(t_k h)^{\frac{1}{4}}\right)^{1-\theta}.$$

Since for $t \in [t_k, t_{k+1})$ it holds $W_t^n = W_{\tau_k}$ and $\|W_t - W_{t_k}\|_p \leq C(p)h^{\frac{1}{2}}$, we have for any $p > 0$ that

$$\sup_{0 \leq t \leq T} \|W_t^n - W_t\|_p = O(h^{\frac{1}{4}}).$$

2.1.3 Some properties of solutions to BSDEs and associated PDEs

In the following we recall and prove results for FBSDEs with a general forward process. Let us consider the following SDE started in (t, x) ,

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dW_r, \quad 0 \leq t \leq s \leq T, \quad (2.6)$$

where b and σ satisfy

Hypothesis 2.1

- i. $b, \sigma \in C_b^{0,2}([0, T] \times \mathbb{R})$, in the sense that the derivatives of order $k = 0, 1, 2$ w.r.t. the space variable are continuous and bounded on $[0, T] \times \mathbb{R}$,
- ii. the first and second derivatives of b and σ w.r.t. the space variable are assumed to be γ -Hölder continuous (for some $\gamma \in (0, 1]$, w.r.t. the parabolic metric $d((x, t), (x', t')) = (|x - x'|^2 + |t - t'|)^{\frac{1}{2}}$ on all compact subsets of $[0, T] \times \mathbb{R}$),
- iii. b, σ are $\frac{1}{2}$ -Hölder continuous in time, uniformly in space,
- iv. $\sigma(t, x) \geq \delta > 0$ for all (t, x) .

2.1.3.1 Malliavin weights

In this section we recall the Malliavin weights and some of their properties (see [GGG12, Subsection 1.1 and Remark 3]). We denote $\nabla X^{t,x} = (\nabla X_s^{t,x})_{s \in [t,T]}$ the variational process of $X^{t,x}$ i.e. it solves

$$\nabla X_s^{t,x} = 1 + \int_t^s b_x(r, X_r^{t,x}) \nabla X_r^{t,x} dr + \int_t^s \sigma_x(r, X_r^{t,x}) \nabla X_r^{t,x} dW_r, \quad (2.7)$$

with $(X_s^{t,x})_{s \in [t,T]}$ given in (2.6).

Lemma 2.1.3 *Let $H : \mathbb{R} \rightarrow \mathbb{R}$ be a polynomially bounded Borel function. If Assumption 2.1 holds and $X^{t,x}$ is given by (2.6), then setting*

$$G(t, x) := \mathbb{E}H(X_R^{t,x})$$

implies that $G \in C^{1,2}([0, R] \times \mathbb{R})$. Especially it holds for $0 \leq t \leq r < R \leq T$ that

$$\partial_x G(r, X_r^{t,x}) = \mathbb{E}[H(X_R^{t,x}) N_R^{r,1,(t,x)} | \mathcal{F}_r^t], \quad \text{and} \quad \partial_x^2 G(r, X_r^{t,x}) = \mathbb{E}[H(X_R^{t,x}) N_R^{r,2,(t,x)} | \mathcal{F}_r^t],$$

where $(\mathcal{F}_r^t)_{r \in [t,T]}$ is the augmented natural filtration of $(W_r^{t,0})_{r \in [t,T]}$,

$$N_R^{r,1,(t,x)} = \frac{1}{R-r} \int_r^R \frac{\nabla X_s^{t,x}}{\sigma(s, X_s^{t,x}) \nabla X_r^{t,x}} dW_s \quad \text{and} \quad N_R^{r,2,(t,x)} = \frac{N_R^{\rho,1,(t,x)} \nabla X_R^{t,x} N_\rho^{r,1,(t,x)} + \nabla N_R^{\rho,1,(t,x)}}{\nabla X_r^{t,x}},$$

with $\rho := \frac{r+R}{2}$. Moreover, for $q \in (0, \infty)$ it holds a.s.

$$(\mathbb{E}[|N_R^{r,i,(t,x)}|^q | \mathcal{F}_r^t])^{\frac{1}{q}} \leq \frac{\kappa_q}{(R-r)^{\frac{i}{2}}}, \quad (2.8)$$

and $\mathbb{E}[N_R^{r,i,(t,x)} | \mathcal{F}_r^t] = 0$ a.s. for $i = 1, 2$. Finally, we have

$$\|\partial_x G(r, X_r^{t,x})\|_{L_p(\mathbb{P})} \leq \kappa_q \frac{\|H(X_R^{t,x}) - \mathbb{E}[H(X_R^{t,x}) | \mathcal{F}_r^t]\|_{L_p(\mathbb{P})}}{\sqrt{R-r}}$$

and

$$\|\partial_x^2 G(r, X_r^{t,x})\|_{L_p(\mathbb{P})} \leq \kappa_q \frac{\|H(X_R^{t,x}) - \mathbb{E}[H(X_R^{t,x}) | \mathcal{F}_r^t]\|_{L_p(\mathbb{P})}}{R-r}$$

for $1 < q, p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$.

2.1.3.2 Regularity of solutions to BSDEs

Let us now consider the FBSDE

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r, \quad 0 \leq t \leq s \leq T, \quad (2.9)$$

where $X^{t,x}$ is the process satisfying (2.6) and the following hypothesis on f and g .

Hypothesis 2.2

- i.* g is locally Hölder continuous with order $\varepsilon \in (0, 1]$ and polynomially bounded ($p_0 \geq 0, C_g > 0$) in the following sense

$$\forall (x, y) \in \mathbb{R}^2, \quad |g(x) - g(y)| \leq C_g(1 + |x|^{p_0} + |y|^{p_0})|x - y|^\varepsilon. \quad (2.10)$$

ii. The function $[0, T] \times \mathbb{R}^3 : (t, x, y, z) \mapsto f(t, x, y, z)$ satisfies

$$|f(t, x, y, z) - f(t', x', y', z')| \leq L_f(\sqrt{t - t'} + |x - x'| + |y - y'| + |z - z'|). \quad (2.11)$$

Notice that (2.10) implies

$$|g(x)| \leq K(1 + |x|^{p_0+1}) =: \Psi(x), \quad x \in \mathbb{R}, \quad (2.12)$$

for some $K > 0$. From the continuity of f we conclude that

$$K_f := \sup_{0 \leq t \leq T} |f(t, 0, 0, 0)| < \infty.$$

The following result is taken from [GGG12, Theorem 1]. We reformulate it here for the simple situation where we need it. On the other hand, we will use $\mathbb{P}_{t,x}$ and are interested in an estimate for all $(t, x) \in [0, T] \times \mathbb{R}$.

Theorem 2.1.4 *Let Assumption 2.1 and 2.2 hold. Then for any $p \in [2, \infty)$ the following assertions are true.*

1. There exists a constant $C_{2.1.4}^y > 0$ such that for $0 \leq t < s < T$ and $x \in \mathbb{R}$,

$$\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \leq C_{2.1.4}^y \Psi(x) \left(\int_t^s (T - r)^{\varepsilon-1} dr \right)^{\frac{1}{2}},$$

2. There exists a constant $C_{2.1.4}^z > 0$ such that for $0 \leq t < s < T$ and $x \in \mathbb{R}$,

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \leq C_{2.1.4}^z \Psi(x) \left(\int_t^s (T - r)^{\varepsilon-2} dr \right)^{\frac{1}{2}}.$$

The constants $C_{2.1.4}^y$ and $C_{2.1.4}^z$ depend on $K_f, L_f, C_g, c_{2.1.5}^{1,2}, T, p_0, b, \sigma, \kappa_q$ and p .

2.1.3.3 Properties of the associated PDE

We collect in the theorem below properties of the solution to the PDE which are mainly known. The new part concerns Δu . For Lipschitz continuous function g , the behavior of Δu has been studied in [Zha01]. General results related to this topic can be found in [CD12].

Theorem 2.1.5 *Consider the FBSDE (2.9) and let Assumptions 2.1 and 2.2 hold. Then for the solution u to the associated PDE*

$$\begin{cases} \partial_t u(t, x) + \frac{\sigma^2(t, x)}{2} \Delta u(t, x) + b(t, x) \nabla_x u(t, x) + f(t, x, u(t, x), \sigma(t, x) \nabla_x u(t, x)) = 0, \\ u(T, x) = g(x), \quad x \in \mathbb{R} \end{cases} \quad t \in [0, T], x \in \mathbb{R},$$

we have (using the notations of Lemma 2.1.3)

1. $Y_t = u(t, X_t)$ where $u(t, x) = \mathbb{E} \left(g(X_T^{t,x}) + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr \right)$ and $|u(t, x)| \leq c_{2.1.5}^1 \Psi(x)$ with Ψ given in (2.12), where $c_{2.1.5}^1$ depends on C_g, T, p_0, L_f, K_f and on the bounds and Lipschitz constants of b and σ .

2. $\nabla_x u$ exists,

$$\nabla_x u(t, x) = \mathbb{E} \left(g(X_T^{t,x}) N_T^{t,1,(t,x)} + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) N_r^{t,1,(t,x)} dr \right), \quad (2.13)$$

and

(a) $\nabla_x u$ is continuous in $[0, T) \times \mathbb{R}$,

(b) $Z_s^{t,x} = \nabla_x u(s, X_s^{t,x}) \sigma(s, X_s^{t,x})$,

(c) $|\nabla_x u(t, x)| \leq \frac{c_{2.1.5}^2 \Psi(x)}{(T-t)^{1-\frac{\varepsilon}{2}}}$,

where $c_{2.1.5}^2$ depends on $C_g, T, p_0, \kappa_2, L_f, K_f$ and on the bounds and Lipschitz constants of b and σ .

3. u_{xx} exists,

$$\Delta u(t, x) = \mathbb{E} \left(g(X_T^{t,x}) N_T^{t,2,(t,x)} + \int_t^T [f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) - f(r, x, Y_t^{t,x}, Z_t^{t,x})] N_r^{t,2,(t,x)} dr \right), \quad (2.14)$$

and

(a) Δu is continuous in $[0, T) \times \mathbb{R}$,

(b) $|\Delta u(t, x)| \leq \frac{c_{2.1.5}^3 \Psi(x)}{(T-t)^{1-\frac{\varepsilon}{2}}}$,

where $c_{2.1.5}^3$ depends on $C_g, T, p_0, \kappa_2, L_f, C_{2.1.4}^y, C_{2.1.4}^z$ and on the bounds and Lipschitz constants of b and σ .

In the following $c_{2.1.5}$ represents $(c_{2.1.5}^1, c_{2.1.5}^2, c_{2.1.5}^3)$ and $c_{2.1.5}^{i,j}$ ($i \neq j$) represents $(c_{2.1.5}^i, c_{2.1.5}^j)$, $(i, j) \in \{1, 2, 3\}$.

2.1.4 Case of Markovian BSDE with Hölder terminal condition

2.1.4.1 Introduction

This Section presents the results established in [L-6]. We are concerned with the discretization of solutions to BSDEs of the following form

$$Y_t = g(W_T) + \int_t^T f(s, W_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (2.15)$$

where W is a standard Brownian motion, f is Lipschitz continuous and g is a Hölder continuous function. We are interested in studying the rate of convergence of (Y^n, Z^n) , defined by

$$Y_{t_k}^n = g(W_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_m}^n \xi_{m+1}, \quad 0 \leq k \leq n-1 \quad (2.16)$$

to (Y, Z) , defined by (2.15). To do this, we construct the random walk from the Brownian motion W by using the Skorohod embedding (see Section 2.1.2). In this case, $(\xi_k)_k$ and W are defined on the

same probability space and the convergence of (Y^n, Z^n) to (Y, Z) can be studied in L_2 (see Theorem 2.1.8). Basically, we get that the L_2 -norm of the error on Y is of order $n^{-\frac{\varepsilon}{4}}$ and the L_2 -norm of the error on Z is of order $\frac{n^{-\frac{\varepsilon}{4}}}{\sqrt{T-v}}$. The proof of this result is based on several ingredients. In particular, we need some estimates on the bound of the first and second derivatives of the solution to the PDE associated to the BSDE (2.15). We use the bounds established in Section 2.1.3.3.

In this section the study of the error $(Y^n - Y, Z^n - Z)$ either relies on (2.16) or on its integral version:

$$Y_t^n = g(W_T^n) + \int_{]t, T]} f(s, W_{s-}^n, Y_{s-}^n, Z_{s-}^n) d\langle W^n \rangle_s + \int_{]t, T]} Z_s^n dW_s^n, \quad 0 \leq t \leq T, \quad (2.17)$$

where the backward equation (2.17) arises from (2.16) by setting $Y_r^n := Y_{t_m}^n$ and $Z_r^n := Z_{t_m}^n$ for $r \in [t_m, t_{m+1})$.

Remark 2.1.6 It was shown in [BDM01] that, as soon as $h \|f\|_{\text{Lip}} < 1$, the BSDE (2.17) has a unique square integrable solution (Y^n, Z^n) , and $(Y_{t_m}^n, Z_{t_m}^n)_{m=0}^{n-1}$ is adapted to the filtration \mathcal{G} .

Let us now introduce the Malliavin representations for Z and Z^n . They are the cornerstone of our study of the error on Z .

2.1.4.2 Representations for Z and Z^n

As recalled in Theorem 2.1.5 and applied in our special case $X = W$, we have the following representation for Z (see Ma and Zhang [MZ02, Theorem 4.2] or Theorem 2.1.5)

$$Z_t = \mathbb{E} \left(g(W_T) N_T^t + \int_t^T f(s, W_s, Y_s, Z_s) N_s^t ds \middle| \mathcal{F}_t \right), \quad 0 \leq t \leq T, \quad (2.18)$$

where for all $s \in (t, T]$ we have

$$N_s^t := \frac{W_s - W_t}{s - t}.$$

Lemma 2.1.7 *Suppose that Assumption 2.2 holds. Then the process Z^n given by (2.16) has the following representation*

$$Z_{t_k}^n = \mathbb{E} \left(g(W_T^n) \frac{W_{t_n}^n - W_{t_k}^n}{t_n - t_k} \middle| \mathcal{G}_k \right) + \mathbb{E} \left(h \sum_{m=k+1}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) \frac{W_{t_m}^n - W_{t_k}^n}{t_m - t_k} \middle| \mathcal{G}_k \right) \quad (2.19)$$

for $k = 0, 1, \dots, n-1$.

2.1.4.3 Main result

Theorem 2.1.8 *Let Assumption 2.2 hold. If W^n is built by using the Skorokhod embedding (i.e. it satisfies (2.5)) then we have, for sufficiently large n , that*

$$\begin{aligned} \mathbb{E} |Y_v - Y_v^n|^2 &\leq C_0 h^{\frac{\varepsilon}{2}} \quad \text{for } v \in [0, T), \\ \mathbb{E} |Z_v - Z_v^n|^2 &\leq C_0 \frac{h^{\frac{\varepsilon}{2}}}{T - t_k} + C_1 \frac{h^{\frac{\varepsilon}{2}}}{(T - v)^{1 - \frac{\varepsilon}{2}}} \mathbf{1}_{v \neq t_k} \quad \text{for } v \in [t_k, t_{k+1}), \quad k = 0, \dots, n-1, \end{aligned}$$

where we have the dependencies $C_0 = C(T, p_0, L_f, C_g, C_{2.1.4}^y, C_{2.1.4}^z, K_f, c_{2.1.5}, \varepsilon)$, $C_1 = C(T, p_0, C_{2.1.4}^z, \varepsilon)$ and $K_f := \sup_{0 \leq t \leq T} |f(t, 0, 0, 0)|$.

A detailed proof of this result can be found in [L-6]. We give here an intuition of it.

Proof. The proof of Theorem 2.1.8 is based on the following inequalities. Let $v \in [t_k, t_{k+1})$, we have

$$\begin{aligned}\|Y_v - Y_v^n\|_2 &\leq \|Y_v - Y_{t_k}\|_2 + \|Y_{t_k} - Y_{t_k}^n\|_2, \\ \|Z_v - Z_v^n\|_2 &\leq \|Z_v - Z_{t_k}\|_2 + \|Z_{t_k} - Z_{t_k}^n\|_2.\end{aligned}$$

$\|Y_v - Y_{t_k}\|_2$ and $\|Z_v - Z_{t_k}\|_2$ are bounded by using Theorem 2.1.4 and $\|Y_{t_k} - Y_{t_k}^n\|_2$ is bounded by studying the difference between the conditional expectation of (2.15) at time $t = t_k$ and the conditional expectation of (2.16) :

$$\begin{aligned}\|Y_{t_k} - Y_{t_k}^n\|_2 &\leq \|\mathbb{E}_{t_k}(g(W_T)) - \mathbb{E}_{\tau_k}(g(W_T^n))\|_2 \\ &\quad + \|\mathbb{E}_{t_k} \int_{t_k}^T f(s, W_s, Y_s, Z_s) ds - h \mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n)\|_2.\end{aligned}$$

$\|\mathbb{E}_{t_k}(g(W_T)) - \mathbb{E}_{\tau_k}(g(W_T^n))\|_2$ is bounded by using the properties of the Skorokhod embedding (see Lemma 2.1.1) and $\|\mathbb{E}_{t_k} \int_{t_k}^T f(s, W_s, Y_s, Z_s) ds - h \mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n)\|_2$ is bounded by using Lemma 2.1.1, Theorem 2.1.4 and the properties of $F(s, x) := f(s, x, u(s, x), u_x(s, x))$ which satisfies

$$|F(t, x_1) - F(t, x_2)| \leq C(1 + |x_1|^{p_0+1} + |x_2|^{p_0+1}) \frac{|x_1 - x_2|}{(T-t)^{1-\frac{\alpha}{2}}}$$

thanks to Theorem 2.1.5. Computing the difference between (2.18) at time $t = t_k$ and (2.19) enables to bound $\|Z_{t_k} - Z_{t_k}^n\|_2$ by using again Lemma 2.1.1, Theorem 2.1.4 and the properties of $F(s, x)$. A Gronwall Lemma ends the proof. \blacksquare

Remark 2.1.9 Theorem 2.1.8 implies that

$$\sup_{v \in [0, T)} \mathbb{E} |Y_v - Y_v^n|^2 \leq C_0 h^{\frac{\alpha}{2}} \quad \text{and} \quad \mathbb{E} \int_0^T |Z_v - Z_v^n|^2 dv \leq C(C_0, C_1, \beta) h^\beta \quad \text{for } \beta \in (0, \frac{\alpha}{2}).$$

2.1.5 Case of a decoupled FBSDE

2.1.5.1 Introduction

This Section presents the results established in [L-3]. The aim is to study the rate of the L_2 -approximation of (Y^n, Z^n) to (Y, Z) when X is a diffusion. This means that we consider the following FBSDE

$$\begin{cases} X_s = x + \int_0^s b(r, X_r) dr + \int_0^s \sigma(r, X_r) dW_r, \\ Y_s = g(X_T) + \int_s^T f(r, X_r, Y_r, Z_r) dr - \int_s^T Z_r dW_r, \end{cases} \quad 0 \leq s \leq T. \quad (2.20)$$

For this, as in the previous section, we generate the random walk W^n by Skorokhod embedding from the Brownian motion W (see Section 2.1.2 and Equation (2.3)). Let (Y^n, Z^n) be the solution to the FBSDE when the Brownian motion W is replaced by W^n . We consider the following discretized BSDE:

$$\begin{cases} X_s^n = x + \int_{(0, s]} b(r, X_{r-}^n) d\langle W^n \rangle_r + \int_{(0, s]} \sigma(r, X_{r-}^n) dW_r^n, \\ Y_s^n = g(X_T^n) + \int_{(s, T]} f(r, X_{r-}^n, Y_{r-}^n, Z_{r-}^n) d\langle W^n \rangle_r - \int_{(s, T]} Z_{r-}^n dW_r^n, \end{cases} \quad 0 \leq s \leq T. \quad (2.21)$$

The special case $X = W$ has already been studied in the previous section, assuming a locally ε -Hölder continuous terminal function g and a Lipschitz continuous generator. In this case the rate of convergence is of order $h^{\frac{\varepsilon}{4}}$ for the L_2 -norm of $Y_t^n - Y_t$, and of order $\frac{h^{\frac{\varepsilon}{4}}}{\sqrt{T-t}}$ for the L_2 -norm of $Z_t^n - Z_t$.

Here we assume that X is a solution to the SDE in (2.20), rather strong conditions on the smoothness and boundedness on f and g and also on b and σ are needed. In Theorem 2.1.22, the main result of the paper, we show that the convergence rate for (Y_t^n, Z_t^n) to (Y_t, Z_t) in L_2 is of order $h^{\frac{1}{4} \wedge \frac{\alpha}{2}}$ provided that g'' exists and is locally α -Hölder continuous.

Remark 2.1.10 For the diffusion setting – in contrast to the case $X = W$ – we can derive the convergence rate for (Y_t^n, Z_t^n) to (Y_t, Z_t) in L_2 only under strong smoothness conditions on the coefficients which include also that g'' is locally α -Hölder continuous (see Assumption 2.3 below). These requirements appear to be necessary. This becomes visible in Subsection 2.1.5.3 where we introduce a discretized Malliavin weight to obtain a representation \hat{Z}^n for Z^n . While it holds that $\hat{Z}^n = Z^n$ when $X = W$, in our case \hat{Z}^n does not coincide with Z^n . However, one can show that the difference $\hat{Z}_t^n - Z_t^n$ converges to 0 in L_2 as $n \rightarrow \infty$ using a Hölder continuity property for the space derivative of the generator in (2.21) (see Proposition 2.1.21). For this Hölder continuity property to hold one needs enough smoothness in space from the solution u^n to the finite difference equation associated to the discretized FBSDE (2.21). Provided that Assumption 2.3 holds we show the smoothness properties for u^n in Proposition 2.1.25.

2.1.5.2 Preliminaries

Let X be a diffusion given by

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad 0 \leq t \leq T$$

and its discretized counterpart

$$X_{t_k}^n = x + h \sum_{j=1}^k b(t_j, X_{t_{j-1}}^n) + \sqrt{h} \sum_{j=1}^k \sigma(t_j, X_{t_{j-1}}^n) \xi_j, \quad k = 1, \dots, n, \quad (2.22)$$

where $(\xi_i)_{i=1,2,\dots}$ is a sequence of i.i.d. Rademacher random variables. \mathcal{G}_k still denotes $\sigma(\xi_i : 1 \leq i \leq k)$ with $\mathcal{G}_0 := \{\emptyset, \Omega\}$. If we extend the sequence $(X_{t_k}^n)_{k \geq 0}$ to a process in continuous time by defining $X_t^n := X_{t_k}^n$ for $t \in [t_k, t_{k+1})$, it is the solution to the forward SDE of (2.21). We introduce the notation $\mathcal{X}_{\tau_k} := X_{t_k}^n$ for all $k = 0, \dots, n$, so that (2.22) turns into

$$\mathcal{X}_{\tau_k} = x + \sum_{j=1}^k b(t_j, \mathcal{X}_{\tau_{j-1}}) h + \sum_{j=1}^k \sigma(t_j, \mathcal{X}_{\tau_{j-1}}) (W_{\tau_j} - W_{\tau_{j-1}}), \quad 0 \leq k \leq n.$$

Lemma 2.1.11 *Under Assumption 2.1 on b and σ it holds for $p \geq 2$ that there exists a constant $C = C(b, \sigma, T, p) > 0$ such that*

1. $\mathbb{E} |X_T^{s,y} - X_T^{t,x}|^p \leq C(|y - x|^p + |s - t|^{\frac{p}{2}})$, $x, y \in \mathbb{R}, s, t \in [0, T]$,
2. $\mathbb{E} \sup_{\tau_l \wedge t_m \leq r \leq \tau_{l+1} \wedge t_m} |X_{t_k+r}^{t_k,x} - X_{t_k+\tau_l \wedge t_m}^{t_k,x}|^p \leq C h^{\frac{p}{4}}$, $0 \leq k \leq n, 0 \leq l \leq n - k - 1, 0 \leq m \leq n - k$,
3. $\mathbb{E} |\nabla X_T^{s,y} - \nabla X_T^{t,x}|^p \leq C(|y - x|^p + |s - t|^{\frac{p}{2}})$, $x, y \in \mathbb{R}, s, t \in [0, T]$,

4. $\mathbb{E} \sup_{0 \leq l \leq m} |\nabla X_{t_k+t_l}^{n,t_k,x}|^p \leq C, \quad 0 \leq k \leq n, 0 \leq m \leq n-k,$
5. $\mathbb{E} |X_{t_k+t_m}^{t_k,x} - \mathcal{X}_{\tau_k+\tau_m}^{\tau_k,y}|^p \leq C(|x-y|^p + h^{\frac{p}{4}}), \quad 0 \leq k \leq n, 0 \leq m \leq n-k,$
6. $\mathbb{E} |\nabla X_{t_k+t_m}^{t_k,x} - \nabla \mathcal{X}_{\tau_k+\tau_m}^{\tau_k,y}|^p \leq C(|x-y|^p + h^{\frac{p}{4}}), \quad 0 \leq k \leq n, 0 \leq m \leq n-k.$

2.1.5.3 The FBSDE and its approximation scheme

Recall the FBSDE (2.20) and its approximation (2.21). The backward equation in (2.21) can equivalently be written in the form

$$Y_{t_k}^n = g(X_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_m}^n \xi_{m+1}, \quad 0 \leq k \leq n, \quad (2.23)$$

if one puts $X_r^n := X_{t_m}^n$, $Y_r^n := Y_{t_m}^n$ and $Z_r^n := Z_{t_m}^n$ for $r \in [t_m, t_{m+1})$.

One can derive an equation for $Z^n = (Z_{t_k}^n)_{k=0}^{n-1}$ if one multiplies (2.23) by ξ_{k+1} and takes the conditional expectation w.r.t. \mathcal{G}_k , so that

$$Z_{t_k}^n = \frac{1}{\sqrt{h}} \mathbb{E} (g(X_T^n) \xi_{k+1} | \mathcal{G}_k) + \mathbb{E} \left(\sqrt{h} \sum_{m=k+1}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) \xi_{k+1} | \mathcal{G}_k \right), \quad 0 \leq k \leq n-1. \quad (2.24)$$

Remark 2.1.12 For n large enough, the BSDE (2.23) has a unique solution (Y^n, Z^n) (see [Tol06, Proposition 1.2]), and $(Y_{t_k}^n, Z_{t_k}^n)_{k=0}^{n-1}$ is adapted to the filtration $(\mathcal{G}_k)_{k=0}^{n-1}$.

Representation for Z We will use the following representation for Z , due to Ma and Zhang (see [MZ02, Theorem 4.2] or Theorem 2.1.5)

$$Z_t = \mathbb{E} \left(g(X_T) N_T^t + \int_t^T f(s, X_s, Y_s, Z_s) N_s^t ds \middle| \mathcal{F}_t \right) \sigma(t, X_t), \quad 0 \leq t \leq T \quad (2.25)$$

where

$$N_s^t = \frac{1}{s-t} \int_t^s \frac{\nabla X_r}{\sigma(r, X_r) \nabla X_t} dW_r. \quad (2.26)$$

Remark 2.1.13 In the following we will assume that g'' exists. In such a case we have the following representation for Z :

$$Z_t = \mathbb{E} \left(g'(X_T) \nabla X_T + \int_t^T f(s, X_s, Y_s, Z_s) N_s^t ds \middle| \mathcal{F}_t \right) \sigma(t, X_t), \quad 0 \leq t \leq T. \quad (2.27)$$

Approximation for Z^n Let us state the discrete counterpart to (2.25), which, in the general case of a forward process X , does not coincide with Z^n (given by (2.24)). In contrast to the continuous-time case, where the variational process and the Malliavin derivative are connected by $\frac{\nabla X_t}{\nabla X_s} = \frac{D_s X_t}{\sigma(s, X_s)}$ ($s \leq t$), we can not expect equality for the corresponding expressions if we use the discretized version of the processes $(\nabla X_t)_t$ and $(D_s X_t)_{s \leq t}$ introduced in (2.28). This counterpart \hat{Z}^n to Z is a key tool in the proof of the convergence of Z^n to Z . As we will see in the proof of Theorem 2.1.22, the study of $\|Z_{t_k}^n - Z_{t_k}\|_2$ goes through the study of $\|Z_{t_k}^n - \hat{Z}_{t_k}^n\|_2$ and $\|\hat{Z}_{t_k}^n - Z_{t_k}\|_2$.

Before defining the discretized version of $(\nabla X_t)_t$ and $(D_s X_t)_{s \leq t}$, we shortly introduce the discretized Malliavin derivative.

Definition 2.1.14 (Definition of $T_{m,+}$, $T_{m,-}$ and \mathcal{D}_m^n) For any function $F : \{-1, 1\}^n \rightarrow \mathbb{R}$, the mappings $T_{m,+}$ and $T_{m,-}$ are defined by

$$T_{m,\pm} F(\xi_1, \dots, \xi_n) := F(\xi_1, \dots, \xi_{m-1}, \pm 1, \xi_{m+1}, \dots, \xi_n), \quad 1 \leq m \leq n.$$

For any $\zeta = F(\xi_1, \dots, \xi_n)$, the discretized Malliavin derivative is defined by

$$\mathcal{D}_m^n \zeta := \frac{\mathbb{E}[\zeta \xi_m | \sigma((\xi_l)_{l \in \{1, \dots, n\} \setminus \{m\}})]}{\sqrt{h}} = \frac{T_{m,+} \zeta - T_{m,-} \zeta}{2\sqrt{h}}, \quad 1 \leq m \leq n.$$

Definition 2.1.15 (Definition of $\phi_x^{(k,l)}$) Let ϕ be a $C^{0,1}([0, T] \times \mathbb{R})$ function. We denote

$$\phi_x^{(k,l)} := \frac{\mathcal{D}_k^n \phi(t_l, X_{t_{l-1}}^n)}{\mathcal{D}_k^n X_{t_{l-1}}^n} := \int_0^1 \phi_x(t_l, \vartheta T_{k,+} X_{t_{l-1}}^n + (1 - \vartheta) T_{k,-} X_{t_{l-1}}^n) d\vartheta.$$

If $\mathcal{D}_k^n X_{t_{\ell-1}}^n \neq 0$ the second ' $:=$ ' holds as an identity.

We are now able to define the discretized version of $(\nabla X_t)_t$ and $(D_s X_t)_{s \leq t}$.

Definition 2.1.16 (Discretized processes $(\nabla X_{t_m}^{n,t_k,x})_{m \in \{k, \dots, n\}}$ and $(\mathcal{D}_k^n X_{t_m}^n)_{m \in \{k, \dots, n\}}$) For all m in $\{k, \dots, n\}$ we define

$$\begin{aligned} \nabla X_{t_m}^{n,t_k,x} &= 1 + h \sum_{l=k+1}^m b_x(t_l, X_{t_{l-1}}^{n,t_k,x}) \nabla X_{t_{l-1}}^{n,t_k,x} + \sqrt{h} \sum_{l=k+1}^m \sigma_x(t_l, X_{t_{l-1}}^{n,t_k,x}) \nabla X_{t_{l-1}}^{n,t_k,x} \xi_l, \quad 0 \leq k \leq n, \\ \mathcal{D}_k^n X_{t_m}^n &= \sigma(t_k, X_{t_{k-1}}^n) + h \sum_{l=k+1}^m b_x^{(k,l)} \mathcal{D}_k^n X_{t_{l-1}}^n + \sqrt{h} \sum_{l=k+1}^m \sigma_x^{(k,l)} (\mathcal{D}_k^n X_{t_{l-1}}^n) \xi_l, \quad 0 < k \leq n. \end{aligned} \tag{2.28}$$

Remark 2.1.17

1. Although $\nabla X_{t_m}^{n,t_k,X_{t_k}^n}$ is not equal to $\frac{\mathcal{D}_{k+1}^n X_{t_m}^n}{\sigma(t_{k+1}, X_{t_k}^n)}$, we can show that the difference of these terms converges in L_p (see Lemma 2.1.18).
2. With the notation introduced above, (2.24) rewrites to

$$Z_{t_k}^n = \mathbb{E}(\mathcal{D}_{k+1}^n g(X_T^n) | \mathcal{G}_k) + \mathbb{E} \left(h \sum_{m=k+1}^{n-1} \mathcal{D}_{k+1}^n f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) | \mathcal{G}_k \right). \tag{2.29}$$

Lemma 2.1.18 Under Assumption 2.1, and for $p \geq 2$, we have

1. $\mathbb{E}|X_{t_l}^n - T_{m,\pm} X_{t_l}^n|^p \leq C(b, \sigma, T, p) h^{\frac{p}{2}}, \quad 1 \leq l, m \leq n,$
2. $\mathbb{E} \left| \nabla X_{t_m}^{n,t_k,X_{t_k}^n} - \frac{\mathcal{D}_{k+1}^n X_{t_m}^n}{\sigma(t_{k+1}, X_{t_k}^n)} \right|^p \leq C(b, \sigma, T, p) h^{\frac{p}{2}}, \quad 0 \leq k < m \leq n.$
3. $\mathbb{E}|\mathcal{D}_k^n X_{t_m}^n|^p \leq C(b, \sigma, T, p), \quad 0 \leq k \leq m \leq n.$

In order to define the discrete counterpart to (2.25), we first define the discrete counterpart to $(N_s^t)_{s \in [t, T]}$ given in (2.26):

$$N_{t_\ell}^{n, t_k} := \sqrt{h} \sum_{m=k+1}^{\ell} \frac{\nabla X_{t_{m-1}}^{n, t_k, X_{t_k}^n} \xi_m}{\sigma(t_m, X_{t_{m-1}}^n) t_\ell - t_k}, \quad k < \ell \leq n. \quad (2.30)$$

Notice that there is some constant $\widehat{\kappa}_2 > 0$ depending on b, σ, T, δ such that

$$\left(\mathbb{E}(|N_{t_\ell}^{n, t_k}|^2 | \mathcal{G}_k) \right)^{\frac{1}{2}} \leq \frac{\widehat{\kappa}_2}{(t_\ell - t_k)^{\frac{1}{2}}}, \quad 0 \leq k < \ell \leq n. \quad (2.31)$$

Definition 2.1.19 (Discrete counterpart to (2.27)) Let the process $\widehat{Z}^n = (\widehat{Z}_{t_k}^n)_{k=0}^{n-1}$ be defined by

$$\widehat{Z}_{t_k}^n := \mathbb{E}(\mathcal{D}_{k+1}^n g(X_T^n) | \mathcal{G}_k) + \mathbb{E} \left(h \sum_{m=k+1}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) N_{t_m}^{n, t_k} | \mathcal{G}_k \right) \sigma(t_{k+1}, X_{t_k}^n), \quad (2.32)$$

Remark 2.1.20 In (2.32) we could have used also the approximate expression $\mathbb{E}(g(X_T^n) N_{t_n}^{n, t_k} \sigma(t_{k+1}, X_{t_k}^n) | \mathcal{G}_k)$, but since we will assume that g'' exists, we work with the correct term.

The study of the convergence $\mathbb{E}_{0,x}^{\mathcal{G}} |Z_{t_k}^n - \widehat{Z}_{t_k}^n|^2$, where $\mathbb{E}_{0,x}^{\mathcal{G}} := \mathbb{E}(\cdot | \mathcal{G}_0, X_0 = x)$, requires stronger assumptions on the coefficients b, σ, f and g .

Hypothesis 2.3 Assumptions 2.1 and 2.2 hold. Additionally, we assume that all first and second derivatives w.r.t. the variables x, y, z of $b(t, x), \sigma(t, x)$ and $f(t, x, y, z)$ exist and are bounded Lipschitz functions w.r.t. these variables, uniformly in time. Moreover, g'' exists and satisfies (2.10).

Proposition 2.1.21 If Assumption 2.3 holds, then

$$\mathbb{E}_{0,x}^{\mathcal{G}} |Z_{t_k}^n - \widehat{Z}_{t_k}^n|^2 \leq C_{2.1.21} \widehat{\Psi}^2(x) h^\alpha,$$

where the function $\widehat{\Psi}(x) := 1 + |x|^{6p_0+8}$ and $C_{2.1.21}$ depends on b, σ, f, g, T, p_0 and δ .

Proof. The proof of Proposition 2.1.21 is based on the following representations (see [BDM01, Proposition 5.1])

$$Y_{t_m}^n = u^n(t_m, X_{t_m}^n), \quad \text{and} \quad Z_{t_m}^n = \mathcal{D}_{m+1}^n u^n(t_{m+1}, X_{t_{m+1}}^n),$$

where u^n is the solution of the finite difference equation (2.35) with terminal condition $u^n(t_n, x) = g(x)$, on Proposition 2.1.25 and Lemma 2.1.18. \blacksquare

2.1.5.4 Main result

Theorem 2.1.22 Let Assumption 2.3 be satisfied. Then for all $v \in [0, T)$ and large enough n , we have

$$\mathbb{E}_{0,x} |Y_v - Y_v^n|^2 + \mathbb{E}_{0,x} |Z_v - Z_v^n|^2 \leq C_{2.1.22} \widehat{\Psi}(x)^2 h^{\frac{1}{2} \wedge \alpha}$$

with $C_{2.1.22} = C(b, \sigma, f, g, T, p_0, \delta)$ and $\widehat{\Psi}$ is given in Proposition 2.1.21.

Remark 2.1.23 As noticed above, the filtration \mathcal{G}_k coincides with \mathcal{F}_{τ_k} , for all $k = 0, \dots, n$. The expectation $\mathbb{E}_{0,x}$ appearing in Theorem 2.1.22 is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. As for the case $X = W$, we give an overview of the proof of Theorem 2.1.22. It is based on the following inequalities. Let $v \in [t_k, t_{k+1})$, we have

$$\begin{aligned}\|Y_v - Y_v^n\|_2 &\leq \|Y_v - Y_{t_k}\|_2 + \|Y_{t_k} - Y_{t_k}^n\|_2, \\ \|Z_v - Z_v^n\|_2 &\leq \|Z_v - Z_{t_k}\|_2 + \|Z_{t_k} - Z_{t_k}^n\|_2.\end{aligned}$$

$\|Y_v - Y_{t_k}\|_2$ and $\|Z_v - Z_{t_k}\|_2$ are bounded by using Theorem 2.1.4 and Proposition 2.1.24 and $\|Y_{t_k} - Y_{t_k}^n\|_2$ is bounded by studying the difference between the conditional expectation of (2.20) at time $t = t_k$ and the conditional expectation of (2.23) :

$$\begin{aligned}\|Y_{t_k} - Y_{t_k}^n\|_2 &\leq \|\mathbb{E}_{t_k}(g(X_T)) - \mathbb{E}_{\tau_k}(g(X_T^n))\|_2 \\ &\quad + \|\mathbb{E}_{t_k} \int_{t_k}^T f(s, X_s, Y_s, Z_s) ds - h \mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n)\|_2.\end{aligned}$$

$\|\mathbb{E}_{t_k}(g(X_T)) - \mathbb{E}_{\tau_k}(g(X_T^n))\|_2$ is bounded by using the properties of the Skorokhod embedding (see Lemma 2.1.11) and $\|\mathbb{E}_{t_k} \int_{t_k}^T f(s, X_s, Y_s, Z_s) ds - h \mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n)\|_2$ is bounded by using Lemma 2.1.11, Theorem 2.1.4, Proposition 2.1.24 and the properties of $F(s, x) := f(s, x, u(s, x), u_x(s, x))$ which satisfies the following inequality under Assumption 2.3. From Theorem 2.1.5 we get that for all $(x_1, x_2) \in \mathbb{R}^2$ there exists (α_1, α_2) in $[\min(x_1, x_2), \max(x_1, x_2)]$ s.t.

$$|F(t, x_1) - F(t, x_2)| \leq C \left(1 + |\alpha_1|^{p_0+1} + \frac{|\alpha_2|^{p_0+1}}{(T-t)^{\frac{1}{2}}} \right) |x_1 - x_2|.$$

Computing the difference between (2.27) at time $t = t_k$ and (2.29) enables to bound $\|Z_{t_k} - Z_{t_k}^n\|_2$. First we introduce $\pm \hat{Z}_{t_k}^n$. Then we bound $\|\hat{Z}_{t_k}^n - Z_{t_k}^n\|_2$ by using Proposition 2.1.21 and we bound $\|\hat{Z}_{t_k}^n - Z_{t_k}^n\|_2$ by using Lemma 2.1.11, Theorem 2.1.4, Proposition 2.1.24 and the properties satisfied by $F(s, x)$. A Gronwall Lemma ends the proof. \blacksquare

2.1.5.5 Some properties of solutions to BSDEs and finite difference equations

Properties on Z Under Assumption 2.3 we are in the position to improve the bound on $\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})}$ given in Theorem 2.1.4.

Proposition 2.1.24 *If Assumption 2.3 holds, then there exists a constant $C_{2.1.24} > 0$ such that for $0 \leq t < s \leq T$ and $x \in \mathbb{R}$,*

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \leq C_{2.1.24} \Psi(x) (s-t)^{\frac{1}{2}},$$

where Ψ has been introduced in Assumption 2.2, $C_{2.1.24}$ depends on $c_{2.1.5}^{2,3}$ (see Theorem 2.1.5), $b, \sigma, f, g, T, p_0, p$, and hence $C_{2.1.24} = C_{2.1.24}(b, \sigma, f, g, T, p_0, p, \delta)$.

Properties on the solution to the finite difference equation Recall the definition of \mathcal{D}_m^n given in (2.1.14). By (2.22),

$$X_{t_{m+1}}^{n,t_m,x} = x + hb(t_{m+1}, x) + \sqrt{h}\sigma(t_{m+1}, x)\xi_{m+1}, \quad (2.33)$$

so that

$$T_{m+1,\pm} u^n(t_{m+1}, X_{t_{m+1}}^{n,t_m,x}) = u^n(t_{m+1}, x + hb(t_{m+1}, x) \pm \sqrt{h}\sigma(t_{m+1}, x)). \quad (2.34)$$

Even though g is only Hölder continuous, one can observe in Theorem 2.1.5 the well-known smoothing property which implies that u , the solution to the PDE (2.4), is differentiable on $[0, T] \times \mathbb{R}$. In the following proposition, for the solution u^n to the finite difference equation, we have to require from g the same regularity as we want for u^n .

Proposition 2.1.25 *Let Assumption 2.3 hold and assume that u^n is a solution to*

$$\begin{aligned} & u^n(t_m, x) - hf(t_{m+1}, x, u^n(t_m, x), \mathcal{D}_{m+1}^n u^n(t_{m+1}, X_{t_{m+1}}^{n,t_m,x})) \\ = & \frac{1}{2} [T_{m+1,+} u^n(t_{m+1}, X_{t_{m+1}}^{n,t_m,x}) + T_{m+1,-} u^n(t_{m+1}, X_{t_{m+1}}^{n,t_m,x})], \quad m = 0, \dots, n-1, \end{aligned} \quad (2.35)$$

with terminal condition $u^n(t_n, x) = g(x)$. Then, for sufficiently small h , the map $x \mapsto u^n(t_m, x)$ is C^2 and it holds

$$|u^n(t_m, x)| + |u_x^n(t_m, x)| \leq C_{u^n,1} \Psi(x), \quad |u_{xx}^n(t_m, x)| \leq C_{u^n,2} \Psi^2(x)$$

and

$$|u_{xx}^n(t_m, x) - u_{xx}^n(t_m, \bar{x})| \leq C_{u^n,3} (1 + |x|^{6p_0+7} + |\bar{x}|^{6p_0+7}) |x - \bar{x}|^\alpha, \quad (2.36)$$

uniformly in $m = 0, \dots, n-1$. The constants $C_{u^n,1}$, $C_{u^n,2}$ and $C_{u^n,3}$ depend on the bounds of f, g, b, σ and their derivatives and on T and p_0 .

2.2 Donsker type result

This Section presents the results established in [L-1] written with Philippe Briand, Christel Geiss and Stefan Geiss. The objective is to confirm this expected and optimal rate $n^{-\varepsilon/2}$ which has been raised in the introduction of Section 2.1. This improvement is possible by using a weak limit approach. The error is then considered in the Wasserstein distance.

2.2.1 Introduction

In this section we consider the random walk approximation of

$$Y_t = g(W_T) + \int_t^T f(s, W_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (2.37)$$

given by

$$Y_t^n = g(W_T^n) + \int_{]t,T]} f(s, W_{s^-}^n, Y_{s^-}^n, Z_s^n) d\langle W^n \rangle_s - \int_{]t,T]} Z_s^n dW_s^n, \quad t \in [0, T].$$

Remark 2.2.1 This approximation is close to (2.17). The main difference lies in Z^n , which is here a predictable process w.r.t. the filtration \mathcal{G} .

In this section we establish a Donsker type result : assuming that g is ε -Hölder, the objective of our study is to confirm the expected and optimal rate $n^{-\varepsilon/2}$. Our starting point is a result of Emmanuel Rio [Rio09] which directly implies that, when $T = 1$, for all $p \geq 1$, there exists a constant C_p such that, for all $n \geq 1$, $\mathcal{W}_p(W_1^n, \mathcal{G}) \leq C_p n^{-1/2}$, where \mathcal{W}_p is the L_p -Wasserstein distance and \mathcal{G} a standard normal random variable. To do so, we first generalize this result to cover the case where $f \equiv 0$, which corresponds to the heat equation. Then, using the associated PDE, in particular representation formulas in the spirit of [MZ02], we are able to prove that

$$\mathcal{W}_p(Y_t^n, Y_t) \leq C_p n^{-(\alpha \wedge \frac{\varepsilon}{2})} \quad \text{and} \quad \mathcal{W}_p(Z_t^n, Z_t) \leq \frac{C_p}{\sqrt{T-t}} n^{-(\alpha \wedge \frac{\varepsilon}{2})}$$

for $t \in [0, T]$ and $t \in [0, T[$, respectively, when g and $f(t, \cdot, y, z)$ are ε -Hölder continuous and $f(\cdot, x, y, z)$ is α -Hölder continuous. As mentioned in Remark 2.2.15-(i), this rate of convergence is optimal. We refer to Theorem 2.2.14 in Section 2.2.5 for the precise statement. One of the main difficulties in the proof concerns various gradient estimates in order to obtain the estimate for $\mathcal{W}_p(Z_t^n, Z_t)$. The main step in the proof consists in the convergence estimates in Theorem 2.2.16, where we consider the approximation of solutions to the following semilinear heat equation

$$\begin{cases} \partial_t u(t, x) + \frac{1}{2} \Delta u(t, x) + f(t, x, u(t, x), \nabla_x u(t, x)) = 0, & (t, x) \in [0, T[\times \mathbb{R}, \\ u(T, \cdot) = g(\cdot). \end{cases} \quad (2.38)$$

and their gradients by solutions to the associated finite difference schemes and their discretized gradients. We show these estimates by probabilistic means.

The Section is organized as follows: we first introduce some notations (Section 2.2.2), then we recall some properties of the random walk in the Wasserstein distance (Section 2.2.3), we give some regularity results on the associated discrete and continuous PDEs (Section 2.2.4) and finally we establish convergence results (Section 2.2.5).

2.2.2 Notations

Hypothesis 2.4 There exist $0 < \varepsilon \leq 1$ and $0 < \alpha \leq 1$ such that it holds:

- i. The function $g : \mathbb{R} \rightarrow \mathbb{R}$ is ε -Hölder continuous: for all $(x, x') \in \mathbb{R}^2$ one has

$$|g(x) - g(x')| \leq \|g\|_\varepsilon |x - x'|^\varepsilon.$$

- ii. The function $f : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is α -Hölder continuous in time, ε -Hölder continuous in space and Lipschitz continuous with respect to (y, z) : for all (t, x, y, z) and (t', x', y', z') in $[0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ one has

$$\begin{aligned} & |f(t, x, y, z) - f(t', x', y', z')| \\ & \leq \|f_t\|_\alpha |t - t'|^\alpha + \|f_x\|_\varepsilon |x - x'|^\varepsilon + \|f_y\|_{\text{Lip}} |y - y'| + \|f_z\|_{\text{Lip}} |z - z'|. \end{aligned} \quad (2.39)$$

Most of the time, we do not need to distinguish between $\|f_y\|_{\text{Lip}}$ and $\|f_z\|_{\text{Lip}}$ and we let $\|f\|_{\text{Lip}} := \max(\|f_y\|_{\text{Lip}}, \|f_z\|_{\text{Lip}})$.

Convention: Later the phrase *a constant $C > 0$ depends on (T, ε, f, g)* stands for the fact that C can be expressed in terms of $(T, \varepsilon, \|f_x\|_\varepsilon, \|f_y\|_{\text{Lip}}, \|f_z\|_{\text{Lip}}, K_f, \|g\|_\varepsilon, g(0))$ where $K_f := \sup_{t \in [0, T]} |f(t, 0, 0, 0)|$. Similarly, a dependence on $(T, \alpha, \varepsilon, f, g)$ means an additional dependence on $(\alpha, \|f_t\|_\alpha)$.

From [BDH⁺03, Theorem 4.2] it is known that under Hypothesis 2.4, the BSDE (2.37) has a unique L_p -solution (Y, Z) for any $p \in]1, \infty[$. So for $(t, x) \in [0, T[\times \mathbb{R}$ we let $(Y_s^{t,x}, Z_s^{t,x})_{s \in [t, T]}$ be the square integrable solution to the BSDE

$$Y_s^{t,x} = g(W_T^{t,x}) + \int_s^T f(r, W_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r, \quad t \leq s \leq T, \quad (2.40)$$

where $W_r^{t,x} := x + W_r - W_t$, and set, as usual, for $x \in \mathbb{R}$, $u(T, x) := g(x)$, and, for $(t, x) \in [0, T[\times \mathbb{R}$,

$$u(t, x) := Y_t^{t,x} = \mathbb{E} \left[g(W_T^{t,x}) + \int_t^T f(r, W_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr \right].$$

It is well known that the function u is continuous on $[0, T] \times \mathbb{R}$ (see also Lemma 2.2.10 below). Under Lipschitz assumptions in (x, y, z) and for $\alpha \geq \frac{1}{2}$, u is the viscosity solution to (2.38), see [Zha17, Theorem 5.5.8]. Moreover, as recalled in Theorem 2.1.5, in this Markovian setting, for $(t, x) \in [0, T] \times \mathbb{R}$, we have $Y_s^{t,x} = u(s, W_s^{t,x})$ a.s. for all $s \in [t, T]$ and for a generator which is Lipschitz continuous in all space variables and a measurable function g with polynomial growth, J. Zhang proved that u belongs to $\mathcal{C}^{0,1}([0, T[\times \mathbb{R})$, that $Z_s^{t,x} = \nabla_x u(s, W_s^{t,x})$ a.e. on $[t, T[\times \Omega$ and that the following representation holds (see [Zha05, Theorem 3.2])

$$\nabla_x u(t, x) = \mathbb{E} \left[g(W_T^{t,x}) \frac{W_T - W_t}{T - t} + \int_t^T f(r, W_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) \frac{W_r - W_t}{r - t} dr \right], \quad (t, x) \in [0, T[\times \mathbb{R}.$$

If F is the function given by

$$F(s, x) := f(s, x, u(s, x), \nabla_x u(s, x)) \quad \text{for } (s, x) \in [0, T[\times \mathbb{R}, \quad (2.41)$$

we thus have

$$u(t, x) = \mathbb{E} \left[g(W_T^{t,x}) + \int_t^T F(r, W_r^{t,x}) dr \right], \quad (t, x) \in [0, T[\times \mathbb{R}, \quad (2.42)$$

together with

$$\nabla_x u(t, x) = \mathbb{E} \left[g(W_T^{t,x}) \frac{W_T - W_t}{T - t} + \int_t^T F(r, W_r^{t,x}) \frac{W_r - W_t}{r - t} dr \right], \quad (t, x) \in [0, T[\times \mathbb{R}. \quad (2.43)$$

Remark 2.2.2 These formulas play an important role in the sequel. In Lemma 2.2.10, we extend these results to the case where $f(t, \cdot, y, z)$ is ε -Hölder continuous and make the regularity of u and $\nabla_x u$ precise.

As already mentioned before, we are concerned with the approximation of the solution $(Y^{t,x}, Z^{t,x})$ to (2.40) by a solution to the BSDE driven by a scaled random walk. To do this, let us consider, on some probability space, not necessarily $(\Omega, \mathcal{F}, \mathbb{P})$, an i.i.d. sequence $(\xi_k)_{k \geq 1}$ of Rademacher

random variables. For $n \in \mathbb{N}^* := \{1, 2, 3, \dots\}$ we set $h := T/n$ and we consider the scaled random walk introduced in (2.3):

$$W_t^n := \sqrt{h} \sum_{k=1}^{\lfloor t/h \rfloor} \xi_k, \quad 0 \leq t \leq T.$$

As we did for the Brownian motion, for $x \in \mathbb{R}$ and $0 \leq t \leq s \leq T$ we put

$$W_s^{n,t,x} := x + W_s^n - W_t^n.$$

Let us introduce some further notation. We denote the ceiling function by $\lceil x \rceil := \min\{r \in \mathbb{Z} : r \geq x\}$ for $x \in \mathbb{R}$. Moreover, we set

$$n_t := \lfloor t/h \rfloor, \quad \underline{t} := h \lfloor t/h \rfloor = hn_t \quad \text{and} \quad \bar{t} := h \lceil t/h \rceil, \quad t \in [0, T].$$

By construction, Y^n is a piecewise constant càdlàg process with $Y_t^n = Y_{\bar{t}}^n$. The process Z^n is defined as an element of $L_2(\Omega \times [0, T], d\mathbb{P} \otimes d\langle W^n \rangle)$, where we start with a Z^n defined only on the points $\{t_k := kh : k = 1, \dots, n\}$ and extend it to $]0, T]$ as a càglàd process $(Z_t^n)_{t \in]0, T]}$ by setting $Z_t^n = Z_{\underline{t}}^n$. The previous BSDE is actually a discrete BSDE that can be solved by hand since, for $k = 0, \dots, n-1$, we have

$$Y_{t_k}^n = Y_{t_{k+1}}^n + h f\left(t_{k+1}, W_{t_k}^n, Y_{t_k}^n, Z_{t_{k+1}}^n\right) - \sqrt{h} Z_{t_{k+1}}^n \xi_{k+1}, \quad Y_{t_n}^n = g(W_T^n). \quad (2.44)$$

Thus, if $Y_{t_{k+1}}^n$ is given,

$$Z_{t_{k+1}}^n = h^{-1/2} \mathbb{E}\left(Y_{t_{k+1}}^n \xi_{k+1} \mid \mathcal{G}_k\right), \quad (2.45)$$

$$\begin{aligned} Y_{t_k}^n &= Y_{t_{k+1}}^n + h f\left(t_{k+1}, W_{t_k}^n, Y_{t_k}^n, Z_{t_{k+1}}^n\right) - \sqrt{h} Z_{t_{k+1}}^n \xi_{k+1} \\ &= \mathbb{E}\left(Y_{t_{k+1}}^n \mid \mathcal{G}_k\right) + h f\left(t_{k+1}, W_{t_k}^n, Y_{t_k}^n, Z_{t_{k+1}}^n\right), \end{aligned} \quad (2.46)$$

where the last equality follows by taking the conditional equation w.r.t. \mathcal{G}_k in the second line.

Remark 2.2.3 If we sum (2.44) we get

$$Y_{t_k}^n = g(W_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_{m+1}}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_{m+1}}^n \xi_{m+1}, \quad 0 \leq k \leq n-1.$$

which is almost (2.2). The only difference is due to the predictable property of the Z^n process considered in this section.

Since we are in a Markovian setting, there is also an analogue of the Feynman-Kac formula. If u is a given function we set

$$D_+^n u(x) := \frac{1}{2} \left(u(x + \sqrt{h}) + u(x - \sqrt{h}) \right), \quad D_-^n u(x) := \frac{1}{2} \left(u(x + \sqrt{h}) - u(x - \sqrt{h}) \right),$$

and

$$\nabla^n u(x) := h^{-1/2} D_-^n u(x). \quad (2.47)$$

Remark 2.2.4 From the definition of D_+^n and D_-^n , we get that if u is ε -Hölder, $D_+^n u$ and $D_-^n u$ are also ε -Hölder with constant $\|u\|_\varepsilon$.

Let U^n be the solution to the finite difference equation, where for $x \in \mathbb{R}$ and $k = 0, \dots, n-1$ we require

$$\begin{cases} U^n(t_k, x) = D_+^n U^n(t_{k+1}, x) + hf(t_{k+1}, x, U^n(t_k, x), h^{-\frac{1}{2}} D_-^n U^n(t_{k+1}, x)), \\ U^n(t_n, x) = g(x). \end{cases} \quad (2.48)$$

This scheme is monotone in the sense of Barles and Souganides [BS91]. We obtain from (2.45) and (2.46) (cf [BDM01, Proposition 5.1]) that

$$\begin{aligned} Y_{t_k}^n &= U^n \left(t_k, \sqrt{h} \sum_{i=1}^k \xi_i \right), \quad k = 0, \dots, n, \\ Z_{t_k}^n &= \nabla^n U^n \left(t_k, \sqrt{h} \sum_{i=1}^{k-1} \xi_i \right), \quad k = 1, \dots, n. \end{aligned}$$

These formulas rewrite in continuous time to

$$Y_t^n = Y_{\underline{t}}^n = U^n(\underline{t}, W_t^n), \quad t \in [0, T], \quad \text{and} \quad Z_t^n = Z_{\bar{t}}^n = \nabla^n U^n(\bar{t}, W_{t-}^n), \quad t \in]0, T].$$

If we set, for $0 \leq t \leq T$ and $x \in \mathbb{R}$, $U^n(t, x) := U^n(\underline{t}, x)$, we have $Y_t^n = U^n(t, W_t^n)$.

More generally, for $0 \leq t < T$, we define $(Y^{n,t,x}, Z^{n,t,x})$ as the solution $Y^{n,t,x} = (Y_s^{n,t,x})_{s \in]t, T]}$ and $Z^{n,t,x} = (Z_s^{n,t,x})_{s \in]t, T]}$ to the BSDE

$$Y_s^{n,t,x} = g(W_T^{n,t,x}) + \int_{]s, T]} f(r, W_{r-}^{n,t,x}, Y_{r-}^{n,t,x}, Z_r^{n,t,x}) d\langle W^n \rangle_r - \int_{]s, T]} Z_r^{n,t,x} dW_r^n, \quad s \in]t, T]. \quad (2.49)$$

We set $Y_T^{n,T,x} = g(x)$. Then,

$$Y_s^{n,t,x} = Y_{\underline{s}}^{n,t,x} = U^n(s, W_s^{n,t,x}), \quad 0 \leq t \leq s \leq T. \quad (2.50)$$

Let us observe that $Z^{n,t,x}$ is first defined at the points $t = t_k$, $k = n_t + 1, \dots, n$. As before we let $Z_s^{n,t,x} := Z_{\bar{s}}^{n,t,x}$ for $s \in]t, T]$. We have

$$Z_s^{n,t,x} = Z_{\bar{s}}^{n,t,x} = \nabla^n U^n(\bar{s}, W_{s-}^{n,t,x}) \quad \text{for } s \in]t, T].$$

In particular,

$$Z_s^{n,t,x} = \nabla^n U^n(\underline{s} + h, W_s^{n,t,x}) \quad \text{whenever } s \in]t, T] \setminus \{t_k : k = n_t + 1, \dots, n\}. \quad (2.51)$$

Of course, we have

$$U^n(t, x) = Y_t^{n,t,x} = Y_{\underline{t}}^{n,t,x} \quad \text{for } t \in [0, T].$$

Similarly, we define, for $(t, x) \in [0, T[\times \mathbb{R}$,

$$\Delta^n(t, x) := \nabla^n U^n(\underline{t} + h, x) = Z_{\underline{t}+h}^{n,t,x}. \quad (2.52)$$

With this notation, (2.51) rewrites as

$$Z_s^{n,t,x} = \Delta^n(s, W_s^{n,t,x}) \quad \text{whenever } s \in]t, T] \setminus \{t_k : k = n_t + 1, \dots, n\}. \quad (2.53)$$

It follows that

$$\begin{aligned} U^n(t, x) &= \mathbb{E} \left[g \left(W_T^{n,t,x} \right) + h \sum_{k=n_t+1}^n f \left(t_k, W_{t_{k-1}}^{n,t,x}, Y_{t_{k-1}}^{n,t,x}, Z_{t_k}^{n,t,x} \right) \right] \\ &= \mathbb{E} \left[g \left(W_T^{n,t,x} \right) + \int_{\underline{t}}^T f \left(\bar{s}, W_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x} \right) ds \right], \end{aligned}$$

which rewrites, taking into account (2.50) and (2.53), to

$$\begin{aligned} U^n(t, x) &= \mathbb{E} \left[g \left(W_T^{n,t,x} \right) + \int_{\underline{t}}^T f \left(\bar{s}, W_s^{n,t,x}, U^n(s, W_s^{n,t,x}), \Delta^n(s, W_s^{n,t,x}) \right) ds \right] \\ &= \mathbb{E} \left[g \left(W_T^{n,t,x} \right) + \int_{\underline{t}}^T f \left(\bar{s}, \Theta_s^{n,t,x} \right) ds \right], \end{aligned} \quad (2.54)$$

where

$$\Theta_s^{n,t,x} := \left(W_s^{n,t,x}, U^n(s, W_s^{n,t,x}), \Delta^n(s, W_s^{n,t,x}) \right). \quad (2.55)$$

We prove in Section 2.2.5 that (U^n, Δ^n) converges to $(u, \nabla_x u)$.

From now on we assume that $n \geq n_0(T, \|f\|_{\text{Lip}})$ where $n_0(T, \|f\|_{\text{Lip}}) \in \mathbb{N}^*$ is the integer given in [L-1, Lemma 14] and which automatically implies also existence and uniqueness of solutions because $n_0 > T\|f\|_{\text{Lip}}$.

2.2.3 Scaled random walk and Wasserstein distance

One starting point of this section is the following result of Emmanuel Rio [Rio09] (Theorem 2.1) (see also [Rio11]). We then extend this result in Proposition 2.2.7 and Corollary 2.2.8 to cover the case where the generator vanishes, i.e. $f \equiv 0$.

For $r > 0$ let ψ_r be defined by $\psi_r(x) = e^{|x|^r} - 1$, and for a real random variable X we let

$$\|X\|_{\psi_r} := \inf \{ a > 0 : \mathbb{E} [\psi_r(X/a)] \leq 1 \}, \quad \inf \emptyset := +\infty.$$

Let us recall that, for any $p > 0$,

$$\sup_{x>0} \left\{ \frac{x^{p \vee r}}{\psi_r(x)} \right\} < +\infty, \quad \|X\|_{L_p} \leq \left(\sup_{x>0} \left\{ \frac{x^{p \vee r}}{\psi_r(x)} \right\} \right)^{1/(p \vee r)} \|X\|_{\psi_r}. \quad (2.56)$$

Let X and X' be two random variables and let us denote by μ the law of X and by ν the law of X' . With the usual abuse of notation, the Wasserstein distance associated to ψ_r is defined by

$$\mathcal{W}_{\psi_r}(\mu, \nu) = \mathcal{W}_{\psi_r}(X, X') := \inf \{ \|X - X'\|_{\psi_r} : \text{law}(X) = \mu, \text{law}(X') = \nu \}.$$

Lemma 2.2.5 *The following results hold*

1. $\| |X|^s \|_{\psi_r} = \|X\|_{\psi_{sr}}^s$ for $r, s > 0$.
2. $\|X\|_{\psi_r} \leq c_{r,s} \|X\|_{\psi_s}$ for $0 < r < s < \infty$ and some constant $c_{r,s} > 0$.
3. There is a constant $c > 0$ such that for i.i.d. Rademacher variables $(\xi_k)_{k \geq 1}$ it holds $\|\xi_1 + \dots + \xi_n\|_{\psi_2} \leq c\sqrt{n}$ for all $n \in \mathbb{N}^*$.

4. Let $\varepsilon \in]0, 1]$, $g : \mathbb{R} \rightarrow \mathbb{R}$ be ε -Hölder continuous, and $r > 0$. Then one has

$$\|g(X) - g(X')\|_{\psi_r} \leq \|g\|_\varepsilon \|X - X'\|_{\psi_{\varepsilon r}}^\varepsilon$$

for random variables $X, X' : \Omega \rightarrow \mathbb{R}$, and

$$\mathcal{W}_{\psi_r}(g(X), g(X')) \leq \|g\|_\varepsilon \mathcal{W}_{\psi_{\varepsilon r}}(X, X')^\varepsilon.$$

for random variables $X : \Omega \rightarrow \mathbb{R}$ and $X' : \Omega' \rightarrow \mathbb{R}$.

5. For $r \geq 1$ one has that $\mathcal{W}_{\psi_r}(\cdot, \cdot)$ is a distance.

Let $(X_k)_{k \geq 1}$ be an i.i.d. sequence of random variables with $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, and such that, for some $\sigma > 0$, $\mathbb{E}[e^{\sigma|X|}] < +\infty$. Let \mathcal{G} be a standard normal random variable. In [Rio09, Theorem 2.1], Emmanuel Rio proved that there exists a constant $C > 0$ such that, for $n \geq 1$,

$$\mathcal{W}_{\psi_1}(n^{-1/2}S_n, \mathcal{G}) \leq C n^{-1/2}, \quad \text{where } S_n = X_1 + \dots + X_n.$$

As a byproduct, for any $p \geq 1$, there exists a constant $c_p > 0$ such that

$$\mathcal{W}_p(n^{-1/2}S_n, \mathcal{G}) \leq c_p n^{-1/2}. \quad (2.57)$$

where \mathcal{W}_p stands for the L_p -Wasserstein distance

$$\mathcal{W}_p(\mu, \nu) = \mathcal{W}_p(X, X') := \inf \left\{ \mathbb{E}[|X - X'|^p]^{1/p} : \text{law}(X) = \mu, \text{law}(X') = \nu \right\}. \quad (2.58)$$

We have also the result of Kantorovich-Rubinstein, i.e.

$$\mathcal{W}_1(\mu, \nu) = \mathcal{W}_1(X, X') = \sup \{ \mathbb{E}[f(X)] - \mathbb{E}[f(X')] : \|f\|_{\text{Lip}} \leq 1 \}. \quad (2.59)$$

Remark 2.2.6 We could also consider here the case where $0 < p < 1$ by using the fact that, in this case, $\mathbb{E}[|X - X'|^p]$ is a distance (see the arguments in [LAS05, Section 7.1]). In general, we have $\mathcal{W}_p(\mu, \nu) \leq \mathcal{W}_q(\mu, \nu)$ for $0 < p < q < \infty$.

Let us start with a straightforward generalization of Rio's result.

Proposition 2.2.7 *There exists a $C > 0$ such that, for all $x \in \mathbb{R}$ and all $0 \leq t \leq s \leq T$,*

$$\mathcal{W}_{\psi_1}(W_s^{n,t,x}, W_s^{t,x}) \leq C \left(\frac{T}{n} \right)^{1/2}.$$

Then, taking into account (2.56), for any $p \geq 1$, there exists a $c_p > 0$ such that, for all $x \in \mathbb{R}$ and all $0 \leq t \leq s \leq T$,

$$\mathcal{W}_p(W_s^{n,t,x}, W_s^{t,x}) \leq c_p \left(\frac{T}{n} \right)^{1/2}. \quad (2.60)$$

Let us finish with a simple consequence of this result that we will use in the sequel.

Corollary 2.2.8 *Let $0 < \varepsilon \leq 1$ and let $g : \mathbb{R} \rightarrow \mathbb{R}$ be an ε -Hölder continuous function. Then there exists a $C > 0$ depending on T such that, for all $x \in \mathbb{R}$ and all $0 \leq t \leq s \leq T$,*

$$|\mathbb{E}[g(W_s^{n,t,x})] - \mathbb{E}[g(W_s^{t,x})]| \leq C \|g\|_\varepsilon n^{-\varepsilon/2},$$

and, setting $\delta(t, s) := \max(s - t, \underline{s} - \underline{t})$,

$$|\mathbb{E}[g(W_s^{n,t,x})(W_s^{n,t,x} - x)] - \mathbb{E}[g(W_s^{t,x})(W_s^{t,x} - x)]| \leq C \|g\|_\varepsilon \delta(t, s)^{1/2} n^{-\varepsilon/2}.$$

2.2.4 Regularity results on u , U^n , ∇u and Δ^n

Let us start by known regularity properties of the function u that follow from classical a priori estimates for BSDEs.

Lemma 2.2.9 *Under Assumption 2.4 there exists a constant $C > 0$ depending on (T, ε, f, g) such that, for all $(t, x) \in [0, T] \times \mathbb{R}$,*

$$|u(t, x)| \leq C(1 + |x|)^\varepsilon, \quad \|u(t, \cdot)\|_\varepsilon \leq C, \quad \|u(\cdot, x)\|_{\varepsilon/2} \leq C(1 + |x|)^\varepsilon.$$

The following Lemma extends [Zha05, Theorem 3.2] and Theorem 2.1.5 to the case where $f(t, \cdot, y, z)$ is ε -Hölder continuous and $f(\cdot, x, y, z)$ is α -Hölder continuous. We also assume that g is ε -Hölder continuous. Comparing to regularity results stated in Theorem 2.1.5, we also provide a bound for $\|\nabla_x u(t, \cdot)\|_\varepsilon$.

Lemma 2.2.10 *Recall the notation (2.41) and let Assumption 2.4 hold.*

1. *The function u belongs to $C^{0,1}([0, T[\times \mathbb{R})$ and, for all $(t, x) \in [0, T[\times \mathbb{R}$, we have,*

$$Z_s^{t,x} = \nabla_x u(s, W_s^{t,x}) \quad \text{for a.e. } (s, \omega) \in [t, T[\times \Omega, \quad (2.61)$$

as well as (2.43) i.e.

$$\nabla_x u(t, x) = \mathbb{E} \left(g(W_T^{t,x}) \frac{W_T - W_t}{T - t} \right) + \mathbb{E} \left(\int_t^T F(s, W_s^{t,x}) \frac{W_s - W_t}{s - t} ds \right).$$

Consequently, for $\mathbb{E}_r := \mathbb{E}[\cdot | \mathcal{F}_r]$,

$$\nabla_x u(r, W_r^{t,x}) = \mathbb{E}_r \left(g(W_T^{t,x}) \frac{W_T - W_r}{T - r} \right) + \mathbb{E}_r \left(\int_r^T F(s, W_s^{t,x}) \frac{W_s - W_r}{s - r} ds \right) \quad \text{a.s. for } r \in [t, T[. \quad (2.62)$$

2. *Moreover, there exists a constant $C > 0$ depending on (T, ε, f, g) such that*

- (a) $\|\nabla_x u(t, \cdot)\|_\varepsilon \leq \frac{C}{\sqrt{T-t}}$ for all $t \in [0, T[$,
- (b) $|\nabla_x u(t, x)| \leq \frac{C}{(T-t)^{(1-\varepsilon)/2}}$ for all $(t, x) \in [0, T[\times \mathbb{R}$.

Remark 2.2.11 From now on we always use the continuous version of $Z_s^{t,x}$ given by $\nabla u(s, W_s^{t,x})$.

Lemma 2.2.12 *For all $(t, x) \in [0, T[\times \mathbb{R}$ and for $n \geq n_0 \in \mathbb{N}^*$, with n_0 defined as in [L-1, Lemma 14], we have*

- 1. $|U^n(t, x)| \leq C(1 + |x|)^\varepsilon$,
- 2. $|\Delta^n(t, x)| \leq \frac{C_n}{(T-t)^{(1-\varepsilon)/2}}$,

where $C > 0$ depends on (T, ε, f, g) and $C_n > 0$ depends on $(T, \varepsilon, f, g, n)$.

Proposition 2.2.13 *Under Hypothesis 2.4, there exists a constant $C > 0$ depending on (T, ε, f, g) such that, for all $x \in \mathbb{R}$,*

$$|\nabla_x u(t, x) - \nabla_x u(r, x)| \leq C \frac{(t-r)^{\varepsilon/2}}{\sqrt{T-t}} \quad \text{for all } 0 \leq r < t < T.$$

2.2.5 Main results

Theorem 2.2.14 *Assume that Hypothesis 2.4 holds. Then there exists a constant $C_\varepsilon > 0$, depending at most on $(T, \alpha, \varepsilon, f, g)$ such that for all $x \in \mathbb{R}$,*

1. $\mathcal{W}_{\psi_{1/\varepsilon}}(Y_s^{n,t,x}, Y_s^{t,x}) \leq C_\varepsilon (1 + |x|)^\varepsilon n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $0 \leq t \leq s \leq T$,
2. $\mathcal{W}_{\psi_{1/\varepsilon}}(Z_s^{n,t,x}, Z_s^{t,x}) \leq C_\varepsilon \frac{(1+|x|)^\varepsilon}{\sqrt{T-s}} n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $s \in [t, T[$.

In particular, for any $p \in [1, \infty[$, there exists a constant $C_p > 0$, depending at most on $(T, \alpha, \varepsilon, f, g, p)$ such that for all $x \in \mathbb{R}$,

1. $\mathcal{W}_p(Y_s^{n,t,x}, Y_s^{t,x}) \leq C_p (1 + |x|)^\varepsilon n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $0 \leq t \leq s \leq T$,
2. $\mathcal{W}_p(Z_s^{n,t,x}, Z_s^{t,x}) \leq C_p \frac{(1+|x|)^\varepsilon}{\sqrt{T-s}} n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $s \in [t, T[$.

Proof. We give here an overview of the proof of Theorem 2.2.14.

$$\begin{aligned} \mathcal{W}_{\psi_r}(Y_s^{t,x}, Y_s^{n,t,x}) &= \mathcal{W}_{\psi_r}(u(s, W_s^{t,x}), U^n(s, W_s^{n,t,x})) \\ &\leq \mathcal{W}_{\psi_r}(u(s, W_s^{t,x}), u(s, W_s^{n,t,x})) + \mathcal{W}_{\psi_r}(u(s, W_s^{n,t,x}), U^n(s, W_s^{n,t,x})). \end{aligned}$$

The first term of the r.h.s. is bounded by Lemma 2.2.5-4, Lemma 2.2.9 and Proposition 2.2.7. The second term of the r.h.s. is bounded by using Theorem 2.2.16. As for the error on Y , we have

$$\begin{aligned} \mathcal{W}_{\psi_r}(\nabla u(s, W_s^{t,x}), \Delta^n(s, W_s^{n,t,x})) \\ \leq \mathcal{W}_{\psi_r}(\nabla u(s, W_s^{t,x}), \nabla u(s, W_s^{n,t,x})) + \mathcal{W}_{\psi_r}(\nabla u(s, W_s^{n,t,x}), \Delta^n(s, W_s^{n,t,x})). \end{aligned}$$

Lemma 2.2.5-4, Lemma 2.2.10 and Proposition 2.2.7 enable to bound the first term of the r.h.s.. The second term of the r.h.s. is bounded by using Theorem 2.2.16. If s does not belong to the time grid we have $\mathcal{W}_{\psi_r}(Z_s^{t,x}, Z_s^{n,t,x}) = \mathcal{W}_{\psi_r}(\nabla u(s, W_s^{t,x}), \Delta^n(s, W_s^{n,t,x}))$ and the result follows. If s belongs to the time grid we have $Z_s^{n,t,x} = \nabla^n U^n(s, W_{s-h/2}^{n,t,x}) = \Delta^n(s - h/2, W_{s-h/2}^{n,t,x})$. Then we first write

$$\mathcal{W}_{\psi_r}(Z_s^{t,x}, Z_s^{n,t,x}) = \mathcal{W}_{\psi_r}(Z_s^{t,x}, Z_{s-h/2}^{n,t,x}) \leq \mathcal{W}_{\psi_r}(Z_s^{t,x}, Z_{s-h/2}^{t,x}) + \mathcal{W}_{\psi_r}(Z_{s-h/2}^{t,x}, Z_{s-h/2}^{n,t,x}). \quad \blacksquare$$

Since $s - \frac{h}{2}$ does not belong to the time grid, the second term of the r.h.s. is bounded by using the previous result. We use Lemma 2.2.10 and Proposition 2.2.13 to bound the first term.

Remark 2.2.15

1. Esseen [Ess58] proved (we write this using the notation introduced in (2.57)) that

$$\lim_{n \rightarrow \infty} \sqrt{n} \mathcal{W}_1(n^{-1/2} S_n, \mathcal{G}) = A,$$

where A is a constant depending on the distribution of X_1 . Zolotarev provided in [Zol76, Theorem 1] a lower estimate for A . If $X_1 = \xi_1$, that is for Rademacher variables, one has $A(\xi_1) \geq 1/2$. This implies that for $g(x) = x$ and $f \equiv 0$ we have

$$\lim_{n \rightarrow \infty} \sqrt{n} \mathcal{W}_1(Y_T^{n,0,0}, Y_T^{0,0}) \geq \frac{1}{2} T^{\frac{1}{2}}.$$

In particular this result ensures that the rate we get is optimal.

2. Dong and Krylov have shown in [DK05, Theorem 2.18], if f is linear in y and z , that the rate in Theorem 2.2.16 (1) can be of order $\frac{1}{n}$ provided that g and f are bounded and sufficiently smooth. Their results also include the case where the forward process is a diffusion, again with bounded and sufficiently smooth coefficients.
3. Theorem 2.2.16 gives a probabilistic proof of the point-wise convergence of the finite difference scheme, for u and $\nabla_x u$, when u solves a semi-linear and non-degenerate equation, with Laplacian smoothing. Deterministic proofs in a general setting of this convergence can be found in the literature for the solution u . For example, Crandall and Lions [CL84] get a convergence order of $1/2$ in the case of Hamilton-Jacobi equations, which is consistent with our result.

Theorem 2.2.14 is a consequence of Theorem 2.2.16, which gives the rate of the point-wise convergence of U^n , the solution to (2.48), towards the solution u to the semilinear heat equation (2.38), as well as the point-wise convergence of the gradients. Theorem 2.2.16 is a key step to our main result and might be of interest in its own right, in particular for its probabilistic proof.

Theorem 2.2.16 *Under Assumption 2.4 there exists a constant $C > 0$ depending at most on $(T, \alpha, \varepsilon, f, g)$ such that*

1. $|u(t, x) - U^n(t, x)| \leq C(1 + |x|)^\varepsilon n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $(t, x) \in \mathbb{R} \times [0, T]$,
2. $|\nabla_x u(t, x) - \Delta^n(t, x)| \leq C \frac{(1+|x|)^\varepsilon}{\sqrt{T-t}} n^{-(\alpha \wedge \frac{\varepsilon}{2})}$ for all $(t, x) \in \mathbb{R} \times [0, T[$.

Proof. We split the proof into three parts. We begin by studying $|u - U^n|$, we proceed by obtaining an estimate for $\nabla u - \Delta^n$, and then we conclude with a Gronwall argument.

From (2.42) and (2.54) we conclude that

$$\begin{aligned} |u(t, x) - U^n(t, x)| &\leq \left| \mathbb{E} \left[g \left(W_T^{t,x} \right) \right] - \mathbb{E} \left[g \left(W_T^{n,t,x} \right) \right] \right| \\ &\quad + \left| \mathbb{E} \left[\int_t^T F(s, W_s^{t,x}) ds \right] - \mathbb{E} \left[\int_t^T f(\bar{s}, \Theta_s^{n,t,x}) ds \right] \right|. \end{aligned} \quad (2.63)$$

We use Corollary 2.2.8 to bound the first term. To bound the second term of the r.h.s., we introduce the function F_n given by

$$F_n(s, x) := f(s, x, U^n(s, x), \Delta^n(s, x)) \quad \text{for } (s, x) \in [0, T[\times \mathbb{R}. \quad (2.64)$$

Using the notation (2.55) we also have that $F_n(s, W_s^{n,t,x}) = f(s, \Theta_s^{n,t,x})$. Since f is α -Hölder continuous in time and thanks to a priori estimates on $(Y^{n,t,x}, Z^{n,t,x})$ (see [L-1, Lemma 14]), bounding the second term of the r.h.s. of (2.63) boils down to bound $\left| \int_t^T \mathbb{E} \left[F(s, W_s^{t,x}) \right] ds - \int_t^T \mathbb{E} \left[F_n(s, W_s^{n,t,x}) \right] ds \right|$. We have

$$\begin{aligned} &\left| \int_t^T \mathbb{E} \left[F(s, W_s^{t,x}) \right] ds - \int_t^T \mathbb{E} \left[F_n(s, W_s^{n,t,x}) \right] ds \right| \\ &\leq \int_t^T \left| \mathbb{E} \left[F(s, W_s^{t,x}) \right] - \mathbb{E} \left[F(s, W_s^{n,t,x}) \right] \right| ds + \int_t^T \mathbb{E} \left[|F - F_n|(s, W_s^{n,t,x}) \right] ds. \end{aligned}$$

Since $\|F(s, \cdot)\|_\varepsilon \leq C \frac{1}{\sqrt{T-s}}$, Corollary 2.2.8 bounds the first term. Coming back to the definition of F and F_n and using the Lipschitz continuity of f , we have

$$|F - F_n|(s, W_s^{n,t,x}) \leq \|f\|_{\text{Lip}} \left[|u - U^n|(s, W_s^{n,t,x}) + |\nabla u - \Delta^n|(s, W_s^{n,t,x}) \right].$$

Setting for simplicity, for $s \in [0, T]$,

$$\beta_n(s) := \sup_{x \in \mathbb{R}} \left\{ \frac{|u - U^n|(s, x)}{(1 + |x|)^\epsilon} \right\} \quad \text{and} \quad \gamma_n(s) := \sup_{x \in \mathbb{R}} \left\{ \frac{|\nabla u - \Delta^n|(s, x)}{(1 + |x|)^\epsilon} \right\}$$

for $s \in [0, T]$, (2.63) becomes by using Lemmas 2.2.9, 2.2.10 and 2.2.12

$$\beta_n(t) \leq C \left(n^{-(\alpha \wedge \frac{\epsilon}{2})} + \int_t^T (\beta_n(s) + \gamma_n(s)) ds \right), \quad t \in [0, T],$$

and since γ_n belongs to $L^1[0, T]$, Gronwall's inequality gives

$$\beta_n(t) \leq C \left(n^{-(\alpha \wedge \frac{\epsilon}{2})} + \int_t^T \gamma_n(s) ds \right), \quad t \in [0, T]. \quad (2.65)$$

In order to take advantage of the previous inequality, we need to estimate $\gamma_n(s)$. To do this, we use the representation (2.43) and the following equality (see [L-1, Equation (33)])

$$\Delta^n(t, x) = \mathbb{E} \left[g(B_T^{n,t,x}) \frac{W_T^{n,t,x} - x}{T - t} \right] + \mathbb{E} \left[\int_{\underline{t}+h}^T f(\bar{s}, \Theta_s^{n,t,x}) \frac{W_s^{n,t,x} - x}{\underline{s} - \underline{t}} ds \right].$$

Then we divide the study into two parts

$$|\nabla u(t, x) - \Delta^n(t, x)| \leq |g \text{ difference}| + |f \text{ difference}|.$$

The g difference is bounded by using Corollary 2.2.8. The f difference is bounded thanks to the regularity of F and Corollary 2.2.8. This boils down to

$$\gamma_n(t) \leq C \left(\frac{n^{-(\alpha \wedge \frac{\epsilon}{2})}}{\sqrt{T-t}} + \int_{(\underline{t}+h) \wedge T}^T (\beta_n(s) + \gamma_n(s)) \frac{ds}{\sqrt{\underline{s} - \underline{t}}} \right), \quad t \in [0, T].$$

Using (2.65) and Gronwall's lemma ends the proof. \blacksquare

2.3 Case of Doubly reflected BSDEs with RCLL obstacles

This Section presents the results established in [L-8] and [L-9] written with Roxana Dumitrescu. We study a discrete time approximation scheme for the solution to doubly reflected Backward Stochastic Differential Equations (DBBSDEs in short) driven by a Brownian motion and a Poisson random process which are mutually independent. We consider the non-markovian case. Moreover, the barriers are supposed to be right-continuous and left-limited (RCLL in short) processes, whose jumps are arbitrary, they can be either predictable or inaccessible. We consider the following equation:

$$\left\{ \begin{array}{l} \text{(i)} \ Y_t = \zeta_T + \int_t^T f(s, Y_s, Z_s, U_s) ds + (A_T - A_t) - (K_T - K_t) - \int_t^T Z_s dW_s - \int_t^T U_s d\tilde{N}_s, \\ \text{(ii)} \ \forall t \in [0, T], \zeta_t \leq Y_t \leq \bar{\zeta}_t \text{ a.s.}, \\ \text{(iii)} \ \int_0^T (Y_{t-} - \zeta_{t-}) dA_t^c = 0 \text{ a.s. and } \int_0^T (\bar{\zeta}_{t-} - Y_{t-}) dK_t^c = 0 \text{ a.s.} \\ \text{(iv)} \ \forall \tau \text{ predictable stopping time, } \Delta A_\tau^d = \Delta A_\tau^d \mathbf{1}_{Y_{\tau-} = \zeta_{\tau-}} \text{ and } \Delta K_\tau^d = \Delta K_\tau^d \mathbf{1}_{Y_{\tau-} = \bar{\zeta}_{\tau-}}. \end{array} \right. \quad (2.66)$$

Here, A^c (resp. K^c) denotes the continuous part of A (resp. K) and A^d (resp. K^d) its discontinuous part, $\{W_t : 0 \leq t \leq T\}$ is a one dimensional standard Brownian motion and

$\{\tilde{N}_t := N_t - \lambda t, 0 \leq t \leq T\}$ is a compensated Poisson process. Both processes are independent and they are defined on the probability space $(\Omega, \mathcal{F}_T, \mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$. The processes A and K have the role to keep the solution between the two obstacles $\underline{\zeta}$ and $\bar{\zeta}$. Since we consider that the jumps of the obstacles can be either predictable or totally inaccessible, A and K are also discontinuous.

In the case of a Brownian filtration, non-linear backward stochastic differential equations were introduced by Pardoux and Peng [PP90]. One barrier reflected BSDEs were firstly studied by El Karoui et al in [EKKP⁺97]. In their setting, one of the components of the solution was forced to stay above a given barrier which is a continuous adapted stochastic process. The main motivation of such equations is the pricing of American options especially in constrained markets. The generalization to the case of two reflecting barriers was carried out by Cvitanic and Karatzas in [CK96]. It is also well known that doubly reflected BSDEs are related to Dynkin games and in finance to the pricing of Israeli options (or Game options). The case of standard BSDEs with jump processes driven by a compensated Poisson random measure was first considered by Tang and Li in [TL94]. The extension to the case of reflected BSDEs and one reflecting barrier with only inaccessible jumps was established by Hamadène and Ouknine [HO03]. Later on, Essaky in [Ess08] and Hamadène and Ouknine in [HO16] had extended these results to a RCLL obstacle with predictable and inaccessible jumps. Results concerning existence and uniqueness of the solution for doubly reflected BSDEs with jumps can be found in [CM08],[DQS16], [HH06], [HW09] and [EHO05].

Numerical schemes for DBBSDEs driven by the Brownian motion and based on a random tree method were proposed by Xu in [Xu11] (see also [MPX08] and [PX11]) and, in the Markovian framework, by Chassagneux in [Cha09]. In the case of a filtration driven also by a Poisson process, some results were provided only in the non-reflected case. In [BE08], the authors proposed a scheme for Forward-Backward SDEs based on the dynamic programming equation and in [LMT13] the authors proposed a fully implementable scheme based on a random binomial tree.

Our aim is to propose an implementable numerical method to approximate the solution to DBBSDEs with jumps and RCLL obstacles (2.66). As for standard BSDEs, the computation of conditional expectations is an important issue. Since we consider reflected BSDEs, we also have to modelize the constraints. In the first part of this work (corresponding to [L-8]) we consider the following approximations

- we approximate the Brownian motion and the Poisson process by two independent random walks,
- we introduce a sequence of penalized BSDEs to approximate the reflected BSDE.

These approximations enable us to provide a fully implementable scheme, called *explicit penalized discrete scheme* in the following : $\bar{y}_n^{p,n} := \underline{\zeta}_n^n$ and for $j = n - 1, \dots, 0$

$$\begin{cases} \bar{y}_j^{p,n} = \mathbb{E}[\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n] + f(t_j, \mathbb{E}(\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n), \bar{z}_j^{p,n}, \bar{u}_j^{p,n})h + \bar{a}_j^{p,n} - \bar{k}_j^{p,n}, \\ \bar{a}_j^{p,n} = \frac{ph}{1+ph} \left(\mathbb{E}[\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n], \bar{z}_j^{p,n}, \bar{u}_j^{p,n}) - \underline{\zeta}_j^n \right)^-, \\ \bar{k}_j^{p,n} = \frac{ph}{1+ph} \left(\bar{\zeta}_j^n - \mathbb{E}[\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n] - f(t_j, \mathbb{E}[\bar{y}_{j+1}^{p,n} | \mathcal{F}_j^n], \bar{z}_j^{p,n}, \bar{u}_j^{p,n}) \right)^-, \\ \bar{z}_j^{p,n} = \frac{1}{\sqrt{h}} \mathbb{E}(\bar{y}_{j+1}^{p,n} \xi_{j+1} | \mathcal{F}_j^n), \\ \bar{u}_j^{p,n} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(\bar{y}_{j+1}^{p,n} \eta_{j+1} | \mathcal{F}_j^n). \end{cases}$$

We prove in Theorem 2.3.12 that the scheme weakly converges to the solution to (2.66). Moreover, in order to prove the convergence of our scheme, we prove, in the case of jump processes driven by a general Poisson random measure, that the solutions to the penalized equations converge to the solution of the doubly reflected BSDE in the case of a driver depending on the solution, which was not the case in the previous literature (see [EHO05], [HH06], [HW09]). This gives another proof for the existence of a solution to DBBSDEs with jumps and RCLL barriers. Our method is based on a combination of penalization, Snell envelope theory, comparison theorem for BSDEs with jumps (see [QS13], [QS14]) and a generalized monotonic theorem under the Mokobodski's condition. It extends [LX07] to the case of DBBSDEs with totally inaccessible jumps. Finally, we illustrate our theoretical results with some numerical simulations in the case of general jumps.

In a second part of this work (corresponding to [L-9]), we propose an alternative scheme to [L-8] to solve (2.66). The scheme proposed here takes the following form: $\bar{y}_n^n := \underline{\zeta}_n^n$ and for $j = n-1, \dots, 0$

$$\begin{cases} \bar{y}_j^n = \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) \delta + \bar{a}_j^n - \bar{k}_j^n, \\ \bar{a}_j^n \geq 0, \bar{k}_j^n \geq 0, \bar{a}_j^n \bar{k}_j^n = 0, \\ \underline{\zeta}_j^n \leq \bar{y}_j^n \leq \bar{\zeta}_j^n, (\bar{y}_j^n - \underline{\zeta}_j^n) \bar{a}_j^n = (\bar{y}_j^n - \bar{\zeta}_j^n) \bar{k}_j^n = 0. \end{cases} \quad (2.67)$$

It generalizes the scheme proposed by [Xu11] to the case of jumps. Compared to the scheme proposed in [L-8], the scheme proposed here —called *reflected scheme* in the following—is based on the direct discretization of (2.66). In particular, there is no penalization step. Then, this method only depends on one parameter of approximation (the number of time steps n), contrary to the scheme proposed in [L-8], which also depends on the penalization parameter. We provide here an *explicit reflected scheme* and an *implicit reflected scheme* and we show the convergence of both schemes. We illustrate numerically the theoretical results and show they coincide with the ones obtained by using the penalized scheme presented in [L-8], for large values of the penalization parameter.

The rest of the Section is organized as follows: Section 2.3.1 presents the general framework, Section 2.3.2 presents the discrete time approximation, Section 2.3.3 presents the penalized scheme and establish its convergence and Section 2.3.4 presents the reflected scheme and establish its convergence.

2.3.1 General framework

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and \mathcal{P} be the predictable σ -algebra on $[0, T] \times \Omega$. Let W be a one-dimensional Brownian motion and $N(dt, de)$ be a Poisson random measure with compensator $\nu(de)dt$ such that ν is a σ -finite measure on \mathbb{R}^* , equipped with its Borel field $\mathcal{B}(\mathbb{R}^*)$. Let $\tilde{N}(dt, du)$ be its compensated process. Let $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$ be the natural filtration associated with W and N .

For each $T > 0$, we use the following notations:

- L_ν^2 is the set of Borelian functions $\ell : \mathbb{R}^* \rightarrow \mathbb{R}$ such that $\int_{\mathbb{R}^*} |\ell(u)|^2 \nu(du) < +\infty$.
The set L_ν^2 is a Hilbert space equipped with the scalar product $\langle \delta, \ell \rangle_\nu := \int_{\mathbb{R}^*} \delta(u) \ell(u) \nu(du)$ for all $\delta, \ell \in L_\nu^2$, and the norm $\|\ell\|_\nu^2 := \int_{\mathbb{R}^*} |\ell(u)|^2 \nu(du)$.
- $\mathcal{B}(\mathbb{R}^2)$ (resp $\mathcal{B}(L_\nu^2)$) is the Borelian σ -algebra on \mathbb{R}^2 (resp. on L_ν^2).
- H^2 is the set of real-valued predictable processes ϕ such that $\|\phi\|_{H^2}^2 := \mathbb{E}(\int_0^T \phi_t^2 dt) < +\infty$.
With the notations given in the Introduction we have $H^2 = H_T^2(\mathbb{R})$.

- H_ν^2 is the set of processes l which are *predictable*, that is, measurable

$$l : ([0, T] \times \Omega \times \mathbb{R}^*, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^*)) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})); \quad (\omega, t, u) \mapsto l_t(\omega, u)$$

$$\text{such that } \|l\|_{H_\nu^2}^2 := \mathbb{E} \left[\int_0^T \|l_t\|_\nu^2 dt \right] < \infty.$$

- \mathcal{S}^2 is the set of real-valued RCLL adapted processes ϕ such that $\|\phi\|_{\mathcal{S}^2}^2 := \mathbb{E}(\sup_{0 \leq t \leq T} |\phi_t|^2) < \infty$. With the notations given in the Introduction we have $\mathcal{S}^2 = \mathcal{S}_T^2(\mathbb{R})$.
- \mathcal{A}^2 is the set of real-valued non decreasing RCLL predictable processes A with $A_0 = 0$ and $\mathbb{E}(A_T^2) < \infty$.
- \mathcal{T}_0 is the set of stopping times τ such that $\tau \in [0, T]$ a.s
- For S in \mathcal{T}_0 , \mathcal{T}_S is the set of stopping times τ such that $S \leq \tau \leq T$ a.s.
- For each RCLL adapted process $\phi = (\phi_t)_{0 \leq t \leq T}$ with $\phi^- \in \mathcal{S}^2$, we denote by $\mathcal{R}(\phi)$ the Snell envelope of ϕ , defined as the minimal RCLL supermartingale greater or equal to ϕ a.s.

Definition 2.3.1 (Driver, Lipschitz driver) A function f is said to be a driver if

- $f : \Omega \times [0, T] \times \mathbb{R}^2 \times L_\nu^2 \rightarrow \mathbb{R}$
 $(\omega, t, y, z, \kappa(\cdot)) \mapsto g(\omega, t, y, z, \kappa(\cdot))$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^2) \otimes \mathcal{B}(L_\nu^2)$ -measurable,
- $\|f(\cdot, 0, 0, 0)\|_\infty < \infty$.

A driver f is called a Lipschitz driver if moreover there exists a constant $C_f \geq 0$ and a bounded, non-decreasing continuous function Λ with $\Lambda(0) = 0$ such that $d\mathbb{P} \otimes dt$ -a.s., for each $(s_1, y_1, z_1, k_1), (s_2, y_2, z_2, k_2)$,

$$|f(\omega, s_1, y_1, z_1, k_1) - f(\omega, s_2, y_2, z_2, k_2)| \leq \Lambda(|s_2 - s_1|) + C_f(|y_1 - y_2| + |z_1 - z_2| + \|k_1 - k_2\|_\nu).$$

Definition 2.3.2 (Mokobodski's condition) Let $\underline{\zeta}, \bar{\zeta}$ in \mathcal{S}^2 . There exist two nonnegative RCLL supermartingales H and H' in \mathcal{S}^2 such that

$$\forall t \in [0, T], \quad \underline{\zeta}_t \mathbf{1}_{t < T} \leq H_t - H'_t \leq \bar{\zeta}_t \mathbf{1}_{t < T} \text{ a.s.}$$

Hypothesis 2.5 A Lipschitz driver g is said to satisfy Assumption 2.5 if the following holds : $d\mathbb{P} \otimes dt$ a.s. for each $(y, z, k_1, k_2) \in \mathbb{R}^2 \times (L_\nu^2)^2$, we have

$$f(t, y, z, k_1) - f(t, y, z, k_2) \geq \langle \theta_t^{y, z, k_1, k_2}, k_1 - k_2 \rangle_\nu,$$

with

$$\begin{aligned} \theta : \Omega \times [0, T] \times \mathbb{R}^2 \times (L_\nu^2)^2 &\mapsto L_\nu^2; \\ (\omega, t, y, z, k_1, k_2) &\mapsto \theta_t^{y, z, k_1, k_2}(\omega, \cdot) \end{aligned}$$

$\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^2) \times \mathcal{B}((L_\nu^2)^2)$ -measurable, bounded, and satisfying $d\mathbb{P} \otimes dt \otimes d\nu(u)$ -a.s., for each $(y, z, k_1, k_2) \in \mathbb{R}^2 \times (L_\nu^2)^2$,

$$\theta_t^{y, z, k_1, k_2}(u) \geq -1 \text{ and } |\theta_t^{y, z, k_1, k_2}(u)| \leq \psi(u),$$

where $\psi \in L_\nu^2$.

Remark 2.3.3 This assumption ensures the comparison theorem for BSDEs with jumps (see [QS13, Theorem 4.2]), which extends the result of [Roy06].

Hypothesis 2.6 $\underline{\zeta}$ and $\bar{\zeta}$ are two adapted RCLL processes with $\underline{\zeta}_T = \bar{\zeta}_T$ a.s., $\underline{\zeta} \in \mathcal{S}^2$, $\bar{\zeta} \in \mathcal{S}^2$, $\underline{\zeta}_t \leq \bar{\zeta}_t$ for all $t \in [0, T]$, the Mokobodski's condition holds and g is a Lipschitz driver satisfying Assumption 2.5.

We introduce the following general reflected BSDE with jumps and two RCLL obstacles

Definition 2.3.4 Let $T > 0$ be a fixed terminal time and g be a Lipschitz driver. Let $\underline{\zeta}$ and $\bar{\zeta}$ be two adapted RCLL processes with $\underline{\zeta}_T = \bar{\zeta}_T$ a.s., $\underline{\zeta} \in \mathcal{S}^2$, $\bar{\zeta} \in \mathcal{S}^2$, $\underline{\zeta}_t \leq \bar{\zeta}_t$ for all $t \in [0, T]$ a.s. A process (Y, Z, U, α) is said to be a solution to the double barrier reflected BSDE (DBBSDE) associated with driver f and barriers $\underline{\zeta}, \bar{\zeta}$ if

$$\begin{cases} (i) Y \in \mathcal{S}^2, Z \in \mathbb{H}^2, U \in \mathbb{H}_v^2 \text{ and } \alpha \in \mathcal{S}^2, \text{ where } \alpha = A - K \text{ with } A, K \text{ in } \mathcal{A}^2 \\ (ii) Y_t = \underline{\zeta}_T + \int_t^T f(s, Y_s, Z_s, U_s) ds + (A_T - A_t) - (K_T - K_t) - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbb{R}^*} U_s(e) \tilde{N}(ds, de), \\ (iii) \forall t \in [0, T], \underline{\zeta}_t \leq Y_t \leq \bar{\zeta}_t \text{ a.s.}, \\ (iv) \int_0^T (Y_{t-} - \underline{\zeta}_{t-}) dA_t = 0 \text{ a.s. and } \int_0^T (\bar{\zeta}_{t-} - Y_{t-}) dK_t = 0 \text{ a.s.} \end{cases} \quad (2.68)$$

Remark 2.3.5 Condition (iv) is equivalent to the following condition : if $K = K^c + K^d$ and $A = A^c + A^d$, where K^c (resp. K^d) represents the continuous (resp. the discontinuous) part of K (the same notation holds for A), then

$$\int_0^T (Y_t - \underline{\zeta}_t) dA_t^c = 0 \text{ a.s.}, \quad \int_0^T (\bar{\zeta}_t - Y_t) dK_t^c = 0 \text{ a.s.}$$

and

$$\forall \tau \in \mathcal{T}_0 \text{ predictable, } \Delta A_\tau^d = \Delta A_\tau^d \mathbf{1}_{Y_{\tau-} = \underline{\zeta}_{\tau-}} \text{ and } \Delta K_\tau^d = \Delta K_\tau^d \mathbf{1}_{Y_{\tau-} = \bar{\zeta}_{\tau-}}.$$

Theorem 2.3.6 ([DQS16, Theorem 4.1]) Suppose $\underline{\zeta}$ and $\bar{\zeta}$ are RCLL adapted processes in \mathcal{S}^2 such that for all $t \in [0, T]$, $\underline{\zeta}_t \leq \bar{\zeta}_t$ and Mokobodski's condition holds. Then, DBBSDE (2.68) admits a unique solution (Y, Z, U, α) in $\mathcal{S}^2 \times \mathbb{H}^2 \times \mathbb{H}_v^2 \times \mathcal{A}^2$.

Remark 2.3.7 As said in [DQS16, Remark 4.3], if for all $t \in]0, T]$ $\underline{\zeta}_{t-} < \bar{\zeta}_{t-}$ a.s., [DQS16, Proposition 4.2] gives the uniqueness of $A, K \in (\mathcal{A}^2)^2$.

2.3.2 Discrete time Approximation

We adopt the framework of [LMT13], presented below.

Random walk approximation of (W, \tilde{N}) For $n \in \mathbb{N}$, we introduce $h := \frac{T}{n}$ and the regular grid $(t_j)_{j=0, \dots, n}$ with step size h (i.e. $t_j := jh$) to discretize $[0, T]$. In order to approximate W , we recall the random walk approximation

$$\begin{cases} W_0^n = 0 \\ W_t^n = \sqrt{h} \sum_{i=1}^{\lfloor t/h \rfloor} \xi_i \end{cases} \quad (2.69)$$

where $\xi_1, \xi_2, \dots, \xi_n$ are independent identically distributed random variables with the following Rademacher law:

$$P(\xi_1 = 1) = P(\xi_1 = -1) = \frac{1}{2}.$$

To approximate \tilde{N} , we introduce a second random walk

$$\begin{cases} \tilde{N}_0^n = 0 \\ \tilde{N}_t^n = \sum_{i=1}^{\lfloor t/h \rfloor} \eta_i \end{cases} \quad (2.70)$$

where $\eta_1, \eta_2, \dots, \eta_n$ are independent and identically distributed random variables with law

$$P(\eta_1 = \kappa_n - 1) = 1 - P(\eta_1 = \kappa_n) = \kappa_n$$

where $\kappa_n = e^{-\frac{\lambda}{n}}$. We assume that both sequences ξ_1, \dots, ξ_n and $\eta_1, \eta_2, \dots, \eta_n$ are defined on the original probability space $(\Omega, \mathbb{F}, \mathbb{P})$. The (discrete) filtration in the probability space is $\mathbb{F}^n = \{\mathcal{F}_j^n : j = 0, \dots, n\}$ with $\mathcal{F}_0^n = \{\Omega, \emptyset\}$ and $\mathcal{F}_j^n = \sigma\{\xi_1, \dots, \xi_j, \eta_1, \dots, \eta_j\}$ for $j = 1, \dots, n$.

The following result states the convergence of (W^n, \tilde{N}^n) to (W, \tilde{N}) for the J_1 -Skorokhod topology, and the convergence of W^n to W in any L^p , $p \geq 1$, for the topology of uniform convergence on $[0, T]$. We refer to [LMT13, Section 3] for more results on the convergence in probability of \mathcal{F}^n -martingales.

Lemma 2.3.8 ([LMT13, Lemma3, (III)], and [BDM01, Proof of Corollary 2.2]) *The couple (W^n, \tilde{N}^n) converges in probability to (W, \tilde{N}) for the J_1 -Skorokhod topology, and*

$$\sup_{0 \leq t \leq T} |W_t^n - W_t| \rightarrow 0 \text{ as } n \rightarrow \infty$$

in probability and in L^p , for any $1 \leq p < \infty$.

Martingale representation Let y_{j+1} denote a \mathcal{F}_{j+1}^n -measurable random variable. As said in [LMT13], we need a set of three strongly orthogonal martingales to represent the martingale difference $m_{j+1} := y_{j+1} - \mathbb{E}(y_{j+1} | \mathcal{F}_j^n)$. We introduce a third martingale increments sequence $\{\mu_j = \xi_j \eta_j, j = 0, \dots, n\}$. In this context there exists a unique triplet (z_j, u_j, v_j) of \mathcal{F}_j^n -random variables such that

$$m_{j+1} := y_{j+1} - \mathbb{E}(y_{j+1} | \mathcal{F}_j^n) = \sqrt{h} z_j \xi_{j+1} + u_j \eta_{j+1} + v_j \mu_{j+1},$$

and

$$\begin{cases} z_j = \frac{1}{\sqrt{h}} \mathbb{E}(y_{j+1} \xi_{j+1} | \mathcal{F}_j^n), \\ u_j = \frac{\mathbb{E}(y_{j+1} \eta_{j+1} | \mathcal{F}_j^n)}{\mathbb{E}((\eta_{j+1})^2 | \mathcal{F}_j^n)} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(y_{j+1} \eta_{j+1} | \mathcal{F}_j^n), \\ v_j = \frac{\mathbb{E}(y_{j+1} \mu_{j+1} | \mathcal{F}_j^n)}{\mathbb{E}((\mu_{j+1})^2 | \mathcal{F}_j^n)} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(y_{j+1} \mu_{j+1} | \mathcal{F}_j^n) \end{cases} \quad (2.71)$$

Remark 2.3.9 (Computing the conditional expectations) Let Φ denote a function from \mathbb{R}^{2j+2} to \mathbb{R} . We use the following formula to compute the conditional expectations

$$\begin{aligned} \mathbb{E}(\Phi(\xi_1, \dots, \xi_{j+1}, \eta_1, \dots, \eta_{j+1}) | \mathcal{F}_j^n) &= \frac{\kappa_n}{2} \Phi(\xi_1, \dots, \xi_j, 1, \eta_1, \dots, \eta_j, \kappa_n - 1) \\ &+ \frac{\kappa_n}{2} \Phi(\xi_1, \dots, \xi_j, -1, \eta_1, \dots, \eta_j, \kappa_n - 1) \\ &+ \frac{1 - \kappa_n}{2} \Phi(\xi_1, \dots, \xi_j, 1, \eta_1, \dots, \eta_j, \kappa_n) \\ &+ \frac{1 - \kappa_n}{2} \Phi(\xi_1, \dots, \xi_j, -1, \eta_1, \dots, \eta_j, \kappa_n). \end{aligned}$$

2.3.3 Penalized scheme

In order to propose an implementable numerical scheme we consider that the Poisson random measure is simply generated by the jumps of a Poisson process. We consider a Poisson process $\{N_t : 0 \leq t \leq T\}$ with intensity λ and jumps times $\{\tau_k : k = 0, 1, \dots\}$. The random measure is then

$$\tilde{N}(dt, de) = \sum_{k=1}^{N_t} \delta_{\tau_k, 1}(dt, de) - \lambda dt \delta_1(de)$$

where δ_a denotes the Dirac function at the point a . In the following, $\tilde{N}_t := N_t - \lambda t$. Then, the unknown fonction $U_s(e)$ does not depend on the magnitude e anymore, and we write $U_s := U_s(1)$.

In this particular case, (2.68) becomes:

$$\left\{ \begin{array}{l} \text{(i) } Y \in \mathcal{S}^2, Z \in \mathbb{H}^2, U \in \mathbb{H}^2 \text{ and } \alpha \in \mathcal{S}^2, \text{ where } \alpha = A - K \text{ with } A, K \text{ in } \mathcal{A}^2 \\ \text{(ii) } Y_t = \underline{\zeta} + \int_t^T f(s, Y_s, Z_s, U_s) ds + (A_T - A_t) - (K_T - K_t) - \int_t^T Z_s dW_s - \int_t^T U_s d\tilde{N}_s, \\ \text{(iii) } \forall t \in [0, T], \underline{\zeta}_t \leq Y_t \leq \bar{\zeta}_t \text{ a.s.}, \\ \text{(iv) } \int_0^T (Y_{t-} - \underline{\zeta}_{t-}) dA_t = 0 \text{ a.s. and } \int_0^T (\bar{\zeta}_{t-} - Y_{t-}) dK_t = 0 \text{ a.s.} \end{array} \right. \quad (2.72)$$

In view of the proof of the convergence of the numerical scheme, we also introduce the penalized version of (2.72):

$$Y_t^p = \underline{\zeta} + \int_t^T f(s, Y_s^p, Z_s^p, U_s^p) ds + A_T^p - A_t^p - (K_T^p - K_t^p) - \int_t^T Z_s^p dW_s - \int_t^T U_s^p d\tilde{N}_s, \quad (2.73)$$

with $A_t^p := p \int_0^t (Y_s^p - \underline{\zeta}_s)^- ds$ and $K_t^p := p \int_0^t (\bar{\zeta}_s - Y_s^p)^- ds$, and $\alpha_t^p := A_t^p - K_t^p$ for all $t \in [0, T]$.

2.3.3.1 Numerical scheme

The basic idea is to approximate the Brownian motion and the Poisson process by random walks based on the binomial tree model. As explained in Section 2.3.2, these approximations enable to get a martingale representation whose coefficients, involving conditional expectations, can be easily computed. Then, we approximate (W, \tilde{N}) in the penalized version of our DBBSDE (i.e. in (2.73)) by using these random walks. Taking conditional expectation and using the martingale representation leads to the *explicit penalized discrete scheme* (2.78). In view of the proof of the convergence of this explicit scheme, we introduce an implicit intermediate scheme (2.75).

Fully implementable numerical scheme In this Section we present two numerical schemes to approximate the solution to the penalized equation (2.73): the first one, (2.75), is an implicit intermediate scheme, useful for the proof of convergence. We denote it $(y_j^{p,n}, z_j^{p,n}, u_j^{p,n}, v_j^{p,n}, a_j^{p,n}, k_j^{p,n})_{j=0, \dots, n}$. We also introduce the main scheme (2.78), which is explicit. We denote it $(\bar{y}_j^{p,n}, \bar{z}_j^{p,n}, \bar{u}_j^{p,n}, \bar{v}_j^{p,n}, \bar{a}_j^{p,n}, \bar{k}_j^{p,n})_{j=0, \dots, n}$ in the following. The implicit scheme (2.75) is not easy to solve numerically, since it involves to inverse a function, as we will see below. However, it plays an important role in the proof of the convergence of the explicit scheme, that's why we introduce it.

In both schemes, we approximate the barrier $(\zeta_t)_t$ (resp. $(\bar{\zeta}_t)_t$) by $(\zeta_j^n)_{j=0,\dots,n}$ (resp. $(\bar{\zeta}_j^n)_{j=0,\dots,n}$). We also introduce their continuous time versions:

$$\widehat{\zeta}_t^n := \zeta_{[t/\delta]}^n, \quad \widehat{\bar{\zeta}}_t^n := \bar{\zeta}_{[t/\delta]}^n.$$

These approximations satisfy

Hypothesis 2.7

(i) For some $r > 2$, $\sup_{n \in \mathbb{N}} \max_{j \leq n} \mathbb{E}(|\zeta_j^n|^r) + \sup_{n \in \mathbb{N}} \max_{j \leq n} \mathbb{E}(|\bar{\zeta}_j^n|^r) + \sup_{t \leq T} \mathbb{E}|\zeta_t|^r + \sup_{t \leq T} \mathbb{E}|\bar{\zeta}_t|^r < \infty$

(ii) $\widehat{\zeta}_t^n$ (resp. $\widehat{\bar{\zeta}}_t^n$) converges in probability to ζ (resp. $\bar{\zeta}$) for the J_1 -Skorokhod topology,

i.e. there exists a family $(\psi^n)_{n \in \mathbb{N}}$ of one-to-one random time changes from $[0, T]$ to $[0, T]$ such that $\sup_{t \in [0, T]} |\psi^n(t) - t| \xrightarrow[n \rightarrow \infty]{} 0$ almost surely and $\sup_{t \in [0, T]} |\widehat{\zeta}_{\psi^n(t)}^n - \zeta_t| \xrightarrow[n \rightarrow \infty]{} 0$ in probability.

Remark 2.3.10 Hypothesis 2.7 implies that for all t in $[0, T]$ $\widehat{\zeta}_{\psi^n(t)}^n$ (resp. $\widehat{\bar{\zeta}}_{\psi^n(t)}^n$) converges to ζ_t (resp. $\bar{\zeta}_t$) in L^2 .

Remark 2.3.11 Let us give different examples of barriers in \mathcal{S}^2 satisfying Assumption 2.7. In this Remark, X represents either ζ or $\bar{\zeta}$.

1. X satisfies the following SDE

$$X_t = X_0 + \int_0^t b_X(X_{s-}) ds + \int_0^t \sigma_X(X_{s-}) dW_s + \int_0^t c_X(X_{s-}) d\tilde{N}_s$$

where b_X, σ_X and c_X are Lipschitz functions. We approximate it by

$$\bar{X}_t^n = \bar{X}_0^n + \int_0^t b_X(\bar{X}_{s-}^n) dc_s^n + \int_0^t \sigma_X(\bar{X}_{s-}^n) dW_s^n + \int_0^t c_X(\bar{X}_{s-}^n) d\tilde{N}_s^n$$

where $c_s^n = h[s/h]$. Since (W^n, \tilde{N}^n) converges in probability to (W, \tilde{N}) for the J_1 -topology, [Slo89, Corollary 1] gives that \bar{X}^n converges to X in probability for the J_1 -topology (for more details on the convergence of sequences of stochastic integrals on the space of RCLL functions endowed with the J_1 -Skorokhod topology, we refer to [JMP89]). Then, \bar{X}^n satisfies Assumption 2.7 (ii). We deduce from Doob and Burkholder-Davis-Gundy inequalities that X and \bar{X}^n satisfy Assumption 2.7 (i) and that X belongs to \mathcal{S}^2 .

2. X is defined by $X_t := \Phi(t, W_t, \tilde{N}_t)$, where Φ satisfies the following assumptions

(a) Φ is uniformly continuous in (t, y) uniformly in x , i.e. there exist two continuous non decreasing functions $g_0(\cdot)$ and $g_1(\cdot)$ from \mathbb{R}_+ to \mathbb{R}_+ with linear growth and satisfying $g_0(0) = g_1(0) = 0$ such that

$$\forall (t, t', x, y, y'), \quad |\Phi(t, x, y) - \Phi(t', x, y')| \leq g_0(|t - t'|) + g_1(|y - y'|).$$

We denote a_0 (resp. a_1) the constant of linear growth for g_0 (resp. g_1) i.e. $\forall (t, y) \in (\mathbb{R}_+)^2, 0 \leq g_0(t) + g_1(y) \leq a_0(1 + t) + a_1(1 + y)$,

(b) Φ is “strongly” locally Lipschitz in x uniformly in (t, y) , i.e. there exists a constant K_0 and an integer p_0 such that

$$\forall (t, x, x', y), \quad |\Phi(t, x, y) - \Phi(t, x', y)| \leq K_0(1 + |x|^{p_0} + |x'|^{p_0})|x - x'|.$$

Then, $\forall (t, x, y)$ we have $|\Phi(t, x, y)| \leq a_0|t| + a_1|y| + K_0(1 + |x|^{p_0})|x| + |\Phi(0, 0, 0)| + a_0 + a_1$. From this inequality, we prove that X satisfies Assumption 2.7 (i) by standard computations. Since (\tilde{N}^n) converges in probability to (\tilde{N}) for the J_1 -topology and $\lim_{n \rightarrow \infty} \sup_t |W_t^n - W_t| = 0$ in L^p for any p (see Lemma 2.3.8), we get that $(X_t^n)_t := (\Phi(c_n(t), W_t^n, \tilde{N}_t^n))_t$ converges in probability to X for the J_1 -topology.

Intermediate penalized implicit discrete scheme After the discretization of the penalized equation (2.73) on time intervals $[t_j, t_{j+1}]_{0 \leq j \leq n-1}$, we get the following discrete backward equation. For all j in $\{0, \dots, n-1\}$

$$\begin{cases} y_j^{p,n} = y_{j+1}^{p,n} + f(t_j, y_j^{p,n}, z_j^{p,n}, u_j^{p,n})h + a_j^{p,n} - k_j^{p,n} - (z_j^{p,n} \sqrt{h} \xi_{j+1} + u_j^{p,n} \eta_{j+1} + v_j^{p,n} \mu_{j+1}) \\ a_j^{p,n} = ph(y_j^{p,n} - \zeta_j^n)^-; k_j^{p,n} = ph(\bar{\zeta}_j^n - y_j^{p,n})^-, \\ y_n^{p,n} := \zeta_n^n. \end{cases} \quad (2.74)$$

Following (2.71), the triplet $(z_j^{p,n}, u_j^{p,n}, v_j^{p,n})$ can be computed as follows

$$\begin{cases} z_j^{p,n} = \frac{1}{\sqrt{h}} \mathbb{E}(y_{j+1}^{p,n} \xi_{j+1} | \mathcal{F}_j^n), \\ u_j^{p,n} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(y_{j+1}^{p,n} \eta_{j+1} | \mathcal{F}_j^n), \\ v_j^{p,n} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(y_{j+1}^{p,n} \mu_{j+1} | \mathcal{F}_j^n), \end{cases}$$

where we refer to Remark 2.3.9 for the computation of conditional expectations. By taking the conditional expectation w.r.t. \mathcal{F}_j^n in (2.74), we get the following scheme, called *implicit penalized discrete scheme*: $y_n^{p,n} := \zeta_n^n$ and for $j = n-1, \dots, 0$

$$\begin{cases} y_j^{p,n} = (\Theta^{p,n})^{-1}(\mathbb{E}(y_{j+1}^{p,n} | \mathcal{F}_j^n)), \\ a_j^{p,n} = ph(y_j^{p,n} - \zeta_j^n)^-; k_j^{p,n} = ph(\bar{\zeta}_j^n - y_j^{p,n})^-, \\ z_j^{p,n} = \frac{1}{\sqrt{h}} \mathbb{E}(y_{j+1}^{p,n} \xi_{j+1} | \mathcal{F}_j^n), \\ u_j^{p,n} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(y_{j+1}^{p,n} \eta_{j+1} | \mathcal{F}_j^n), \end{cases} \quad (2.75)$$

where $\Theta^{p,n}(y) = y - f(jh, y, z_j^{p,n}, u_j^{p,n}) - ph(y - \zeta_j^n)^- + ph(\bar{\zeta}_j^n - y)^-$.

We also introduce the continuous time version $(Y_t^{p,n}, Z_t^{p,n}, U_t^{p,n}, A_t^{p,n}, K_t^{p,n})_{0 \leq t \leq T}$ of the solution to (2.75):

$$Y_t^{p,n} := y_{\lfloor t/h \rfloor}^{p,n}, Z_t^{p,n} := z_{\lfloor t/h \rfloor}^{p,n}, U_t^{p,n} := u_{\lfloor t/h \rfloor}^{p,n}, A_t^{p,n} := \sum_{i=0}^{\lfloor t/h \rfloor} a_i^{p,n}, K_t^{p,n} := \sum_{i=0}^{\lfloor t/h \rfloor} k_i^{p,n}. \quad (2.76)$$

We also introduce $\alpha_t^{p,n} := A_t^{p,n} - K_t^{p,n}$, for all $t \in [0, T]$.

Main scheme As said before, the numerical inversion of the operator $\Theta^{p,n}$ is not easy and is time consuming. If we replace $y_j^{p,n}$ by $\mathbb{E}(y_{j+1}^{p,n}|\mathcal{F}_j^n)$ in f , (2.74) becomes

$$\begin{cases} \bar{y}_j^{p,n} = \bar{y}_{j+1}^{p,n} + f(t_j, \mathbb{E}(\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n), \bar{z}_j^{p,n}, \bar{u}_j^{p,n})h + \bar{a}_j^{p,n} - \bar{k}_j^{p,n} - (\bar{z}_j^{p,n}\sqrt{h}\xi_{j+1} + \bar{u}_j^{p,n}\eta_{j+1} + \bar{v}_j^{p,n}\mu_{j+1}) \\ \bar{a}_j^{p,n} = ph(\bar{y}_j^{p,n} - \zeta_j^n)^-; \bar{k}_j^{p,n} = ph(\zeta_j^n - \bar{y}_j^{p,n})^-, \\ \bar{y}_n^{p,n} := \zeta_n^n. \end{cases} \quad (2.77)$$

Now, by taking the conditional expectation in the above equation, we obtain:

$$\bar{y}_j^{p,n} = \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n], \bar{z}_j^{p,n}, \bar{u}_j^{p,n})h + \bar{a}_j^{p,n} - \bar{k}_j^{p,n}.$$

Solving this equation, we get the following scheme, called *explicit penalized scheme*: $\bar{y}_n^{p,n} := \zeta_n^n$ and for $j = n-1, \dots, 0$

$$\begin{cases} \bar{y}_j^{p,n} = \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n] + f(t_j, \mathbb{E}(\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n), \bar{z}_j^{p,n}, \bar{u}_j^{p,n})h + \bar{a}_j^{p,n} - \bar{k}_j^{p,n}, \\ \bar{a}_j^{p,n} = \frac{ph}{1+ph} \left(\mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n], \bar{z}_j^{p,n}, \bar{u}_j^{p,n}) - \zeta_j^n \right)^-, \\ \bar{k}_j^{p,n} = \frac{ph}{1+ph} \left(\zeta_j^n - \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n] - f(t_j, \mathbb{E}[\bar{y}_{j+1}^{p,n}|\mathcal{F}_j^n], \bar{z}_j^{p,n}, \bar{u}_j^{p,n}) \right)^-, \\ \bar{z}_j^{p,n} = \frac{1}{\sqrt{h}} \mathbb{E}(\bar{y}_{j+1}^{p,n}\xi_{j+1}|\mathcal{F}_j^n), \\ \bar{u}_j^{p,n} = \frac{1}{\kappa_n(1-\kappa_n)} \mathbb{E}(\bar{y}_{j+1}^{p,n}\eta_{j+1}|\mathcal{F}_j^n). \end{cases} \quad (2.78)$$

As for the implicit scheme, we define the continuous time version $(\bar{Y}_t^{p,n}, \bar{Z}_t^{p,n}, \bar{U}_t^{p,n}, \bar{A}_t^{p,n}, \bar{K}_t^{p,n})_{0 \leq t \leq T}$ of the solution to (2.78):

$$\bar{Y}_t^{p,n} = \bar{y}_{[t/h]}^{p,n}, \quad \bar{Z}_t^{p,n} = \bar{z}_{[t/h]}^{p,n}, \quad \bar{U}_t^{p,n} = \bar{u}_{[t/h]}^{p,n}, \quad \bar{A}_t^{p,n} = \sum_{j=0}^{[t/h]} \bar{a}_j^{p,n}, \quad \bar{K}_t^{p,n} = \sum_{j=0}^{[t/h]} \bar{k}_j^{p,n}. \quad (2.79)$$

We also introduce $\bar{\alpha}_t^{p,n} := \bar{A}_t^{p,n} - \bar{K}_t^{p,n}$, for all $t \in [0, T]$.

2.3.3.2 Convergence result

The following result states the convergence of $\bar{\Theta}^{p,n} := (\bar{Y}^{p,n}, \bar{Z}^{p,n}, \bar{U}^{p,n}, \bar{\alpha}^{p,n})$ to $\Theta := (Y, Z, U, \alpha)$, the solution to the DBBSDE (2.72).

Theorem 2.3.12 *Assume that Assumptions 2.6 and 2.7 hold. The sequence $(\bar{Y}^{p,n}, \bar{Z}^{p,n}, \bar{U}^{p,n})$ defined by (2.79) converges to (Y, Z, U) , the solution to the DBBSDE (2.72), in the following sense: $\forall r \in [1, 2]$*

$$\lim_{p \rightarrow \infty} \lim_{n \rightarrow \infty} \left(\mathbb{E} \left[\int_0^T |\bar{Y}_s^{p,n} - Y_s|^2 ds \right] + \mathbb{E} \left[\int_0^T |\bar{Z}_s^{p,n} - Z_s|^r ds \right] + \mathbb{E} \left[\int_0^T |\bar{U}_s^{p,n} - U_s|^r ds \right] \right) \rightarrow 0. \quad (2.80)$$

Moreover, $\bar{Z}^{p,n}$ (resp. $\bar{U}^{p,n}$) weakly converges in H^2 to Z (resp. to U) and for $0 \leq t \leq T$, $\bar{\alpha}_{\psi^n(t)}^{p,n}$ converges weakly to α_t in $L^2(\mathcal{F}_T)$ as $n \rightarrow \infty$ and $p \rightarrow \infty$, where $(\psi^n)_{n \in \mathbb{N}}$ is a one-to-one random map from $[0, T]$ to $[0, T]$ such that $\sup_{t \in [0, T]} |\psi^n(t) - t| \xrightarrow[n \rightarrow \infty]{} 0$ a.s..

Proof. In order to prove this result, we split the error in three terms, by introducing $\Theta_t^{p,n} := (Y_t^{p,n}, Z_t^{p,n}, U_t^{p,n}, \alpha_t^{p,n})$, the solution to the implicit penalized discrete scheme (2.76) and $\Theta_t^p := (Y_t^p, Z_t^p, U_t^p, \alpha_t^p)$, the penalized version of (2.72), defined by (2.73). For the error on Y , we get

$$\mathbb{E}\left[\int_0^T |\bar{Y}_s^{p,n} - Y_s|^2 ds\right] \leq 3 \left(\mathbb{E}\left[\int_0^T |\bar{Y}_s^{p,n} - Y_s^{p,n}|^2 ds\right] + \mathbb{E}\left[\int_0^T |Y_s^{p,n} - Y_s^p|^2 ds\right] + \mathbb{E}\left[\int_0^T |Y_s^p - Y_s|^2 ds\right] \right),$$

and the same splitting holds for $|\bar{Z}^{p,n} - Z|^r$ and $|\bar{U}^{p,n} - U|^r$. For the increasing processes, we have:

$$\mathbb{E}[|\bar{\alpha}_{\psi^n(t)}^{p,n} - \alpha_t|^2] \leq 3 \left(\mathbb{E}[|\bar{\alpha}_{\psi^n(t)}^{p,n} - \alpha_{\psi^n(t)}^{p,n}|^2] + \mathbb{E}[|\alpha_{\psi^n(t)}^{p,n} - \alpha_t^p|^2] + \mathbb{E}[|\alpha_t^p - \alpha_t|^2] \right). \quad (2.81)$$

The proof of Theorem 2.3.12 ensues from [L-8, Proposition 4.2], [L-8, Corollary 4.4] and [L-8, Proposition 4.5]. [L-8, Proposition 4.2] states the convergence of the error between $\bar{\Theta}^{p,n}$, the explicit penalization scheme defined in (2.79), and $\Theta^{p,n}$, the implicit penalization scheme. It generalizes the results of [PX11]. [L-8, Corollary 4.4] states the convergence (in n) of $\Theta^{p,n}$ to Θ^p . This is based on the convergence of a standard BSDE with jumps in discrete time setting to the associated BSDE with jumps in continuous time setting, which is proved in [LMT13]. Finally, [L-8, Proposition 4.5] proves the convergence (in p) of the penalized BSDE with jumps Θ^p to Θ , the solution of the DBBSDE (2.72). In fact, we prove a more general result since we show the convergence of penalized BSDEs to (2.68) in the case of jumps driven by a general Poisson random measure. \blacksquare

2.3.4 Reflected scheme

Let us now introduce an additional assumption on f , which ensures the comparison theorem for BSDEs with jumps (see [QS13, Theorem 4.2]). The comparison theorem plays a key role in the proof of the convergence of the penalized scheme (see Section 2.3.3), which is useful to prove the convergence of the reflected scheme.

In this special case where the Poisson random measure is simply generated by the jumps of a Poisson process, Hypothesis 2.5 becomes the following Hypothesis.

Hypothesis 2.8 A Lipschitz driver f is said to satisfy Assumption 2.8 if the following holds : $d\mathbb{P} \otimes dt$ a.s. for each $(y, z, u_1, u_2) \in \mathbb{R}^4$, we have

$$f(t, y, z, u_1) - f(t, y, z, u_2) \geq \theta(u_1 - u_2), \text{ with } -1 \leq \theta \leq \theta_0.$$

We also assume the following hypothesis on the barriers.

Hypothesis 2.9 $\underline{\zeta}$ and $\bar{\zeta}$ are Itô processes of the following form

$$\underline{\zeta}_t = \underline{\zeta}_0 + \int_0^t b_s^\underline{\zeta} ds + \int_0^t \sigma_s^\underline{\zeta} dW_s + \int_0^t \beta_{s-}^\underline{\zeta} d\tilde{N}_s \quad (2.82)$$

$$\bar{\zeta}_t = \bar{\zeta}_0 + \int_0^t b_s^{\bar{\zeta}} ds + \int_0^t \sigma_s^{\bar{\zeta}} dW_s + \int_0^t \beta_{s-}^{\bar{\zeta}} d\tilde{N}_s \quad (2.83)$$

where $b^\underline{\zeta}$, $b^{\bar{\zeta}}$, $\sigma^\underline{\zeta}$, $\sigma^{\bar{\zeta}}$, $\beta^\underline{\zeta}$ and $\beta^{\bar{\zeta}}$ are adapted RCLL processes such that there exist $r > 2$ and a constant $C_{\underline{\zeta}, \bar{\zeta}}$ such that $\mathbb{E}(\sup_{s \leq T} |b_s^\underline{\zeta}|^r) + \mathbb{E}(\sup_{s \leq T} |b_s^{\bar{\zeta}}|^r) + \mathbb{E}(\sup_{s \leq T} |\sigma_s^\underline{\zeta}|^r) + \mathbb{E}(\sup_{s \leq T} |\sigma_s^{\bar{\zeta}}|^r) + \mathbb{E}(\sup_{s \leq T} |\beta_s^\underline{\zeta}|^r) + \mathbb{E}(\sup_{s \leq T} |\beta_s^{\bar{\zeta}}|^r) \leq C_{\underline{\zeta}, \bar{\zeta}}$. We also assume $\underline{\zeta}_T = \bar{\zeta}_T$ a.s., $\underline{\zeta}_t \leq \bar{\zeta}_t$ for all $t \in [0, T]$.

The barriers $\underline{\zeta}$ and $\bar{\zeta}$ given in Assumption 2.9 are approximated in the following way: for all $k \in \{1, \dots, n\}$

$$\underline{\zeta}_k^n = \underline{\zeta}_0 + \sum_{i=0}^{k-1} b_{t_i}^{\underline{\zeta}} h + \sum_{i=0}^{k-1} \sigma_{t_i}^{\underline{\zeta}} \sqrt{h} \xi_{i+1} + \sum_{i=0}^{k-1} \beta_{t_i}^{\underline{\zeta}} \eta_{i+1}, \quad (2.84)$$

$$\bar{\zeta}_k^n = \bar{\zeta}_0 + \sum_{i=0}^{k-1} b_{t_i}^{\bar{\zeta}} h + \sum_{i=0}^{k-1} \sigma_{t_i}^{\bar{\zeta}} \sqrt{h} \xi_{i+1} + \sum_{i=0}^{k-1} \beta_{t_i}^{\bar{\zeta}} \eta_{i+1}. \quad (2.85)$$

Lemma 2.3.13 *Under Assumption 2.9, there exists a constant $C_{\underline{\zeta}, \bar{\zeta}, T, \lambda}$ depending on $C_{\underline{\zeta}, \bar{\zeta}}$, T and λ such that*

$$(i) \sup_n \max_j \mathbb{E}(|\underline{\zeta}_j^n|^r) + \sup_n \max_j \mathbb{E}(|\bar{\zeta}_j^n|^r) + \sup_{t \leq T} \mathbb{E}(|\underline{\zeta}_t|^r) + \sup_{t \leq T} \mathbb{E}(|\bar{\zeta}_t|^r) \leq C_{\underline{\zeta}, \bar{\zeta}, T, \lambda}$$

$$(ii) \underline{\zeta}^n \text{ (resp. } \bar{\zeta}^n) \text{ converges in probability to } \underline{\zeta} \text{ (resp. } \bar{\zeta}) \text{ in } J_1\text{-Skorokhod topology.}$$

The explicit reflected scheme is introduced by replacing y_j^n by $\mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n]$ in f . We obtain

$$\begin{cases} \bar{y}_j^n = \bar{y}_{j+1}^n + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) h + \bar{a}_j^n - \bar{k}_j^n - \bar{z}_j^n \sqrt{h} \xi_{j+1} - \bar{u}_j^n \eta_{j+1} - \bar{v}_j^n \mu_{j+1}, \\ \bar{a}_j^n \geq 0, \bar{k}_j^n \geq 0, \bar{a}_j^n \bar{k}_j^n = 0, \\ \underline{\zeta}_j^n \leq \bar{y}_j^n \leq \bar{\zeta}_j^n, (\bar{y}_j^n - \underline{\zeta}_j^n) \bar{a}_j^n = (\bar{y}_j^n - \bar{\zeta}_j^n) \bar{k}_j^n = 0. \end{cases} \quad (2.86)$$

with terminal condition $\bar{y}_n^n = \underline{\zeta}_n^n$. By taking the conditional expectation in (2.86) with respect to \mathcal{F}_j^n , we derive that:

$$(\bar{\mathcal{S}}_1) \begin{cases} \bar{y}_n^n = \underline{\zeta}_n^n, \\ \bar{y}_j^n = \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) h + \bar{a}_j^n - \bar{k}_j^n, \bar{a}_j^n \geq 0, \bar{k}_j^n \geq 0, \bar{a}_j^n \bar{k}_j^n = 0, \\ \underline{\zeta}_j^n \leq \bar{y}_j^n \leq \bar{\zeta}_j^n, (\bar{y}_j^n - \underline{\zeta}_j^n) \bar{a}_j^n = (\bar{y}_j^n - \bar{\zeta}_j^n) \bar{k}_j^n = 0. \end{cases}$$

We get that $(\bar{\mathcal{S}}_1)$ is equivalent to $(\bar{\mathcal{S}}_2)$

$$(\bar{\mathcal{S}}_2) \begin{cases} \bar{y}_n^n = \underline{\zeta}_n^n, \\ \bar{y}_j^n = \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) h + \bar{a}_j^n - \bar{k}_j^n, \\ \bar{a}_j^n = (\mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) h - \underline{\zeta}_j^n)^-, \\ \bar{k}_j^n = (\mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n] + f(t_j, \mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n], \bar{z}_j^n, \bar{u}_j^n) h - \bar{\zeta}_j^n)^+. \end{cases}$$

We also introduce the continuous time version $(\bar{Y}_t^n, \bar{Z}_t^n, \bar{U}_t^n, \bar{A}_t^n, \bar{K}_t^n)_{0 \leq t \leq T}$ of $(\bar{y}_j^n, \bar{z}_j^n, \bar{u}_j^n, \bar{a}_j^n, \bar{k}_j^n)_{j \leq n}$:

$$\bar{Y}_t^n := \bar{y}_{[t/h]}^n, \bar{Z}_t^n := \bar{z}_{[t/h]}^n, \bar{U}_t^n := \bar{u}_{[t/h]}^n, \bar{A}_t^n := \sum_{i=0}^{[t/h]} \bar{a}_i^n, \bar{K}_t^n := \sum_{i=0}^{[t/h]} \bar{k}_i^n. \quad (2.87)$$

In the following $\bar{\Theta}^n := (\bar{Y}^n, \bar{Z}^n, \bar{U}^n, \bar{A}^n - \bar{K}^n)$ and $\bar{\alpha}^n := \bar{A}^n - \bar{K}^n$.

We prove in [L-9, Section 4] that $\bar{\Theta}^n$ converges to $\bar{\Theta} := (Y_t, Z_t, U_t, A_t - K_t)_{0 \leq t \leq T}$, the solution to the DRBSDE (2.66). The main result is stated in the following Theorem.

Theorem 2.3.14 *Suppose that Assumption 2.9 holds and f is a Lipschitz driver satisfying Assumption 2.8. Then we have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_0^T |\bar{Y}_t^n - Y_t|^2 dt + \int_0^T |\bar{Z}_t^n - Z_t|^2 dt + \int_0^T |\bar{U}_t^n - U_t|^2 dt \right] = 0.$$

Moreover, $\bar{\alpha}_{\psi^n(t)}^n$ converges weakly to α_t in $L^2(\mathcal{F}_T)$.

Proof. The proof is done by splitting the error in three terms (the same inequality holds for the errors on Z and U)

$$\mathbb{E} \left[\int_0^T |\bar{Y}_t^n - Y_t|^2 dt \right] \leq 3\mathbb{E} \left[\int_0^T |\bar{Y}_t^n - Y_t^n|^2 dt \right] + 3\mathbb{E} \left[\int_0^T |Y_t^n - Y_t^{p,n}|^2 dt \right] + 3\mathbb{E} \left[\int_0^T |Y_t^{p,n} - Y_t|^2 dt \right].$$

For the increasing processes, we have:

$$\mathbb{E} [|\bar{\alpha}_{\psi^n(t)}^n - \alpha_t|^2] \leq 3 \left(\mathbb{E} [|\bar{\alpha}_{\psi^n(t)}^n - \alpha_{\psi^n(t)}^n|^2] + \mathbb{E} [|\alpha_{\psi^n(t)}^n - \alpha_t^{p,n}|^2] + \mathbb{E} [|\alpha_t^{p,n} - \alpha_t|^2] \right).$$

[L-9, Proposition 4.5] states the convergence of the error $\bar{\Theta}^n - \Theta^n$, where Θ^n represents the solution given by an implicit reflected scheme (see [L-9, Eq (3.6)]) which consists in replacing $\mathbb{E}[\bar{y}_{j+1}^n | \mathcal{F}_j^n]$ by y_j^n in f in the scheme (2.86). [L-9, Proposition 4.6] states the convergence of $\Theta^n - \Theta^{p,n}$, where $\Theta^{p,n} := (Y^{p,n}, Z^{p,n}, U^{p,n}, A^{p,n} - K^{p,n})$ represents the solution given by the implicit penalization scheme (see (2.74)) and the result of the convergence of $\Theta^{p,n} - \Theta$ is recalled in [L-9, Theorem 3.5] (it has already been proved in [L-8]). \blacksquare

Chapter 3

Monte Carlo method and Wiener Chaos expansion

3.1 Introduction

In this Chapter we present an algorithm based on Picard iterations and Wiener chaos expansion to approximate the solutions of different kind of BSDEs. This forward method involves initially applying Picard iterations, followed by Wiener chaos expansion to approximate conditional expectations. Unlike most methods outlined in Section 1.2, the primary advantage of this approach is its applicability to non-Markovian BSDEs. Additionally, Wiener chaos expansion provides an approximation of Z without incurring any additional costs. The convergence of this algorithm was rigorously established in [L-10]. Subsequently, this method was extended to encompass BSDEs with jumps in [L-7] and to the realm of McKean-Vlasov BSDEs as detailed in [L-25].

Let us first present the principle of the algorithm in the case of standard BSDEs and a real-valued Brownian motion. Let (Y, Z) denote the solution to

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \quad (3.1)$$

when ζ belongs to $L^2(\mathcal{F}_T)$. As already used in several quoted papers (see for example [BD07, L-14, BS12]), our starting point is the use of Picard iterations: $(Y^0, Z^0) = (0, 0)$ and for $q \in \mathbb{N}$,

$$Y_t^{q+1} = \zeta + \int_t^T f(s, Y_s^q, Z_s^q) ds - \int_t^T Z_s^{q+1} \cdot dW_s, \quad 0 \leq t \leq T.$$

It is well-known that the sequence (Y^q, Z^q) converges exponentially fast towards the solution (Y, Z) to BSDE (3.1). First, we write this Picard scheme in a forward way

$$\begin{aligned} Y_t^{q+1} &= \mathbb{E} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds \mid \mathcal{F}_t \right) - \int_0^t f(s, Y_s^q, Z_s^q) ds, \\ Z_t^{q+1} &= D_t Y_t^{q+1} = D_t \mathbb{E} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds \mid \mathcal{F}_t \right), \end{aligned}$$

where $D_t X$ stands for the Malliavin derivative of the random variable X .

Then, in order to compute the previous conditional expectation, we use a Wiener chaos expansion of the random variable

$$F^q = \zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds.$$

More precisely, we use the following orthogonal decomposition of the random variable F^q

$$F^q = \mathbb{E}[F^q] + \sum_{k \geq 1} \sum_{|n|=k} d_k^n \prod_{i \geq 1} K_{n_i} \left(\int_0^T g_i(s) dW_s \right),$$

where K_l denotes the Hermite polynomial of degree l , $(g_i)_{i \geq 1}$ is an orthonormal basis of $L^2([0, T]; \mathbb{R})$ and, if $n = (n_i)_{i \geq 1}$ is a sequence of integers, $|n| = \sum_{i \geq 1} n_i$. $(d_k^n)_{k \geq 1, |n|=k}$ is the sequence of coefficients ensuing from the decomposition of F^q . Of course, from a practical point of view, we only keep a finite number of terms in this expansion:

- we work with a finite number of chaos p ;
- we choose a finite number of functions g_1, \dots, g_N .

This leads to the following approximation with $n = (n_1, \dots, n_N)$

$$F^q \simeq \mathbb{E}[F^q] + \sum_{1 \leq k \leq p} \sum_{|n|=k} d_k^n \prod_{1 \leq i \leq N} K_{n_i} \left(\int_0^T g_i(s) dW_s \right).$$

One of the key points in using such a decomposition is that, for choices of simple functions g_1, \dots, g_N , there exist explicit formulas for both

$$\mathbb{E}(F^q | \mathcal{F}_t) \quad \text{and} \quad Z_t^{q+1} = D_t \mathbb{E}(F^q | \mathcal{F}_t); \quad (3.2)$$

this plays a crucial role in our algorithm. Using these formulas and starting from M trajectories of the underlying Brownian motion enable to construct M trajectories of the solution (Y, Z) to the BSDE.

In the following, the functions g_i are chosen as step functions:

$$g_i = \mathbf{1}_{] \bar{t}_{i-1}, \bar{t}_i]}(t) / \sqrt{h}, \quad i = 1, \dots, N, \quad \text{where } \bar{t}_i := ih, \quad h = \frac{T}{N}$$

and the previous formulas are really simple (see (3.10)-(3.11) and Proposition 3.2.7). Eventually, the main advantage of this method is that only one decomposition has to be computed per Picard iteration: the decomposition of F^q . Therein lies a main difference between our approach and the approach based on regression technics developed by C. Bender and R. Denk in [BD07]. In their paper, for a given Picard iteration q and for each time t_i of the mesh grid, two projections have to be computed, one for $Y_{t_i}^q$ and one for $Z_{t_i}^q$. The second difference comes from the way of computing Z^q . In our method, once the decomposition of F^q is computed, Z^q is given explicitly as the Malliavin derivative of Y^q .

Our algorithm can handle fully path dependent terminal conditions, and was extended to the case of BSDEs with jumps (see Section 3.3) and to McKean-Vlasov BSDEs (see Section 3.4). Section 3.2 presents the results obtained in [L-10], Section 3.3 presents the results obtained in [L-7] and Section 3.4 presents the results obtained in [L-25].

3.1.1 Definitions and Notations

- $\|\partial_{sp}^j f\|_\infty^2$, the norm of the derivatives of $f([0, T] \times \mathbb{R}^d, \mathbb{R})$ w.r.t. all the space variables x which sum equals j : $\|\partial_{sp}^j f\|_\infty^2 := \sum_{|k|=j} \|\partial_{x_1}^{k_1} \dots \partial_{x_d}^{k_d} f\|_\infty^2$, where $|k| = k_1 + \dots + k_d$.
- C_p^∞ , the set of smooth functions $f : \mathbb{R}^n \mapsto \mathbb{R}$ with partial derivatives of polynomial growth.

- $\|(\cdot, \cdot)\|_{L^p}^p$, $p \in \mathbb{N}$, $p \geq 2$, the norm on the space $S_T^p(\mathbb{R}) \times H_T^p(\mathbb{R}^d)$ defined by

$$\|(Y, Z)\|_{L^p}^p := \mathbb{E} \left(\sup_{t \in [0, T]} |Y_t|^p \right) + \int_0^T \mathbb{E}(|Z_t|^p) dt. \quad (3.3)$$

We also recall some useful definitions related to Malliavin calculus. We use the notations of [Nua06].

- \mathcal{S} denotes the class of random variables of the form $F = f(W(h_1), \dots, W(h_n))$, where $f \in C_p^\infty(\mathbb{R}^{n \times d}, \mathbb{R})$, for all $j \leq n$ $h_j = (h_j^1, \dots, h_j^d) \in L^2([0, T]; \mathbb{R}^d)$ and for all $i \leq d$ $W^i(h_j^i) = \int_0^T h_j^i(t) dW_t^i$.
- $\mathbb{D}^{r,2}$ denotes the closure of \mathcal{S} w.r.t. the following norm on \mathcal{S}

$$\|F\|_{\mathbb{D}^{r,2}}^2 := \mathbb{E}|F|^2 + \sum_{q=1}^r \sum_{|\alpha|_1=q} \mathbb{E} \left(\int_0^T \cdots \int_0^T |D_{(t_1, \dots, t_q)}^\alpha F|^2 dt_1 \cdots dt_q \right)$$

where α is a multi-index $(\alpha_1, \dots, \alpha_q) \in \{1, \dots, d\}^q$ $|\alpha|_1 := \sum_{i=1}^q \alpha_i = q$ and D^α represents the multi-index Malliavin derivative operator. We recall $\mathbb{D}^{\infty,2} = \bigcap_{r=1}^\infty \mathbb{D}^{r,2}$.

Remark 3.1.1 When $d = 1$, $\|F\|_{\mathbb{D}^{r,2}}^2 := \mathbb{E}|F|^2 + \sum_{q=1}^r \mathbb{E} \left(\int_0^T \cdots \int_0^T |D_{(t_1, \dots, t_q)}^{(q)} F|^2 dt_1 \cdots dt_q \right) = \mathbb{E}|F|^2 + \sum_{q=1}^r \|D^{(q)} F\|_{L^2(\Omega \times [0, T]^q)}^2$.

Let $m \in \mathbb{N}^*$ and $j \in \mathbb{N}$, $j \geq 2$. We also introduce the following notations

- $\mathcal{D}^{m,j}$ denotes the space of all \mathcal{F}_T -measurable r.v. such that

$$\|F\|_{\mathcal{D}^{m,j}}^j := \sum_{1 \leq l \leq m} \sum_{|\alpha|_1=l} \sup_{t_1 \leq \dots \leq t_l} \mathbb{E}[|D_{t_1, \dots, t_l}^\alpha F|^j] < \infty$$

where $\sup_{t_1 \leq \dots \leq t_l}$ means $\sup_{(t_1, \dots, t_l): t_1 \leq \dots \leq t_l}$.

- $\mathcal{S}^{m,j}$ denotes the space of all couple of processes (Y, Z) belonging to $S_T^j(\mathbb{R}) \times H_T^j(\mathbb{R}^d)$ and such that

$$\|(Y, Z)\|_{\mathcal{S}^{m,j}}^j := \sum_{1 \leq l \leq m} \sum_{|\alpha|_1=l} \sup_{t_1 \leq \dots \leq t_l} \|(D_{t_1, \dots, t_l}^\alpha Y, D_{t_1, \dots, t_l}^\alpha Z)\|_{L^j}^j < \infty.$$

We recall

$$\|(Y, Z)\|_{\mathcal{S}^{m,j}}^j = \sum_{1 \leq l \leq m} \sum_{|\alpha|_1=l} \sup_{t_1 \leq \dots \leq t_l} \left\{ \mathbb{E} \left[\sup_{t_i \leq r \leq T} |D_{t_1, \dots, t_l}^\alpha Y_r|^j \right] + \int_{t_l}^T \mathbb{E} [|D_{t_1, \dots, t_l}^\alpha Z_r|^j] dr \right\}.$$

We also denote $\mathcal{S}^{m,\infty} := \bigcap_{j \geq 2} \mathcal{S}^{m,j}$.

3.2 Standard case

In this Section we present the results [L-10] written with Philippe Briand. As presented in the introduction of this Chapter, they concern the case of standard BSDEs (3.1). In the following sections we present the Wiener chaos expansion and its properties (Sections 3.2.1 and 3.2.2), the main steps of the algorithm (Section 3.2.3), prove its convergence and establish the rate of this convergence w.r.t. the different parameters (Section 3.2.4). We also present numerical results (Section 3.2.5).

3.2.1 Wiener Chaos Expansion

3.2.1.1 Notations and useful results

We refer to [Nua06] for more details on this section. Let us briefly recall the Wiener chaos expansion in the simple case of a real-valued Brownian motion. It is well known that every random variable $F \in L^2(\mathcal{F}_T)$ has an expansion of the following form:

$$\begin{aligned} F = & \mathbb{E}[F] + \int_0^T u_1(s_1) dW_{s_1} \\ & + \int_0^T \int_0^{s_2} u_2(s_2, s_1) dW_{s_1} dW_{s_2} + \dots + \int_0^T \int_0^{s_n} \dots \int_0^{s_2} u_n(s_n, \dots, s_1) dW_{s_1} \dots dW_{s_n} + \dots \end{aligned} \quad (3.4)$$

where the functions $(u_n, n \geq 1)$ are deterministic functions. There is an ambiguity for the definition of these functions u_n . We adopt in this paper the following point of view: the function u_n is defined on the simplex

$$\mathcal{S}_n(T) := \{(s_1, \dots, s_n) \in [0, T]^n : 0 < s_1 < \dots < s_n < T\}.$$

We define the iterated integral for a deterministic function $f \in L^2(\mathcal{S}_n(T))$ as

$$J_n(f) := \int_0^T \int_0^{s_n} \dots \int_0^{s_2} f(s_n, \dots, s_1) dW_{s_1} \dots dW_{s_n}.$$

Due to the Itô isometry, $\|J_n(f)\|^2 = \|f\|_{L^2(\mathcal{S}_n(T))}^2$ and $\mathbb{E}[J_n(f)J_m(g)] = \delta_{nm} \langle f, g \rangle_{L^2(\mathcal{S}_n(T))}$. Then, $\|F\|^2 = \sum_{n \geq 0} \|u_n\|_{L^2(\mathcal{S}_n(T))}^2$.

Definition 3.2.1 Let F be a random variable in $L^2(\mathcal{F}_T)$ whose chaos expansion is given by (3.4). We introduce

- $P_n(F) := J_n(u_n)$ the Wiener chaos of order n of F .
- $\mathcal{C}_p(F) := \sum_{n \leq p} P_n(F)$ the chaos decomposition of F up to order p , i.e.

$$\begin{aligned} \mathcal{C}_p(F) = & \mathbb{E}[F] + \int_0^T u_1(s_1) dW_{s_1} + \int_0^T \int_0^{s_2} u_2(s_2, s_1) dW_{s_1} dW_{s_2} \\ & + \dots + \int_0^T \int_0^{s_p} \dots \int_0^{s_2} u_p(s_p, \dots, s_1) dW_{s_1} \dots dW_{s_p}. \end{aligned} \quad (3.5)$$

We state two Lemmas useful for the sequel.

Lemma 3.2.2 (Nualart) $F \in \mathbb{D}^{m,2}$ if and only if $\|D^m F\|_{L^2(\Omega \times [0, T]^m)}^2 = \sum_{n \geq 0} (n + m - 1) \times \dots \times n \times \mathbb{E}[|P_n(F)|^2] < \infty$. In this case, we have

$$\sum_{n \geq 0} (n + m - 1) \times \dots \times n \times \mathbb{E}[|P_n(F)|^2] \leq \|F\|_{\mathbb{D}^{m,2}}^2.$$

From Lemma 3.2.2, we deduce

Lemma 3.2.3 ([L-7, Lemma 2.4]) Let $1 \leq m \leq p + 1$ and $F \in \mathbb{D}^{m,2}$. We have

$$\mathbb{E}[|F - \mathcal{C}_p(F)|^2] \leq \frac{\|F\|_{\mathbb{D}^{m,2}}^2}{(p + 2 - m) \dots (p + 1)}.$$

The following Lemma gives some useful properties of the chaos decomposition.

Lemma 3.2.4

- Let F be a r.v. in $L^2(\mathcal{F}_T)$. $\forall p \geq 1$, we have $\mathbb{E}(|\mathcal{C}_p(F)|^2) \leq \mathbb{E}(|F|^2)$. If F belongs to $L^j(\mathcal{F}_T)$, $\forall j > 2$, $\mathbb{E}(|\mathcal{C}_p(F)|^j) \leq (1 + p(j - 1)^{\frac{p}{2}})^j \mathbb{E}(|F|^j)$.
- Let H be in $H_T^2(\mathbb{R})$. We have $\mathcal{C}_p\left(\int_0^T H_s ds\right) = \int_0^T \mathcal{C}_p(H_s) ds$.
- For all $F \in \mathbb{D}^{1,2}$ and for all $t \leq r$, $D_t \mathbb{E}_r[\mathcal{C}_p(F)] = \mathbb{E}_r[\mathcal{C}_{p-1}(D_t F)]$.

The first result ensues from the fact that for $j > 2$ $\|P_n(F)\|_j \leq (j - 1)^{\frac{n}{2}} \|F\|_j$ (see [Nua06, page 63]).

3.2.1.2 Wiener chaos expansion and Hermite polynomials

Another approach to Wiener chaos expansion uses Hermite polynomials. This approach can be easily generalized when considering d -dimensional Brownian motions, this is then the one we consider in the following. We present it for $d = 1$. Let $\{g_i\}_{i \geq 1}$ be an orthonormal basis of $L^2([0, T]; \mathbb{R})$. The Wiener chaos of order n , $P_n(F)$, is the L^2 -closure of the vector field spanned by

$$\left\{ \prod_{i \geq 1} \sqrt{n_i!} K_{n_i} \left(\int_0^T g_i(s) dW_s \right) : |(n_i)_{i \geq 1}| := \sum n_i = n \right\}$$

where K_n is the Hermite polynomial of order n defined by the expansion:

$$e^{xt - t^2/2} = \sum_{n \geq 0} K_n(x) t^n.$$

It is well known that

$$\begin{aligned} H'_n(x) &= nH_{n-1}(x) \\ H_n(x) &= xH_{n-1}(x) - H'_{n-1}(x) = xH_{n-1}(x) - (n-1)H_{n-2}(x). \end{aligned}$$

In this paper, we will work with the following normalization of the Hermite polynomials:

$$\forall n \in \mathbb{N}, \quad K_n(x) = \frac{1}{n!} H_n(x),$$

with the convention $K_{-1} \equiv 0$. With this normalization, we have $K'_n(x) = K_{n-1}(x)$ for any integer n .

$$\begin{aligned} K'_n(x) &= K_{n-1}(x) \\ K_n(x) &= \frac{1}{n} (xK_{n-1}(x) - K'_{n-1}(x)) = \frac{1}{n} (xK_{n-1}(x) - K_{n-2}(x)). \end{aligned}$$

It is well-known that $(K_n)_{n \geq 0}$ is a sequence of orthogonal polynomials in $L^2(\mathbb{R}, \mu)$, where μ denotes the reduced centered Gaussian measure. Moreover, we have

$$\int_{\mathbb{R}} K_n^2(x) \mu(dx) = \frac{1}{n!}.$$

Every square integrable random variable F , measurable with respect to \mathcal{F}_T , admits the following orthogonal decomposition

$$F = d_0 + \sum_{k \geq 1} \sum_{|n|=k} d_k^n \prod_{i \geq 1} K_{n_i} \left(\int_0^T g_i(s) dW_s \right), \quad (3.6)$$

where $n = (n_i)_{i \geq 1}$ is a sequence of positive integers and where $|n|$ stands for $\sum_{i \geq 1} n_i$. Taking into account the normalization of the Hermite polynomials we use, we get

$$d_0 = \mathbb{E}[F], \quad d_k^n = n! \mathbb{E} \left[F \times \prod_{i \geq 1} K_{n_i} \left(\int_0^T g_i(s) dW_s \right) \right],$$

where $n! = \prod_{i \geq 1} n_i!$. Before describing the chaos decomposition formulas we use in the algorithm, we give a Lemma useful in the sequel.

Lemma 3.2.5 *Let $g \in L^2([0, T]; \mathbb{R})$ and let $U_t = \int_0^t g^2(s) ds$. For $n \in \mathbb{N}$, let us define*

$$M_t^n = U_t^{n/2} K_n \left(W(g)_t / \sqrt{U_t} \right), \quad W(g)_t = \int_0^t g(s) dW_s.$$

Then $\{M_t^n\}_{0 \leq t \leq T}$ is a martingale and

$$dM_t^n = g(t) M_t^{n-1} dW_t.$$

3.2.2 Chaos decomposition approximations

These formulas are based on the decomposition (3.6). To get tractable formulas, we consider a finite number of chaos and a finite number of functions (g_1, \dots, g_N) . The $(g_i)_{1 \leq i \leq N}$ functions are chosen such that we can quickly compute $\mathbb{E}(F|\mathcal{F}_t)$ and $D_t \mathbb{E}(F|\mathcal{F}_t)$ (as required in (3.2)). We develop in this Section the case $d = 1$, we refer to [L-10, Section B.2] when $d > 1$.

The first step consists in considering a finite number of chaos. In order to approximate the random variable F , we consider its projection $\mathcal{C}_p(F)$ onto the first p chaos, introduced in Definition 3.2.1. We have

$$\mathcal{C}_p(F) = d_0 + \sum_{1 \leq k \leq p} \sum_{|n|=k} d_k^n \prod_{i \geq 1} K_{n_i} \left(\int_0^T g_i(s) dW_s \right). \quad (3.7)$$

Of course, we still have an infinite number of terms in the previous sum and the second step consists in working with only the first N functions g_1, \dots, g_N of an orthonormal basis of $L^2([0, T]; \mathbb{R})$.

Let us consider a regular mesh grid of N time steps $\mathcal{T} = \{\bar{t}_i = i \frac{T}{N}, i = 0, \dots, N\}$ and the N step functions

$$g_i = \mathbf{1}_{] \bar{t}_{i-1}, \bar{t}_i]}(t) / \sqrt{h}, \quad i = 1, \dots, N, \quad \text{where } h := \frac{T}{N}. \quad (3.8)$$

We complete these N functions g_1, \dots, g_N into an orthonormal basis of $L^2([0, T]; \mathbb{R})$, $(g_i)_{i \geq 1}$. For instance, one can consider the Haar basis on each interval $(\bar{t}_{i-1}, \bar{t}_i)$, $i = 1, \dots, N$. We implicitly assume that $N \geq p$. This leads to the following approximation

$$\mathcal{C}_p^N(F) = d_0 + \sum_{1 \leq k \leq p} \sum_{|n|=k} d_k^n \prod_{1 \leq i \leq N} K_{n_i} \left(\int_0^T g_i(s) dW_s \right), \quad (3.9)$$

where $n = (n_1, \dots, n_N)$ and $|n| = n_1 + \dots + n_N$. Due to the simplicity of the functions g_i , $i = 1, \dots, N$, we can compute explicitly

$$\int_0^T g_i(s) dW_s = G_i, \quad \text{where } G_i = \frac{W_{\bar{t}_i} - W_{\bar{t}_{i-1}}}{\sqrt{h}}.$$

Roughly speaking this means that P_k , the k^{th} chaos, is generated by

$$\{K_{n_1}(G_1) \dots K_{n_N}(G_N) : n_1 + \dots + n_N = k\}.$$

Thus, the approximation we use for the random variable F is

$$\mathcal{C}_p^N(F) = d_0 + \sum_{k=1}^p \sum_{|n|=k} d_k^n K_{n_1}(G_1) \dots K_{n_N}(G_N) = d_0 + \sum_{k=1}^p \sum_{|n|=k} d_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i), \quad (3.10)$$

where the coefficients d_0 and d_k^n are given by

$$d_0 = \mathbb{E}[F], \quad d_k^n = n! \mathbb{E}[F K_{n_1}(G_1) \dots K_{n_N}(G_N)]. \quad (3.11)$$

The following Lemma, similar to Lemma 3.2.4, gives some useful properties of the operator \mathcal{C}_p^N

Lemma 3.2.6 *Let F be a r.v. in $L^2(\mathcal{F}_T)$ and H be in $\mathbb{H}_T^2(\mathbb{R})$. Then*

- $\forall (p, N) \in (\mathbb{N}^*)^2$, $\mathbb{E}(|\mathcal{C}_p^N(F)|^2) \leq \mathbb{E}(|\mathcal{C}_p(F)|^2) \leq \mathbb{E}(|F|^2)$,
- $\mathcal{C}_p^N\left(\int_0^T H_s ds\right) = \int_0^T \mathcal{C}_p^N(H_s) ds$.
- For all $t \leq r$, $D_t \mathbb{E}_r[\mathcal{C}_p^N(F)] = \mathbb{E}_r[\mathcal{C}_{p-1}^N(D_t F)]$.

From (3.10), we deduce the expressions of $\mathbb{E}_t(\mathcal{C}_p^N F)$ and $D_t \mathbb{E}_t(\mathcal{C}_p^N(F))$, useful for the approximation of (Y, Z) by the chaos decomposition (see (3.2)).

Proposition 3.2.7 *Let F be a real random variable in $L^2(\mathcal{F}_T)$ and let r be an integer in $\{1, \dots, N\}$. For all $\bar{t}_{r-1} < t \leq \bar{t}_r$, we have*

$$\begin{aligned} \mathbb{E}_t(\mathcal{C}_p^N F) &= d_0 + \sum_{k=1}^p \sum_{|n(r)|=k} d_k^n \prod_{i < r} K_{n_i}(G_i) \times \left(\frac{t - \bar{t}_{r-1}}{h}\right)^{\frac{nr}{2}} K_{n_r} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}}\right), \\ D_t \mathbb{E}_t(\mathcal{C}_p^N(F)) &= h^{-1/2} \sum_{k=1}^p \sum_{\substack{|n(r)|=k \\ n_r > 0}} d_k^n \prod_{i < r} K_{n_i}(G_i) \times \left(\frac{t - \bar{t}_{r-1}}{h}\right)^{\frac{n_r-1}{2}} K_{n_r-1} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}}\right), \end{aligned}$$

where, if $r \leq N$ and $n = (n_1, \dots, n_N)$, $n(r)$ stands for (n_1, \dots, n_r) .

Remark 3.2.8 For $t = \bar{t}_r$ and $r \geq 1$, Proposition 3.2.7 leads to

$$\begin{aligned} \mathbb{E}_{\bar{t}_r}(\mathcal{C}_p^N F) &= d_0 + \sum_{k=1}^p \sum_{|n(r)|=k} d_k^n \prod_{i \leq r} K_{n_i}(G_i) \\ D_{\bar{t}_r} \mathbb{E}_{\bar{t}_r}(\mathcal{C}_p^N F) &= h^{-1/2} \sum_{k=1}^p \sum_{\substack{|n(r)|=k \\ n_r > 0}} d_k^n \prod_{i < r} K_{n_i}(G_i) \times K_{n_r-1}(G_r). \end{aligned}$$

When $r = 0$, we get $\mathbb{E}_{\bar{t}_0}(\mathcal{C}_p^N F) = d_0$ and we define $D_{\bar{t}_0} \mathbb{E}_{\bar{t}_0}(\mathcal{C}_p^N F) = \frac{1}{\sqrt{h}} d_1^{e_1}$ (which is the limit of $D_t \mathbb{E}_t(\mathcal{C}_p^N F)$ when t tends to 0).

3.2.3 Approximation steps of the BSDE

The algorithm is based on four types of approximations : Picard iterations, a Wiener chaos expansion up to a finite order, the truncation of an $L^2([0, T]; \mathbb{R})$ basis in order to apply formulas of Proposition 3.2.7, and a Monte Carlo method to approximate the coefficients d_0 and d_k^n defined in (3.11). Picard iterations are presented in Section 3.2.3.1, Wiener chaos expansion is presented in Section 3.2.3.2, the truncation is presented in Section 3.2.3.3 and the Monte Carlo method is presented in Section 3.2.3.4.

3.2.3.1 Picard iterations

The first step consists in approximating (Y, Z) — solution to (3.1) — by Picard sequence $(Y^q, Z^q)_q$, built as follows : $(Y^0 = 0, Z^0 = 0)$ and for all $q \geq 1$

$$Y_t^{q+1} = \zeta + \int_t^T f(s, Y_s^q, Z_s^q) ds - \int_t^T Z_s^{q+1} \cdot dW_s, \quad 0 \leq t \leq T. \quad (3.12)$$

From (3.12), under the assumptions that $\xi \in \mathbb{D}^{1,2}$ and $f \in C_b^{0,1,1}$, we express (Y^{q+1}, Z^{q+1}) as a function of the processes (Y^q, Z^q) :

$$Y_t^{q+1} = \mathbb{E}_t \left(\zeta + \int_t^T f(s, Y_s^q, Z_s^q) ds \right), \quad Z_t^{q+1} = D_t Y_t^{q+1}, \quad (3.13)$$

which can also be written

$$Y_t^{q+1} = \mathbb{E}_t \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds \right) - \int_0^t f(s, Y_s^q, Z_s^q) ds, \quad Z_t^{q+1} = D_t Y_t^{q+1}. \quad (3.14)$$

As recalled in the introduction, the computation of the conditional expectation is the cornerstone in the numerical resolution of BSDEs. Chaos decomposition formulas enable to circumvent this problem.

3.2.3.2 Truncation on the first p chaos

Computing the chaos decomposition of the r.v. $F = \zeta + \int_t^T f(s, Y_s^q, Z_s^q) ds$ (appearing in (3.13)) in order to compute Y_t^{q+1} is not judicious. F depends on t , and then the computation of Y^{q+1} on the grid $\mathcal{T} = \{\bar{t}_i = i \frac{T}{N}, i = 0, \dots, N\}$ would require $N + 1$ calls to the chaos decomposition function. To build an efficient algorithm, we need to call the chaos decomposition function as less as possible, since each call is computationally demanding and brings an approximation error due to the truncation and to the Monte-Carlo approximation (see next Sections). Then, we look for a r.v. F^q independent of t such that Y_t^{q+1} and Z_t^{q+1} can be expressed as functions of $\mathbb{E}_t(F^q)$, $D_t \mathbb{E}_t(F^q)$ and of Y^q and Z^q . Equation (3.14) gives a more tractable expression of Y^{q+1} . Let F^q be defined by $F^q := \zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds$. Then

$$Y_t^{q+1} = \mathbb{E}_t(F^q) - \int_0^t f(s, Y_s^q, Z_s^q) ds, \quad Z_t^{q+1} = D_t \mathbb{E}_t(F^q). \quad (3.15)$$

The second type of approximation consists in computing the chaos decomposition of F^q up to order p . Since F^q does not depend on t , the chaos decomposition function \mathcal{C}_p is called only once per Picard iteration.

Let $(Y^{q,p}, Z^{q,p})$ denote the approximation of (Y^q, Z^q) built at step q using a chaos decomposition with order p : $(Y^{0,p}, Z^{0,p}) = (0, 0)$ and

$$Y_t^{q+1,p} = \mathbb{E}_t [C_p(F^{q,p})] - \int_0^t f(s, Y_s^{q,p}, Z_s^{q,p}) ds, \quad Z_t^{q+1,p} = D_t \mathbb{E}_t [C_p(F^{q,p})], \quad (3.16)$$

where $F^{q,p} = \zeta + \int_0^T f(s, Y_s^{q,p}, Z_s^{q,p}) ds$. In the sequel, we also use the following equality

$$Z_t^{q+1,p} = \mathbb{E}_t [D_t C_p(F^{q,p})]. \quad (3.17)$$

3.2.3.3 Truncation of the basis

The third type of approximation comes from the truncation of the orthonormal $L^2([0, T]; \mathbb{R})$ basis used in the definition of C_p (3.7). Instead of considering a basis of $L^2([0, T]; \mathbb{R})$, we only keep the first N functions (g_1, \dots, g_N) defined by (3.8) to build the chaos decomposition function C_p^N (3.9). Proposition 3.2.7 gives us explicit formulas for $\mathbb{E}_t(C_p^N F)$ and $D_t \mathbb{E}_t(C_p^N F)$. From (3.16), we build $(Y^{q,p,N}, Z^{q,p,N})_q$ in the following way : $(Y^{0,p,N}, Z^{0,p,N}) = (0, 0)$ and

$$Y_t^{q+1,p,N} = \mathbb{E}_t(C_p^N(F^{q,p,N})) - \int_0^t f(s, Y_s^{q,p,N}, Z_s^{q,p,N}) ds, \quad Z_t^{q+1,p,N} = D_t(\mathbb{E}_t(C_p^N(F^{q,p,N}))), \quad (3.18)$$

where $F^{q,p,N} := \zeta + \int_0^T f(s, Y_s^{q,p,N}, Z_s^{q,p,N}) ds$.

Equation (3.18) is tractable as soon as we know closed formulas for the coefficients d_k^n of the chaos decomposition of $\mathbb{E}_t(C_p^N(F^{q,p,N}))$ and $D_t(\mathbb{E}_t(C_p^N(F^{q,p,N})))$ (see Proposition 3.2.7). When it is not the case, we need to use a Monte-Carlo method to approximate these coefficients. The next Section is devoted to this method and to the practical implementation.

3.2.3.4 Monte-Carlo simulations of the chaos decomposition

Let F denote a r.v. of $L^2(\mathcal{F}_T)$. Practically, when we are not able to compute exactly d_0 and/or the coefficients d_k^n of the chaos decomposition (3.10)-(3.11) of F , we use Monte-Carlo simulations to approximate them. Let $(F^m)_{1 \leq m \leq M}$ be a M i.i.d. sample of F and $(G_1^m, \dots, G_N^m)_{1 \leq m \leq M}$ be a M i.i.d. sample of (G_1, \dots, G_N) . We recall that d_0 and the coefficients $(d_k^n)_{1 \leq k \leq p, |n|=k}$ are given by $d_0 = \mathbb{E}[F]$ and $d_k^n = n! \mathbb{E}[F K_{n_1}(G_1) \dots K_{n_N}(G_N)]$ (see (3.11)). Then, they are solutions of

$$\arg \min_{\mathbf{c}=(c_0, (c_k^n)_{1 \leq k \leq p, |n|=k})} \mathbb{E}[|F - \psi(\mathbf{c}, G)|^2], \quad (3.19)$$

where $\psi : (\mathbf{c}, G) \mapsto c_0 + \sum_{k=1}^p \sum_{|n|=k} c_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i)$. We approximate $\mathbf{d} := (d_0, (d_k^n)_{1 \leq k \leq p, |n|=k})$ by approximating the expectations of (3.11) by empirical means $\widehat{\mathbf{d}}_M := (\widehat{d}_0, \widehat{d}_k^n)_{1 \leq k \leq p, |n|=k}$ where

$$\widehat{d}_0 := \frac{1}{M} \sum_{m=1}^M F^m, \quad \widehat{d}_k^n := \frac{n!}{M} \sum_{m=1}^M F^m K_{n_1}(G_1^m) \dots K_{n_N}(G_N^m), \quad (3.20)$$

In the following, $C_p^{N,M}(F)$ denotes the approximation of the chaos decomposition of order p of F when using the first method to approximate the coefficients d_k^n :

$$C_p^{N,M}(F) = \widehat{d}_0 + \sum_{k=1}^p \sum_{|n|=k} \widehat{d}_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i). \quad (3.21)$$

$\mathbb{E}_t(\mathcal{C}_p^{N,M}(F))$ and $D_t(\mathbb{E}_t(\mathcal{C}_p^{N,M}(F)))$ denote the conditional expectations obtained in Proposition 3.2.7 when $((d_0, d_k^n)_{1 \leq k \leq p, |n|=k})$ are replaced by $((\widehat{d}_0, \widehat{d}_k^n)_{1 \leq k \leq p, |n|=k})$:

$$\begin{aligned}\mathbb{E}_t(\mathcal{C}_p^{N,M}F) &:= \widehat{d}_0 + \sum_{k=1}^p \sum_{|n(r)|=k} \widehat{d}_k^n \prod_{i < r} K_{n_i}(G_i) \times \left(\frac{t - \bar{t}_{r-1}}{h}\right)^{\frac{n_r}{2}} K_{n_r} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}}\right), \\ D_t \mathbb{E}_t(\mathcal{C}_p^{N,M}(F)) &:= h^{-1/2} \sum_{k=1}^p \sum_{\substack{|n(r)|=k \\ n_r > 0}} \widehat{d}_k^n \prod_{i < r} K_{n_i}(G_i) \times \left(\frac{t - \bar{t}_{r-1}}{h}\right)^{\frac{n_r-1}{2}} K_{n_r-1} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}}\right),\end{aligned}$$

Remark 3.2.9 When M samples of $\mathcal{C}_p^{N,M}(F)$ are needed, we can either use the same samples as the ones used to compute \widehat{d}_0 and \widehat{d}_k^n : $(\widehat{\mathcal{C}}_p^N(F))^m = \widehat{d}_0 + \sum_{k=1}^p \sum_{|n|=k} \widehat{d}_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$, or use new ones. In the first case, we only require M samples of F and (G_1, \dots, G_N) . The coefficients \widehat{d}_k^n and \widehat{d}_0 are not independent of $\prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$. The notation $\mathbb{E}_t(\mathcal{C}_p^{N,M}(F))$ introduced above cannot be linked to $\mathbb{E}(\mathcal{C}_p^{N,M}F|\mathcal{F}_t)$. In the second case, the coefficients \widehat{d}_k^n and \widehat{d}_0 are independent of $\prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$ and we have $\mathbb{E}_t(\mathcal{C}_p^{N,M}F) = \mathbb{E}(\mathcal{C}_p^{N,M}F|\mathcal{F}_t)$. This second approach requires $2M$ samples of F and (G_1, \dots, G_N) and its variance increases with N . Practically, we use the first technique.

We introduce the processes $(Y^{q+1,p,N,M}, Z^{q+1,p,N,M})$, useful in the following. It corresponds to the approximation of $(Y^{q+1,p,N}, Z^{q+1,p,N})$ when we use $\mathcal{C}_p^{N,M}$ instead of \mathcal{C}_p^N , i.e. when we use a Monte Carlo procedure to compute the coefficients d_k^n .

$$Y_t^{q+1,p,N,M} = \mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M})) - \int_0^t f(\theta_s^{q,p,N,M}) ds, \quad Z_t^{q+1,p,N,M} = D_t(\mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M}))), \quad (3.22)$$

where $F^{q,p,N,M} := \zeta + \int_0^T f(\theta_s^{q,p,N,M}) ds$ and $\theta_s^{q,p,N,M} = (s, Y_s^{q,p,N,M}, Z_s^{q,p,N,M})$.

3.2.4 Convergence results

We aim at bounding the error between (Y, Z) — the solution to (3.1) — and $(Y^{q,p,N,M}, Z^{q,p,N,M})$ defined by (3.22). Before stating the main result of the paper, we introduce some hypotheses.

In the following, (t_1, \dots, t_n) and (s_1, \dots, s_n) denote two vectors such that

$$0 \leq t_1 \leq \dots \leq t_n \leq T, \quad 0 \leq s_1 \leq \dots \leq s_n \leq T \quad \text{and} \quad \forall i, \quad s_i \leq t_i.$$

Hypothesis 3.1 Let $m \in \mathbb{N}^*$. We say that F satisfies Hypothesis \mathcal{H}_m if F satisfies the two following hypotheses

- \mathcal{H}_m^1 : $\forall j \geq 2$ $F \in \mathcal{D}^{m,j}$, i.e. $\|F\|_{m,j}^j < \infty$
- \mathcal{H}_m^2 : $\forall j \geq 2, \forall i \in \{1, \dots, m\}, \forall l_0 \leq i - 1, \forall l_1 \leq m - i, \forall l \in \{1, \dots, d\}$ and for all multi-indices α_0 and α_1 such that $|\alpha_0| = l_0$ and $|\alpha_1| = l_1 + 1$, there exist two positive constants β_F and k_l^F such that

$$\sup_{t_1 \leq \dots \leq t_{l_0}} \sup_{s_{i+1} \leq \dots \leq s_{i+l_1}} \mathbb{E}[|D_{t_1, \dots, t_{l_0}}^{\alpha_0} (D_{t_i, s_{i+1}, \dots, s_{i+l_1}}^{\alpha_1} F - D_{s_i, \dots, s_{i+l_1}}^{\alpha_1} F)|^j] \leq k_l^F(j) (t_i - s_i)^{j\beta_F},$$

where $l = l_0 + l_1 + 1$. In the following, we denote $K_m^F(j) = \sup_{l \leq m} k_l^F(j)$.

Remark 3.2.10 If F satisfies \mathcal{H}_m^2 , for all multi-index α such that $|\alpha| = l$ we have

$$|\mathbb{E}(D_{t_1, \dots, t_l}^\alpha F) - \mathbb{E}(D_{s_1, \dots, s_l}^\alpha F)| \leq K_l^F ((t_1 - s_1)^{\beta_F} + \dots + (t_l - s_l)^{\beta_F}), \quad (3.23)$$

where K_l^F is a constant.

Lemma 3.2.11 Let F denote a r.v. in $L^2(\mathcal{F}_T)$ satisfying (3.23) for $m = p$. We have

$$\mathbb{E}(|(\mathcal{C}_p^N - \mathcal{C}_p)(F)|^2) \leq (K_p^F)^2 \left(\frac{T}{N}\right)^{2\beta_F} T(1+T)e^T,$$

where K_p^F and β_F are defined in Hypothesis 3.1.

Hypothesis 3.2 Let $(p, N) \in \mathbb{N}^2$. We say that a r.v. F satisfies $\mathcal{H}_{p,N}^3$ if

$$V_{p,N}(F) := \mathbb{V}(F) + \sum_{k=1}^p \sum_{|n|=k} n! \mathbb{V} \left(F \prod_{i=1}^N K_{n_i}(G_i) \right) < \infty.$$

Lemma 3.2.12 Let F be a r.v. satisfying Hypothesis $\mathcal{H}_{p,N}^3$. We have

$$\mathbb{E}(|(\mathcal{C}_p^N - \mathcal{C}_p^{N,M})(F)|^2) = \frac{1}{M} V_{p,N}(F).$$

Moreover, we have $\mathbb{E}(|\mathcal{C}_p^{N,M}(F)|^2) \leq \mathbb{E}(|F|^2) + \frac{1}{M} V_{p,N}(F)$.

Remark 3.2.13 If F is bounded by K , we get $V_{p,N}(F) \leq K^2 \sum_{k=0}^p \binom{N}{k}$. Then, every bounded r.v. satisfies $\mathcal{H}_{p,N}^3$.

This Remark ensues from $\mathbb{E} \left(\prod_{i=1}^N K_{n_i}^2(G_i) \right) = \frac{1}{n!}$.

Remark 3.2.14 Let X be the \mathbb{R}^{m_0} -valued process solution of

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

where W is a d -dimensional Brownian motion and $b : [0, T] \times \mathbb{R}^{m_0} \rightarrow \mathbb{R}^{m_0}$ and $\sigma : [0, T] \times \mathbb{R}^{m_0} \rightarrow \mathbb{R}^{m_0 \times d}$ are two $C^{0,m}$ functions uniformly lipschitz w.r.t. x and Hölder continuous of parameter $\frac{1}{2}$ w.r.t. t , with linear growth in x and with bounded derivatives. Then, every random variable ζ of type $g(X_T)$ or $g(\int_0^T X_s ds)$ with $g : \mathbb{R}^n \rightarrow \mathbb{R}$ in C_p^∞ satisfies \mathcal{H}_m and $\mathcal{H}_{p,N}^3$, for all p and N .

Theorem 3.2.15 Let k be an integer s.t. $k \leq p$. Assume that ζ satisfies \mathcal{H}_{p+q} and $\mathcal{H}_{p,N}^3$ and $f \in C_b^{0,p+q-1,p+q-1}$. We have

$$\|(Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M})\|_{L_2}^2 \leq \frac{A_0}{2^q} + \frac{A_1(q, k)}{(p+1)^k} + A_2(q, p) \left(\frac{T}{N}\right)^{2\beta_\zeta \wedge 1} + \frac{A_3(q, p, N)}{M},$$

where A_0, A_1, A_2 and A_3 are given in the following proof. If $f \in C_b^{0,\infty,\infty}$ and ζ satisfies \mathcal{H}_∞ and $\mathcal{H}_{\infty,\infty}^3$, we get

$$\lim_{q \rightarrow \infty} \lim_{p \rightarrow \infty} \lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \|(Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M})\|_{L_2}^2 = 0.$$

Proof. We split the error in 4 terms :

1. Picard iterations : $\mathcal{E}^q = \|(Y - Y^q, Z - Z^q)\|_{L^2}^2$, where (Y^q, Z^q) is defined by (3.12),
2. the truncation of the chaos decomposition : $\mathcal{E}^{q,p} = \|(Y^q - Y^{q,p}, Z^q - Z^{q,p})\|_{L^2}^2$, where $(Y^{q,p}, Z^{q,p})$ is defined by (3.16),
3. the truncation of the $L^2([0, T]; \mathbb{R})$ basis : $\mathcal{E}^{q,p,N} = \|(Y^{q,p} - Y^{q,p,N}, Z^{q,p} - Z^{q,p,N})\|_{L^2}^2$, where $(Y^{q,p,N}, Z^{q,p,N})$ is defined by (3.18),
4. the Monte-Carlo approximation to compute the expectations : $\mathcal{E}^{q,p,N,M} = \|(Y^{q,p,N} - Y^{q,p,N,M}, Z^{q,p,N} - Z^{q,p,N,M})\|_{L^2}^2$, where $(Y^{q,p,N,M}, Z^{q,p,N,M})$ is defined by (3.22).

We have

$$\|(Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M})\|_{L^2}^2 \leq 4(\mathcal{E}^q + \mathcal{E}^{q,p} + \mathcal{E}^{q,p,N} + \mathcal{E}^{q,p,N,M}).$$

From [EKQP97b, Corollary 2.1], we know that under Hypothesis 1.1 we have

$$\mathcal{E}^q := \|(Y - Y^q, Z - Z^q)\|_{L^2}^2 \leq \frac{A_0}{2^q}, \quad \blacksquare$$

where A_0 depends on T , $\|\xi\|_2^2$ and on $\|f(\cdot, 0, 0)\|_{L^2([0, T]; \mathbb{R})}^2$. The same procedure applies in the proof of [L-10, Propositions 4.11, 4.15 and 4.17] which bound $\mathcal{E}^{q,p}$, $\mathcal{E}^{q,p,N}$ and $\mathcal{E}^{q,p,N,M}$. Thanks to Picard iterations $\mathcal{E}^{q+1,\cdot}$ is bounded by $C_{0,\cdot}\mathcal{E}^{q,\cdot} + C_{1,\cdot}$, where $C_{1,p}$ is related to Lemma 3.2.3, $C_{1,p,N}$ is related to Lemma 3.2.11 and $C_{1,p,N,M}$ is related to Lemma 3.2.12. Since $\mathcal{E}^{0,\cdot} = 0$ we get the following upper bounds. From [L-10, Proposition 4.11] we get that $\mathcal{E}^{q,p} \leq \frac{A_1(q,m)}{(p+1)^m}$ where $A_1(q,m) := \frac{(C_1 T(T+1)L_f^2)^{q-1}}{C_1 T(T+1)L_f^2 - 1} K_1(q,m)$ and $K_1(q,m)$ depends on T , m , $\|\xi\|_{m+q, 2^{\frac{(m+q-1)!}{(m-1)!}}}$ and on $(\|\partial_{sp}^k f\|_\infty)_{1 \leq k \leq m+q-1}$. [L-10, Proposition 4.15] gives that $\mathcal{E}^{q,p,N} \leq A_2(q,p) \left(\frac{T}{N}\right)^{1 \wedge 2\beta\xi}$, where $A_2(q,p) := K_2(q,p)T(T+1)e^T \frac{(C_2 T(T+1)L_f^2)^{q-1}}{C_2 T(T+1)L_f^2 - 1}$ and $K_2(q,p)$ depends on K_p^ξ , T , $\|\xi\|_{p+q, 2^{\frac{(p+q-1)!}{(p-1)!}}}$ and on $(\|\partial_{sp}^k f\|_\infty)_{1 \leq k \leq p+q-1}$. [L-10, Proposition 4.17] states that $\mathcal{E}^{q,p,N,M} \leq \frac{A_3(q,p,N)}{M}$, where $A_3(q,p,N) := K_3(p,N) \frac{(C_3 T(T+1)L_f^2)^{q-1}}{C_3 T(T+1)L_f^2 - 1}$ and $K_3(p,N) := 168 \left(V_{p,N}(\xi) + T^2 \|f\|_\infty^2 \sum_{k=0}^p \binom{N}{k} \right)$.

3.2.5 Numerical Examples

The computations have been done on a PC INTEL Core 2 Duo P9600 2.53 GHz with 4Gb of RAM.

3.2.5.1 Nonlinear driver and path-dependent terminal condition

We consider the case of a non linear driver and a path-dependent terminal condition in dimension $d = 1$: $f(t, y, z) = \cos(y)$ and $\zeta = \sup_{0 \leq t \leq 1} W_t$.

- **Convergence in p .** Table 3.1 represents the evolution of $\bar{Y}_0^{q,p,N,M}$ and $\bar{Z}_0^{q,p,N,M}$ w.r.t q (Picard iteration index), when $p = 2$ and $p = 3$. We also give the CPU time needed to get $\bar{Y}_0^{6,p,N,M}$ and $\bar{Z}_0^{6,p,N,M}$. We fix $M = 10^5$ and $N = 20$. The seed of the generator is also fixed.

One notes that the difference between the values of $\bar{Y}_0^{q,2,N,M}$ and $\bar{Y}_0^{q,3,N,M}$ (resp. $\bar{Z}_0^{q,2,N,M}$ and $\bar{Z}_0^{q,3,N,M}$) doesn't exceed 0.2% (resp. 0.6%). This is due to the fast convergence of the algorithm in p . The CPU time is 12 times higher when $p = 3$ than when $p = 2$. Then, the use of order 3 in the chaos decomposition is not necessary. In the following, we take $p = 2$.

iterations	1	2	3	4	5	6	CPU time
$p = 2$	1.656357	1.017117	1.237135	1.186691	1.195462	1.194256	14.06
$p = 3$	1.656357	1.012091	1.234398	1.183544	1.192367	1.191173	174.09

Table 3.1: Evolution of $\bar{Y}_0^{q,p,N,M}$ w.r.t. Picard's iterations, $M = 10^5$, $N = 20$ and CPU time

iterations	1	2	3	4	5	6	CPU time
$p = 2$	0.969128	0.249148	0.525273	0.459326	0.470069	0.469117	14.06
$p = 3$	0.969128	0.242977	0.523846	0.455827	0.466903	0.465939	174.09

Table 3.2: Evolution of $\bar{Z}_0^{q,p,N,M}$ w.r.t. Picard's iterations, $M = 10^5$, $N = 20$ and CPU time

- **Convergence in M .** Figure 3.2.1 illustrates the evolution of $\bar{Y}_0^{q,p,N,M}$ and $\bar{Z}_0^{q,p,N,M}$ w.r.t. q when $p = 2$ and $N = 20$ for different values of M . The seed of the generator is random. When M equals 10^4 and 10^5 the algorithm stabilizes after very few iterations. When $M = 10^3$, there is no convergence.

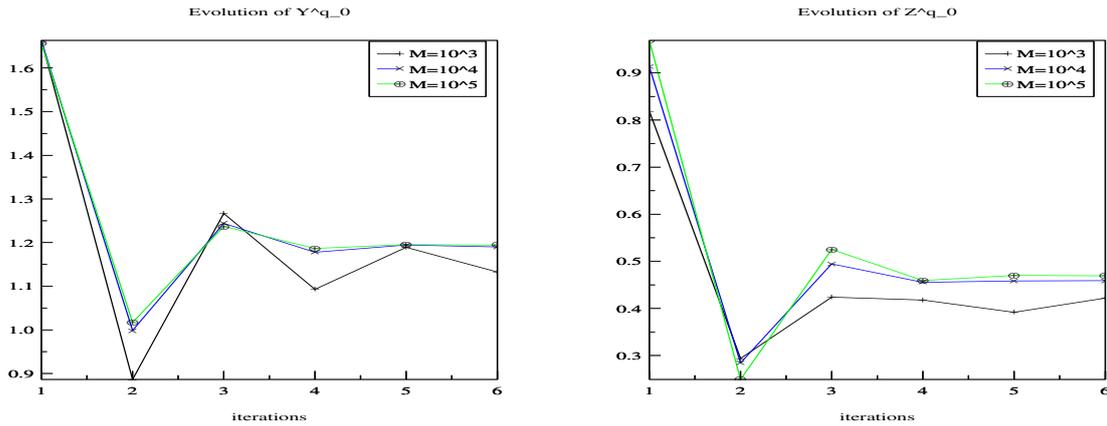


Figure 3.2.1: Evolution of $\bar{Y}_0^{q,p,N,M}$ and $\bar{Z}_0^{q,p,N,M}$ w.r.t. q and M when $N = 20$, $p = 2 - \zeta = \sup_{0 \leq t \leq 1} W_t$, $f(t, y, z) = \cos(y)$.

- **Convergence in N .** Figure 3.2.2 illustrates the evolution of $\bar{Y}_0^{q,p,N,M}$ and $\bar{Z}_0^{q,p,N,M}$ w.r.t. q when $p = 2$ and $M = 10^5$ for different values of N . The seed of the generator is random. The algorithm converges even when $N = 10$, but $\bar{Y}_0^{6,p,10,M}$ is quite below $\bar{Y}_0^{6,p,40,M}$.

3.2.5.2 Linear Driver - Financial Benchmark

We consider the case of pricing and hedging a Discrete Down and Out Barrier Call option, i.e. $f(t, y, z) = -ry$ and $\xi := (S_T - K)_+ \mathbf{1}_{\forall n \in [0, N] S_{t_n} \geq L}$, where S represents the Black-Scholes diffusion

$$S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t}, \quad \forall t \in [0, T].$$

The option parameters are $r = 0.01$, $\sigma = 0.2$, $T = 1$, $K = 0.9$, $L = 0.85$, $S_0 = 1$ and $N = 20$ (N is also the number of time discretizations of the chaos decomposition).

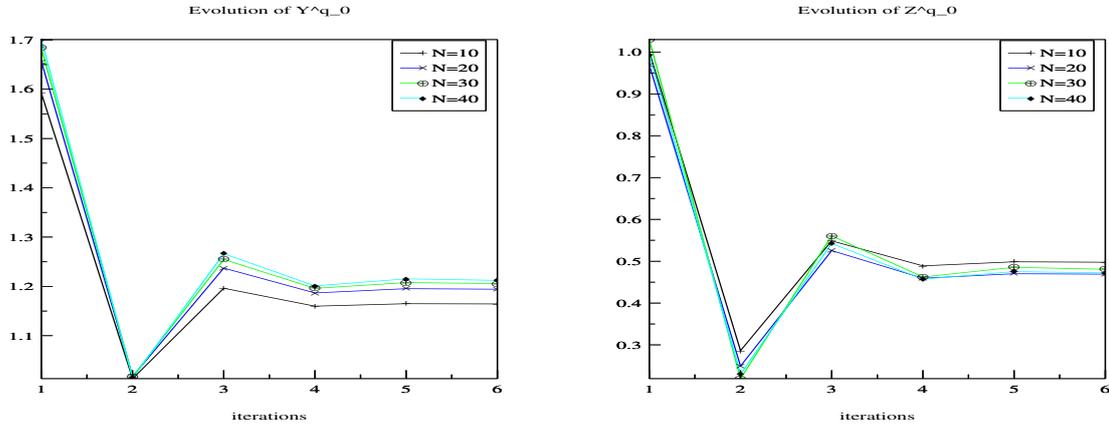


Figure 3.2.2: Evolution of $\bar{Y}_0^{q,p,N,M}$ and $\bar{Z}_0^{q,p,N,M}$ w.r.t. N when $M = 10^5$, $p = 2 - \zeta = \sup_{0 \leq t \leq 1} W_t$, $f(t, y, z) = \cos(y)$

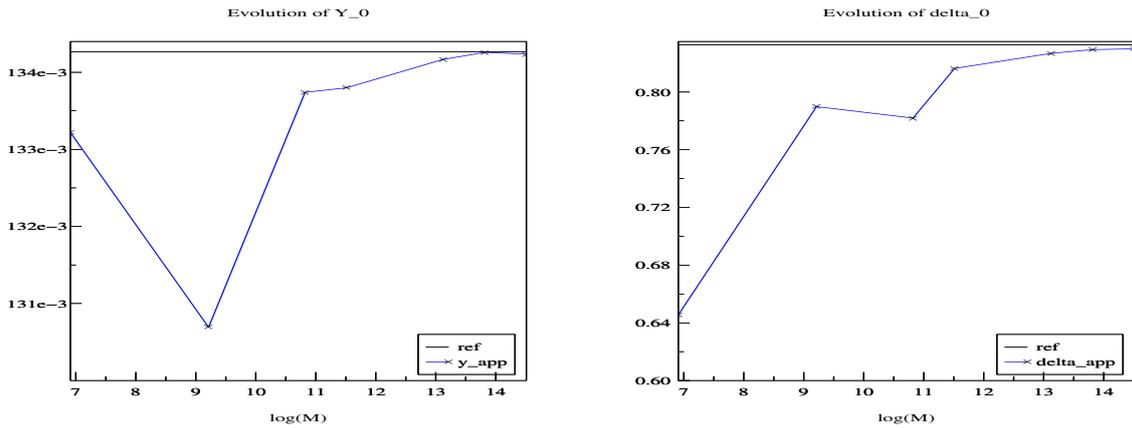


Figure 3.2.3: Evolution of $\bar{Y}_0^{q,p,N,M}$ and $\delta_0 := \frac{\bar{Z}_0^{q,p,N,M}}{\sigma S_0}$ w.r.t. $\log(M)$ when $N = 20$, $p = 2$, $q = 5$ -Discrete Down and Out Barrier Call option

We can get a benchmark for Y_0 and Z_0 by using a variance reduction Monte Carlo method. For this set of parameters, the reference values are $Y_0 = 0.134267$ with a confidence interval $7.9468e-05$ and $\delta_0 = \frac{Z_0}{\sigma S_0} = 0.8327$. We compare them with $\bar{Y}_0^{q,p,N,M}$ and $\frac{\bar{Z}_0^{q,p,N,M}}{\sigma S_0}$ when $N = 20$, $p = 2$, $q = 5$ (we choose the first value of q from which the algorithm has converged) for different values of M . Figure 3.2.3 represents the evolution of $\bar{Y}_0^{5,p,N,M}$ and $\delta_0^{5,p,N,M}$ w.r.t. $\log(M)$. One notices that for $M = 10^6$ the computed values are very close to the reference ones.

3.3 Case with jumps

In this Section we present the results obtained in [L-7] written with Christel Geiss. We are interested in the numerical approximation of solutions (Y, Z, U) to backward stochastic differential equations (BSDEs in the sequel) with jumps of the following form

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s, U_s) ds - \int_t^T Z_s dW_s - \int_{]t,T]} U_s d\tilde{N}_s, \quad 0 \leq t \leq T, \quad (3.24)$$

where W is a 1-dimensional standard Brownian motion and \tilde{N} is a compensated Poisson process independent from W , i.e. $\tilde{N}_t := N_t - \kappa t$ and $(N_t)_{t \geq 0}$ is a Poisson process with intensity $\kappa > 0$. The terminal condition ζ is a real-valued \mathcal{F}_T -measurable random variable where $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ stands for the augmented natural filtration associated with W and N . Under standard Lipschitz assumptions on the driver f , the existence and uniqueness of the solution have been stated by Tang and Li [TL94], generalizing the seminal paper of Pardoux and Peng [PP92].

The main objective of this section is to propose a numerical method to approximate the solution (Y, Z, U) of (3.24) by extending the results presented in Section 3.2 to the case of jumps. In case of BSDEs driven by a Poisson random measure, Bouchard and Elie [BE08] have proposed a scheme based on the dynamic programming equation and studied the rate of convergence of the method when the terminal condition is given by $\zeta = g(X_T)$, where g is a Lipschitz function and X is a forward process. More recently, Geiss and Steinicke [GS16] have extended this result to the case of a terminal condition which may be a Borel function of finitely many increments of the Lévy forward process X which is not necessarily Lipschitz but only satisfies a fractional smoothness condition. In the case of jumps driven by a compensated Poisson process, Lejay, Mordecki and Torres [LMT13] have developed a fully implementable scheme based on a random binomial tree, following the approach proposed by Briand, Delyon and Mémin [BDM01].

As said before, we extend here the algorithm based on Picard iterations and Wiener chaos expansion introduced in [L-10] to the case of BSDEs with jumps. Our starting point is the use of Picard iterations: $(Y^0, Z^0, U^0) = (0, 0, 0)$ and for $q \in \mathbb{N}$,

$$Y_t^{q+1} = \zeta + \int_t^T f(s, Y_s^q, Z_s^q, U_s^q) ds - \int_t^T Z_s^{q+1} \cdot dW_s - \int_{]t,T]} U_s^{q+1} d\tilde{N}_s, \quad 0 \leq t \leq T.$$

Writing this Picard scheme in a forward way gives

$$\begin{aligned} Y_t^{q+1} &= \mathbb{E} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q, U_s^q) ds \mid \mathcal{F}_t \right) - \int_0^t f(s, Y_s^q, Z_s^q, U_s^q) ds, \\ Z_t^{q+1} &= \mathbb{E} \left(D_t^{(0)} Y_t^{q+1} \mid \mathcal{F}_{t-} \right) = \mathbb{E} \left(D_t^{(0)} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q, U_s^q) ds \right) \mid \mathcal{F}_{t-} \right), \\ U_t^{q+1} &= \mathbb{E} \left(D_t^{(1)} Y_t^{q+1} \mid \mathcal{F}_{t-} \right) = \mathbb{E} \left(D_t^{(1)} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q, U_s^q) ds \right) \mid \mathcal{F}_{t-} \right), \end{aligned}$$

where $D_t^{(0)} X$ (resp. $D_t^{(1)} X$) stands for the Malliavin derivative of the random variable X with respect to the Brownian motion (resp. w.r.t. the Poisson process).

In order to compute the previous conditional expectation, we use a Wiener chaos expansion of the random variable

$$F^q = \zeta + \int_0^T f(s, Y_s^q, Z_s^q, U_s^q) ds.$$

More precisely, we use the following orthogonal decomposition of the random variable F^q (see [L-7, Proposition 2.6])

$$F^q = \mathbb{E}[F^q] + \sum_{k=1}^{\infty} \sum_{l=0}^k \sum_{\mathbf{k}_l \in \mathbb{N}^l} \sum_{\mathbf{j}_{k-l} \in \mathbb{N}^{k-l}} d_{\mathbf{k}_l, \mathbf{j}_{k-l}} L_l^{0, \dots, 0}(\tilde{e}[k_1, \dots, k_l]) L_{k-l}^{1, \dots, 1}(\tilde{e}[j_1, \dots, j_{k-l}]).$$

where $L_m^{0, \dots, 0}(g)$ (resp. $L_m^{1, \dots, 1}(g)$) denotes the iterated integral of order m of g w.r.t. the Brownian motion (resp. w.r.t. the compensated Poisson process), $(\tilde{e}[k_1, \dots, k_m])_{\mathbf{k}_m \in \mathbb{N}^m}$ is an orthogonal basis of $(\tilde{L}^2)^{\otimes m}([0, T]; \mathbb{R})$, the subspace of symmetric functions from $(L^2)^{\otimes m}([0, T]; \mathbb{R})$. The sequence of coefficients $\{d_{\mathbf{k}_l, \mathbf{j}_{k-l}}\}_{\mathbf{k}_l \in \mathbb{N}^l, \mathbf{j}_{k-l} \in \mathbb{N}^{k-l}}$ ensues from the Wiener chaos decomposition of F^q .

The point to get an implementable scheme is that we only keep a finite number of terms in this expansion: we use a finite number of chaos and we choose a finite number of functions $\{e_1, \dots, e_N\}$ to build $\{\tilde{e}[k_1, \dots, k_m]\}_{\mathbf{k}_m \in \{1, \dots, N\}^m}$. More precisely, if we choose $e_i := \frac{1}{\sqrt{h}} \mathbf{1}_{] \bar{t}_{i-1}, \bar{t}_i]}$ where $\bar{t}_i = ih$ and $h := \frac{T}{N}$, we obtain

$$F^q \sim \mathbb{E}[F^q] + \sum_{k=1}^p \sum_{|n|=k} d_k^n \prod_{i=1}^N K_{n_i^P} \left(\frac{W_{\bar{t}_i} - W_{\bar{t}_{i-1}}}{\sqrt{h}} \right) C_{n_i^P}(N_{\bar{t}_i} - N_{\bar{t}_{i-1}}, \kappa h),$$

where K_i (resp. C_i) denotes the Hermite (resp. Charlier) polynomial of degree i , $n = (n_1^W, \dots, n_N^W, n_1^P, \dots, n_N^P)$ is a vector of integers and $|n| = \sum_{i=1}^N (n_i^W + n_i^P)$. By using this approximation of F^q we can easily compute $\mathbb{E}(F^q | \mathcal{F}_t)$, $\mathbb{E}(D_t^{(0)} F^q | \mathcal{F}_{t-})$ and $\mathbb{E}(D_t^{(1)} F^q | \mathcal{F}_{t-})$, which gives us $(Y_t^{q+1}, Z_t^{q+1}, U_t^{q+1})$. To get a fully implementable algorithm, it remains to approximate $\mathbb{E}(F^q)$ and the coefficients $\{d_k^n\}_{n, k}$ by Monte Carlo.

There is no hypercontractivity property in the Poisson chaos decomposition case. This property plays an important role in the proof of the convergence in the Brownian case. To circumvent this problem, we exploit a result of Last, Penrose, Schulte and Thäle [LPST14], which gives a formula to compute the expectation of products of Poisson multiple integrals, and the according result for the Brownian case from Peccati and Taqqu [PT11]. In fact, in [L-7, Proposition 2.9] we get an explicit expression for

$$\mathbb{E}(I_{n_1}(f_{n_1}) \cdots I_{n_l}(f_{n_l}))$$

in terms of a combinatoric sum of tensor products of the chaos kernels f_{n_i} . Here $I_{n_i}(f_{n_i})$ denotes the multiple integral of order n_i with respect to the process $W + \tilde{N}$. By this expression one gets the required estimates for the truncated chaos without the hypercontractivity property. Therefore, to prove the convergence of the method we may proceed similarly to [L-10], and split the error into four terms:

- the error due to Picard iterations
- the error due to the truncation onto the chaos up to order p
- the error due to the finite number of basis functions $\{e_1, \dots, e_N\}$ for each chaos
- the error due to the Monte Carlo simulations to approximate the expectations appearing in the coefficients $\{d_k^n\}_{n,k}$.

The rest of the Section is organized as follows: Section 3.3.1 presents some preliminaries, Sections 3.3.2 and 3.3.3 introduce Wiener chaos expansion and its approximation in case of jumps, Section 3.3.4 presents the approximation steps of the BSDE with jumps, Section 3.3.5 states the convergence results and finally Section 3.3.6 presents numerical examples.

3.3.1 Preliminaries

Definition 3.3.1 $\|(\cdot, \cdot, \cdot)\|_{L^p}^p$, $p \geq 1$, the norm on the space $S_T^p(\mathbb{R}) \times H_T^p(\mathbb{R}) \times H_T^p(\mathbb{R})$ is defined by

$$\|(Y, Z, U)\|_{L^p}^p := \mathbb{E} \left(\sup_{t \in [0, T]} |Y_t|^p \right) + \int_0^T \mathbb{E}(|Z_t|^p) dt + \kappa \int_0^T \mathbb{E}(|U_t|^p) dt. \quad (3.25)$$

Hypothesis 3.3 We assume

- the terminal condition ζ belongs to $L^2(\mathcal{F}_T)$;
- the generator $f \in C([0, T] \times \mathbb{R}^3; \mathbb{R})$ is Lipschitz continuous in space, uniformly in t : there exists a constant L_f such that

$$|f(t, y_1, z_1, u_1) - f(t, y_2, z_2, u_2)| \leq L_f (|y_1 - y_2| + |z_1 - z_2| + |u_1 - u_2|).$$

Lemma 3.3.2 If Hypothesis (3.3) is satisfied and $\zeta \in \mathbb{D}^{1,2}$ (defined below) we get from [GS16, Theorem 3.4] that for a.e. $t \in [0, T]$

$$Z_t = \mathbb{E}[D_t^{(0)} Y_t | \mathcal{F}_{t-}], \quad U_t = \mathbb{E}[D_t^{(1)} Y_t | \mathcal{F}_{t-}] \quad \mathbb{P} - a.s. \quad (3.26)$$

where $D_t^{(0)} X$ stands for the Malliavin derivative w.r.t. the Brownian motion of the random variable X , and $D_t^{(1)} X$ stands for the Malliavin derivative w.r.t. the Poisson process of the random variable X . Here $\mathbb{E}[\cdot | \mathcal{F}_{t-}]$ should be understood as the predictable projection, and since the paths $s \mapsto D_t^{(i)} Y_s$ are a.s. càdlàg we define $D_t^{(i)} Y_t := \lim_{s \downarrow t} D_t^{(i)} Y_s$ if the limit exists, and zero otherwise.

3.3.2 Wiener chaos expansion

We refer to [LJUV02] and [Pri09] for more details on this section. Let us briefly recall the Wiener chaos expansion in the case of a real-valued Brownian motion and an independent Poisson process with intensity $\kappa > 0$.

We define

$$G_0(t) = B_t, \quad G_1(t) = N_t - \kappa t,$$

and $L_k^{i_1, \dots, i_k}(f)$ the iterated integral of f with respect to G_0 and G_1

$$L_k^{i_1, \dots, i_k}(f) = \int_0^T \left(\int_0^{t_k^-} \cdots \left(\int_0^{t_2^-} f(t_1, \dots, t_k) dG_{i_1}(t_1) \right) \cdots dG_{i_{k-1}}(t_{k-1}) \right) dG_{i_k}(t_k).$$

We have the following chaotic representation property.

Proposition 3.3.3 ([LJUV02, Proposition 2.1]) For $k \in \mathbb{N}^*$ define

$$\mathbf{i}_k := (i_1, \dots, i_k) \in \{0, 1\}^k.$$

Any $F \in L^2(\mathcal{F}_T)$ has a unique representation of the form

$$F = \mathbb{E}(F) + \sum_{k=1}^{\infty} \sum_{\mathbf{i}_k \in \{0, 1\}^k} L_k^{\mathbf{i}_k}(f_{\mathbf{i}_k}), \quad (3.27)$$

where $f_{\mathbf{i}_k} \in L^2(\Sigma_k)$ and $\Sigma_k = \{(t_1, \dots, t_k) \in [0, T]^k : 0 < t_1 < \dots < t_k < T\}$ is the simplex of $[0, T]^k$.

Let $|\mathbf{i}_k| := \sum_{j=1}^k i_j$. Due to the isometry property it holds

$$\|L_k^{\mathbf{i}_k}(f)\|^2 = \kappa^{|\mathbf{i}_k|} \|f\|_{\Sigma_k}^2,$$

and for any $f \in L^2(\Sigma_k)$, $g \in L^2(\Sigma_m)$, $\mathbf{i}_k \in \{0, 1\}^k$, and $\mathbf{j}_m \in \{0, 1\}^m$ we have (see [LJUV02, Proposition 1.1])

$$\mathbb{E}[L_k^{\mathbf{i}_k}(f)L_m^{\mathbf{j}_m}(g)] = \begin{cases} \kappa^{|\mathbf{i}_k|} \int_{\Sigma_k} f(t_1, \dots, t_k) g(t_1, \dots, t_k) dt_1 \cdots dt_k & \text{if } \mathbf{i}_k = \mathbf{j}_m \\ 0 & \text{otherwise.} \end{cases}$$

Then, $\|F\|_2^2 = \mathbb{E}[F]^2 + \sum_{k \geq 1} \sum_{\mathbf{i}_k} \kappa^{|\mathbf{i}_k|} \|f_{\mathbf{i}_k}\|_{L^2(\Sigma_k)}^2$. The chaos approximation of F up to order p is defined by

$$\mathcal{C}_p(F) := \mathbb{E}(F) + \sum_{k=1}^p \sum_{\mathbf{i}_k} L_k^{\mathbf{i}_k}(f_{\mathbf{i}_k}). \quad (3.28)$$

- Let $f \in L^2(\Sigma_k)$ and $j \in \{0, 1\}$. Following [LJUV02], we define the derivative of $L_k^{\mathbf{i}_k}(f)$ w.r.t. the Brownian motion and the Poisson process as the element of $L^2(\Omega \times [0, T])$ given by

$$D_t^{(j)} L_k^{\mathbf{i}_k}(f) = \sum_{l=1}^k \mathbf{1}_{\{i_l=j\}} L_{k-1}^{i_1, \dots, \widehat{i_l}, \dots, i_k}(f(\underbrace{\dots}_{l-1}, t, \dots)), \quad (3.29)$$

where \widehat{i} means that the i -th index is omitted.

- Let $j \in \{0, 1\}$. We extend the definition of $D^{(j)}$ to

$$\text{Dom } D^{(j)} := \left\{ F \in L^2(\mathcal{F}_T) \text{ satisfying (3.27) and } \sum_{k=1}^{\infty} \sum_{\mathbf{i}_k} \sum_{l=1}^k \mathbf{1}_{\{i_l=j\}} \kappa^{|\mathbf{i}_k|} \|f_{\mathbf{i}_k}\|_{\Sigma_k}^2 < \infty \right\}.$$

If $F \in \text{Dom } D^{(j)}$ then

$$\|F\|_{\text{Dom } D^{(j)}}^2 := \mathbb{E}|F|^2 + \kappa^j \mathbb{E} \int_0^T |D_t^{(j)} F|^2 dt < \infty.$$

- F with chaotic representation (3.27) belongs to $\text{Dom } D =: \mathbb{D}^{1,2}$ if F belongs to $\text{Dom } D^{(0)} \cap \text{Dom } D^{(1)}$, i.e.

$$\|F\|_{\mathbb{D}^{1,2}}^2 := \mathbb{E}|F|^2 + \sum_{k=1}^{\infty} k \sum_{\mathbf{i}_k} \kappa^{|\mathbf{i}_k|} \|f_{\mathbf{i}_k}\|_{\Sigma_k}^2 < \infty.$$

More generally, we define $\mathbb{D}^{m,2}$ as follows:

- Let $m \geq 1$. We say that F satisfying (3.27) belongs to $\mathbb{D}^{m,2}$ if it holds

$$\|F\|_{\mathbb{D}^{m,2}}^2 := \mathbb{E}|F|^2 + \sum_{l=1}^m \sum_{k=l}^{\infty} \frac{k!}{(k-l)!} \sum_{\mathbf{i}_k} \kappa^{|\mathbf{i}_k|} \|f_{\mathbf{i}_k}\|_{\Sigma_k}^2 < \infty.$$

We recall

$$\mathbb{D}^{\infty,2} = \bigcap_{m=1}^{\infty} \mathbb{D}^{m,2}.$$

Lemma 3.3.4 *Let $1 \leq m \leq p+1$ and $F \in \mathbb{D}^{m,2}$. We have*

$$\mathbb{E}[|F - \mathcal{C}_p(F)|^2] \leq \frac{\|F\|_{\mathbb{D}^{m,2}}^2}{(p+2-m) \cdots (p+1)}.$$

Equation (3.27) can be rewritten in the following form

Proposition 3.3.5 *Any $F \in L^2(\mathcal{F}_T)$ can be represented as*

$$F = \mathbb{E}[F] + \sum_{k=1}^{\infty} \sum_{l=0}^k \sum_{\mathbf{k}_l \in \mathbb{N}^l} \sum_{\mathbf{j}_{k-l} \in \mathbb{N}^{k-l}} d_{\mathbf{k}_l, \mathbf{j}_{k-l}} L_l^{0, \dots, 0}(\tilde{e}[k_1, \dots, k_l]) L_{k-l}^{1, \dots, 1}(\tilde{e}[j_1, \dots, j_{k-l}]),$$

where $d_{\mathbf{k}_l, \mathbf{j}_{k-l}}$ are some coefficients depending on F and where we use the notation

$$\tilde{e}[k_1, \dots, k_m] := \frac{1}{m!} \sum_{\pi \in \mathcal{S}_m} e_{k_{\pi(1)}} \otimes \cdots \otimes e_{k_{\pi(m)}}, \quad k_j \in \mathbb{N},$$

and \mathcal{S}_m stands for the set of all permutations of $\{1, \dots, m\}$.

Remark 3.3.6 We deduce from Proposition 3.3.5 that

$$\mathcal{C}_p(F) = \mathbb{E}[F] + \sum_{k=1}^p \sum_{l=0}^k \sum_{\mathbf{k}_l} \sum_{\mathbf{j}_{k-l}} d_{\mathbf{k}_l, \mathbf{j}_{k-l}} L_l^{0, \dots, 0}(\tilde{e}[k_1, \dots, k_l]) L_{k-l}^{1, \dots, 1}(\tilde{e}[j_1, \dots, j_{k-l}]). \quad (3.30)$$

The following Lemma, analogous to Lemma 3.2.3, gives some useful properties of the chaos decomposition.

Lemma 3.3.7

- Let F be a r.v. in $L^2(\mathcal{F}_T)$. $\forall p \geq 1$, we have $\mathbb{E}(|\mathcal{C}_p(F)|^2) \leq \mathbb{E}(|F|^2)$.
- Let H be in $H_T^2(\mathbb{R})$. We deduce from [L-7, Remark 2.8] that $\mathcal{C}_p\left(\int_0^T H_s ds\right) = \int_0^T \mathcal{C}_p(H_s) ds$.
- For all $F \in \mathbb{D}^{1,2}$, for all $i \in \{0, 1\}$ and for all $t \leq r$, $D_t^{(i)} \mathbb{E}_r[\mathcal{C}_p(F)] = \mathbb{E}_r[\mathcal{C}_{p-1}(D_t^{(i)} F)]$.

3.3.3 Chaos decomposition approximations

Following the same steps as in Section 3.2.2, we choose $N \in \mathbb{N}$ and let $\{\bar{t}_0, \bar{t}_1, \dots, \bar{t}_N\}$ be a regular grid of $[0, T]$, i.e. $\forall i \in \{0, \dots, N\}$, $\bar{t}_i = ih$ where $h = \frac{T}{N}$. From now on we will use a fixed orthonormal basis $\{e_i\}_{i \geq 1}$ of $L^2([0, T]; \mathbb{R})$: we set

$$e_i(t) := \frac{1}{\sqrt{h}} \mathbf{1}_{] \bar{t}_{i-1}, \bar{t}_i]}(t), \quad i \in \{0, \dots, N\} \tag{3.31}$$

and complete this sequence to a basis in $L^2([0, T]; \mathbb{R})$, for example, by using the Haar basis on each interval $] \bar{t}_{i-1}, \bar{t}_i]$. Let $\mathbf{n}^W = (n_1^W, \dots, n_N^W)$ be the vector of non-negative integers such that $|\mathbf{n}^W| = k$. Then (see [Pri09, Proposition 5.1.3])

$$L_k^{0, \dots, 0}(e_1^{\otimes n_1^W} \circ \dots \circ e_N^{\otimes n_N^W}) = \frac{\mathbf{n}^W!}{|\mathbf{n}^W|!} \prod_{i=1}^N K_{n_i^W} \left(\frac{\Delta W_i}{\sqrt{h}} \right), \tag{3.32}$$

where $\Delta W_i = W_{\bar{t}_i} - W_{\bar{t}_{i-1}}$ and \circ stands for the symmetric tensor product.

Definition 3.3.8 The Charlier polynomial of order $m \in \mathbb{N}$ and of parameter $t \geq 0$ is defined by

$$C_0(x, t) = 1, \quad C_1(x, t) = x - t, \quad x \in \mathbb{R}$$

and by the relation

$$C_{m+1}(x, t) = (x - m - t)C_m(x, t) - mtC_{m-1}(x, t).$$

The sequence $\left\{ \frac{1}{\sqrt{m!(\kappa t)^m}} C_m(\cdot, \kappa t) \right\}_{m \in \mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{N}, \nu_{\kappa t})$, where $\nu_{\kappa t}$ denotes the law of a Poisson random variable with parameter κt . Let $\mathbf{n}^P = (n_1^P, \dots, n_N^P)$ be the vector of non-negative integers such that $|\mathbf{n}^P| = k$. Using the grid (3.31), we have (see [Pri09, Proposition 6.2.9])

$$L_k^{1, \dots, 1}(e_1^{\otimes n_1^P} \circ \dots \circ e_N^{\otimes n_N^P}) = \frac{1}{|\mathbf{n}^P|! h^{\frac{|\mathbf{n}^P|}{2}}} \prod_{i=1}^N C_{n_i^P}(\Delta N_i, \kappa h), \tag{3.33}$$

where $\Delta N_i = N_{\bar{t}_i} - N_{\bar{t}_{i-1}}$.

Instead of summing over all $\mathbf{k}_l \in \mathbb{N}^l$ and $\mathbf{j}_{k-l} \in \mathbb{N}^{k-l}$ in (3.30), we only consider the first N functions $\{e_1, \dots, e_N\}$ of the basis $\{e_i\}_i$ defined in (3.31). This gives (together with the orthogonal projection onto the chaos up to order p) the following approximation of F

$$\begin{aligned} \mathcal{C}_p^N(F) &= \mathbb{E}[F] \\ &+ \sum_{k=1}^p \sum_{l=0}^k \sum_{\mathbf{k}_l \in \{1, \dots, N\}^l} \sum_{\mathbf{j}_{k-l} \in \{1, \dots, N\}^{k-l}} d_{\mathbf{k}_l, \mathbf{j}_{k-l}} L_l^{0, \dots, 0}(\tilde{e}[k_1, \dots, k_l]) L_{k-l}^{1, \dots, 1}(\tilde{e}[j_1, \dots, j_{k-l}]) \\ &:= \mathbb{E}[F] + \sum_{k=1}^p P_k^N(F). \end{aligned} \quad (3.34)$$

Let us now write $\mathcal{C}_p^N(F)$ ($p \leq N$) in terms of Hermite and Charlier polynomials. From [L-7, Lemma 2.7], we get

$$\mathcal{C}_p^N(F) = d_0 + \sum_{k=1}^p \sum_{|\mathbf{n}|=k} d_k^{\mathbf{n}} \prod_{i=1}^N K_{n_i^W}(G_i) C_{n_i^P}(Q_i, \kappa h), \quad (3.35)$$

where $\mathbf{n} = (\mathbf{n}^W, \mathbf{n}^P) = (n_1^W, \dots, n_N^W, n_1^P, \dots, n_N^P)$, $d_0 = \mathbb{E}(F)$ and

$$d_k^{\mathbf{n}} := \frac{\mathbf{n}^W!}{\mathbf{n}^P!(\kappa h)^{|\mathbf{n}^P|}} \mathbb{E} \left(F \prod_{i=1}^N K_{n_i^W}(G_i) C_{n_i^P}(Q_i, \kappa h) \right). \quad (3.36)$$

The following Proposition is similar to Proposition 3.2.7.

Proposition 3.3.9 *Let F be a real random variable in $L^2(\mathcal{F}_T)$ and let r be an integer in $\{1, \dots, N\}$. For all $\bar{t}_{r-1} < t \leq \bar{t}_r$, we have*

$$\begin{aligned} \mathbb{E}_t(\mathcal{C}_p^N(F)) &= d_0 + \\ &\sum_{k=1}^p \sum_{|\mathbf{n}(r)|=k} d_k^{\mathbf{n}} \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W}{2}} K_{n_r^B} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) C_{n_r^P}(N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) \\ &\quad \times \underbrace{\left(\prod_{i < r} K_{n_i^W}(G_i) C_{n_i^P}(Q_i, \kappa h) \right)}_{:= A_r}, \end{aligned}$$

$$\begin{aligned} D_t^{(0)} \mathbb{E}_t(\mathcal{C}_p^N(F)) &= h^{-1/2} \sum_{k=1}^p \sum_{\substack{|\mathbf{n}(r)|=k \\ n_r^W > 0}} d_k^{\mathbf{n}} \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W - 1}{2}} K_{n_r^W - 1} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) C_{n_r^P}(N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) A_r, \end{aligned}$$

$$\begin{aligned} D_t^{(1)} \mathbb{E}_t(\mathcal{C}_p^N(F)) &= \sum_{k=1}^p \sum_{\substack{|\mathbf{n}(r)|=k \\ n_r^P > 0}} d_k^{\mathbf{n}} \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W}{2}} K_{n_r^W} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) n_r^P C_{n_r^P - 1}(N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) A_r, \end{aligned}$$

where for $r \leq N$ $\mathbf{n}(r) = (\mathbf{n}^W(r), \mathbf{n}^P(r))$, and $\mathbf{n}^A(r)$ stands for (n_1^A, \dots, n_r^A) , where $A = W$ or P and $\mathbf{n}_r = (\mathbf{n}_r^W, \mathbf{n}_r^P)$.

3.3.4 Approximation steps of the BSDE

3.3.4.1 Picard's approximation

Picard's iterations: $(Y^0, Z^0, U^0) = (0, 0, 0)$ and for $q \in \mathbb{N}$,

$$Y_t^{q+1} = \zeta + \int_t^T f(s, Y_s^q, Z_s^q, U_s^q) ds - \int_t^T Z_s^{q+1} dW_s - \int_{]t, T]} U_s^{q+1} d\tilde{N}_s, \quad 0 \leq t \leq T.$$

It is well-known that the sequence (Y^q, Z^q, U^q) converges exponentially fast towards the solution (Y, Z, U) to BSDE (3.24). We write this Picard scheme in a forward way. Let F^q denote $F^q := \zeta + \int_0^T f(s, Y_s^q, Z_s^q, U_s^q) ds$. We define

$$Y_t^{q+1} = \mathbb{E} \left(F^q \mid \mathcal{F}_t \right) - \int_0^t f(s, Y_s^q, Z_s^q, U_s^q) ds, \quad (3.37)$$

$$Z_t^{q+1} = \mathbb{E} \left(D_t^{(0)} F^q \mid \mathcal{F}_{t-} \right), \quad U_t^{q+1} = \mathbb{E} \left(D_t^{(1)} F^q \mid \mathcal{F}_{t-} \right). \quad (3.38)$$

3.3.4.2 Chaos approximation

Let $(Y^{q,p}, Z^{q,p}, U^{q,p})$ denote the approximation of (Y^q, Z^q, U^q) built at step q using a chaos decomposition up to order p : $(Y^{0,p}, Z^{0,p}, U^{0,p}) = (0, 0, 0)$ and

$$Y_t^{q+1,p} = \mathbb{E} \left[\mathcal{C}_p(F^{q,p}) \mid \mathcal{F}_t \right] - \int_0^t f(s, Y_s^{q,p}, Z_s^{q,p}, U_s^{q,p}) ds, \quad (3.39)$$

$$Z_t^{q+1,p} = \mathbb{E} \left[D_t^{(0)} \mathcal{C}_p(F^{q,p}) \mid \mathcal{F}_{t-} \right], \quad U_t^{q+1,p} = \mathbb{E} \left[D_t^{(1)} \mathcal{C}_p(F^{q,p}) \mid \mathcal{F}_{t-} \right], \quad (3.40)$$

where $F^{q,p} = \zeta + \int_0^T f(s, Y_s^{q,p}, Z_s^{q,p}, U_s^{q,p}) ds$.

3.3.4.3 Truncation of the basis

The third type of approximation comes from the truncation of the orthonormal $L^2([0, T]; \mathbb{R})$ basis $\{e_i\}_{i \geq 1}$ defined in (3.31). Instead of considering the whole basis we only keep the first N functions $\{e_1, \dots, e_N\}$ to build the chaos decomposition projections \mathcal{C}_p^N . Proposition 3.3.9 gives us explicit formulas for $\mathbb{E}_t(\mathcal{C}_p^N(F))$, $D_t^{(0)} \mathbb{E}_t(\mathcal{C}_p^N(F))$ and $D_t^{(1)} \mathbb{E}_t(\mathcal{C}_p^N(F))$. From (3.39) and (3.40), we build $(Y^{q,p,N}, Z^{q,p,N}, U^{q,p,N})_q$ in the following way: $(Y^{0,p,N}, Z^{0,p,N}, U^{0,p,N}) = (0, 0, 0)$ and

$$Y_t^{q+1,p,N} = \mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N})) - \int_0^t f(s, Y_s^{q,p,N}, Z_s^{q,p,N}, U_s^{q,p,N}) ds, \quad (3.41)$$

$$Z_t^{q+1,p,N} = D_t^{(0)}(\mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N}))), \quad U_t^{q+1,p,N} = D_t^{(1)}(\mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N}))), \quad (3.42)$$

where $F^{q,p,N} := \zeta + \int_0^T f(s, Y_s^{q,p,N}, Z_s^{q,p,N}, U_s^{q,p,N}) ds$.

It is not necessary here to use predictable projections of $Z^{q+1,p,N}$ and $U^{q+1,p,N}$. In fact, $Z^{q+1,p,N}$ and $U^{q+1,p,N}$ are adapted and càdlàg, and from their explicit representation given above one concludes that the predictable projections are the left-continuous modifications: $\mathbb{E}_{t-} Z_t^{q+1,p,N} = Z_{t-}^{q+1,p,N}$ and $\mathbb{E}_{t-} U_t^{q+1,p,N} = U_{t-}^{q+1,p,N}$. Moreover, the integral in (3.41) does not change if one uses left-continuous modifications.

3.3.4.4 Monte Carlo approximation

Let F denote a r.v. of $L^2(\mathcal{F}_T)$. In practise, when we are not able to compute exactly d_0 and/or the coefficients d_k^n of the chaos decomposition (3.35)-(3.36) of F , we use Monte-Carlo simulations to approximate them. Let $\{F^m\}_{1 \leq m \leq M}$ be a M i.i.d. sample of F and $\{(G_1^m, Q_1^m), \dots, (G_N^m, Q_N^m)\}_{1 \leq m \leq M}$ be a M i.i.d. sample of $\{(G_1, Q_1), \dots, (G_N, Q_N)\}$.

We approximate the expectations of (3.36) by empirical means

$$\widehat{d}_0 := \frac{1}{M} \sum_{m=1}^M F^m, \quad \widehat{d}_k^n := \frac{\mathbf{n}^W!}{\mathbf{n}^P!(\kappa h)^{|\mathbf{n}^P|} M} \sum_{m=1}^M \left(F^m \prod_{i=1}^N K_{n_i^W} (G_i^m) C_{n_i^P} (Q_i^m, \kappa h) \right). \quad (3.43)$$

In the following, we denote

$$\mathcal{C}_p^{N,M}(F) = \widehat{d}_0 + \sum_{k=1}^p \sum_{|\mathbf{n}|=k} \widehat{d}_k^n \prod_{1 \leq i \leq N} K_{n_i^W} (G_i) C_{n_i^P} (Q_i, \kappa h). \quad (3.44)$$

$\mathbb{E}_t(\mathcal{C}_p^{N,M}(F))$ and $D_t(\mathbb{E}_t(\mathcal{C}_p^{N,M}(F)))$ denote the conditional expectations obtained in Proposition 3.3.9 when $(d_0, \{d_k^n\}_{1 \leq k \leq p, |\mathbf{n}|=k})$ are replaced by $(\widehat{d}_0, \{\widehat{d}_k^n\}_{1 \leq k \leq p, |\mathbf{n}|=k})$:

$$\begin{aligned} \mathbb{E}_t(\mathcal{C}_p^{N,M}(F)) &= \widehat{d}_0 + \\ &\sum_{k=1}^p \sum_{|\mathbf{n}(r)|=k} \widehat{d}_k^n \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W}{2}} K_{n_r^W} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) C_{n_r^P} (N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) \\ &\quad \times \underbrace{\left(\prod_{i < r} K_{n_i^W} (G_i) C_{n_i^P} (Q_i, \kappa h) \right)}_{:= A_r}, \end{aligned}$$

$$\begin{aligned} D_t^{(0)} \mathbb{E}_t(\mathcal{C}_p^{N,M}(F)) &= \\ &= h^{-1/2} \sum_{k=1}^p \sum_{\substack{|\mathbf{n}(r)|=k \\ n_r^W > 0}} \widehat{d}_k^n \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W - 1}{2}} K_{n_r^W - 1} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) C_{n_r^P} (N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) A_r, \end{aligned}$$

$$\begin{aligned} D_t^{(1)} \mathbb{E}_t(\mathcal{C}_p^{N,M}(F)) &= \\ &= \sum_{k=1}^p \sum_{\substack{|\mathbf{n}(r)|=k \\ n_r^P > 0}} \widehat{d}_k^n \left(\frac{t - \bar{t}_{r-1}}{h} \right)^{\frac{n_r^W}{2}} K_{n_r^W} \left(\frac{W_t - W_{\bar{t}_{r-1}}}{\sqrt{t - \bar{t}_{r-1}}} \right) n_r^P C_{n_r^P - 1} (N_t - N_{\bar{t}_{r-1}}, \kappa(t - \bar{t}_{r-1})) A_r. \end{aligned}$$

We introduce the processes $(Y^{q+1,p,N,M}, Z^{q+1,p,N,M}, U^{q+1,p,N,M})$, useful in the following. It corresponds to the approximation of $(Y^{q+1,p,N}, Z^{q+1,p,N}, U^{q+1,p,N})$ when we use $\mathcal{C}_p^{N,M}$ instead of \mathcal{C}_p^N , i.e. when we use a Monte Carlo procedure to compute the coefficients d_k^n .

$$Y_t^{q+1,p,N,M} = \mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M})) - \int_0^t f(\theta_s^{q,p,N,M}) ds, \quad (3.45)$$

$$Z_t^{q+1,p,N,M} = D_t^{(0)}(\mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M}))), \quad U_t^{q+1,p,N,M} = D_t^{(1)}(\mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M}))), \quad (3.46)$$

where $F^{q,p,N,M} := \zeta + \int_0^T f(\theta_s^{q,p,N,M}) ds$ and $\theta_s^{q,p,N,M} = (s, Y_s^{q,p,N,M}, Z_s^{q,p,N,M}, U_s^{q,p,N,M})$.

3.3.5 Convergence results

We aim at bounding the error between (Y, Z) — the solution of (3.24) — and $(Y^{q,p,N,M}, Z^{q,p,N,M})$ defined by (3.45)-(3.46). Before stating the main result of the paper, we introduce some hypotheses and some intermediate results on $\mathcal{C}_p, \mathcal{C}_p^N$ and $\mathcal{C}_p^{N,M}$.

Hypothesis 3.4 Let $m \in \mathbb{N}^*$. We say that F satisfies Hypothesis \mathcal{H}_m if F satisfies the two following hypotheses

- $\mathcal{H}_m^1 : \forall j \in \mathbb{N}^* F \in \mathcal{D}^{m,j}$, i.e. $\|F\|_{m,j}^j < \infty$.
- $\mathcal{H}_m^2 : \forall j \in \mathbb{N}^* \forall l_0, l_1 \in \mathbb{N}$ such that $l = l_0 + l_1 + 1 \leq m$ there exist two positive constants β_F and $k_l^F(j)$ such that for all multi-indices $\alpha = (\alpha_1, \dots, \alpha_{l_0}) \in \{0, 1\}^{l_0}$, $\gamma = (\gamma_1, \dots, \gamma_{l_1+1}) \in \{0, 1\}^{l_1+1}$ and for a.e. $t_i \in [0, T], s_i \in [0, T]$ it holds

$$\text{ess sup}_{t_1, \dots, t_{l_0}} \text{ess sup}_{s_{i+1}, \dots, s_{i+l_1}} \mathbb{E} |D_{t_1, \dots, t_{l_0}}^\alpha (D_{t_i, s_{i+1}, \dots, s_{i+l_1}}^\gamma F - D_{s_i, \dots, s_{i+l_1}}^\gamma F)|^j \leq k_l^F(j) |t_i - s_i|^{j\beta_F}.$$

In the following, we denote $K_m^F(j) = \max_{l \leq m} k_l^F(j)$.

Remark 3.3.10 If F satisfies \mathcal{H}_m , for all $l \leq m$ and for all multi-indices $\alpha = (\alpha_1, \dots, \alpha_l) \in \{0, 1\}^l$ we have for a.e. $(t_1, \dots, t_l) \in [0, T]^l$ and $(s_1, \dots, s_l) \in [0, T]^l$ that

$$|\mathbb{E}(D_{t_1, \dots, t_l}^\alpha F) - \mathbb{E}(D_{s_1, \dots, s_l}^\alpha F)| \leq K_m^F(1) (|t_1 - s_1|^{\beta_F} + \dots + |t_l - s_l|^{\beta_F}). \quad (3.47)$$

Lemma 3.3.11 Assume F satisfies (3.47) with $m = p$. Then

$$\mathbb{E} |(\mathcal{C}_p^N - \mathcal{C}_p)(F)|^2 \leq \overline{K_p^F} \left(\frac{T}{N} \right)^{2\beta_F} T(1+T)e^T.$$

where $\overline{K_p^F} := \sum_{j=1}^p (K_j^F)^2$ (with $K_j^F := K_j^F(1)$ from (3.47)).

Hypothesis 3.5 Let $(p, N) \in \mathbb{N}^2$. We say that a r.v. F satisfies $\mathcal{H}_{p,N}^3$ if

$$V_{p,N}(F) := \mathbb{V}(F) + \sum_{k=1}^p \sum_{|\mathbf{n}|=k} \frac{(\mathbf{n}^W)!}{(\mathbf{n}^P)!(\kappa h)^{|\mathbf{n}^P|}} \mathbb{V} \left(F \prod_{i=1}^N K_{n_i^W}(G_i) C_{n_i^P}(Q_i, \kappa h) \right) < \infty,$$

where $\mathbb{V}(\zeta)$ denotes the variance of a r.v. ζ .

Lemma 3.3.12 Let F be a r.v. satisfying Hypothesis $\mathcal{H}_{p,N}^3$. We have

$$\mathbb{E} |(\mathcal{C}_p^N - \mathcal{C}_p^{N,M})(F)|^2 = \frac{1}{M} V_{p,N}(F).$$

Moreover, we have $\mathbb{E} |(\mathcal{C}_p^{N,M})(F)|^2 \leq \mathbb{E} |F|^2 + \frac{1}{M} V_{p,N}(F)$.

Theorem 3.3.13 Let m be an integer s.t. $1 \leq m \leq p+1$. Assume that ζ satisfies \mathcal{H}_{p+q+1} and $\mathcal{H}_{p,N}^3$ and $f \in C_b^{0,p+q+1,p+q+1,p+q+1}$. We have

$$\begin{aligned} & \| (Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M}, U - U^{q,p,N,M}) \|_{L_2}^2 \\ & \leq \frac{A_0}{2^q} + \frac{A_1(q, m)}{(p+2-m) \cdots (p+1)} + A_2(q, p) \left(\frac{T}{N} \right)^{2\beta_\zeta \wedge 1} + \frac{A_3(q, p, N)}{M}, \end{aligned}$$

where A_0, A_1, A_2 and A_3 are given in the following proof. If $f \in C_b^{0,\infty,\infty,\infty}$ and ζ satisfies \mathcal{H}_m for all $m \in \mathbb{N}^*$ and $\mathcal{H}_{p,N}^3$ for all $(p, N) \in \mathbb{N}^2$, we get

$$\lim_{q \rightarrow \infty} \lim_{p \rightarrow \infty} \lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \|(Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M}, U - U^{q,p,N,M})\|_{L_2}^2 = 0.$$

Proof. We split the error into 4 terms :

1. Picard iterations : $\mathcal{E}^q = \|(Y - Y^q, Z - Z^q, U - U^q)\|_{L_2}^2$, where (Y^q, Z^q, U^q) is defined by (3.37)-(3.38),
2. the truncation of the chaos decomposition : $\mathcal{E}^{q,p} = \|(Y^q - Y^{q,p}, Z^q - Z^{q,p}, U^q - U^{q,p})\|_{L_2}^2$, where $(Y^{q,p}, Z^{q,p}, U^{q,p})$ is defined by (3.39)-(3.40),
3. the truncation of the $L^2([0, T]; \mathbb{R})$ basis : $\mathcal{E}^{q,p,N} = \|(Y^{q,p} - Y^{q,p,N}, Z^{q,p} - Z^{q,p,N}, U^{q,p} - U^{q,p,N})\|_{L_2}^2$, where $(Y^{q,p,N}, Z^{q,p,N}, U^{q,p,N})$ is defined by (3.41)-(3.42),
4. the Monte-Carlo approximation to compute the expectations : $\mathcal{E}^{q,p,N,M} = \|(Y^{q,p,N} - Y^{q,p,N,M}, Z^{q,p,N} - Z^{q,p,N,M}, U^{q,p,N} - U^{q,p,N,M})\|_{L_2}^2$, where $(Y^{q,p,N,M}, Z^{q,p,N,M}, U^{q,p,N,M})$ is defined by (3.45)-(3.46).

We have

$$\|(Y - Y^{q,p,N,M}, Z - Z^{q,p,N,M}, U - U^{q,p,N,M})\|_{L_2}^2 \leq 4(\mathcal{E}^q + \mathcal{E}^{q,p} + \mathcal{E}^{q,p,N} + \mathcal{E}^{q,p,N,M}).$$

From [TL94, Lemma 2.4], we know that under Hypothesis 3.3, the sequence $(Y^q, Z^q, U^q)_q$ defined by (3.37)-(3.38) converges to $(Y, Z, U) d\mathbb{P} \times dt$ a.e. and in $S_T^2(\mathbb{R}) \times H_T^2(\mathbb{R}) \times H_T^2(\mathbb{R})$. Moreover, we have

$$\mathcal{E}^q := \|(Y - Y^q, Z - Z^q, U - U^q)\|_{L_2}^2 \leq \frac{A_0}{2^q}, \quad (3.48)$$

where the constant A_0 depends on $T, \|\xi\|_{L^2}^2$ and on $\|f(\cdot, 0, 0, 0)\|_{L^2([0, T]; \mathbb{R})}^2$. As in the proof of Theorem 3.2.15, the same procedure applies in the proof of [L-7, Propositions 4.9, 4.13 and 4.15] to bound $\mathcal{E}^{q,p}, \mathcal{E}^{q,p,N}$ and $\mathcal{E}^{q,p,N,M}$. Thanks to Picard iterations $\mathcal{E}^{q+1, \cdot}$ is bounded by $C_0, \mathcal{E}^{q, \cdot} + C_{1, \cdot}$, where $C_{1,p}$ is related to Lemma 3.3.4, $C_{1,p,N}$ is related to Lemma 3.3.11 and $C_{1,p,N,M}$ is related to Lemma 3.3.12. Since $\mathcal{E}^{0, \cdot} = 0$ we get the following upper bounds. From [L-7, Proposition 4.9] we get that $\mathcal{E}^{q,p} \leq \frac{A_1(q,m)}{(p+2-m) \cdots (p+1)}$ where $A_1(q, m) := \frac{(C_1 T(T+1)L_f^2)^{q-1}}{C_1 T(T+1)L_f^2 - 1} \max_{1 \leq l \leq q} K_1(l, m)$ and $K_1(l, m)$ depends on $T, m, \|\xi\|_{m+q, 2}^{\frac{(m+q-1)!}{(m-1)!}}$ and on $(\|\partial_{sp}^k f\|_{\infty})_{1 \leq k \leq m+q}$. [L-7, Proposition 4.13] gives that $\mathcal{E}^{q,p,N} \leq A_2(q, p) \left(\frac{T}{N}\right)^{1 \wedge 2\beta\xi}$, where $A_2(q, p) := K_2(q, p)T(T+1)e^{T \frac{(C_2 T(T+1)L_f^2)^{q-1}}{C_2 T(T+1)L_f^2 - 1}}$ and $K_2(p)$ depends on $\overline{K_p^\xi}, T, \|\xi\|_{p,1}$ and on $(\|\partial_{sp}^k f\|_{\infty})_{1 \leq k \leq p}$. [L-7, Proposition 4.15] states that $\mathcal{E}^{q,p,N,M} \leq \frac{A_3(q,p,N)}{M}$, where $A_3(q, p, N) := K_3(p, N) \frac{(C_3 T(T+1)L_f^2)^{q-1}}{C_3 T(T+1)L_f^2 - 1}$ and $K_3(p, N) := C \left(V_{p,N}(\xi) + T^2 \|f\|_{\infty}^2 \sum_{k=0}^p \binom{2N}{k} \right)$ for some $C > 0$. \blacksquare

3.3.6 Numerical examples

The following example is borrowed from [LMT13]. We consider a Poisson process N with $\kappa = 1$ and the following BSDE

$$\begin{aligned} dY_t &= -cU_t dt + Z_t dW_t + U_t (dN_t - dt), \\ \zeta &= N_T. \end{aligned}$$

The explicit solution is given by

$$(Y_t, Z_t, U_t) = (N_t + (1 + c)(T - t), 0, 1).$$

Figure 3.3.1 represents the evolution of $(Y_0^{q,p,N,M}, Z_0^{q,p,N,M}, U_0^{q,p,N,M})$ with respect to M when $q = 5, p = 2$ and $N = 20$.

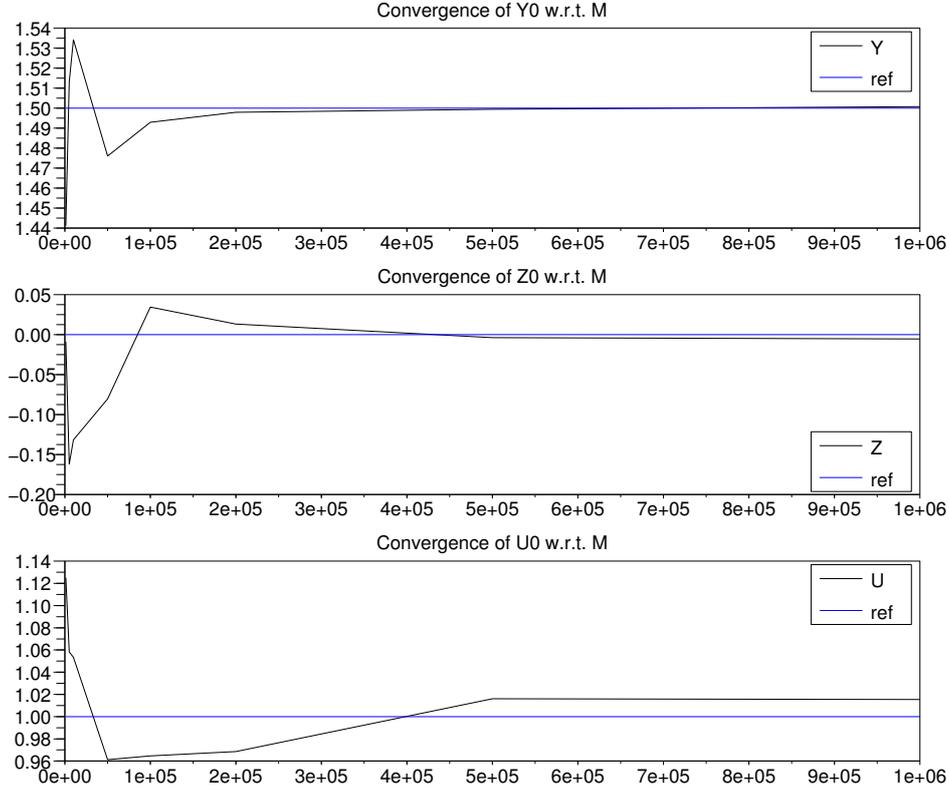


Figure 3.3.1: Evolution of $(Y_0^{q,p,N,M}, Z_0^{q,p,N,M}, U_0^{q,p,N,M})$ with respect to M when $p = 2, N = 20, q = 5, c = 0.5, T = 1$

Table 3.3 gives the computational time needed by the algorithm with this choice for q, p, N and for different values of M . We notice from Figure 3.3.1 that the value of $(Y_0^{q,p,N,M}, Z_0^{q,p,N,M}, U_0^{q,p,N,M})$ is close to the true solution from $M = 2 \times 10^5$. When $M = 2 \times 10^5$, the CPU time is about 1 minute, which is quite small.

M	10^3	5×10^3	10^4	5×10^4	10^5	2×10^5	5×10^5	10^6
CPU time (in s)	0.253	1.277	2.567	13.24	26.81	56.91	142.75	283.65

Table 3.3: CPU time w.r.t. M when $p = 2, N = 20, q = 5, c = 0.5, T = 1$

3.4 McKean-Vlasov Case

The aim of this section is to extend the results of [L-10] to the case of McKean-Vlasov BSDEs, i.e. to provide an algorithm based on Wiener chaos expansion to solve BSDEs of the following type

$$Y_t = \zeta + \int_t^T f(s, Y_s, Z_s, [Y_s], [Z_s]) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \quad (3.49)$$

where $[\theta]$ is the notation for the law of a random variable θ and f is a map from $[0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}^d)$ into \mathbb{R} . It is based on [L-25] and was written with Céline Acary-Robert, Philippe Briand and Abir Ghannoum.

Such type of BSDEs have been introduced in [BDLP09] and [BLP09] in a more particular framework: in [BDLP09], the authors study the mean field problem in a Markovian setting and prove the existence and the uniqueness of the solution when the terminal condition is of type $\zeta = \mathbb{E}[g(x, X_T)]|_{x=X_T}$ where X is a driving adapted stochastic process, and the generator is defined by $\mathbb{E}[f(s, \lambda, \Lambda_s)]|_{\lambda=\Lambda_s}$ where $\Lambda_s = (X_s, Y_s, Z_s)$. In [BLP09], the authors extend the result of existence and uniqueness to a more general framework and link the mean-field BSDE to non local partial differential equation.

The first study of numerical methods for McKean-Vlasov BSDEs goes back to a few years (see [Ala15], [CdRGT15], [CCD19]). Usually, forward McKean-Vlasov SDEs are solved by using particle algorithms (see [AKH02], [TV03], [Bos05]) in which the McKean term is approximated by the empirical measure of a large number of interacting particles with independent noise. Adapting such algorithms to the backward problem is not obvious as the high dimension of the involved Brownian motion (given by the number of particles) induces, a priori, a high dimension backward problem with bad consequences for the numerical implementation. The above mentioned papers on numerical methods for McKean-Vlasov BSDEs do not use particle systems. In [CdRGT15], the authors present a method based on cubature for decoupled McKean-Vlasov forward backward SDE. In [CCD19], the authors consider the case of strongly coupled forward-backward SDE of McKean-Vlasov type. They propose a scheme whose principle is to implement recursively Picard iterations on small time intervals, since Picard Theorem only applies in small time for fully coupled problems.

In this section we propose a method based on Wiener chaos expansion and particle system approximation which is neither more complex nor more costly than solving a standard BSDE of type (3.1) when $f(y, z, \mu, \nu)$ is a separated variable function in (y, z) and (μ, ν) . We recall that the method based on Wiener chaos expansion to solve standard BSDEs consists in writing the Picard scheme of (3.1) in a forward way

$$\begin{aligned} Y_t^{q+1} &= \mathbb{E} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds \mid \mathcal{F}_t \right) - \int_0^t f(s, Y_s^q, Z_s^q) ds, \\ Z_t^{q+1} &= D_t Y_t^{q+1} = D_t \mathbb{E} \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q) ds \mid \mathcal{F}_t \right), \end{aligned}$$

and to use Wiener chaos expansion to easily compute conditional expectations and their Malliavin derivatives. More precisely, all r.v. F in L^2 can be written

$$F = \mathbb{E}(F) + \sum_{k \geq 1} \sum_{|n|=k} d_k^n \prod_{i \geq 1} K_{n_i} \left(\int_0^T g_i(s) dW_s \right),$$

where K_l denotes the Hermite polynomial of degree l , $(g_i)_{i \geq 1}$ is an orthonormal basis of $L^2([0, T]; \mathbb{R})$ and, if $n = (n_i)_{i \geq 1}$ is a sequence of integers, $|n| = \sum_{i \geq 1} n_i$. $(d_k^n)_{k \geq 1, |n|=k}$ is the sequence of coefficients ensuing from the decomposition of F . The numerical method consists in working with a finite number of chaos, a finite number of functions $(g_i)_i$ and in using Monte-Carlo approximation

to compute the coefficients $(d_k^n)_{k,n}$. In case of McKean-Vlasov BSDE, the generator depends on the laws of the processes. The idea is to use M particles which will serve both to approximate the law of (Y, Z) and to compute the coefficients $(d_k^n)_{k,n}$ by Monte Carlo. By doing this, we manage to get a computational cost which is of the same order as the one obtained in case of standard BSDEs for many cases of the driver f . However, this pooling of particles costs the independence in the Monte Carlo approximation, making the proof of the convergence more difficult and leading to a slower speed of convergence in M .

The rest of the Section is organized as follows: Section 3.4.1 presents some preliminaries, Section 3.4.2 presents the approximation steps of the McKean-Vlasov BSDE, Section 3.4.3 states the convergence results and finally Section 3.4.4 presents some numerical examples.

3.4.1 Preliminaries

Hypothesis 3.6 We assume:

- the generator $f : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is Lipschitz continuous: there exists a constant L_f such that for all $t \in \mathbb{R}^+$, $y_1, y_2 \in \mathbb{R}$, $z_1, z_2 \in \mathbb{R}^d$, $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R})$ and $\nu_1, \nu_2 \in \mathcal{P}_2(\mathbb{R}^d)$

$$|f(t, y_1, z_1, \mu_1, \nu_1) - f(t, y_2, z_2, \mu_2, \nu_2)| \leq L_f \left(|y_1 - y_2| + |z_1 - z_2| + \mathcal{W}_2(\mu_1, \mu_2) + \mathcal{W}_2(\nu_1, \nu_2) \right).$$

- $\mathbb{E}(|\zeta|^2 + \int_0^T |f(s, 0, 0, [\delta_0], [\delta_0])|^2 ds) < \infty$.

Theorem 3.4.1 *Given standard parameters (f, ζ) , there exists a unique pair $(Y, Z) \in \mathbb{S}_T^2(\mathbb{R}) \times \mathbb{H}_T^2(\mathbb{R}^d)$ which solves (3.49).*

A proof of this Theorem can be found in [L-25, Theorem 9].

3.4.2 Approximation steps of the BSDE

The algorithm is based on five types of approximations: Picard iterations, a Wiener chaos expansion up to a finite order, the truncation of an $L^2([0, T]; \mathbb{R})$ basis in order to apply formulas of Proposition 3.2.7, a Monte Carlo method to approximate the coefficients d_0 and d_k^n defined in (3.11) and the particle system.

3.4.2.1 Picard approximations

The first step consists in approximating (Y, Z) —the solution to (3.49)—by Picard's sequence $(Y^q, Z^q)_q$, built as follows: $(Y^0 = 0, Z^0 = 0)$ and for all $q \geq 1$

$$Y_t^{q+1} = \zeta + \int_t^T f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds - \int_t^T Z_s^{q+1} \cdot dW_s, \quad 0 \leq t \leq T. \quad (3.50)$$

From (3.50), under the assumptions that $\zeta \in \mathbb{D}^{1,2}$ and $f \in C_b^{0,1,1,0,0}$, we express (Y^{q+1}, Z^{q+1}) as a function of the processes (Y^q, Z^q) ,

$$Y_t^{q+1} = \mathbb{E}_t \left(\zeta + \int_t^T f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds \right), \quad Z_t^{q+1} = D_t Y_t^{q+1}, \quad (3.51)$$

which can also be written

$$\begin{aligned}
Y_t^{q+1} &= \mathbb{E}_t \left(\zeta + \int_0^T f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds \right) - \int_0^t f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds, \\
Z_t^{q+1} &= D_t Y_t^{q+1}.
\end{aligned} \tag{3.52}$$

As recalled in the Introduction, the computation of the conditional expectation is the cornerstone in the numerical resolution of BSDEs. Chaos decomposition formulas enable us to circumvent this problem.

3.4.2.2 Chaos approximation

Computing the chaos decomposition of the r.v. $F = \zeta + \int_t^T f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds$ (appearing in (3.51)) in order to compute Y_t^{q+1} is not judicious. F depends on t , and then the computation of Y^{q+1} on the grid $\mathcal{T} = \{\tilde{t}_i = i\frac{T}{N}, i = 0, \dots, N\}$ would require $N + 1$ calls to the chaos decomposition function. To build an efficient algorithm, we need to call the chaos decomposition function as infrequently as possible, since each call is computationally demanding and brings an approximation error due to the truncation, the Monte Carlo approximation and to the particle approximation (see next sections). Then we look for a r.v. F^q independent of t such that Y_t^{q+1} and Z_t^{q+1} can be expressed as functions of $\mathbb{E}_t(F^q)$, $D_t \mathbb{E}_t(F^q)$ and of Y_q and Z_q . Equation (3.52) gives a more tractable expression of Y^{q+1} . Let F^q be defined by $F^q := \zeta + \int_0^T f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds$. Then

$$Y_t^{q+1} = \mathbb{E}_t(F^q) - \int_0^t f(s, Y_s^q, Z_s^q, [Y_s^q], [Z_s^q]) ds, \quad Z_t^{q+1} = D_t \mathbb{E}_t(F^q). \tag{3.53}$$

The second type of approximation consists in computing the chaos decomposition of F^q up to order p . Since F^q does not depend on t , the chaos decomposition function \mathcal{C}_p (see (3.7)) is called only once per Picard iteration.

Let $(Y^{q,p}, Z^{q,p})$ denote the approximation of (Y^q, Z^q) built at step q using a chaos decomposition with order p : $(Y^{0,p}, Z^{0,p}) = (0, 0)$ and

$$\begin{aligned}
Y_t^{q+1,p} &= \mathbb{E}_t(\mathcal{C}_p(F^{q,p})) - \int_0^t f(s, Y_s^{q,p}, Z_s^{q,p}, [Y_s^{q,p}], [Z_s^{q,p}]) ds, \\
Z_t^{q+1,p} &= D_t \mathbb{E}_t(\mathcal{C}_p(F^{q,p})),
\end{aligned} \tag{3.54}$$

where $F^{q,p} = \zeta + \int_0^T f(s, Y_s^{q,p}, Z_s^{q,p}, [Y_s^{q,p}], [Z_s^{q,p}]) ds$. In the sequel, we also use the following equality:

$$Z_t^{q+1,p} = \mathbb{E}_t(D_t \mathcal{C}_p(F^{q,p})). \tag{3.55}$$

3.4.2.3 Truncation of the basis

The third type of approximation comes from the truncation of the orthonormal $L^2([0, T]; \mathbb{R})$ basis used in the definition of \mathcal{C}_p (3.7). Instead of considering a basis of $L^2([0, T]; \mathbb{R})$, we only keep the first N functions (g_1, \dots, g_N) defined by (3.8) to build the chaos decomposition function \mathcal{C}_p^N (3.9). Proposition 3.2.7 gives us explicit formulas for $\mathbb{E}_t(\mathcal{C}_p^N F)$ and $D_t \mathbb{E}_t(\mathcal{C}_p^N F)$. From (3.54), we build $(Y^{q,p,N}, Z^{q,p,N})_q$ in the following way: $(Y^{0,p,N}, Z^{0,p,N}) = (0, 0)$ and

$$\begin{aligned}
Y_t^{q+1,p,N} &= \mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N})) - \int_0^t f(s, Y_s^{q,p,N}, Z_s^{q,p,N}, [Y_s^{q,p,N}], [Z_s^{q,p,N}]) ds, \\
Z_t^{q+1,p,N} &= D_t \mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N})),
\end{aligned} \tag{3.56}$$

where $F^{q,p,N} = \zeta + \int_0^T f(s, Y_s^{q,p,N}, Z_s^{q,p,N}, [Y_s^{q,p,N}], [Z_s^{q,p,N}]) ds$.

Equation (3.56) is tractable as soon as we know closed formulas for the coefficients d_k^n of the chaos decomposition of $\mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N}))$ and $D_t \mathbb{E}_t(\mathcal{C}_p^N(F^{q,p,N}))$ (see Proposition 3.2.7). When it is not the case, we need to use a Monte Carlo method to approximate these coefficients. The next section is devoted to this method.

3.4.2.4 Monte Carlo approximation

Let F denote a r.v. of $L^2(\mathcal{F}_T)$. Practically, when we are not able to compute exactly $d_0 = \mathbb{E}(F)$ and/or the coefficients $d_n^k = n! \mathbb{E}(F K_{n_1}(G_1) \cdots K_{n_N}(G_N))$ of the chaos decomposition (3.10)-(3.11) of F , we use Monte Carlo simulations to approximate them. Let $(F^m)_{1 \leq m \leq M}$ be a M i.i.d. sample of F and $(G_1^m, \dots, G_N^m)_{1 \leq m \leq M}$ be a M i.i.d. sample of (G_1, \dots, G_N) . We propose a method which consists in approximating the expectations $\mathbf{d} := (d_0, (d_k^n)_{1 \leq k \leq p, |n|=k})$ by empirical means $\widehat{\mathbf{d}}_M := (\hat{d}_0, (\hat{d}_k^n)_{1 \leq k \leq p, |n|=k})$ where

$$\hat{d}_0 := \frac{1}{M} \sum_{m=1}^M F^m, \quad \hat{d}_k^n := \frac{n!}{M} \sum_{m=1}^M F^m K_{n_1}(G_1^m) \cdots K_{n_N}(G_N^m). \tag{3.57}$$

Definition 3.4.2 Let F be a r.v. of $L^2(\mathcal{F}_T)$ and (F^1, \dots, F^M) be M identically distributed r.v. with the law of F . We denote $\mathcal{C}_p^{N,M}((F^m)_{1 \leq m \leq M})$ the following approximation of $\mathcal{C}_p^N(F)$

$$\mathcal{C}_p^{N,M}((F^m)_{1 \leq m \leq M}) = \hat{d}_0 + \sum_{k=1}^p \sum_{|n|=k} \hat{d}_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i). \tag{3.58}$$

where $(\hat{d}_0, (\hat{d}_k^n)_{1 \leq k \leq p, |n|=k})$ are defined by (3.57). We also introduce the filtration $(\mathcal{G}_t)_{0 \leq t \leq T}$ defined by $\mathcal{G}_t = \sigma((B_s)_{s \leq t}, (G_1^m, \dots, G_N^m)_{1 \leq m \leq M})$.

Remark 3.4.3 When (F^1, \dots, F^M) is an independent and identically distributed sample of F , we adopt the short notation $\mathcal{C}_p^{N,M}(F)$ to refer to $\mathcal{C}_p^{N,M}((F^m)_{1 \leq m \leq M})$. The notation $\mathcal{C}_p^{N,M}((F^m)_{1 \leq m \leq M})$ is only used when the r.v. F^1, \dots, F^M are not independent. This is the case in the next paragraph, when we introduce the particle system to approximate the law of (Y, Z) (see (3.64)).

Before introducing the processes $(Y^{q+1,p,N,M}, Z^{q+1,p,N,M})$, we define $\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m))$ the conditional expectation of $\mathcal{C}_p^{N,M}((F^m)_m)$ w.r.t. \mathcal{G}_t which corresponds to the conditional expectation obtained in Proposition 3.2.7 when $(d_0, d_k^n)_{1 \leq k \leq p, |n|=k}$ are replaced by $(\hat{d}_0, \hat{d}_k^n)_{1 \leq k \leq p, |n|=k}$, i.e.

$$\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m)) := \hat{d}_0 + \sum_{k=1}^p \sum_{|n(r)|=k} \hat{d}_k^n \prod_{i < r} K_{n_i}(G_i) \times \left(\frac{t - \tilde{t}_{r-1}}{h} \right)^{(n_r)/2} K_{n_r} \left(\frac{W_t - W_{\tilde{t}_{r-1}}}{\sqrt{t - \tilde{t}_{r-1}}} \right). \tag{3.59}$$

We also introduce

$$D_t \mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m)) := h^{-1/2} \sum_{k=1}^p \sum_{\substack{|n(r)|=k \\ n(r)>0}} \hat{d}_k^n \prod_{i<r} K_{n_i}(G_i) \times \left(\frac{t - \tilde{t}_{r-1}}{h} \right)^{(n_r-1)/2} K_{n_r-1} \left(\frac{W_t - W_{\tilde{t}_{r-1}}}{\sqrt{t - \tilde{t}_{r-1}}} \right). \quad (3.60)$$

which is linked to $\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m))$ by the following relation

$$\forall t \leq T, \hat{d}_0 + \int_0^t D_s \mathbb{E}_s(\mathcal{C}_p^{N,M}((F^m)_m)) dW_s = \mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m)). \quad (3.61)$$

We keep the notation ' D_t ' since the above equality is nothing else than the Clark-Ocone formula when considering W as a \mathcal{G} -Brownian motion.

Remark 3.4.4 When M samples of $(\mathcal{C}_p^{N,M}((F^m)_m))_m$ are needed, we can use the same samples as the ones used to compute \hat{d}_0 and \hat{d}_k^n : $(\mathcal{C}_p^{N,M}((F^m)_m))_m = \hat{d}_0 + \sum_{k=1}^p \sum_{|n|=k} \hat{d}_k^n \prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$, or use new ones. In the first case, we only require M samples (G_1, \dots, G_N) . We built F^1, \dots, F^M from these samples. The coefficients \hat{d}_k^n and \hat{d}_0 are not independent of $\prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$. In this case the notation $\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m))$ introduced above is not equal to $\mathbb{E}(\mathcal{C}_p^{N,M}((F^m)_m) | \mathcal{G}_t)$. In the second case, the coefficients \hat{d}_k^n and \hat{d}_0 are independent of $\prod_{1 \leq i \leq N} K_{n_i}(G_i^m)$ and we have $\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^m)_m)) = \mathbb{E}(\mathcal{C}_p^{N,M}((F^m)_m) | \mathcal{G}_t)$. This second approach requires $2M$ samples of F and (G_1, \dots, G_N) and its variance increases with N . Practically, we use the first technique.

Let us now introduce the couple of processes $(Y^{q+1,p,N,M}, Z^{q+1,p,N,M})$, which corresponds to the approximation of $(Y^{q+1,p,N}, Z^{q+1,p,N})$ when we use $\mathcal{C}_p^{N,M}$ instead of \mathcal{C}_p^N , that is, when we use a Monte Carlo procedure to compute the coefficients d_k^n .

$$\begin{aligned} Y_t^{q+1,p,N,M} &:= \mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M})) - \int_0^t f(\theta_s^{q,p,N,M}) ds, \\ Z_t^{q+1,p,N,M} &:= D_t \mathbb{E}_t(\mathcal{C}_p^{N,M}(F^{q,p,N,M})), \end{aligned} \quad (3.62)$$

where $F^{q,p,N,M} := \zeta + \int_0^T f(\theta_s^{q,p,N,M}) ds$ and $\theta_s^{q,p,N,M} := (s, Y_s^{q,p,N,M}, Z_s^{q,p,N,M}, [Y_s^{q,p,N,M}], [Z_s^{q,p,N,M}])$.

3.4.2.5 Particle system

In this section, we introduce an interacting particle system associated to (3.62) to approximate the law of $Y_s^{q,p,N,M}$ and $Z_s^{q,p,N,M}$. Indeed we replace one stochastic differential equation with unknown processes $Y_s^{q,p,N,M}$ and $Z_s^{q,p,N,M}$ with a system of M ordinary stochastic differential equations whose solution consists in a system of particles $(Y_s^{q,p,N,M,m}, Z_s^{q,p,N,M,m})_{1 \leq m \leq M}$, replacing the law of the processes $Y_s^{q,p,N,M}$ and $Z_s^{q,p,N,M}$ by the empirical mean law

$$\begin{aligned} [Y_s^{q,p,N,M,m}]^M &= \frac{1}{M} \sum_{m=1}^M \delta_{Y_s^{q,p,N,M,m}}, \\ [Z_s^{q,p,N,M,m}]^M &= \frac{1}{M} \sum_{m=1}^M \delta_{Z_s^{q,p,N,M,m}}. \end{aligned}$$

Our candidates are the particles:

$$\begin{aligned}
Y_t^{q+1,p,N,M,m} &:= \mathbb{E}_t(\mathcal{C}_p^{N,M}((F^{q,p,N,M,m})_m)) \\
&\quad - \int_0^t f(s, Y_s^{q,p,N,M,m}, Z_s^{q,p,N,M,m}, [Y_s^{q,p,N,M,m}]^M, [Z_s^{q,p,N,M,m}]^M) ds, \quad (3.63) \\
Z_t^{q+1,p,N,M,m} &:= D_t \mathbb{E}_t(\mathcal{C}_p^{N,M}((F^{q,p,N,M,m})_m)),
\end{aligned}$$

where, for all $m \in \{1, \dots, M\}$,

$$F^{q,p,N,M,m} = \zeta + \int_0^T f(s, Y_s^{q,p,N,M,m}, Z_s^{q,p,N,M,m}, [Y_s^{q,p,N,M,m}]^M, [Z_s^{q,p,N,M,m}]^M) ds. \quad (3.64)$$

Remark 3.4.5 Looking at (3.63) and (3.64), we notice that we use the M particles to compute $(\hat{d}_0, (\hat{d}_k^n)_{1 \leq k \leq p, |n|=k})$. Taking the same M drawings for the Monte Carlo simulations and the particle system has two impacts, one positive and one negative :

- the algorithm is not more costly than the one solving standard BSDEs when f is a separated variable function (see Remark 3.4.6),
- the samples $(F^{q,p,N,M,m})_{1 \leq m \leq M}$ being not independant, the speed of convergence of the algorithm will be badly impacted (see [L-25, Remark 22]).
- As for (3.62), $\mathbb{E}_t(\mathcal{C}_p^{N,M}((F^{q,p,N,M,m})_m))$ corresponds to $\mathbb{E}(\mathcal{C}_p^{N,M}((F^{q,p,N,M,m})_m) | \mathcal{G}_t)$ when we consider the second approach of [L-25, Remark 14].

Remark 3.4.6 The computational effort regarding the particle system of the algorithm depends on the form of the driver. To simplify we omit the z -terms in f .

- If $f(y, \mu) = f^1(y) + \int \phi(v) d\mu$ then $f(Y_t, [Y_t]) = f^1(Y_t) + \mathbb{E}(\phi(Y_t))$. Using a particle system $(Y_t^m)_{1 \leq m \leq M}$ we approximate $\mathbb{E}(\phi(Y_t))$ by $\frac{1}{M} \sum_{m=1}^M \phi(Y_t^m)$. This is done once and used for all the particles Y_t^m to evaluate $f(Y_t^m, [Y_t^m]^M)$. the complexity is of order M and then it is the same as the one we get when applying the algorithm for classical BSDEs.
- If the interaction is more complicated, say for example $f(y, \mu) = \int b(y, v) d\mu(v)$ the approximation of $f(Y_t, [Y_t])$ has to be recomputed for each particle and is approximated by

$$f(Y_t^m, [Y_t^m]^M) = \frac{1}{M} \sum_{m'=1}^M b(Y_t^m, [Y_t^{m'}]).$$

The computational cost is of $O(M^2)$ and thus different from the algorithm for classical BSDEs.

3.4.3 Convergence results

We aim at bounding the error between (Y, Z) — the solution to (3.49) — and $(Y^{q,p,N,M,m}, Z^{q,p,N,M,m})$ defined by (3.63). In this Section we consider the second approach presented in Remark 3.4.4. Before stating the main result of the paper, we recall some hypotheses introduced in [L-10].

In the following, (t_1, \dots, t_n) and (s_1, \dots, s_n) denote two vectors such that

$$0 \leq t_1 \leq \dots \leq t_n \leq T, \quad 0 \leq s_1 \leq \dots \leq s_n \leq T \quad \text{and} \quad \forall i, s_i \leq t_i.$$

Hypothesis 3.7 (Hypothesis \mathcal{H}_m). Let $m \in \mathbb{N}^*$. We say that F satisfies Hypothesis \mathcal{H}_m if F satisfies the two following hypotheses:

- \mathcal{H}_m^1 : $\forall j \geq 2$, $F \in \mathcal{D}^{m,j}$, that is $\|F\|_{m,j}^j < \infty$;
- \mathcal{H}_m^2 : $\forall j \geq 2$, $\forall i \in \{1, \dots, m\}$, $\forall l_0 \leq i - 1$, $\forall l_1 \leq m - i$, $\forall l \in \{1, \dots, d\}$ and for all multi-indices α_0 and α_1 such that $|\alpha_0| = l_0$ and $|\alpha_1| = l_1 + 1$, there exist two positive constants β_F and k_l^F such that

$$\sup_{t_1 \leq \dots \leq t_{l_0}} \sup_{s_{i+1} \leq \dots \leq s_{i+l_1}} \mathbb{E}(|D_{t_1, \dots, t_{l_0}}^{\alpha_0} (D_{t_i, s_{i+1}, \dots, s_{i+l_1}}^{\alpha_1} F - D_{s_i, \dots, s_{i+l_1}}^{\alpha_1} F)|^j) \leq k_l^F(j) (t_i - s_i)^{j\beta_F},$$

where $l = l_0 + l_1 + 1$. In the following, we denote $K_m^F(j) = \sup_{l \leq m} k_l^F(j)$.

Remark 3.4.7 If F satisfies \mathcal{H}_m^2 , for all multi-index α such that $|\alpha| = l$, we have

$$|\mathbb{E}(D_{t_1, \dots, t_l}^\alpha F) - \mathbb{E}(D_{s_1, \dots, s_l}^\alpha F)| \leq K_l^F((t_1 - s_1)^{\beta_F} + \dots + (t_l - s_l)^{\beta_F}), \quad (3.65)$$

where K_l^F is a constant.

Hypothesis 3.8 (Hypothesis $\mathcal{H}_{p,N}^3$). Let $(p, N) \in \mathbb{N}^2$. We say that a r.v. F satisfies $\mathcal{H}_{p,N}^3$ if

$$V_{p,N}(F) := \mathbb{V}(F) + \sum_{k=1}^p \sum_{|n(N)|_1=k} n(N)! \mathbb{V} \left(F \prod_{i=1}^N K_{n_i}(G_i) \right) < \infty,$$

where \mathbb{V} denotes the variance.

Lemma 3.4.8 Let F be a r.v. satisfying Hypothesis $\mathcal{H}_{p,N}^3$ and let $(F^m)_{1 \leq m \leq M}$ be M identically distributed r.v. with law F . We get

$$\mathbb{E}(|(C_p^N - C_p^{N,M})(F^m)_m|^2) \leq \mathbb{E}(|F|^2) + \sum_{k=1}^p \sum_{|n(N)|_1=k} \frac{n(N)!}{2} \mathbb{E} \left(|F|^2 \prod_{i=1}^N K_{n_i}^2(G_i) \right),$$

and so, we obtain

$$\mathbb{E}(|C_p^{N,M}(F^m)_m|^2) \leq 2\mathbb{E}(|F|^2) + \sum_{k=1}^p \sum_{|n(N)|_1=k} \frac{n(N)!}{2} \mathbb{E} \left(|F|^2 \prod_{i=1}^N K_{n_i}^2(G_i) \right).$$

Remark 3.4.9 If F is bounded by K , we get $V_{p,N}(F) \leq K^2 \sum_{k=0}^p \binom{k}{N}$. Then every bounded r.v. satisfies $\mathcal{H}_{p,N}^3$.

This remark ensues from $\mathbb{E}(\prod_{i=1}^N K_{n_i}^2(G_i)) = \frac{1}{n!}$.

Theorem 3.4.10 Assume that ζ satisfies \mathcal{H}_{p+q} and there exists a real $r > 4$ s.t. $\zeta \in (\mathbb{L}^r \cap \mathcal{D}^{1,r}) \subset \mathcal{H}_{p,N}^3$ and f is Lipschitz and in $C_b^{0,p+q}$. We have

$$\begin{aligned} & \| (Y - Y^{q,p,N,M,m}, Z - Z^{q,p,N,M,m}) \|_{L_2}^2 \\ & \leq \frac{A_0}{2^q} + \frac{A_1(q,p)}{(p+1)!} + A_2(q,p) \left(\frac{T}{N} \right)^{2\beta_\zeta \wedge 1} + \frac{A_3(q,p,N)}{M} + \frac{A_4(q,p,N)(\log M)^{\frac{\alpha}{2}}}{M(\frac{1}{4} \wedge \frac{\beta}{2})^{\frac{1}{2q}}} \end{aligned}$$

where A_0, A_1, A_2, A_3 and A_4 are given in the proof below. and (α, β) depends on the dimension d and is defined by [L-25, Equation (5.14)].

If $f \in C_b^{0,\infty}$ and ζ satisfies \mathcal{H}_∞ and $\mathcal{H}_{\infty,\infty}^3$, we obtain that, for each $m \geq 1$,

$$\lim_{q \rightarrow \infty} \lim_{p \rightarrow \infty} \lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \| (Y - Y^{q,p,N,M,m}, Z - Z^{q,p,N,M,m}) \|_{L_2}^2 = 0.$$

Proof. We split the error into 5 terms:

1. Picard iterations: $\varepsilon^q = \|(Y - Y^q, Z - Z^q)\|_{L_2}^2$, where (Y^q, Z^q) is defined by (3.50);
2. the truncation of the chaos decomposition: $\varepsilon^{q,p} = \|(Y^q - Y^{q,p}, Z^q - Z^{q,p})\|_{L_2}^2$, where $(Y_{q,p}, Z_{q,p})$ is defined by (3.54);
3. the truncation of the $L^2([0, T]; \mathbb{R})$ basis: $\varepsilon^{q,p,N} = \|(Y^{q,p} - Y^{q,p,N}, Z^{q,p} - Z^{q,p,N})\|_{L_2}^2$, where $(Y_{q,p,N}, Z_{q,p,N})$ is defined by (3.56);
4. the Monte Carlo approximation to compute the expectations: $\varepsilon^{q,p,N,M} = \|(Y^{q,p,N} - Y^{q,p,N,M}, Z^{q,p,N} - Z^{q,p,N,M})\|_{L_2}^2$, where $(Y_{q,p,N,M}, Z_{q,p,N,M})$ is defined by (3.62).
5. the particle system: $\varepsilon^{q,p,N,M,m} = \|(Y^{q,p,N,M} - Y^{q,p,N,M,m}, Z^{q,p,N,M} - Z^{q,p,N,M,m})\|_{L_2}^2$, where $(Y_{q,p,N,M,m}, Z_{q,p,N,M,m})$ is defined by (3.63). We have

$$\|(Y - Y^{q,p,N,M,m}, Z - Z^{q,p,N,M,m})\|_{L_2}^2 \leq 5(\varepsilon^q + \varepsilon^{q,p} + \varepsilon^{q,p,N} + \varepsilon^{q,p,N,M} + \varepsilon^{q,p,N,M,m}).$$

From [L-25, Remark 11], we know that under Hypothesis 3.6, the sequence $(Y^q, Z^q)_q$ converges to $(Y, Z) d\mathbb{P} \times dt$ a.s. and in $S_T^2(\mathbb{R}) \times H_T^2(\mathbb{R}^d)$. Moreover, we have

$$\varepsilon^q := \|(Y - Y^q, Z - Z^q)\|_{L_2}^2 \leq \frac{A_0}{2^q},$$

where A_0 depends on T , $\|\xi\|_2^2$ and on $\|f(\cdot, 0, 0, [0], [0])\|_{L_2^2([0, T]; \mathbb{R})}^2$. As in the proof of Theorem 3.2.15, the same procedure applies in the proof of [L-25, Propositions 30, 34, 36 and 38] to bound $\mathcal{E}^{q,p}$, $\mathcal{E}^{q,p,N}$, $\mathcal{E}^{q,p,N,M}$ and $\mathcal{E}^{q,p,N,M,m}$. Thanks to Picard iterations $\mathcal{E}^{q+1,\cdot}$ is bounded by $C_0 \mathcal{E}^{q,\cdot} + C_{1,\cdot}$, where $C_{1,p}$ is related to Lemma 3.2.3, $C_{1,p,N}$ is related to Lemma 3.2.11, $C_{1,p,N,M}$ is related to Lemma 3.2.12 and $C_{1,p,N,M,m}$ is related to Lemma 3.4.8. Since $\mathcal{E}^{0,\cdot} = 0$ we get the following upper bounds. From [L-25, Proposition 30] we get that $\mathcal{E}^{q,p} \leq \frac{A_1(q,m)}{(p+2-m)\dots(p+1)}$ where $A_1(q,m) := \frac{(C_1 T(T+1)L_f^2)^{q-1}}{C_1 T(T+1)L_f^2-1} K_1(q,m)$ and $K_1(q,m)$ depends on T , m , $\|\xi\|_{m+q, 2}^{\frac{(m+q-1)!}{(m-1)!}}$ and on $(\|\partial_{sp}^k f\|_\infty)_{1 \leq k \leq m+q}$. [L-25, Proposition 34] gives that $\mathcal{E}^{q,p,N} \leq A_2(q,p) \left(\frac{T}{N}\right)^{1 \wedge 2\beta\xi}$, where $A_2(q,p) := K_2(p)T(T+1)e^T \frac{(C_2 T(T+1)L_f^2)^{q-1}}{C_2 T(T+1)L_f^2-1}$ and $K_2(p)$ depends on $\overline{K_p^\xi}$, T , $\|\xi\|_{p,1}$ and on $(\|\partial_{sp}^k f\|_\infty)_{1 \leq k \leq p}$. [L-25, Proposition 36] states that $\mathcal{E}^{q,p,N,M} \leq \frac{A_3(q,p,N)}{M}$, where $A_3(q,p,N) := K_3(p,N) \frac{(C_3 T(T+1)L_f^2)^{q-1}}{C_3 T(T+1)L_f^2-1}$ and $K_3(p,N) := C \left(V_{p,N}(\xi) + T^2 \|f\|_\infty^2 \sum_{k=0}^p \binom{2N}{k} \right)$ for some $C > 0$. [L-25, Proposition 38] states that $\mathcal{E}^{q,p,N,M,m} \leq \frac{A_4(q,p,N)(\log M)^{\frac{\alpha}{2}}}{M \left(\frac{1}{4} \wedge \frac{\beta}{2}\right)^{\frac{1}{2q}}}$, where $A_4(q,p,N)$ depends on q, p, N, T and $\|f\|_\infty$. ■

Remark 3.4.11 Compared to Theorem 3.2.15 we notice that the additional term $\frac{A_4(q,p,N)(\log M)^{\frac{\alpha}{2}}}{M \left(\frac{1}{4} \wedge \frac{\beta}{2}\right)^{\frac{1}{2q}}}$ appears in the error bound. This term corresponds to the error approximation due to the particle system. It is clearly worse than $\frac{A_3(q,p,N)}{M}$ which corresponds to the error due to the Monte Carlo approximation. As seen in Lemma 3.4.8, introducing some dependency between M identically distributed r.v. gives a worse control of the error $C_p^N - C_p^{N,M}$ than the one obtained when considering i.i.d. r.v. (see Lemma 3.2.12).

Remark 3.4.12 As in the standard case, we could have handled the case of path-dependent generators. To do so, we should have assumed some regularity on the Malliavin derivatives of f with respect to the path-dependent component (see [L-10, Remark 4.7]).

3.4.4 Numerical examples

In this section, we illustrate the algorithm by presenting some explicit computations.

We consider on $[0, T]$ the following sort of processes

$$Y_t = \zeta + \int_t^T \left(\alpha Y_s + \beta \mathbb{E}(Y_s) + \gamma \mathbb{E}(Z_s) \right) ds - \int_t^T Z_s \cdot dW_s, \quad 0 \leq t \leq T, \quad (3.66)$$

where α, β and $\gamma \in \mathbb{R}$. First, we study the solution of (3.66) when $\gamma = 0$. Then, we study the solution of (3.66) in the general case.

Case (1) $\gamma = 0$. We have

$$Y_t = e^{\alpha(T-t)} \left(\mathbb{E}_t(\zeta) + \mathbb{E}(\zeta)(e^{\beta(T-t)} - 1) \right),$$

and

$$Z_t = e^{\alpha(T-t)} \mathbb{E}_t(D_t \zeta).$$

Now, if $\zeta = W_T$, we obtain

$$(Y_t, Z_t) = \left(e^{\alpha(T-t)} W_t, e^{\alpha(T-t)} \right),$$

and if $\zeta = W_T^2$, we get

$$(Y_t, Z_t) = \left(e^{\alpha(T-t)} (W_t^2 - t + T e^{\beta(T-t)}), 2e^{\alpha(T-t)} W_t \right).$$

Case (2) $\gamma \neq 0$. We consider two different values of ζ :

If $\zeta = W_T$ and $\beta \neq 0$ we have

$$(Y_t, Z_t) = \left(e^{\alpha(T-t)} \left(W_t + \frac{\gamma}{\beta} (e^{\beta(T-t)} - 1) \right), e^{\alpha(T-t)} \right).$$

If $\zeta = W_T^2$ we have

$$(Y_t, Z_t) = \left(e^{\alpha(T-t)} (W_t^2 - t + T e^{\beta(T-t)}), 2e^{\alpha(T-t)} W_t \right).$$

If $\zeta = \sqrt{|W_T|}$ we don't have any analytical formula for (Y_t, Z_t) . The exact solution will be approached by running the algorithm with a lot of drawings.

- **Convergence in p .** Table 1 represents the evolution of $\tilde{Y}_0^{q,p,N,M}$ w.r.t q (Picard iteration index), when $p = 2$ and $p = 3$. We also give the CPU time needed to get $\tilde{Y}_0^{8,p,N,M}$ and $\tilde{Z}_0^{8,p,N,M}$. We fix $M = 10^7$, $N = 20$, $\zeta = \sqrt{|W_T|}$ and $f(t, Y, Z) = Y + \mathbb{E}(Z)$. The seed of the generator is also fixed. Note that the difference between the values of $\tilde{Y}_0^{8,2,N,M}$ and $\tilde{Y}_0^{8,3,N,M}$ does not exceed 0.12%. This is due to the fast convergence of the algorithm in p . CPU time is 5 times higher when $p = 3$ than when $p = 2$.
- **Convergence in M .** Figure 3.4.1 represents the evolution of $\tilde{Y}_0^{q,p,N,M}$ and $\tilde{Z}_0^{q,p,N,M}$ w.r.t. q when $p = 2$, $N = 20$, $f(t, Y, Z) = Y + \mathbb{E}(Z)$ and $\zeta = W_T^2$ for different values of M . For this set of parameters, the exact solution is $Y_0 = 2.7183$ and $Z_0 = 0$. We notice that $\tilde{Y}_0^{q,p,N,M}$ (resp. $\tilde{Z}_0^{q,p,N,M}$) converges to the exact solution when $M \geq 10^5$ (resp. $M \geq 10^6$). For Y and Z , the algorithm stabilizes after very few iterations.
- **Convergence in N .** Figure 3.4.2 represents the evolution of $\tilde{Y}_0^{q,p,N,M}$ and $\tilde{Z}_0^{q,p,N,M}$ w.r.t. q when $p = 2$, $M = 10^6$, $f(t, Y, Z) = Y + \mathbb{E}(Z)$ and $\zeta = W_T^2$ for different values of N . We notice that (Y_0^q, Z_0^q) converges to the exact solution (Y_0, Z_0) when N increases.

Iterations	1	2	3	4	5	6	7	8	Real time (s)
p=2	0.8223	1.6440	2.0337	2.1501	2.1746	2.1784	2.1788	2.1788	49.596
p=3	0.8222	1.6444	2.0350	2.1521	2.1771	2.1811	2.1816	2.1816	284.947

Table 3.4: Evolution of $\tilde{Y}_0^{q,p,N,M}$ ($p = 2$ and $p = 3$) w.r.t. Picard iterations, $M = 10^7$, $N = 20$, $\zeta = \sqrt{|W_T|}$, $f(t, Y, Z) = Y + \mathbb{E}(Z)$ and the real time (in seconds) of calculation.

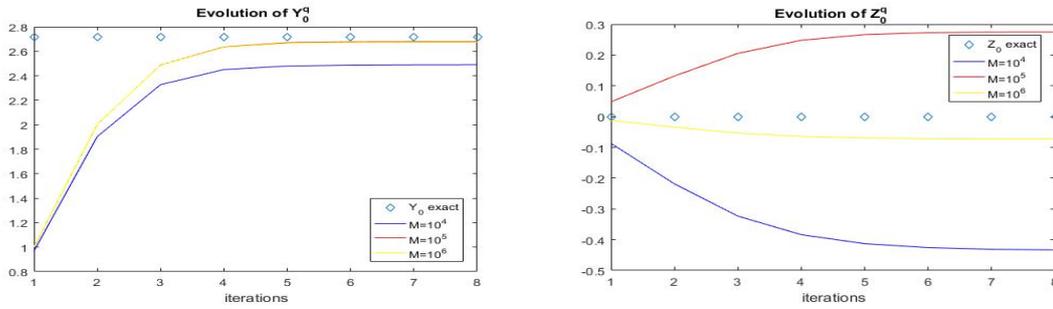


Figure 3.4.1: Evolution of $\tilde{Y}_0^{q,p,N,M}$ and $\tilde{Z}_0^{q,p,N,M}$ w.r.t. q for different values of M when $N = 20$, $p = 2$, $\zeta = W_T^2$, $f(t, Y, Z) = Y + \mathbb{E}(Z)$.

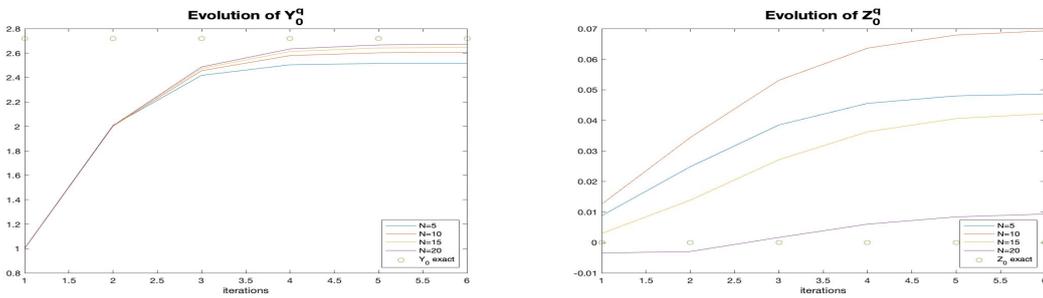


Figure 3.4.2: Evolution of $\tilde{Y}_0^{q,p,N,M}$ and $\tilde{Z}_0^{q,p,N,M}$ w.r.t. q for different values of N when $M = 10^6$, $p = 2$, $\zeta = W_T^2$, $f(t, Y, Z) = Y + \mathbb{E}(Z)$.

Bibliography

- [Aaz13] S. Aazizi. Discrete-time approximation of decoupled forward–backward stochastic differential equations driven by pure jump Lévy processes. *Advances in Applied Probability*, 45(3):791 – 821, 2013.
- [AKH00] F. Antonelli and A. Kohatsu-Higa. Filtration stability of backward SDEs. *Stochastic Analysis and Applications*, 18(1):11–37, 2000.
- [AKH02] F. Antonelli and A. Kohatsu-Higa. Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *The Annals of Applied Probability*, 12(2):423–476, 2002.
- [AKU21] S. Ankirchner, T. Kruse, and M. Urusov. Wasserstein convergence rates for random bit approximations of continuous Markov processes. *Journal of Mathematical Analysis and Applications*, 493(2):423–476, 2021.
- [Ala15] S. Alanko. *Regression-based Monte Carlo methods for solving nonlinear PDEs*. PhD thesis, New York University, 2015.
- [Bal97] V. Bally. Approximation scheme for solutions of BSDE. *Backward stochastic differential equations (N. El Karoui and L. Mazliak, eds)*, Pitman Res. Notes Math. Ser., 364, Longman, Harlow:177–191, 1997.
- [BBC⁺21] Ch. Beck, S. Becker, P. Cheridito, A. Jentzen, and A. Neufeld. Deep Splitting Method for Parabolic PDEs. *SIAM Journal on Scientific Computing*, 43(5):A3135–A3154, 2021.
- [BC08] B. Bouchard and J.-F. Chassagneux. Discrete-time approximation for continuously and discretely reflected BSDEs. *Stochastic Processes and their Applications*, 118(12):2269–2293, 2008.
- [BCS98] Mark Broadie, Jakša Cvitanić, and H Mete Soner. Optimal replication of contingent claims under portfolio constraints. *The Review of Financial Studies*, 11(1):59–79, 1998.
- [BD07] Ch. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Process. Appl.*, 117(12):1793–1812, 2007.
- [BDH⁺03] Ph. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica. l_p solutions of backward stochastic differential equations. *Stochastic Processes and their Applications*, 108(1):109–129, 2003.
- [BDLP09] R. Buckdahn, B. Djehiche, J. Li, and S. Peng. Mean-field backward stochastic differential equations: a limit approach. *The Annals of Probability*, 37(4):1524–1565, 2009.

- [BDM01] Ph. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for BSDEs. *Electronic Communications in Probability*, 6:1–14, 2001.
- [BDM02] Ph. Briand, B. Delyon, and J. Mémin. On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications*, 97(2):229 – 253, 2002.
- [BE08] B. Bouchard and R. Elie. Discrete-time approximation of decoupled Forward-Backward SDE with jumps. *Stochastic Processes and their Applications*, 118:53–75, 2008.
- [BEM18] B. Bouchard, R. Elie, and L. Moreau. Regularity of BSDEs with a convex constraint on the gains-process. *Bernoulli*, 24(3):1613–1635, 2018.
- [BET04] B. Bouchard, I. Ekeland, and N. Touzi. On the Malliavin approach to Monte Carlo approximation of conditional expectations. *Finance and Stochastics*, 8(1):45–71, 2004.
- [BHJK23] Ch. Beck, M. Hutzenthaler, A. Jentzen, and B. Kuckuck. An overview on deep learning-based approximation methods for partial differential equations. *Discrete and Continuous Dynamical Systems - B*, 28(6):3697–3746, 2023.
- [Bis73] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.*, 44:384–404, 1973.
- [BLP09] R. Buckdahn, J. Li, and S. Peng. Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Processes and their Applications*, 119(10):3133–3154, 2009.
- [BM10] Ch. Bender and T. Moseler. Importance Sampling for Backward SDEs. *Stochastic Analysis and Applications*, 28(2):226–253, 2010.
- [Bos05] M. Bossy. Some stochastic particle methods for nonlinear parabolic PDEs. *ESAIM Proc.*, 15:18–57, 2005.
- [BP03a] V. Bally and G. Pagès. A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.
- [BP03b] V. Bally and G. Pagès. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Processes and their Applications*, 106(1):1–40, 2003.
- [BPJ01] V. Bally, G. Pagès, and Printems; J. A stochastic quantization method for nonlinear problems. *Monte Carlo Methods and Applications*, 7(1-2):21–34, 2001.
- [BS91] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.*, 4:271–283, 1991.
- [BS12] Ch. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. In René A. Carmona, Pierre Del Moral, Peng Hu, and Nadia Oudjane, editors, *Numerical Methods in Finance*, pages 257–289, 2012.
- [BT04] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2):175–206, 2004.
- [BTW17] B. Bouchard, X. Tan, and X. Warin. Numerical approximation of general Lipschitz BSDEs with branching processes. *ESAIM: Proceedings and Surveys*, 65, 10 2017.

- [BTZW17] B. Bouchard, X. Tan, Y. Zou, and X. Warin. Numerical approximation of BSDEs using local polynomial drivers and branching processes. *Monte Carlo Methods and Applications*, 23(4):241–263, 2017.
- [BZ08] Ch. Bender and J. Zhang. Time discretization and Markovian iteration for coupled FBSDEs. *The Annals of Probability*, 18(1):1003–1049, 2008.
- [CC14] J.-F. Chassagneux and D. Crisan. Runge-Kutta schemes for backward stochastic differential equations. *The Annals of Applied Probability*, 24(2):679–720, 2014.
- [CCD19] J.-F. Chassagneux, D. Crisan, and F. Delarue. Numerical Method for FBSDEs of McKean-Vlasov type. *Annals of Applied Probability*, 29(3):1640–1684, 2019.
- [CCSJ21] S. Campbell, Y. Chen, A. Shrivats, and S. Jaimungal. Deep Learning for Principal-Agent Mean Field Games. <https://arxiv.org/abs/2110.01127>, 2021.
- [CD12] D. Crisan and F. Delarue. Sharp Derivative Bounds for Solutions of Degenerate Semi-Linear Partial Differential Equations. *Journal Functional Analysis*, 263(10):3024–3101, 2012.
- [CdRGT15] P.-E. Chaudru de Raynal and C.A. Garcia Trillos. A cubature based algorithm to solve decoupled McKean-Vlasov forward-backward stochastic differential equations. *Stochastic Processes and their Applications*, 125(6):2206–2255, 2015.
- [CÉK12] J.-F. Chassagneux, R. Élie, and I. Kharroubi. Discrete-time Approximation of Multi-dimensional BSDEs with oblique reflections. *Annals of Applied Probability*, 22:971–1007, 2012.
- [CGG23] G. Callegaro, A. Gnoatto, and M. Grasselli. A fully quantization-based scheme for FBSDEs. *Applied Mathematics and Computation*, 441:127666, 2023.
- [CGT20] J.-F. Chassagneux and C.A. Garcia Trillos. Cubature methods to solve BSDEs: Error expansion and complexity control. *Mathematics of Computation*, 89:1895–1932, 2020.
- [Cha09] J.-F. Chassagneux. A discrete-time approximation for doubly reflected BSDEs. *Adv. in Appl. Probab.*, 41(1):101–130, 2009.
- [Cha14] J.-F. Chassagneux. Linear multistep schemes for BSDEs. *SIAM Journal on Numerical Analysis*, 52(6):2815–2836, 2014.
- [Che97] D. Chevance. Résolution numérique des équations différentielles stochastiques rétrogrades. PhD. thesis, Université de Provence-Aix-Marseille I, Marseille, 1997.
- [CK96] J. Cvitanic and I. Karatzas. Backward stochastic differential equations with reflection and Dynkin games. *The Annals of Probability*, 24(4):2024–2056, 1996.
- [CKS98] J. Cvitanic, I. Karatzas, and H.M. Soner. Backward stochastic differential equations with constraints on the gains-process. *Annals of Probability*, pages 1522–1551, 1998.
- [CKSY23] J. Chessari, R. Kawai, Y. Shinozaki, and T Yamada. Numerical methods for backward stochastic differential equations: A survey. *Probability Surveys*, 20, 2023.
- [CL84] M.G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Mathematics of Computation*, 43:1–19, 1984.

- [CM08] S Crépey and A Matoussi. Reflected and doubly reflected BSDEs with jumps: a priori estimates and comparison. *Ann. Appl. Probab.*, 18(5):2041–2069, 2008.
- [CM12] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing. *SIAM Journal on Financial Mathematics*, 3(1):534–571, 2012.
- [CM14] D. Crisan and K. Manolarakis. Second order discretization of backward SDEs and simulation with the cubature method. *The Annals of Applied Probability*, 24(2):652 – 678, 2014.
- [CMM99] F. Coquet, V. Mackevičius, and J. Mémin. Corrigendum to stability in \mathbf{d} of martingales and backward equations under discretization of filtration [stochastic process. appl., 75, 1998]. *Stochastic Process. Appl.*, 82(2):335–338, 1999.
- [CMS01] F. Coquet, J. Mémin, and L. Slominski. On weak convergence of filtrations. *Séminaire de probabilités de Strasbourg*, 35:306–328, 2001.
- [CMT10] D. Crisan, K. Manolarakis, and N. Touzi. On the Monte-Carlo simulation of BSDEs : An improvement on the Malliavin weights. *Stochastic Processes and their Applications*, 120(7):1133–1158, 2010.
- [CR15] J.-F. Chassagneux and A. Richou. Numerical Stability Analysis of the Euler Scheme for BSDEs. *SIAM Journal on Numerical Analysis*, 53(2):1172–1193, 2015.
- [CR16] J.-F. Chassagneux and A. Richou. Numerical Simulation of quadratic BSDEs. *The Annals of Applied Probability*, 26(1):262–304, 2016.
- [DK05] H. Dong and N.V. Krylov. Rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C_1 and C_2 coefficients. *Electronic Journal of Differential Equations*, 102:1–25, 2005.
- [DMP96] J. Douglas, J. Ma, and Ph. Protter. Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability*, 6(3):940 – 968, 1996.
- [DQS16] R. Dumitrescu, M.C. Quenez, and A. Sulem. Generalized Dynkin games and doubly reflected BSDEs with jumps. *Electronic Journal of Probability*, 21:1–32, 2016.
- [EHJ17] W. E, J. Han, and A. Jentzen. Deep Learning-Based Numerical Methods for High-Dimensional Parabolic Partial Differential Equations and Backward Stochastic Differential Equations. *Communications in Mathematics and Statistics*, 5(4):349–380, 2017.
- [EHJ20] W. E, J. Han, and A. Jentzen. Algorithms for solving high dimensional PDEs: from nonlinear Monte Carlo to machine learning. *Nonlinearity*, 35:278 – 310, 2020.
- [EHJK19] W. E, M. Hutzenthaler, A. Jentzen, and T. Kruse. On Multilevel Picard Numerical Approximations for High-Dimensional Nonlinear Parabolic Partial Differential Equations and High-Dimensional Nonlinear Backward Stochastic Differential Equations. *Journal of Scientific Computing*, 79(3):1534–1571, 2019.
- [EHO05] E.H. Essaky, N. Harraj, and Y. Ouknine. Backward stochastic differential equation with two reflecting barriers and jumps. *Stochastic Analysis and Applications*, 23:921–938, 2005.

- [EKH03] N. El-Karoui and S. Hamadène. Bsdés and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations. *Stochastic Processes and their Applications*, 107(1):145–169, 2003.
- [EKKP⁺97] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M.C. Quenez. Reflected solutions of Backward SDE’s and related obstacle problems for PDE’s. *The Annals of Probability*, 25(2):702–737, 1997.
- [EKPQ97a] N. El Karoui, E. Pardoux, and M.C. Quenez. *Reflected Backward SDEs and American Options*, page 215–231. Numerical Methods in Finance, Cambridge University Press, 1997.
- [EKPQ97b] N. El Karoui, S. Peng, and M.-C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
- [ENP21] R. El Nmeir and G. Pagès. Quantization-based approximation of reflected BSDEs with extended upper bounds for recursive quantization, 2021. <https://arxiv.org/abs/2105.07684>.
- [Ess58] C.G. Esseen. *On mean central limit theorems*. Kungl. Tekn. Högsk. Handl. Stockholm **121**, 1958.
- [Ess08] E.H. Essaky. Reflected backward stochastic differential equation with jumps and RCLL obstacle. *Bulletin des Sciences Mathématiques*, 132(8):690–710, 2008.
- [FLLL01] E. Fournié, J.-M. Lasry, J. Lebuchoux, and P.-L. Lions. Applications of Malliavin calculus to Monte Carlo methods in finance II. *Finance and Stochastics*, 5(2):201–236, 2001.
- [FT19] M. Fujii and A. Takahashi. Asymptotic expansion for forward-backward SDEs with jumps. *Stochastics*, 91(2):175–214, 2019.
- [FYZ18] Y. Fu, J. Yang, and W. Zhao. Prediction-Correction Scheme for Decoupled Forward Backward Stochastic Differential Equations with Jumps. *East Asian Journal on Applied Mathematics*, 6(3):253–277, 2018.
- [FZ20] J.-P. Fouque and Z. Zhang. Deep Learning Methods for Mean Field Control Problems With Delay. *Frontiers in Applied Mathematics and Statistics*, 6, 2020.
- [GGG12] Ch. Geiss, S. Geiss, and E. Gobet. Generalized fractional smoothness and L_p -variation of BSDEs with non-Lipschitz terminal condition. *Stochastic Processes and their Applications*, 122(5):2078–2116, 2012.
- [GLSTV16] E. Gobet, J. G. López-Salas, P. Turkedjiev, and C. Vázquez. Stratified Regression Monte-Carlo Scheme for Semilinear PDEs and BSDEs with Large Scale Parallelization on GPUs. *SIAM Journal on Scientific Computing*, 38(6):C652–C677, 2016.
- [GLW06] E. Gobet, J.-P. Lemor, and X. Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12(5):889–916, 2006.
- [GMW22] M. Germain, J. Mikael, and X. Warin. Numerical Resolution of McKean-Vlasov FB-SDEs Using Neural Networks. *Methodology and Computing in Applied Probability*, 24(4):2557–2586, 2022.

- [GP15] E. Gobet and S. Pagliarani. Analytical Approximations of BSDEs with Nonsmooth Driver. *SIAM Journal on Financial Mathematics*, 6(1):919–958, 2015.
- [GPP22] A. Gnoatto, M. Patacca, and A. Picarelli. A deep solver for BSDEs with jumps. <https://arxiv.org/abs/2211.04349>, 2022.
- [GPW22] M. Germain, H. Pham, and X. Warin. Approximation Error Analysis of Some Deep Backward Schemes for Nonlinear PDEs. *SIAM Journal on Scientific Computing*, 44(1):A28–A56, 2022.
- [GS16] C. Geiss and A. Steinicke. L_2 -variation of Lévy driven BSDEs with non-smooth terminal conditions. *Bernoulli*, 22(2):995–1025, 2016.
- [GT16] E. Gobet and P. Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation*, 85(299):1359–1391, 2016.
- [HH06] S. Hamadène and M. Hassani. BSDEs with two reacting barriers driven by a Brownian motion and an independent Poisson noise and related Dynkin game. *Electronic Journal of Probability*, 11:121–145, 2006.
- [HHI⁺14] U. Horst, Y. Hu, P. Imkeller, A. Réveillac, and J. Zhang. Forward–backward systems for expected utility maximization. *Stochastic Processes and their Applications*, 124(5):1813–1848, 2014.
- [HHL23] J. Han, R. Hu, and J. Long. Learning High-Dimensional McKean-Vlasov Forward-Backward Stochastic Differential Equations with General Distribution Dependence, 2023. <https://arxiv.org/abs/2204.11924>.
- [HIM05] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15(3):1691–1712, 2005.
- [HJE18] J. Han, A. Jentzen, and W. E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.
- [HJKN23] M. Hutzenthaler, A. Jentzen, T. Kruse, and T.A. Nguyen. Overcoming the curse of dimensionality in the numerical approximation of backward stochastic differential equations. *Journal of Numerical Mathematics*, 31(1):1–28, 2023.
- [HKN22] M. Hutzenthaler, T. Kruse, and T.A. Nguyen. On the speed of convergence of Picard iterations of backward stochastic differential equations. *Probability, Uncertainty and Quantitative Risk*, 7(2):133–150, 2022.
- [HL12] P. Henry-Labordère. Cutting CVA’s complexity. *Risk*, 25:67–73, 01 2012.
- [HL20] J. Han and J. Long. Convergence of the deep BSDE method for coupled FBSDEs. *Probability, Uncertainty and Quantitative Risk*, 5(0):5, 2020.
- [HLTT14] P. Henry-Labordère, X. Tan, and N. Touzi. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Processes and their Applications*, 124(2):1112–1140, 2014.

- [HN22] M. Hutzenthaler and T.A. Nguyen. Multilevel Picard approximations for high-dimensional decoupled forward-backward stochastic differential equations, 2022. <https://arxiv.org/abs/2204.08511>.
- [HO03] S. Hamadène and Y. Ouknine. Reflected Backward stochastic differential equation with jumps and random obstacle. *Elec. Journ. of Probab.*, 8:1–20, 2003.
- [HO16] S. Hamadène and Y. Ouknine. Reflected Backward SDEs with general jumps. *Theory of Probability & Its Applications*, 60(2):263–280, 2016.
- [HPW19] C. Huré, H. Pham, and X. Warin. Deep backward schemes for high-dimensional nonlinear PDEs. *Math. Comput.*, 89:1547–1579, 2019.
- [HT10] Y. Hu and S. Tang. Multi-dimensional BSDE with oblique reflection and optimal switching. *Probability Theory and Related Fields*, 147:89–121, 2010.
- [HW09] S. Hamadène and H. Wang. BSDEs with two RCLL Reflecting Obstacles driven by a Brownian Motion and Poisson Measure and related Mixed Zero-Sum Games. *Stochastic Processes and their Applications*, 119:2881–2912, 2009.
- [JMP89] A. Jakubowski, J. Mémin, and G. Pagès. Convergence en loi des suites d’intégrales stochastiques sur l’espace \mathbb{D}^1 de Skorokhod. *Prob. Th. and Rel. Fields*, 81:111–137, 1989.
- [JPPZ20] S. Ji, S. Peng, Y. Peng, and X. Zhang. Three Algorithms for Solving High-Dimensional Fully Coupled FBSDEs Through Deep Learning. *IEEE Intelligent Systems*, 35:71–84, 2020.
- [JPPZ22] S. Ji, S. Peng, Y. Peng, and X. Zhang. A deep learning method for solving stochastic optimal control problems driven by fully-coupled FBSDEs, 2022. <https://arxiv.org/abs/2204.05796>.
- [Kob00] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28(2):558 – 602, 2000.
- [LAS05] N. Gigli L. Ambrosio and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser, Basel, Boston, Berlin, 2005.
- [LGW05] J.-P. Lemor, E. Gobet, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005.
- [LJUV02] J.A. León, Solé J.L., F. Utzet, and J. Vives. On Lévy processes, Malliavin calculus and market models with jumps. *Finance and Stochastics*, 6:197–225, 2002.
- [LMT13] A. Lejay, E. Mordecki, and S. Torres. Numerical approximation of Backward Stochastic Differential Equations with Jumps. <http://hal.archives-ouvertes.fr/docs/00/85/95/89/PDF/lejay-mordecki-torres-v3.pdf>, 2013.
- [LPST14] G. Last, M. Penrose, M. Schulte, and C. Thäle. Moments and central limit theorems for some multivariate poisson functionals. *Advances in Appl. Probab.*, 46(2):348–364, 2014.

- [LSZ20] Y. Liu, Y. Sun, and W. Zhao. A Fully Discrete Explicit Multistep Scheme for Solving Coupled Forward Backward Stochastic Differential Equations. *Advances in Applied Mathematics and Mechanics*, 12(3):643–663, 2020.
- [LX07] J.P. Lepeltier and M. Xu. Reflected backward stochastic differential equations with two RCLL barriers. *ESAIM: Probability and Statistics*, 11:3–22, 2007.
- [Mos10] T. Moseler. A Picard-type Iteration for Backward Stochastic Differential Equations: Convergence and Importance Sampling. PhD. thesis, University of Konstanz, 2010.
- [MPSMT02] J. Ma, Ph. Protter, J. San Martín, and S. Torres. Numerical Method for Backward Stochastic Differential Equations. *Ann. Appl. Probab.*, 12(1):302–316, 02 2002.
- [MPX08] J. Mémin, S. Peng, and M. Xu. Convergence of solutions of discrete Reflected backward SDE's and simulations. *Acta Mathematica Sinica*, 24(1):1–18, 01 2008.
- [MPY94] J. Ma, Ph. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly: a four step scheme. *Probability Theory and Related Fields*, 98:339–359, 1994.
- [MSMT11] M. Martínez, J. San Martín, and S. Torres. Numerical Method for Reflected Backward Stochastic Differential Equations. *Stochastic Analysis and Applications*, 29(6):1008–1032, 2011.
- [MZ02] J. Ma and J. Zhang. Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.*, 12(4):1390–1418, 11 2002.
- [MZ05] J. Ma and J. Zhang. Representations and regularities for solutions to BSDEs with reflections. *Stochastic Processes and their Applications*, 115(4):539–569, 2005.
- [NS19] S. Ninomiya and Y. Shinozaki. Higher-order Discretization Methods of Forward-backward SDEs Using KLVN-scheme and Their Applications to XVA Pricing. *Applied Mathematical Finance*, 26(3):257–292, 2019.
- [Nua06] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
- [PP90] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.
- [PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pages 200–217. Springer, 1992.
- [Pri09] N. Privault. *Stochastic Analysis in Discrete and Continuous settings. With normal martingales*. Lecture notes in Mathematics. Springer-Verlag, Berlin, 2009.
- [PS18] G. Pagès and A. Sagna. Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering. *Stochastic Processes and their Applications*, 128(3):847–883, 2018.
- [PT11] G Peccati and M. Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams*. Springer-Verlag, 2011.

- [PX11] S. Peng and M. Xu. Numerical algorithms for bsdes with 1-d Brownian motion: convergence and simulation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45:335–360, 2011.
- [QS13] M.C. Quenez and A. Sulem. BSDEs with jumps, optimization and applications to dynamic risk measures. *Stochastic Processes and their Applications*, 123(8):3328–3357, 2013.
- [QS14] M.C. Quenez and A. Sulem. Reflected BSDEs and robust optimal stopping for dynamic risk measures with jumps. *Stochastic Processes and their Applications*, 124(9):3031–3054, 2014.
- [REK00] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.
- [Ric11] A. Richou. Numerical Simulation of BSDEs with drivers of quadratic growth. *The Annals of Applied Probability*, 21(5):1933–1964, 2011.
- [Rio09] E. Rio. Upper bounds for minimal distances in the central limit theorem. *Annales de l’Institut Henri Poincaré (B) Probabilités et Statistiques*, 45(3):802–817, 2009.
- [Rio11] E. Rio. Asymptotic constants for minimal distance in the central limit theorem. *Electron. Commun. Probab.*, 16:96–103, 2011.
- [Roy06] M. Royer. Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic Process. Appl.*, 116(10):1358–1376, 2006.
- [Slo89] L. Slomiński. Stability of strong solutions of stochastic differential equations. *Stochastic Process. Appl.*, 31(2):173–202, 1989.
- [Tao10] T. Tao. *An Epsilon of Room, I: Real Analysis*. An Epsilon of Room. American Mathematical Society, 2010.
- [TL94] S.H. Tang and X Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Cont. and Optim.*, 32:1447–1475, 1994.
- [Tol06] S. Toldo. Stability of solutions of BSDEs with random terminal time. *ESAIM: Probability and Statistics*, 10:141–163, 2006.
- [TV03] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations. *The Annals of Applied Probability*, 13(1):140–180, 2003.
- [TY15] A. Takahashi and T. Yamada. An asymptotic expansion of forward-backward SDEs with a perturbed driver. *International Journal of Financial Engineering*, 2(2):1550020, 2015.
- [Wal03] J. Walsh. The Rate of Convergence of the Binomial Tree Scheme. *Finance Stoch.*, 7:337–361, 07 2003.
- [Xu11] M. Xu. Numerical algorithms and Simulations for Reflected Backward Stochastic Differential Equations with Two Continuous Barriers. *Journal of Computational and Applied Mathematics*, 236:1137–1154, 2011.

- [YZZ20] J. Yang, W. Zhao, and T. Zhou. A Unified Probabilistic Discretization Scheme for FB-SDEs: Stability, Consistency, and Convergence Analysis. *SIAM Journal on Numerical Analysis*, 58(4):2351–2375, 2020.
- [ZCP06] W. Zhao, L. Chen, and S. Peng. A new kind of accurate numerical method for backward stochastic differential equations. *SIAM Journal on Scientific Computing*, 28(4):1563–1581, 2006.
- [ZFFZ14] W. Zhao, Y. Fu, and T. Zhou. New Kinds of High-Order Multistep Schemes for Coupled Forward Backward Stochastic Differential Equations. *SIAM Journal on Scientific Computing*, 36(4):A1731–A1751, 2014.
- [Zha01] J. Zhang. fine properties of backward stochastic differential equations, with applications. PhD thesis, Purdue University, 2001.
- [Zha04] J. Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 02 2004.
- [Zha05] J. Zhang. Representation of solutions to BSDEs associated with a degenerate FSDE. *The Annals of Applied Probability*, 15(3):1798–1831, 2005.
- [Zha17] J. Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*. Probability theory and stochastic modelling. Springer, 2017.
- [ZLZ12] W. Zhao, Y. Li, and G. Zhang. A generalized θ -scheme for solving backward stochastic differential equations. *Discrete and Continuous Dynamical Systems - B*, 17:1585–1603, 2012.
- [ZM23] W. Zhang and H. Min. Lp-error estimates for numerical schemes for solving certain kinds of mean-field backward stochastic differential equations. *Journal of Theoretical Probability*, 36(2):762–778, 2023.
- [Zol76] V.M. Zolotarev. Metric distances in spaces of random variables and their distributions. *Math. USSR Sbornik*, 30:373–401, 1976.
- [ZWP09] W. Zhao, J-L. Wang, and S. Peng. Error estimates of the theta-scheme for backward stochastic differential equations. *Discrete and Continuous Dynamical Systems-series B*, 12:905–924, 2009.
- [ZZJ10] W. Zhao, G. Zhang, and L. Ju. A stable multistep scheme for solving backward stochastic differential equations. *SIAM Journal on Numerical Analysis*, 48(4):1369–1394, 2010.