Topics on Backward Stochastic Differential Equations.

Theoretical and practical aspects.

Arnaud LIONNET

St John's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Mathematics

Tuesday 12th November 2013

To my dedicated readers.

Abstract.

This doctoral thesis is concerned with some theoretical and practical questions related to backward stochastic differential equations (BSDEs) and more specifically their connection with some parabolic partial differential equations (PDEs). The thesis is made of three parts.

In the first part, we study the probabilistic representation for a class of multidimensional PDEs with quadratic nonlinearities of a special form. We obtain a representation formula for the PDE solution in terms of the solutions to a Lipschitz BSDE. We then use this representation to obtain an estimate on the gradient of the PDE solutions by probabilistic means. In the course of our analysis, we are led to prove some results for the associated multidimensional quadratic BSDEs, namely an existence result and a partial uniqueness result.

In the second part, we study the well-posedness of a very general quadratic reflected BSDE driven by a continuous martingale. We obtain the comparison theorem, the special comparison theorem for reflected BSDEs (which allows to compare the increasing processes of two solutions), the uniqueness and existence of solutions, as well as a stability result. The comparison theorem (from which uniqueness follows) and the special comparison theorem are obtained through natural techniques and minimal assumptions. The existence is based on a perturbative procedure, and holds for a driver whis is Lipschitz, or slightly-superlinear, or monotone with arbitrary growth in y. Finally, we obtain a stability result, which gives in particular a local Lipschitz estimate in BMO for the martingale part of the solution.

In the third and last part, we study the time-discretization of BSDEs having nonlinearities that are monotone but with polynomial growth in the primary variable. We show that in that case, the explicit Euler scheme is likely to diverge, while the implicit scheme converges. In fact, by studying the family of θ-schemes, which are mixed explicit-implicit, θ characterizing the degree of implicitness, we find that the scheme converges when the implicit component is dominant ($\theta \geq \frac{1}{2}$). We then propose a tamed explicit scheme, which converges. We show that the implicit-dominant schemes with $\theta > \frac{1}{2}$ and our tamed explicit scheme converge with order $\frac{1}{2}$, while the trapezoidal scheme ($\theta = \frac{1}{2}$) converges with order $\frac{7}{4}$.

Acknowledgements.

In the first place I would like to thank my supervisor, Zhongmin Qian, who proposed to a younger me to come to Oxford for the doctorate, made sure I would work in good conditions (funding, a place in one of the best Colleges, a desk in the best Institute in town) and took me with him on a project in my first year, which lead me into the field of BSDEs. I would also very much like to thank Terry Lyons. While I haven't had the privilege of being his student, I could benefit from his advice as a young researcher looking forward to making a career in academia.

I am most grateful to my examiners, Ying Hu and Hanqing Jin, who accepted to examine me within a very short time frame.

Some very special thanks go to my colleagues, in particular Gonçalo dos Reis and Lukasz Szpruch with whom I wrote the paper I am the most happy with so far. It was a long and bumpy road, but the outcome was well worth the effort. I also thank Samuel Cohen and Gechun Liang with whom I discussed on many occasions (mostly not about research, but sometimes about work) and co-organised a conference.

The majority of my time over the past four years (and this count includes the time spent sleeping, cooking, eating, etc) was spent in the Oxford-Man Institute. My time there would definitely have been much less happy without the other young ones. My thoughts go in particular to Youness, Bahman, Nathan and Cavit, all people with whom I wish I had spent more time.

While a thesis is a personal journey from student to researcher, and while my current interests have been shaped to a great extent by my stay at the Oxford–Man Institute, I feel that the mathemacian I have become has been influenced in great part by my time at the Ecole Normale Supérieure de Lyon. I would therefore like to address my thanks to my teachers. To Cédric Villani, for his high-level, clear and motivated lectures. To Vincent Beffara, for his Introduction to Probability which surely went far beyond what was initially meant by "introduction". To Christophe Sabot, for his very cool attitude. To Francis Clarke, for his inimitable style and his dry humour. To Stéphane Attal, not only for his teaching, but to whom I am very grateful for having been ready to co-supervise me with Cédric Bernardin. Finally to Etienne Ghys: his lectures were casually amazing, but not as amazing as his qualities as a speaker in general. Etienne

Ghys, as well Cédric Villani, definitely contributed to making me aware that speaking is a whole art, of which they are masters. Doing mathematics is one thing, and it is indeed the first thing, but when the mathematics is done, telling those mathematics is a totally other —but very important— thing. Everyone should be telling good stories.

I have received financial support from a number of different sources during my doctoral years. The ENS Lyon gave me a stipend for my first year (and for the three years before that). The EPSRC doctoral training grant EP/P505216/1 covered my university fees and some research allowance. The Lamb And Flag Scholarship of St John's College provided me with a stipend for my second and third years. The Mathematical Institute Prizes Fund gave me an extra six months of stipend. Finally, the Oxford-Man Institute gave me a generous research support, in addition to the day-to-day facilities.

I would also like to thank a number of administrators who helped making a number of things run very smoothly for me and who work, invisible to most, to make and manage our working environment. So thanks to Sandy, Lucy, Ali, Helen, Wendy, Kelly, Laura and Caroline.

Last, and by no means least, there is a number people who have had nothing to do with my thesis, and indeed who often have nothing to with academia. I will not enumerate them all here, but they are very important in my life. They will recognize themselves.

How to read this thesis.

This thesis is made of 4 chapters. Chapters 2, 3 and 4 report on research results that are independent and can be read separately. The first chapter is a general introduction to the themes with which the thesis is concerned, and contains a brief overview of what is discussed in each subsequent chapter.

Statement of authorship.

I wrote this thesis on my own. However some of the research work reported was carried collaboratively.

Chapter 2 presents results which were obtained jointly with Zhongmin Qian (my supervisor) and Gechun Liang. Chapter 4 presents results which were obtained jointly with Gonçalo dos Reis and Lukasz Szpruch. In both cases, those results led to an article, in which the writing was also done collaboratively. These chapters use some of the technical material prepared in that view (theorems, computations, etc), but are otherwise my own writing.

Papers on which this thesis is based.

This thesis is based on 3 articles which have already been written and whose status is the following.

- G. Liang, A. Lionnet and Z. Qian. *On Girsanov's transform for BSDEs*
  (arXiv:1011.3228 ; under revision)
- A. Lionnet. *Some results on general quadratic reflected BSDEs driven by a continous martingale*
  (accepted for publication in *Stochastic Processes and their Applications*)
- A. Lionnet, G. dos Reis and L. Szpruch. *Time discretization of FBSDEs with polynomial growth drivers and reaction-diffusion PDEs*
  (submitted ; arXiv:1309.2865)

Contents

Chapter 1

Introduction.

In this thesis, we present results obtained in the course of some of the research projects carried by the author. This research lies in the area of backward stochastic differential equations (abbreviated as BSDEs).

BSDEs constitute a relatively recent, dynamic and exciting area in stochastic analysis. Indeed, while BSDEs are concerned with stochastic processes, the techniques involved are very analytic by nature, as will be seen clearly in the following. In this regard, BSDEs lie well within stochastic analysis.

Linear BSDEs were first introduced by Bismut [6]. However their systematic study really started with the seminal paper [65] of Pardoux and Peng. These equations and their solutions can be understood in a number of different ways depending on the motivations for studying them. In the context of optimal stochastic control (which was the motivation of Bismut), they are the adjoint equations in the stochastic version of Pontryagin's maximum principle. It was realized very early on, since Peng [69, 70], that BSDE solutions provide probabilistic representation for the solutions to a large number of semilinear partial differential equations (PDEs thereafter). In fact, systems of forward and backward stochastic differential equations (FBSDEs thereafter) really are the probabilistic counterpart of parabolic PDEs. BSDEs also arise naturally in mathematical finance as the most straightforward language in which to express the replication of derivatives. Of course, BSDEs also appear often in mathematical finance due to the fact that this field was already making great use of PDEs and stochastic control. Finally, BSDEs are deeply connected to nonlinear expectations and dynamic risk measures. These numerous applications make the case for studying in detail these equations and explain the high level of effort devoted to them, of which ours is part.

More specifically, our guiding line in this thesis will be the link between BSDEs and parabolic PDEs. We will first study a particular class of PDEs together with their associated BSDEs, and provide a probabilistic representation in that case. We will then be concerned with obstacle problems. We will formulate reflected BSDEs (the appropriate probabilistic counterpart) in a sufficiently general way so as to allow to deal with initial value, Dirichlet boundary and Neumann boundary problems, and study the well-posedness of reflected BSDEs in such a general context. Finally, we will come back to a more classical setting and turn our attention to some numerical aspects. We will study the convergence or divergence of the standard time-discretization scheme for BSDEs, under some analytical assumptions on the nonlinearities motivated by reaction-diffusion PDEs.

Before presenting further the research of the coming chapters, it is a good time to recall what BSDEs are and the main results known about them.

1.1 Backward stochastic differential equations.

1.1.1 Generalities on BSDEs.

Given a time interval $[0, T]$, a BSDE is an equation

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t \\ Y_T = \xi \end{cases} \tag{1.1.1}$$

In the above, W is a fixed Brownian motion, on some standard filtered probability space $(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}, P)$, and we assume generally that $\mathcal{F}$, which is fixed, is the augmented filtration of W. What is given and constitute the parameters (or data) in this equation are the terminal condition ξ (an $\mathcal{F}_T$-measurable random variable) and the coefficient f (a measurable function of ω, t, y, z, although the dependence in ω is typically not explicitly written). The coefficient f can also called the drift coefficient, nonlinearity coefficient, the generator, or driver. What we look for, and call a solution to the BSDE, is a pair $(Y_t, Z_t)_{t\in[0,T]}$ of stochastic processes, which are required to be adapted to the filtration $\mathcal{F}$.

Essentially, the BSDE describes the evolution backward in time of the state variable

Y, starting from the terminal condition ξ, and whose dynamics is governed by the drift coefficient f. The process Z is called secondary variable, or control variable, or control process (for reasons that will be explained below).

Naturally, the BSDE written in differential form above can be rewritten in integral form (in fact, this is how it should rigorously be understood)

$$Y_t = \xi + \int_t^T f(u, Y_u, Z_u)du - \int_t^T Z_u dW_u \ . \tag{1.1.2}$$

BSDEs are not SDEs with a terminal condition. The adaptedness issue.

The requirement that the solution be adapted puts severe constraints on what kind of equation can be used for the dynamics of Y, and rules out the possibility to conceive BSDEs as SDEs with a terminal condition instead of an initial one. After all, one might be tempted to draw a parallel with ODEs. Considering the autonomous ODE $dy_t = f(y_t)dt$, it does not matter whether the problem is given with an initial condition $(0, y_0)$ or terminal condition (T, y_0). Ultimately, this problem consists in finding the curve(s) tangent to the vector field f and passing through y_0. Viewed otherwise, the two problems are equivalent by a time-reversal. This, however, cannot be done in the context of stochastic processes, where the filtration gives a clear direction to time.

Suppose for a moment that the problem was to find a process Y such that

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt \\ Y_T = \xi \ . \end{cases}$$

This is simply an ODE. Now, let us look at the simple example where $f = 0$ and $\xi = W_T$. The only candidate solution Y would have to be constant on $[0, T]$ and, given the terminal condition, we would have $Y_t = \xi = W_T$ for all t, and therefore would not be adapted. This is why it is sometimes said that the role of the variable Z is to "correct for the adaptedness", or to give more degrees of freedom in the definition of a solution, so as to allow the existence of adapted solutions.

Suppose that we looked more generally at the problem of finding Y, or even (Y, Z), satisfying

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + g(t, Y_t, Z_t)dW_t \\ Y_T = \xi \ . \end{cases}$$

This is an SDE, for each fixed process Z. Again, in the simple example where $f = g = 0$, there can be no adapted solution. The process Z must give enough freedom in the dynamics to correct for the lack of adaptedness coming from the increment $-f(t, Y_t, Z_t)dt$. Interestingly, although SDEs are not exactly the right way to think of BSDEs, in their seminal paper [65] Pardoux and Peng were considering an equation as above, but with the requirement that for fixed (t, y), the map $z \mapsto g(t, y, z)$ is a Lipschitz bijection. That is, Z has full freedom to determine the coefficient in front of dW_t.

BSDEs as stochastic, backward ODEs.

BSDEs can be regarded as the stochastic version of a backward ODE. If one continues the thread of the discussion above, seeing as the solution to the ODE $Y_t = \xi + \int_t^T f(Y_u)du$ may not be adapted, one may be tempted to consider the next best thing and take the conditional expectation, since this is the best projection on the correctly-measurable random variables.

Let us keep this in mind but work from the definition given above of a BSDE (1.1.1)–(1.1.2). If (Y, Z) is a solution, then, since Y is adapted, $Y_t = E\big(Y_t \big| \mathcal{F}_t\big)$. Also, we know that $dM_t = Z_t dW_t$ is the martingale part of the semimartingale Y. So (1.1.2) implies that

$$\begin{cases} Y_t = E\left(\xi + \displaystyle\int_t^T f(u, Y_u, Z_u)du \middle| \mathcal{F}_t\right) \\ \text{where } Z \text{ is the progressively measurable process such that} \\ \displaystyle\int ZdW \text{ is the martingale part of } Y\ . \end{cases} \tag{1.1.3}$$

That (1.1.3) is equivalent to (1.1.1)–(1.1.2) follows from a few of lines of manipulations and the use of the martingale representation theorem.

In this way, we see that BSDEs can be understood as backward ODEs with continuously taken conditional expectations. This view of BSDEs will find more weight when we look at the connection with PDEs.

Martingales and nonlinear martingales.

At one extreme end of BSDEs, when $f = 0$, we see from (1.1.3) that the solution Y is simply the martingale with terminal value ξ. This partly explains why, in the general case, one can consider the solutions to BSDEs as "nonlinear martingales". At the other end of the spectrum, in the case when the terminal condition ξ is not random, and when f is not allowed to be random either, BSDEs boil down to (deterministic, backward) ODEs.

On both ends one has a comparison property : if $\xi \leq \xi'$, and if $f \leq f'$, then $Y \leq Y'$. This holds true also for general BSDEs, so long as the coefficient f satisfies some regularity assumptions, and is called the comparison theorem.

Well-posedness questions, known results.

In view of the many applications of BSDEs, it is a question of central importance to know under what conditions on the coefficient f and the terminal condition ξ these equations are *well-posed* (in the sense of Hadamard). That is, we want to know when a solution exists and is unique, and we also want some continuous dependence of the solution on the data (principally ξ).

Pardoux and Peng [65] proved that when f is Lipschitz in the variable (y, z), uniformly in the others (i.e. ω and t), and when ξ is square integrable, then the BSDE has a unique square-integrable solution. Their original result was reproved in a somehow simpler way in El Karoui, Peng and Quenez [31], where the authors also gave regularity estimates. The comparison theorem under these assumptions was obtained in Peng [70]. Many developments took place in the following 20 years, and a few main sets of analytical assumptions on the data (f, ξ) have been identified under which BSDEs are usually well-behaved and the usual well-posedness results can be obtained.

Kobylanski [50] obtained the first results for quadratic BSDEs, that is to say when the coefficient f can have at most quadratic growth in the z variable. This case occurs naturally in finance when considering problems of utility maximization and indifference pricing (see Hu, Imkeller and Müller [40], Rouge and El Karoui [75]). Under the assumption that ξ is bounded, and that the dependence of f in y is Lipschitz, Kobylanski was able to obtain existence, comparison and uniqueness. These results were later given a new approach by Tevzadze [77], and later on by Briand and Elie [12].

However, in the simple case of the quadratic coefficient $f(z) = \frac{1}{2}\gamma z^2$, the solution of the BSDE can be written explicitly and it is seen that it is not necessary to require that ξ be bounded, rather it should only be required to have some exponential integrability. Briand and Hu [13] and [14] obtained the first results of existence and uniqueness for quadratic BSDEs with unbounded (but exponentially integrable) terminal conditions.

This analytical setting was further studied by Delbaen, Hu and Richou [25] and [26], as well as Barrieu and El Karoui [4].

Another important analytical case is when f satisfies a monotonicy assumption in the y variable, that is to say there exists $\mu > 0$ such that

$$\big\langle y' - y \big| f(t, y', z) - f(t, y, z) \big\rangle \leq \mu |y' - y|^2 \ ,$$

for all t, y, y', z [1].

In many reaction-diffusion PDEs the nonlinearity f only depends on y, typically in a polynomial way, and this assumption applies. Pardoux [64] and Briand and Carmona [8] studied this case when f is Lipschitz in the variable z, uniformly in the others (i.e. ω, t and y), and the terminal condition is in L^p, $p \geq 2$, and showed that the BSDE is well-posed under these assumptions. Under the same assumption on the y-dependence, but with a quadratic growth in z and a bounded terminal condition, the well-posedness of the equation was shown by Briand, Lepeltier and San Martin [15].

All the results mentioned above concerning a quadratic dependence in z apply only to scalar BSDEs. When f is Lipchitz in z, whether Lipschitz or monotone in y, the well-posedness is obtained without difficulty for $\mathbb{R}^n$-valued BSDEs. But when f is quadratic in z, little is known for multidimensional BSDEs. A result in Tevzadze [77] guarantees that if f is purely quadratic in z (no first-order terms) and if the terminal condition is bounded and small enough, then there exists a solution. However, Frei and dos Reis [34] show some counter-examples of BSDEs for which there can be no solutions. See also Frei [33] and Delbaen and Tang [27].

1.1.2 BSDEs and PDEs.

The application of (F)BSDEs to PDEs is the guiding theme of this thesis. Each chapter contains a short presentation of the specific question it treats in this respect, where needed. However, we would like to recall here the general idea, which will be put to intensive use in Chapter 1 but pervades the whole thesis.

[1] We use the notation $\langle x|y\rangle$ for the scarlar product of two elements $x, y, \in \mathbb{R}^n$, and the notation $\langle x, y\rangle$ for the quadratic covariation of two semimartingales x and y.

We will be concerned only with the case of parabolic PDEs,

$$\begin{cases} v_t + \frac{1}{2} v_{xx} \cdot a + v_x \cdot b + f(t, x, v, v_x \sigma) = 0 \\ v(T, \cdot) = \Phi \ , \end{cases} \tag{1.1.4}$$

for $(t, x) \in D =]-\infty, T] \times \mathbb{R}^d$, for a fixed and given time $T > 0$. In the above equation, the unknown v function is from D to $\mathbb{R}^n$, the terminal condition $\Phi : \mathbb{R}^d \to \mathbb{R}^n$. The coefficients are $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $a : [0, T] \times \mathbb{R}^d \to S_d^+(\mathbb{R})$, the set of positive symmetric $d \times d$ matrices. So $v_x \cdot b = \sum_{i=1}^d \partial_i v \, b^i \in \mathbb{R}^n$ and $v_{xx} \cdot a = \sum_{i,j=1}^d \partial_{i,j} v \, a^{i,j} \in \mathbb{R}^n$. The matrix a is suposed to be given by $a = \sigma\sigma^*$ for some matrix σ (typically the unique square root of a in $S_d^+(\mathbb{R})$, but σ does not have to be symmetric, or even a square matrix, it could be a $m \times d$ matrix of rank d, for $m \geq d$).

This PDE is formally equivalent to the system of FBSDE

$$dX_t^{s,x} = b(t, X_t^{s,x})dt + \sigma(t, X_t^{s,x})dW_t \qquad \text{with } X_s^{s,x} = x \tag{1.1.5}$$

$$dY_t^{s,x} = -f(t, X_t^{s,x}, Y_t^{s,x}, Z_t^{s,x})dt + Z_t dW_t \qquad \text{with } Y_T^{s,x} = \Phi(X_T^{s,x}). \tag{1.1.6}$$

where $t \in [s, T]$, and $(s, x) \in D =]-\infty, T] \times \mathbb{R}^d$.

The connection is best understood as being what the method of characteristics gives for 2nd order, parabolic PDEs. To see this, fix $(s, x) \in D$ and let us look at how to determine $v(s, x)$. Consider the path $(\overline{X}_t^{s,x})_{t\in[s,T]}$ given by $\overline{X}_t^{s,x} = (t, X_t^{s,x})$, where $X_s^{s,x} = x$ and $dX_t^{s,x}$ is given by the SDE (1.1.5). We look at the value $Y_t^{s,x} = v(\overline{X}_t^{s,x}) = v(t, X_t^{s,x})$ along the path $\overline{X}^{s,x}$. If v is of class $C^{1,2}$, applying Itô's formula, we see that if we set $Z_t^{s,x} = (v_x\sigma)(t, X_t^{s,x})$, then $(Y_t^{s,x}, Z_t^{s,x})_{t\in[s,T]}$ is a solution to the BSDE (1.1.6) with terminal condition $Y_T^{s,x} = \Phi(X_T^{s,x})$.

In the particular case where $f = 0$, we know that $v(s, x) = Y_s^{s,x} = E\big[\Phi(X_T^{s,x})\big]$; this is the well-known Feynman–Kac formula. In the general case, the relation $v(s, x) = Y_s^{s,x}$ representing the PDE solution in terms of the FBSDE solution is called a nonlinear Feynman–Kac formula.

However, the interesting direction to establish is the converse of the one presented above. Starting from the solution to the BSDE, and setting $v(s, x) := Y_s^{s,x}$, one can show under reasonable assumptions that v so defined is a solution in some sense to the PDE (1.1.4).

The link between PDEs and FBSDEs first appeared in Peng [69], who proved that the classical solution to a PDE induces a solution to the corresponding BSDE (as shown above). In Pardoux and Peng [66] it was shown that if f and Φ are regular enough (of class C^3), then v defined by $v(s,x) := Y_s^{s,x}$ is a classical solution to (1.1.4). It was proved in Peng [70] and Pardoux and Peng [66] that under the standard assumptions that f is Lipschitz and Φ as well, $v(s,x) := Y_s^{s,x}$ defines a viscosity solution to the PDE. See Chapter 1 for more references on this topic.

Although each chapter contains its own introduction, we now give a very brief presentation of each of them, with an overview of the results they contain.

1.2 A representation formula between a class of multidimensional quadratic PDEs and the associated BSDEs, and applications.

In Chapter 2, we are concerned with the connection between PDEs and BSDEs in the case where the PDE is multidimensional, with a quadratic nonlinearity f of a special form.

Problem studied and motivation.

While it is always possible to connect formally a given PDE problem to its equivalent forward-backward stochastic problem, there is then the task of establishing this connection rigorously. In some cases, the difficulty comes from the nature of the problem studied, and the fact that the probabilistic counterpart has not yet been understood (for instance when reflected BSDEs were introduced, and were shown to be the counterpart of obstacle problems for PDEs). In some cases, the nature of the problem is known and understood, but the connection has not yet been established under the particular set of assumptions under consideration (for instance, when the connection was shown between PDEs and BSDEs for nonlinearities f which can be quadratic in z, and a bounded terminal condition : it was formally the well known parabolic PDE case, but the connection was previously proved only for Lipschiz f).

Once the connection between a certain PDE problem and its probabilistic forward-backward stochastic problem is proved, under a given set of assumptions on f and Φ (and b and σ), the situation is often not pushed further. A notable exception to this is when obtaining further representation theorems, which deepen the connection and leads to path-regularity theorems that are needed for analyzing the convergence of numerical methods for BSDEs (which is discussed in Chapter 4).

We are interested in establishing some sort of connection between a PDE problem and its BSDE equivalent, the difficullty in our case coming from the analytical assumptions. We would also like to use this connection in new ways, to derive PDE estimates using probabilitic techniques.

More precisely, we study multidimensional PDEs of the form

$$\begin{cases} v_t + \dfrac{1}{2}\Delta v + v_x\, f(t, v, v_x) + g(t, v, v_x) = 0 \\ v(T, \cdot) = \Phi \ , \end{cases}$$

for Lipschitz function f and g. In the above, $(t, x) \in [0, T] \times \mathbb{R}^d$, v is $\mathbb{R}^n$ valued, and f is $\mathbb{R}^d$-valued. The term $v_x\, f(t, v, v_x)$ is quadratic in v_x. Fixing $(0, x) \in [0, T] \times \mathbb{R}^d$, the equivalent FBSDE is

$$\begin{cases} X_s = x \\ dX_t = dW_t \\ dY_t = -\Big[Z_t f(t, Y_t, Z_t) + g(t, Y_t, Z_t)\Big]dt + Z_t dW_t \\ Y_T = \Phi(X_T). \end{cases}$$

Known results.

An abundant literature has been devoted to the links between PDEs and BSDEs, of which we mention only a few landmarks.

As mentionned above, the link between PDEs and BSDEs was worked out first by Peng [69, 70] and Pardoux and Peng [66]. They studied mainly the case of parabolic PDEs with Lipschitz functions f and Φ. It was shown how a classical solution to the PDE induces a solution to the BSDE. It was also shown that under extra regularity assumptions, the solution to a FBSDE induces a classical solution to the PDE, and that under only the Lipschitz assumption it induces a viscosity solution to the PDE.

We also note that Barles and Lesigne [3] studied the connection with PDE solutions in the sense of distributions.

Then, different problems were studied. For instance, Pardoux and Peng considered stochastic PDEs and their associated backward doubly stochastic differential equations [67]. Hu [39] and Pardoux and Zhang [68] considered problems with Neumann boundary conditions. Obstacle problems for PDEs were connected with BSDEs reflected on one obstacle (El Karoui, Kapoudjian, Pardoux, Peng and Quenez [30]) or two obstacles (Cvitanic and Karatzas [24]). The connection was also established under more general assumptions. Kobylanski [50] proved it for quadratic nonlinearities and bounded terminal conditions, while Briand and Hu [14] proved it for unbounded terminal conditions.

As far as multidimensional quadratic BSDEs are concerned, few positive results are known. A result in Tevzadze [77] ensures that if the nonlinearity is purely quadratic and the terminal condition small enough, there exists a solution. However examples from Frei and dos Reis [34] show that some simple such BSDEs can have no solution at all.

Results obtained.

Multidimensional quadratic BSDEs are little understood. In our case however, we can take advantage of the special structure of the nonlinearity. Using a change of measure, we can "sweep out" the quadratic term and be led to the Lipschitz BSDE

$$\begin{cases} dY_t = -\Big[g(t, Y_t, Z_t)\Big]dt + Z_t dB_t \\ Y_T = \Phi(X_T), \end{cases}$$

where B is again a Brownian motion. We can consider the change of measure for the whole FBSDE system, which leads to the now-coupled equation

$$\begin{cases} X_s = x \\ dX_t = f(t, Y_t, Z_t)dt + dB_t \\ dY_t = -\Big[g(t, Y_t, Z_t)\Big]dt + Z_t dB_t \\ Y_T = \Phi(X_T). \end{cases}$$

The advantage of the measure-changed equation is that is it now Lipschitz, and therefore more amenable to computations and estimates. However, the change of measure in the direction we performed it above depends on the solution of the BSDE of interest. In Chapter 2, we need to perform it in the other direction, which requires a fixed point problem to be solved first. Consequently, we establish the validity of the representation of v in terms of the solution to a Lipschitz BSDE (instead of the natural, quadratic one).

We then use this representation to establish the PDE estimate

$$\int_0^t P_s \left|\nabla u^i\right|^p (s,x)\, ds \leq c e^{pcT} \left(\|\Phi\|^p + |g(0,0)|^p T^p\right) \exp\left[\frac{p}{2(2-p)} t \sup_{|y|\leq K} |f(y)|^2\right]$$

for any $i \in \{1 \dots n\}$, $p \in [1,2[$, and where $(P_t)_{t\geq 0}$ is the heat semigroup and K is a constant made explicit in Chapter 2. For further use of BSDEs to study PDEs, we also mention the papers Hu and Qian [41], Hu, Qian and Zhang [42].

1.3 Well-posedness for reflected BSDEs driven by a continuous martingale.

In Chapter 3, we are concerned with the well-posedness of reflected BSDEs formulated in a very general in setting.

Problem studied and motivation.

"Reflected" BSDEs are, in fact, BSDEs subject to a constraint : the solution process Y is required to remain above a lower obtacle L. In order to achieve this, it is necessary to add to the usual dynamics $dY_s = -f_s ds + Z_s dW_s$ a "force" dK that drives Y upward. One wants that extra term to be minimal, so that K is only active to prevent Y from passing below the obstacle L. This optimality condition (known as the Skorohod condition) is often expressed as $\int_0^T 1_{\{Y_s > L_s\}} dK_s = 0$. So a reflected BSDE takes the following form :

$$\begin{cases} dY_s = -f(s, Y_s, Z_s)ds - dK_s + Z_s dW_s, \\ Y_T = \xi, \\ Y_t \geq L_t \text{ for all } t \in [0, T], \\ K \text{ is continuous, increasing, starts from 0 and } \displaystyle\int_0^T 1_{\{Y_s > L_s\}} dK_s = 0, \end{cases} \tag{1.3.1}$$

where the solution to be determined is now the triple (Y, Z, K).

In the context of PDEs, reflected BSDEs are the probabilistic counterpart to obstacle PDE problems (variational inequalities). That is, instead of the PDE (1.1.4), we now want v, essentially, to satisfy the PDE so long as v is strictly greater than an obstacle l and to always remain above that obstacle. More precisely, we want

$$\max\left\{ v_t + \frac{1}{2} v_{xx} \cdot a + v_x \cdot b + f(t, x, v, v_x\, \sigma)\ ,\ l - v \right\} = 0.$$

The backward stochastic problems associated with PDE problems, whether with initial condition (defined on the whole $\mathbb{R}^d$), with Dirichlet boundary condition or with Neumann boundary conditions, can all be studied with the same formalism if one considers a sufficiently general formulation of backward stochastic problems. To cover these cases, we need to allow that the terminal time T may be a stopping time and that the drift may be of the form $f(t, Y_t, Z_t)dt + g(t, Y_t)dA_t$ where A is an increasing process. We will, in fact, cover an even greater generality and study the well-posedness of the reflected BSDE

$$\begin{cases} dY_t = -dV(Y, N)_t - dK_t + dN_t \\ Y_T = \xi \\ Y \geq L \\ 1_{\{Y_t > L_t\}} dK_t = 0\ , \end{cases}$$

with drift given by

$$dV(Y, N)_t = f(t, Y_t, Z_t \sigma_t) dC_t + d\left\langle \nu, N^{\perp} \right\rangle + g_s d\left\langle N^{\perp} \right\rangle_t,$$

and where the martingale part N of Y has the decomposition $dN_t = Z_t dM_t + dN^{\perp}$ on the reference martingale M, with $N^{\perp}$ orthogonal to M (in the sense that $\langle M, N^{\perp} \rangle$). In this setting, C is a given increasing process and f, ν, g are the coefficients of the driver $dV(Y, N)_t$.

Known results.

Reflected BSDEs were introduced by El Karoui, Kapoudjian, Pardoux, Peng and Quenez in [30], where the well-posedness was studied under the assumptions that f is Lipschitz, the terminal condition is square-integrable and the lower obstacle a continuous square-integrable semimartingale. As far as the analytical assumptions on (f, ξ) are concerned, this corresponds naturally to the setting considered initially in Pardoux and Peng [65] or in El Karoui, Peng and Quenez [31].

As already mentioned, the well-posedness for BSDEs has been extended to a number of other cases.

Kobylanski obtained in [50] the well-posedness of BSDEs when f is quadratic in z, Lipschitz in y, and ξ is bounded. Lepeltier and San Martin [55] relaxed the second assumption and allowed f to have slightly superlinear growth in y. Using essentially the same techniques, Kobylanski, Lepeltier, Quenez and Torrès [51] were then able to prove the analogue results for RBSDEs when the obstacle L is bounded.

Reflected BSDEs when ξ is bounded, f quadratic in z but monotone with arbitrary growth in y were studied by Xu [79], after Briand, Lepeltier and San Martin [15] obtained the well-posedness for BSDEs under these assumptions.

Under the assumptions that f is quadratic in z and Lipschitz in y, Lepeltier and Xu [56] treated the case of reflected BSDEs with unbounded terminal condition ξ and bounded obstacle L, while Bayraktar and Yao [5] removed the condition that L be bounded. These results are based on the progress made on quadratic BSDEs with unbounded terminal conditions by Briand and Hu [13, 14] and Delbaen, Hu and Richou [25] (on which more progress was done afterwards, see Delbaen, Hu and Richou [26] and Barrieu and El Karoui [4]).

However, as already mentioned, the case of quadratic BSDEs with bounded terminal conditions was recently considerably simplified. Tevzadze [77] and Briand and Elie [12] gave simpler approaches, and these ideas have so far not been tested against reflected BSDEs (but for our study in Chapter 3, using the technique from [77]).

The above-mentioned works [30], [51], [79], [56] and [5] dealt with reflected BSDEs in a Brownian setting. However, BSDEs have been studied in a general martingale setting (see El Karoui and Huang [29], Tevzadze [77], Morlais [62], Barrieu and El Karoui [4]), and in a general filtered probability space in Cohen and Elliott [20].

Regarding the stability of the solution with respect to changes in the terminal condition, global Lipschitz estimates for the martingale part in $\mathcal{H}^p$ were obtained in Briand, Delyon, Hu, Pardoux and Stoica [11], Briand and Confortola [9] and Ankirchner, Imkeller and Dos Reis [2]. Kazi-Tani, Possamai and Zhou [48] provide a global $\frac{1}{2}$-Hölder estimate in the smaller space BMO, for quadratic BSDEs with bounded terminal conditions (and with jumps, but obviously their result and technique would hold without the jumps).

Results obtained.

We obtain the well-posedness results for reflected BSDEs driven by a continuous martingale, as written above, under the assumptions that the terminal condition ξ is bounded and that f is quadratic in z, while the dependence in y can be Lipschitz, slightly superlinear, or monotone with arbitrary growth. This extends the results obtained in a Brownian setting in [51] (slightly superlinear growth in y) and [79] (monotonicity and arbitrary growth in y), under the same assumptions on the z dependence and ξ.

While it would be possible to generalize to the continuous martingale setting some of the results known for reflected BSDEs in a Brownian setting by using the same techniques which were used to prove them in the first place, we generally use different ones, with the double aim of providing a treatment as self-contained as possible (not relying on BSDE results for instance) and of providing more morally satisfying proofs.

First, we obtain the comparison theorem in our setting (from which uniqueness follows). To prove it, we adapt to reflected BSDEs (whether set in a Brownian setting or not) the linearization and BMO argument from Hu, Imkeller and Müller [40], rather than relying on an optimal stopping representation and invoking the comparison theorem for BSDEs, as was done in [51]. Prior to doing this, we show, for reflected BSDEs, that for a bounded solution, the martingale part is in BMO (which had not been used even for Brownian reflected BSDEs). In doing so, we also obtain a BMO-style estimate for the increasing process K.

We obtain the special comparison theorem for reflected BSDE, which allows to compare the increasing processes of two solutions. In a Brownian setting this theorem was proved in various cases in [37], [71], [54], [51], via the penalization approach to BSDEs, using the comparison theorem for BSDEs and identifying the quantitites which, at the limit, become the increasing processes. We provide a different proof, more intrinsic, and which works under minimal assumptions.

For the existence result, we adapt to the context of reflected BSDEs the technique introduced by Tevzadze [77]. There is a difficulty caused by the fact that the underlying problem for reflected BSDEs is not linear, and as a consequence the sum of solutions to RBSDEs is in general not the solution to the RBSDE one would want. However, by reinterpreting the technique from [77] as a pertubation procedure, we are able to identify the type of equation that perturbations satisfy and conclude to existence. We also note that the technique, initially used for a Lipschitz dependence in y is generalizable, and in our context we conclude to existence of a solution when f can have slightly superlinear growth in y or be monotone with arbitrary growth.

Finally, we obtain a local regularity result for the martingale part in BMO. That is, given a terminal condition ξ, we show that there exists $R > 0$ and $c(\xi) > 0$ such that, for any other bounded terminal conditions ξ', ξ'' at distance less than or equal to R from ξ, if we denote by $S' = (Y', N', K')$ and $S'' = (Y'', N'', K'')$ the associated solutions, we have

$$\|N'' - N'\|_{BMO} \leq c(\xi)\|\xi'' - \xi'\|_\infty \, .$$

This improves on the $\frac{1}{2}$-Hölder regularity proved in [48], at the expense of holding only for small perturbations.

1.4 Time-discretization of monotone FBSDEs with polynomial growth.

In Chapter 4, we are concerned with the time discretization of FBSDEs under some analytical assumptions relevant for reaction-diffusion PDEs. Indeed, many such equations have nonlinearities which are typically polynomial in v (that is to say y, in the probabilistic counterpart).

Problem studied and motivation.

We have explained how the solution v to the PDE (1.1.4) can be represented as $v(s,x) = Y_s^{s,x}$, and more generally for $t \in [s,T]$

$$v(s, X_t^{s,x}) = Y_t^{s,x} \qquad \text{and} \qquad (v_x\sigma)(s, X_t^{s,x}) = Z_t^{s,x} \ ,$$

where $(X_t^{s,x}, Y_t^{s,x}, Z_t^{s,x})_{t\geq s}$ is the solution to the FBSDE (1.1.5)–(1.1.6). This implies that every numerical method to compute the solution (X,Y,Z) to the FBSDE provides a probabilistic method to approximate the solution v to the PDE.

Let us look more closely at the FBSDE (1.1.5)–(1.1.6). Since the SDE (1.1.5) for the forward component X is not coupled with the backward component (Y,Z), known time-discretization techniques for SDEs are readily available to approximate numerically the process X, with known and proven orders of convergence (see for instance Kloeden and Platen [49]). Considering a partition $\pi^N : 0 = t_0 < t_1 < \ldots < t_N = T$ of the time interval $[0,T]$, with $N+1$ points and a mesh $|\pi^N|$ going to 0 as $N \longrightarrow +\infty$, one can construct a numerically computable $(X_i)_{i=0\ldots N}$ close to $(X_t)_{t\in[0,T]}$ in some relevant sense. More specifically, if

$$\max_{i=0,\ldots,N} E\big[\,|X_{t_i} - X_i|^p\big]^{\frac{1}{p}} \leq c|\pi|^\gamma,$$

for some $\gamma > 0$, we say that there is convergence of order (at least) γ.

The study of numerical methods to compute an approximation of the BSDE solution (Y,Z) is more recent and the subject of a growing effort. As a first step, one needs to obtain a time-discretization for the BSDE (1.1.6). For instance, in Chapter 4 we study the following family of θ-schemes, where the parameter $\theta \in [0,1]$ characterizes the degree of implicitness. Define first $Y_N = \Phi(X_N)$ and $Z_N = 0$, and then for $i = N-1,\ldots,0$ define

$$\begin{aligned} Y_i &= E\Big(Y_{i+1} + \theta f(t_i, X_i, Y_i, Z_i)h_{i+1} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1}\Big|\mathcal{F}_i\Big) \\ Z_i &= E\bigg(\Big\{Y_{i+1} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1}\Big\}\frac{\Delta W_{i+1}}{h_{i+1}}\bigg|\mathcal{F}_i\bigg) \end{aligned}$$

where $h_{i+1} = t_{i+1} - t_i$ and $\Delta W_{i+1} = W_{t_{i+1}} - W_{t_i}$. In the case $\theta = 0$ this corresponds to the explicit backward Euler scheme and in the case $\theta = 1$ this is the implicit backward Euler scheme. The first question is whether there is convergence of $(Y_i, Z_i)_{i=0\dots N}$ to $(Y, Z)_{t\in[0,T]}$ in some sense. More specifically, in chaper 4 we want to obtain the error estimate

$$\mathrm{ERR}_\pi(Y,Z) := \Big(\max_{i=0,\dots,N} E\big[|Y_{t_i} - Y_i|^2 \big] + \sum_{i=0}^{N-1} E\big[|\bar{Z}_{t_i} - Z_i|^2 \big] h_{i+1} \Big)^{\frac{1}{2}} \le c|\pi|^\gamma,$$

for some $\gamma > 0$, where $\bar{Z}_{t_i} = \frac{1}{h_{i+1}} E\big(\int_{t_i}^{t_{i+1}} Z_u du \big| \mathcal{F}_i \big)$.

When the time-discretization converges, a second step is then to approximate numerically the conditional expectations used to define Y_i and Z_i, in order to obtain a fully implementable numerical scheme. However, the error created by the approximation of the conditional expectations will not be studied in this thesis and we refer the reader to Gobet and Turkedjiev [35] for these questions.

The study of the convergence of the above discretization to the solution to the BSDE requires first to obtain a finer theoretical result for (non-approximated) BSDEs known as path-regularity. This states that the trajectories of Y and Z are somehow continuous in some $\mathcal{S}^p \times \mathcal{H}^p$ sense. More precisely, it says that for $|\pi|$ small enough,

$$E\Big[\sup_i \sup_{t_i \le t \le t_{i+1}} |Y_t - Y_{t_i}|^p \Big] + E\Big[\sup_i \sup_{t_i \le t \le t_{i+1}} |Y_t - Y_{t_{i+1}}|^p \Big] \le c|\pi|^{\frac{p}{2}} \qquad \text{and}$$

$$E\Big[\Big(\sum_i \int_{t_i}^{t_{i+1}} |Z_t - Z_{t_i}|^2 dt \Big)^p \Big] + E\Big[\Big(\sum_i \int_{t_i}^{t_{i+1}} |Z_t - Z_{t_{i+1}}|^2 dt \Big)^p \Big] \le c|\pi|^{\frac{p}{2}}$$

Known results.

In the case where f and g are Lipchitz, this path-regularity result was obtained by Zhang [81] for $p = 2$. Once this was acquired, the road was open to study the convergence of the time-discretization for Lipschitz BSDEs, and this was done by Zhang [82] and Bouchard and Touzi [7] for slightly different schemes.

Different time-discretization schemes have then been considered, aiming in particular at higher order of approximation. For instance, Crisan and Manolarakis [22] studied a mixed explicit-implicit scheme (similar in spirit to the scheme above with $\theta = \frac{1}{2}$) and showed that it was of order 2. More recently, adapting ODE methods for the time-discretization, Chassagneux and Crisan [18] introduced for BSDEs the family of Runge-Kutta schemes, while Chassagneux [17] introduced linear multi-step methods, providing further higher order schemes.

A few papers have also studied the numerical approximation of solutions to quadratic BSDEs with bounded terminal condition. Mainly, the idea was to approximate the quadratic driver by a Lipschitz one (Imkeller and dos Reis [46]) or to approximate the bounded terminal condition Φ by a Lipschitz one (Richou [72], see also [73]), which leads to dealing with a Lipschitz BSDE, for which error estimates for the time-discretization are already known. More recently, Chassagneux and Richou [19] have proposed a more straightforward approach.

Results obtained.

We analyse the time-discretization in the case where f is monotone and has polynomial growth in y, as well as Lipschitz dependence in z, and Φ is Lipschitz.

More precisely, we first prove the path-regularity theorem in that setting (together with the differentiability results needed for it), and we prove this for all $p \geq 2$. Having obtained the path-regularity, we study the convergence of the θ-scheme.

We prove that provided that $\theta \geq \frac{1}{2}$ (the implicit component dominates), the scheme converges. However, we give an example in the case $\theta = 0$ (the fully explicit scheme) where the approximation given by the scheme explodes. This indicates that we should expect, in general, that the scheme diverges for $\theta < \frac{1}{2}$. This is the first instance in the literature related to the time-discretization of BSDEs where we observe a different behaviour for the explicit and implicit scheme. It is indeed a rule of thumb, known for the time-discretization of ODEs and SDEs, that implicit schemes are more stable.

However, while the explicit scheme is less stable, and tends to diverge in this setting, it is more easy to implement and computationally faster, seeing as it does not require the extra layer of computation needed to compute Y_i in the implicit scheme. To fix this explosion issue, we propose a tamed version of the explicit scheme for which we can recover convergence.

In all these cases, the convergence occurs with order $\gamma = \frac{1}{2}$. Thanks to the way we study the convergence (and in particular what we call the Fundamental Lemma, which allows to separate the ingredients required for convergence), we can easily show along the way that in the case $\theta = \frac{1}{2}$, the scheme has a higher order of convergence ($\gamma = \frac{7}{4}$) under the assumption that f does not depend on z (a minor assumption if one has in mind the application to reaction-diffusion PDEs with nonlinearities which are polynomial in v).

Chapter 2

A representation formula between a class of multidimensional quadratic PDEs and the associated BSDEs, and applications.

2.1 Introduction

2.1.1 Motivation

In this chapter, we are interested in the probabilistic representation of solutions to the multidimensional PDE

$$\begin{cases} \dfrac{\partial u^i}{\partial t} + \dfrac{1}{2}\Delta u^i + \displaystyle\sum_{j=1}^{d} u^j \dfrac{\partial u^i}{\partial x^j} = 0 \qquad \text{for all } i \in \{1, \ldots, d\} \text{ and for } (t,x) \in [0,T[\times\mathbb{R}^d \\ u(T,\cdot) = h \ . \end{cases} \tag{2.1.1}$$

The quadratic nonlinearity $F(u, \nabla u) = \nabla u.u = (u \cdot \nabla)u$ is typical of equations like the Euler equation and the Navier-Stokes equation in fluid dynamics. In fact, (2.1.1) is the Navier-Stokes equation in which one has dropped the pressure term and the divergence-free condition, keeping only the nonlinear convection term $(u \cdot \nabla)u$. This kind of PDEs has been used as simplified models for phenomena such as turbulence flows.

Due to the special structure of the system (2.1.1), the maximum principle applies to $|u(x,t)|^2$, so a bounded solution exists as long as the terminal data h (note that we orient the time axis in the opposite direction of the diffusion time) is regular and bounded, which makes a distinctive difference from the Navier-Stokes equations. According to Theorem 7.1 on page 596 in [53], if the initial data h is smooth and bounded with bounded derivatives, then a bounded, smooth solution u to the initial/terminal value problem (2.1.1) exists for all time. Our main interest is to establish a convenient probabilistic representation for the solution u, by applying Girsanov's transform to the corresponding BSDEs.

Indeed, we know from Peng [69] that if we are given a Brownian motion $\tilde{B}$, and consider the processes $Y_t^u = u(t, x + \tilde{B}_t)$ and $Z_t^u = \nabla u(t, x + \tilde{B}_t)$, an application of Itô's formula shows that (Y^u, Z^u) solves the BSDE

$$\begin{cases} dY_t = -Z_t Y_t dt + Z_t d\tilde{B}_t \\ Y_T = h(x + \tilde{B}_T) = \xi \ . \end{cases}$$

2.1.2 Literature review

However the above equation is a multidimensional quadratic BSDE. While scalar BSDEs with quadratic nonlinearity are becoming well-understood, whether for bounded terminal conditions (see Kobylanski [50], Tevzadze [77], Briand and Elie [12] for the well-posedness results) or for unbounded terminal conditions (see Briand and Hu [13, 14]), Delbaen, Hu and Richou [25, 26] and Barrieu and El Karoui [4]), very little is known about multidimensional quadratic BSDEs. A result in Tevzadze [77] guarantees the existence of a bounded solution if the nonlinearity F contains only quadratic terms and if the terminal data ξ is small enough in L^∞ (an assumption that we do not make here). However, some counter-examples by Frei and dos Reis [34] show that some simple multidimensional quadratic BSDEs cannot have a solution, even for small terminal condition. Beyond that, there seem to be essentially no well-posedness results for these BSDEs.

2.1.3 Overview of the content of this chapter.

Our eventual goal is to study the PDE solution u through its probabilistic representation (Y^u, Z^u), in particular to obtain estimates on u in terms of the heat semigroup. For this we need to have estimates and a minimum of well-posedness theory for the corresponding BSDE. But, as explained above, there are no general results for multidimensional quadratic BSDEs.

Thanks to the special structure of the nonlinearity, we are able to swap the above BSDE for a Lipschitz BSDEs, by means of a change of measure. Consequently, we obtain a representation for u in terms of the solution (Y, Z) to a Lipschitz BSDE. These BSDEs are very well understood, even when multidimensional, and therefore the representation of u in terms of (Y, Z) allows us to carry the desired study of u via probabilistic techniques.

Our contribution is the following :

- We provide a probabilistic representation for the solution to a multidimensional PDE with a quadratic nonlinearity of the form of (2.1.1) in terms of the solution to a Lipschitz BSDE, more amenable to computations
- We show how, in general, such probabilistic representations allow to obtain estimates for the PDE solution, in particular in terms of the heat semigroup.
- We obtain an existence result and a partial uniqueness result for the corresponding class of multidimensional quadratic BSDEs.

2.2 Setting, notation and some BSDE results.

2.2.1 Setting and notation.

More generally than (2.1.1), we will study the class of equations

$$\begin{cases} \dfrac{\partial u^i}{\partial t} + \dfrac{1}{2}\Delta u^i + g^i(u, \nabla u) + \displaystyle\sum_{j=1}^{d} f^j(u, \nabla u)\dfrac{\partial u^i}{\partial x^j} = 0 \\ u(T, \cdot) = h, \end{cases} \tag{2.2.1}$$

for $i \in \{1, \dots, m\}$ and $(t, x) \in [0, T[\times \mathbb{R}^d$, where the nonlinearity functions $f : \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^d$ and $g : \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$, while the terminal condition $h : \mathbb{R}^d \to \mathbb{R}^m$.

Here m and d are not necessarily equal.

In the whole chapter, we make the following assumptions on the data h, and (f, g) of the problem :

- f and g are Lipschitz, with Lipschitz constants L_f and L_g
- h is bounded and Lipschitz (with constant L_h)

Under these conditions, it is known (see for instance theorem 7.1, p596, in Ladyzhenskaya *et al.* [53]) that there exists a unique bounded classical solution u to (2.2.1), and that ∇u is also bounded.

$T > 0$ being fixed, we consider the space $\Omega = C([0, T], \mathbb{R}^d)$ of continuous paths from $[0, T]$ to $\mathbb{R}^d$, equipped with the canonical filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$. We also consider a probability measure $\mathbb{P}$ and a process B such that $(\Omega, \mathcal{F}, \mathbb{P}, B)$ is a Brownian setting (i.e. B is a d-dimensional Brownian motion under $\mathbb{P}$, and $\mathcal{F}$ is the augmented filtration of B).

2.2.2 Recall of some BSDE results.

We will study the solution u to the above PDE by means of BSDEs. We recall some known results on these equations that we will use in the sequel.

For any random variable $\xi \in L^2(\mathcal{F}_T)$, since g is Lipschitz, we know from Pardoux and Peng [65] that the BSDE

$$\begin{cases} dY_t = -g(Y_t, Z_t)dt + Z_t dB_t \\ Y_T = \xi \end{cases} \tag{2.2.2}$$

admits a unique solution (Y, Z) in $\mathcal{S}^2 \times \mathcal{H}^2$. We recall here that

- $\mathcal{S}^2$ is the space of continuous, progressively measurable processes Y's such that the norm
$$\|Y\|^2_{\mathcal{S}^2} = \boldsymbol{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < +\infty \ ,$$
- $\mathcal{H}^2$ is the space of progressively measurable processes Z's such that the norm
$$\|Z\|^2_{\mathcal{H}^2} = \boldsymbol{E}\left[\int_0^T |Z_t|^2 dt\right] < +\infty \ .$$

When needed, we denote the solution to (2.2.2) by $(Y(\xi), Z(\xi))$.

We have the following known estimate (see El Karoui *et al.* [31]). There exists a constant $c > 0$ such that, for any random variables $\xi, \eta \in L^2$,

$$\|Y(\xi) - Y(\eta)\|_{\mathcal{S}^2}^2 + \|Z(\xi) - Z(\eta)\|_{\mathcal{H}^2}^2 \leq ce^{cT}\boldsymbol{E}\big[|\xi - \eta|^2\big] \tag{2.2.3}$$

Finally, let us also introduce some classical spaces of processes that will be used in the sequel.

- BMO is the space of martingales M such that the norm

  $$\|M\|_{BMO}^2 = \sup_{t \in \mathcal{T}} \left\| E\bigg(\langle M \rangle_T - \langle M \rangle_t \bigg| \mathcal{F}_t \bigg) \right\|_\infty < +\infty \ ,$$

  where $\mathcal{T}$ is the set of stopping times t such that $0 \leq t \leq T$. By extension, we say that a progressively measurable process $Z \in \mathcal{H}^2$ is in BMO if $M = \int ZdB \in BMO$.

- $\mathcal{H}^\infty$ is the space of bounded progressively measurable processes Z.

2.3 Representation formula.

We want to obtain a convenient probabilistic representation for the solution u to the multidimensional PDE (2.2.1).

We consider, for a random variable $\xi \in L^2(\Omega, \mathcal{F}_T)$ to be determined later, the solution $(Y(\xi), Z(\xi))_{t \in [0,T]}$ to the Lipschitz BSDE (2.2.2). We also define for $t \in [0, T]$ the process

$$\tilde{B}_t = \tilde{B}(\xi)_t = B_t + \int_0^t f(Y(\xi)_s, Z(\xi)_s) ds \ .$$

Theorem 2.3.1. *If $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ is a solution to the equation*

$$\xi = h(x + \tilde{B}(\xi)_T)$$

then we have the representation

$$u(t, x + \tilde{B}(\xi)_t) = Y(\xi)_t$$

and

$$\nabla u(t, x + \tilde{B}(\xi)_t) = Z(\xi)_t$$

for all $t \in [0, T]$, almost surely.

Proof. Since ξ is a solution of this fixed point equation and h is bounded, ξ is bounded. Proposition 2.3.3 below ensures that $(Y(\xi), Z(\xi)) \in \mathcal{S}^\infty \times BMO$. Then, since f is Lipschitz, it has at most linear growth and therefore $N(\xi)$ is a BMO martingale. Consequently, the exponential $\mathcal{E}(N(\xi))$, where

$$N(\xi)_t = -\int_0^t f(Y(\xi)_s, Z(\xi)_s) dB_s \ ,$$

is a martingale on $[0, T]$, so we can define a measure $\mathbb{Q}$ by $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = \mathcal{E}(N(\xi))_t$. By Girsanov's theorem (and Lévy's theorem), $\tilde{B}(\xi)$ is then a Brownian motion under $\mathbb{Q}$.

On the one hand, by construction of $\tilde{B}(\xi)$, $(Y(\xi), Z(\xi))$ is a solution to

$$dY_t = -\Big[\sum_{j=1}^d Z_t^j f^j(Y_t, Z_t) + g(t, Y_t, Z_t)\Big] dt + Z_t d\tilde{B}_t$$

with terminal condition $Y_T = \xi$. On the other hand, defining $Y_t^u = u(t, x + \tilde{B}(\xi)_t)$ and $Z_t^u = \nabla u(t, x + \tilde{B}(\xi)_t)$, it follows from Itô's formula that (Y^u, Z^u) solves the above BSDE with terminal condition $Y_T^u = h(x + \tilde{B}(\xi)_T)$. Since ξ satisfies $\xi = h(x + \tilde{B}(\xi)_T)$, $\big(Y(\xi), Z(\xi)\big)$ and (Y^u, Z^u) are solutions to the same BSDE, and because Z^u is bounded, they are equal (see proposition 2.3.2 below), which provides the desired representation. $\square$

The technique used above shows that, for any bounded ξ, we can construct a weak solution $(Y, Z, W, \mathbb{R})$ to the BSDE

$$\begin{aligned} dY_t &= -\Big[g(t, Y_t, Z_t) + \sum_{j=1}^d Z_t^j f^j(Y_t, Z_t)\Big] dt + Z_t dW_t \\ Y_T &= \xi \ , \end{aligned} \tag{2.3.1}$$

Indeed in the proof we showed that $(Y, Z, W, \mathbb{R}) := (Y(\xi), Z(\xi), \tilde{B}, \mathbb{Q})$ was such a solution.

There is no general uniqueness result for such a multidimensional quadratic BSDEs.

However we can prove the following particular result.

Proposition 2.3.2. *If $(Y, Z) \in \mathcal{S}^\infty \times BMO$ and $(Y', Z') \in \mathcal{S}^\infty \times \mathcal{H}^\infty$ are two solutions of the BSDE (2.3.1) (with bounded ξ, necessarily), then $(Y, Z) = (Y', Z')$.*

Proof. Define $\Delta Y = Y' - Y$ and $\Delta Z = Z' - Z$. We have $\Delta Y_T = 0$ and

$$\begin{aligned} d\Delta Y_t &= -\Big[g(Y_t', Z_t') - g(Y_t, Z_t) + Z_t' f(Y_t', Z_t') - Z_t f(Y_t, Z_t)\Big]dt + \Delta Z_t dW_t \\ &= -\Big[g(Y_t', Z_t') - g(Y_t, Z_t) + Z_t'\big(f(Y_t', Z_t') - f(Y_t, Z_t)\big) + \Delta Z_t f(Y_t, Z_t)\Big]dt \\ &\qquad + \Delta Z_t dW_t\ . \end{aligned}$$

Now, since $(Y, Z) \in \mathcal{S}^\infty \times BMO$ and since f has at most linear growth, $\int f(Y, Z)dW$ is a BMO martingale so we can define the measure $\mathbb{R}$ by $\frac{d\mathbb{R}}{d\mathbb{P}} = \mathcal{E}(\int_0^\cdot f(Y_s, Z_s)dW_s)_t$ on $\mathcal{F}_t$. Then $\tilde{W} = W - \int_0^\cdot f(Y_s, Z_s)ds$ is a $\mathbb{R}$-Brownian motion, and we have

$$d\Delta Y_t = -\Big[G(\Delta Y_t, \Delta Z_t)\Big]dt + \Delta Z_t d\tilde{W}_t,$$

with

$$\begin{aligned} G(\Delta Y_t, \Delta Z_t) &= g(Y_t', Z_t') - g(Y_t, Z_t) + Z_t'\big(f(Y_t', Z_t') - f(Y_t, Z_t)\big) \\ &= g(Y_t + \Delta Y_t, Z_t + \Delta Z_t) - g(Y_t, Z_t) \\ &\qquad + Z_t'\Big(f(Y_t + \Delta Y_t, Z_t + \Delta Z_t) - f(Y_t, Z_t)\Big)\ . \end{aligned}$$

Since g and f are Lipschitz and Z' is bounded, G is Lipschitz, and then uniqueness for Lipschitz BSDEs guarantees that $(\Delta Y, \Delta Z) = (0, 0)$. □

The uniqueness result that we proved above applies in our context, in the proof of theorem 2.3.1. Indeed, we know that u is bounded with bounded derivative, so (Y^u, Z^u) is in $\mathcal{S}^\infty \times \mathcal{H}^\infty$. Also, since $\xi = h(x + \tilde{B}(\xi)_T)$ with h bounded, ξ is bounded and the estimate below ensures that the solution $(Y(\xi), Z(\xi))$ to the Lipschitz BSDE (2.2.2) is in $\mathcal{S}^\infty \times BMO$.

Proposition 2.3.3. *For any $\xi \in L^\infty$, for any $g : \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ Lipschitz, and any $(Y, Z) \in \mathcal{S}^2 \times \mathcal{H}^2$ solving the BSDE (2.2.2), we have,*

$$\|Y\|_{\mathcal{S}^\infty}^2 + \|Z\|_{BMO}^2 \leq ce^{cT}\left[\|\xi\|_\infty^2 + |g(0,0)|^2 T\right]$$

where c is a constant > 0 depending only on the Lipschitz constant of g.

Proof. Standard computations, using Itô's formula for $|Y_t|^2$, lead to

$$|Y_t|^2 + \Big(1 - \frac{1}{\nu}\Big) \int_t^T |Z_s|^2 ds$$
$$\leq |\xi| + \int_t^T |g(0,0)|^2 ds + \int_t^T (1 + 2L_g + \nu L_g^2)|Y_s|^2 ds - \int_t^T 2Y_s g(Y_s, Z_s) dW_s,$$

for any $\nu > 1$. Denoting by $b = 1 + 2L_g + \nu L_g^2$, $m(t) = \int_t^T 2Y_s g(Y_s, Z_s) dW_s$ and $A(t) = |\xi| + \int_t^T |g(0,0)|^2 ds$, Gronwall's lemma implies that

$$|Y_t|^2 + \Big(1 - \frac{1}{\nu}\Big) \int_t^T |Z_s|^2 ds \leq A(t) - m(t) + \int_t^T e^{b(u-t)} b\big(A(u) - m(u)\big) du \ .$$

Taking $\boldsymbol{E}\big(\,\cdot\,|\mathcal{F}_t\big)$, noting that $\boldsymbol{E}\big(m(u)|\mathcal{F}_t\big) = 0$ for $u \geq t$ (because $\int_0^{\cdot} 2Y_s g(Y_s, Z_s) dW_s$ is a martingale, since $Y \in S^2$, $Z \in H^2$ and g has at most linear growth), we finally obtain, for $\nu = 2$,

$$|Y_t|^2 + \frac{1}{2}\boldsymbol{E}\bigg(\int_t^T |Z_s|^2 ds \bigg| \mathcal{F}_t \bigg) \leq e^{b(T-t)} \Big(\|\xi\|_\infty^2 + |g(0,0)|^2\, T \Big) \ .$$

So the conclusion follows with $c = \max(1 + 2L_g + 2L_g^2, 3)$. □

By Theorem 2.3.1, in order to provide a convenient probabilistic representation for (2.2.1), the problem is reduced to solving the functional equation

$$\xi = h\left(x + B_T - \int_0^T f(Y(\xi)_s, Z(\xi)_s) ds \right) := \phi(\xi). \tag{2.3.2}$$

Proposition 2.3.4. *There exists a (unique) $\xi \in L^2(\mathcal{F}_T)$ satisfying (2.3.2)*

Proof. The proof is carried in two steps. In the first one, we solve the problem when T is small. In the second one, we split the time interval into small intervals and "patch" the solutions.

Step 1. In this step, we show that there exists a solution if T is sufficiently small that

$$2c\, e^{cT}\, T(T+1) L_h L_f < 1 \ ;$$

The constant c is that from estimate (2.2.3), and only depends on the problem parameter g (not on ξ). For ξ and η in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$,

$$
\begin{aligned}
&|\phi(\xi) - \phi(\eta)| \\
&\quad \leq \left| h\left(x + B_T - \int_0^T f(Y(\xi)_s, Z(\xi)_s)ds\right) - h\left(x + B_T - \int_0^T f(Y(\eta)_s, Z(\eta)_s)ds\right)\right| \\
&\quad \leq L_h \int_0^T |f(Y(\xi)_s, Z(\xi)_s) - f(Y(\eta)_s, Z(\eta)_s)|ds \\
&\quad \leq L_h L_f \int_0^T |Y(\xi)_s - Y(\eta)_s| + |Z(\xi)_s - Z(\eta)_s|ds
\end{aligned}
$$

and therefore

$$
\begin{aligned}
\boldsymbol{E}\left[|\phi(\xi) - \phi(\eta)|^2\right] &\leq L_h^2 L_f^2 \boldsymbol{E}\left[\left(\int_0^T |Y(\xi)_s - Y(\eta)_s| + |Z(\xi)_s - Z(\eta)_s|ds\right)^2\right] \\
&\leq 2L_h^2 L_f^2 T \boldsymbol{E}\left[\int_0^T |Y(\xi)_s - Y(\eta)_s|^2 + |Z(\xi)_s - Z(\eta)_s|^2 ds\right].
\end{aligned}
$$

Using the estimate (2.2.3) we have

$$
\begin{aligned}
\boldsymbol{E}\left[|\phi(\xi) - \phi(\eta)|^2\right] &\leq 2L_h^2 L_f^2 T \boldsymbol{E}\left[T \sup_{0\leq s\leq T} |Y(\xi)_s - Y(\eta)_s|^2 + \int_0^T |Z(\xi)_s - Z(\eta)_s|^2 ds\right] \\
&\leq 2ce^{cT} L_h^2 L_f^2 T(T+1)\boldsymbol{E}[|\xi - \eta|^2]
\end{aligned}
$$

and the claim follows from a simple application of the fixed point theorem.

Step 2. T is now arbitrary. Since it is known that u has bounded derivative, let L_u be a common Lipschitz constant for all the $u(t, \cdot)$, $t \in [0, T]$. For N sufficiently big, $\tau := T/N$ satisfies

$$
2c\, e^{c\tau}\, \tau(\tau + 1) L_u L_f < 1 \; .
$$

We split the time interval $[0, T]$ into time intervals $[i\tau, (i+1)\tau]$ for $i = 0 \ldots N-1$.

We first work on the time interval $[0, \tau]$ (corresponding to $i = 0$). By what was done in *Step 1*, but now with $u(\tau, \cdot)$ in place of h, it is possible to find $\xi = \xi^0$ in $L^2(\mathcal{F}_\tau)$ solving the fixed point equation $\xi^0 = u(\tau, x + \tilde{B}(\xi^0)_\tau)$, where $\tilde{B}(\xi^0)_t = B_t -$

$\int_0^t f(Y_s(\xi^0), Z_s(\xi^0))ds$ for $t \in [0, \tau]$. We therefore have for $t \in [0, \tau]$ the representation

$$
\begin{aligned}
u(t, x + \tilde{B}(\xi^0)_t) &= Y(\xi^0)_t \\
\nabla u(t, x + \tilde{B}(\xi^0)_t) &= Z(\xi^0)_t \ .
\end{aligned}
$$

We then move on to the interval $[\tau, 2\tau]$ (corresponding to $i = 1$). Again, repeating *Step 1*, but now with the function $u(2\tau, \cdot)$ instead of h, and x replaced by $x + \tilde{B}(\xi^0)_\tau$, we find $\xi = \xi^1$ in $L^2(\mathcal{F}_{2\tau})$ such that $\xi^1 = u(2\tau, x + \tilde{B}(\xi^1)_{2\tau})$, where $\tilde{B}(\xi^1)_t = \tilde{B}(\xi^0)_\tau + (B_t - B_\tau) - \int_\tau^t f(Y_s(\xi^1), Z_s(\xi^1))ds$, for $t \in [\tau, 2\tau]$. We therefore have the representation, for $[\tau, 2\tau]$,

$$
\begin{aligned}
u(t, x + \tilde{B}(\xi^1)_t) &= Y(\xi^1)_t \\
\nabla u(t, x + \tilde{B}(\xi^1)_t) &= Z(\xi^1)_t \ .
\end{aligned}
$$

And we go on until the N^{th} step, for $i = N - 1$, where we are on the time interal $[T - \tau, T]$. We obtain the existence of a random variable ξ^{N-1} in $L^2(\mathcal{F}_T)$ such that $\xi^{N-1} = u(0, x + \tilde{B}(\xi^{N-1})_T) = h(x + \tilde{B}(\xi^{N-1})_T)$, where $\tilde{B}(\xi^{N-1})_t = \tilde{B}(\xi^{N-2})_{(N-1)\tau} + (B_t - B_{(N-1)\tau}) - \int_{(N-1)\tau}^t f(Y_s(\xi^{N-1}), Z_s(\xi^{N-1}))ds$ for $t \in [(N-1)\tau, T]$. And this gives the representation, for $t \in [(N-1)\tau, T]$,

$$
\begin{aligned}
u(t, x + \tilde{B}(\xi^{N-1})_t) &= Y(\xi^{N-1})_t \\
\nabla u(t, x + \tilde{B}(\xi^{N-1})_t) &= Z(\xi^{N-1})_t \ .
\end{aligned}
$$

Now, by construction, the process $\tilde{B}(\xi^{N-1}) = \tilde{B}$, defined on $[0, T]$ by $\tilde{B}_t = \tilde{B}(\xi^i)_t$ if $t \in [i\tau, (i+1)\tau]$, is a Brownian motion. Considering now the BSDE solution $(Y(\xi^{N-1}), Z(\xi^{N-1}))$ on $[0, T]$ (not just $[T - \tau, T]$), uniqueness for Lipschitz BSDEs guarantees that it coïncides with $(Y(\xi^i), Z(\xi^i))$ on $[i\tau, (i+1)\tau]$. Consequently, $\tilde{B}(\xi^{N-1})_t = B_t - \int_0^t f(Y(\xi^{N-1})_s, Z(\xi^{N-1})_s)ds$ and ξ^{N-1} is the sought random variable. □

2.4 Application to the obtention of a PDE estimate.

In this section use the representation given by theorem 2.3.1 to establish some explicit gradient estimates for the solution of (2.2.1). We use the same notation as in the previous section, but we now particularize to the case where $f(y,z) = f(y)$, that is to say, f only depends on y. We denote by (P_t) the heat semigroup, defined by $P_t v(x) = \boldsymbol{E}^{\mathbb{R}}[v(W_t^x)]$, for any measure $\mathbb{R}$ and any $\mathbb{R}$-Browian motion W^x (starting at x).

Proposition 2.4.1. *For any $p \in [1,2)$, for any $t \le T$, and $i = 1, \cdots, m$,*

$$\int_0^t P_s \left|\nabla u^i\right|^p (x, T-s)\, ds \le c e^{pcT} \left(\|h\|_\infty^p + \|g(0,0)\|^p\right) \exp\left[\frac{p}{2(2-p)} t \sup_{|y|\le K} |f(y)|^2\right],$$

where c is an explicitly computable constant and $K = ce^{cT}\left[\|h\|_\infty^2 + |g(0,0)|^2 T\right]$ is the constant apprearing in proposition 2.3.3.

Proof. Using the fact that $\tilde{B}(\xi)$ is a Brownian motion under $\mathbb{Q}$, the Hölder inequality, the definition of $R_s = \frac{d\mathbb{Q}}{d\mathbb{P}}$ on $\mathcal{F}_s$ and the Hölder inequality again, we have

$$\begin{aligned}
&\int_0^t P_s \left(\left|\nabla u^i\right|^p (T-s, \cdot)\right)(x) ds \\
&\quad = \int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[\left|\nabla u^i\right|^p (T-s, x + \tilde{B}(\xi)_s)\right] ds \\
&\quad = \int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[\frac{1}{R_s^{\frac{p}{2}}} \left|\nabla u^i\right|^p (T-s, x+\tilde{B}(\xi)_s)\ R_s^{\frac{p}{2}}\right] ds \\
&\quad \le \int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[\frac{1}{R_s} \left|\nabla u^i\right|^2 (T-s, x+\tilde{B}(\xi)_s)\right]^{\frac{p}{2}} \boldsymbol{E}^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right]^{1-\frac{p}{2}} ds \\
&\quad = \int_0^t \boldsymbol{E}^{\mathbb{P}}\left[\left|\nabla u^i\right|^2 (T-s, x+\tilde{B}(\xi)_s)\right]^{\frac{p}{2}} \boldsymbol{E}^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right]^{1-\frac{p}{2}} ds \\
&\quad \le \left(\int_0^t \boldsymbol{E}^{\mathbb{P}}\left[\left|\nabla u^i\right|^2 (T-s, x+\tilde{B}(\xi)_s)\right] ds\right)^{\frac{p}{2}} \left(\int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right] ds\right)^{1-\frac{p}{2}} \\
&\quad = \left(\boldsymbol{E}^{\mathbb{P}}\left[\int_0^t |Z(\xi)_s^i|^2 ds\right]\right)^{\frac{p}{2}} \left(\int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right] ds\right)^{1-\frac{p}{2}} \\
&\quad = \left(\boldsymbol{E}^{\mathbb{P}}\left[\langle Y(\xi)^i\rangle_t\right]\right)^{\frac{p}{2}} \left(\int_0^t \boldsymbol{E}^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right] ds\right)^{1-\frac{p}{2}}
\end{aligned}$$

and the following two lemmas allow to conclude. □

Lemma 2.4.2. *For any $p \in [1,2)$,*

$$E^{\mathbb{Q}}\left[R_s^{\frac{p}{2-p}}\right] = E^{\mathbb{P}}\left[R_s^{\frac{2}{2-p}}\right] \le \exp\left\{\frac{p}{(2-p)^2}\ s \max_{|y|\le K} |f(y)|^2\right\}.$$

where $K = ce^{cT}\left[\|h\|_\infty^2 + |g(0,0)|^2 T\right]$ is the constant apprearing in proposition 2.3.3.

Proof. The first equality results from the definition of $\mathbb{Q}$ and the fact that R_s is $\mathcal{F}_s$-measurable. Then,

$$\begin{aligned}
R_s^{\frac{2}{2-p}} &= \exp\left\{ -\frac{2}{2-p}\,\frac{1}{2}\int_0^s |f(Y_r)|^2 dr + \int_0^s \frac{2}{2-p} f(Y_r) dB_r \right\} \\
&= \exp\left\{ \frac{p}{(2-p)^2}\int_0^s |f(Y_r)|^2 dr \right\} \\
&\quad \exp\left\{ -\frac{1}{2}\int_0^s \frac{4}{(2-p)^2}|f(Y_r)|^2 dr + \int_0^s \frac{2}{2-p} f(Y_r) dB_r \right\} \\
&\le \exp\left\{ \frac{p}{(2-p)^2}\; s \max_{|y|\le K} |f(y)|^2 \right\} M_s\;,
\end{aligned}$$

where we used Proposition 2.3.3, and

$$M_s = \exp\left\{ -\frac{1}{2}\int_0^s \frac{4}{(2-p)^2}|f(Y_r)|^2 dr + \int_0^s \frac{2}{2-p} f(Y_r) dB_r \right\}$$

is a $\mathbb{P}$-martingale. Taking the expectation gives the result. □

Lemma 2.4.3. *For some constant c,*

$$E^{\mathbb{P}}\left[\langle Y(\xi)^i\rangle_t\right]^{\frac{p}{2}} \le ce^{cT}\left(\|h\|_\infty^p + |g(0,0)|^p\right)$$

Proof. This is a direct consequence of the fact that

$$E^{\mathbb{P}}\left[\langle Y(\xi)^i\rangle_t\right]^{\frac{p}{2}} = E^{\mathbb{P}}\left[\int_0^t |Z_s|^2 ds\right]^{\frac{p}{2}} \le E^{\mathbb{P}}\left[\int_0^T |Z_s|^2 ds\right]^{\frac{p}{2}}.$$

Indeed, the quantity $E^{\mathbb{P}}\left[\int_0^T |Z_s|^2 ds\right]$ is bounded above by the BMO norm of Z. So, by the estimate of proposition 2.3.3, we have

$$E^{\mathbb{P}}\left[\int_0^T |Z_s|^2 ds\right]^{\frac{p}{2}} \le \left(ce^{cT}\left(\|h\|_\infty^2 + |g(0,0)|^2 T\right)\right)^{\frac{p}{2}} \le ce^{pcT}\left(\|h\|_\infty^p + |g(0,0)|^p T^{\frac{p}{2}}\right)$$

□

Chapter 3

Well-posedness of general quadratic reflected BSDEs driven by a continuous martingale.

3.1 Introduction.

3.1.1 Motivation.

The BSDEs associated with several types of PDE problems can all be studied under the same formalism.

Consider the parabolic PDE with Neumann boundary conditions

$$\begin{aligned}
& v_t + \frac{1}{2} v_{xx} \cdot a + v_x \cdot b + f(t, x, v, v_x\, \sigma) = 0 && \text{for } (t, x) \in [0, T[\times D\ , \\
& \frac{\partial v}{\partial n} = g && \text{for all } (t, x) \in [0, T[\times \partial D\ , \\
& v(T, \cdot) = \varphi && \text{for all } x \in \overline{D}\ ,
\end{aligned}$$

where D is a domain in $\mathbb{R}^d$ with a smooth boundary ∂D, and n is the inward pointing (usually) normal vector on the boundary. The BSDEs associated with such PDEs are the so-called "generalized BSDEs", introduced by Hu in [39] and Pardoux and Zhang in [68]. More precisely, when doing the connection between these PDEs and BSDEs, we consider a forward dynamics given by $dX_t = b(X_t)dt + n(X_t)dA_t + \sigma(X_t)dW_t$, A being the local time of the reflected diffusion X on the boundary ∂D of the domain. The corresponding BSDE then has the dynamics $dY_t = -f(Y_t, Z_t)dt - g(Y_t)dA_t + Z_t dW_t$, and terminal condition $Y_T = \varphi(X_T)$. We know from El Karoui and Huang [29] that one can enhance the current increasing process (dt) by setting $dC_t = dt + dA_t$, and find h such that $f(Y_t, Z_t)dt + g(Y_t)dA_t = h(Y_t, Z_t)dC_t$. This leads to consider BSDEs driven by a general increasing process $dY_t = -f(Y_t, Z_t)dC_t + Z_t dW_t$.

If instead of a Neumann boundary condition we consider a Dirichlet boundary condition,

$$\begin{aligned}
&v_t + \frac{1}{2} v_{xx} \cdot a + v_x \cdot b + f(t, x, v, v_x\,\sigma) = 0 && \text{for } (t,x) \in [0,T[\times D\ ,\\
&v = \psi && \text{for all } (t,x) \in [0,T[\times \partial D\ ,\\
&v(T,\cdot) = \varphi && \text{for all } x \in \overline{D}\ ,
\end{aligned}$$

then the diffusion $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ should be stopped when either the process X touches the boundary (in which case we know the value to be $\psi(X_\tau)$) or t reaches T (in which case we know the value to be $\varphi(X_\tau)$), where $\tau = \inf\{t \in [0,T] : X_t \in \partial D\} \wedge T$. This means we are considering the BSDE with a standard dynamics $dY_t = -f(Y_t, Z_t)dt + Z_t dW_t$ up to the stopping time τ, with terminal condition $Y_\tau = \Phi(X_\tau)$, where

$$\Phi(x) = \begin{cases} \varphi(x) & \text{if } x \in D \\ \psi(x) & \text{if } x \in \partial D, \end{cases}$$

the first case being for when X did not touch ∂D before T and the terminal value is given by the terminal condition φ of the PDE, the second case being for when X touched ∂D before T and the terminal value is given by the Dirichlet boundary condition ψ. Naturally, for the PDE problem to be a minimum consistent, the boundary condition ψ on ∂D and the terminal condition on $\overline{D}$ are so that for all $x \in \partial D$, $\psi(x) = \varphi(x)$.

There is no difficulty, at least formally, in considering at once the general BSDE

$$\begin{cases} dY_t = -f(Y_t, Z_t)dC_t + Z_t dW_t \\ Y_\tau = \Phi(X_\tau), \end{cases}$$

with both a (possibly) random terminal time τ and a (possibly non dt) increasing process C. Such a general BSDE covers at once the BSDEs associated to parabolic

PDEs set on the whole space $\mathbb{R}^d$, or set on a domain $D \subseteq \mathbb{R}^d$, whether with Dirichet or Neumann boundary condition. In PDEs, it is possible to recast these 3 types of parabolic problems in terms of operator. One essentially imbeds in the operator the behaviour at the boundary, so that formally it is always the same problem, but with a general operator. What we showed above is the equivalent for BSDEs.

One can consider an even greater generality of BSDEs. In many situations in mathematical finance, in the cases of incomplete markets, the randomness is not coming only from the d-dimensional Brownian motion W. Typically, it could be coming from a higher-dimensional Brownian motion but one can only control the exposure to a certain number of components since there are only a certain number of tradable assets. In such a case, W would not have the martingale representation property. This leads to consider a BSDE where the reference martingale is not necessarily W, but say a general martingale M which might not have the representation property. So that the martingale part N of the solution Y would have the decomposition $N = \int ZdM + N^\perp$ on M, with $N^\perp$ the component of N orthogonal to M (in the sense that $\langle M, N^\perp \rangle = 0$). This orthogonal complement $N^\perp$ is interpreted as the tracking error and used in the context of Föllmer-Schweizer strategies in mathematical finance.

Since the martingale part is $N = \int ZdM + N^\perp$, one could consider a drift depending on N not only through Z (that is, the component $\int ZdM$ absolutely continuous with respect to M) but also through $N^\perp$. We will consider such a dependence, of the form $d\langle \nu, N^\perp \rangle + g_s d\langle N^\perp \rangle$, where ν is a martingale orthogonal to M, characterizing the linear term in $N^\perp$, and g is a process, characterizing the quadratic term in $N^\perp$.

Combining all the generality that we have motivated above, we are naturally led to consider the general backward stochastic equation

$$\begin{cases} dY_t = dV(Y,N)_t + dN_t \\ Y_\tau = \xi, \end{cases}$$

where the drift is given by

$$dV(Y,N)_t = f(Y_t, Z_t)dC_t + d\langle \nu, N^\perp \rangle_t + g_t d\langle N^\perp \rangle_t \ .$$

We would like to study backward stochastic problems formulated in such a generality when the drift V has a quadratic dependence on the martingale part N of the solution. However, unlike in the previous chapter, we will remain in dimension $n = 1$. BSDEs in the strict sense (standard BSDEs) have already been studied with a certain level of generality (considering the generality only from a formal point of view for the moment, not looking at the analytical assumptions). For instance, El Karoui and Huang [29] studied BSDEs driven by a martingale when the drift is Lipschitz, Morlais [62] and Tevzadze [77] considered general BSDEs with a quadratic drift and a bounded terminal condition, while Cohen and Elliott [20] looked at BSDEs over a general filtration.

In this chapter, we will look at reflected BSDEs with the same generality as described above, under the assumption that the drift is at most quadratic in the martingale part and that the terminal condition is bounded. More precisely, we will study the theory (mainly the well-posedness questions) for those reflected BSDEs. On the way, we will obtain a new and more intrinsic proof of the special comparison theorem for reflected BSDEs (see below). We will also obtain a local Lipschitz estimate for N in the space BMO, the technique for which applies naturally to (standard) BSDEs in the Brownian case as well, and provide an improvement on previously known regularity.

3.1.2 Introduction to reflected BSDEs

Before moving on to presenting the research work that constitutes the main matter of this chapter, let us do first a brief presentation of reflected BSDEs, since this is what we will be concerned with in this chapter and it has not been presented in chapter 1.

Reflected BSDEs.

Given the name (reflected BSDEs) and the motivations for studying these equations (for us in this chapter, but also in general in the BSDE literature), it is natural to use BSDEs as a starting point.

A reflected BSDE is essentially a BSDEs, with terminal condition ξ and dynamics $dY_t = -f(Y_t, Z_t)dt + Z_t dW_t$, but we now want Y to remain above a lower barrier L : $Y_t \geq L_t$. Naturally, if we want this constraint to be satisfied, we need to allow the solution Y not to follow strictly the BSDE dynamics (characterized by f), and to have the possibility to drift upward. For this, we add a term $dK_t \geq 0$ in the dynamics : $dY_t = -f(Y_t, Z_t)dt - dK_t + Z_t dW_t$. This term is required to be minimal so that it only prevents Y from passing under L. The condition expressing that is the Skorohod optimality condition $1_{Y_t > L_t} dK_t = 0$. So in the end, solving a reflected BSDE consists in finding (Y, Z, K) such that

$$\begin{cases} dY_t = -f(Y_t, Z_t)dt - dK_t + Z_t dW_t \\ Y_T = \xi \\ Y_t \geq L_t \text{ for all } t < T \\ 1_{Y_t > L_t} dK_t = 0 \ , \end{cases}$$

where K is sought as a increasing process ($dK_t \geq 0$), continous, starting from 0 ($K_0 = 0$) and progressively measurable. It can of course be rewritten in integral form, for both the dynamics and the Skorohod condition :

$$\begin{cases} Y_t = \xi + \int_t^T f(Y_u, Z_u)du + (K_T - K_t) - \int_t^T Z_u dW_u \\ Y_t \geq L_t \text{ for all } t < T \\ \int_0^T 1_{Y_t > L_t} dK_t = 0. \end{cases}$$

One sometimes finds the Skorohod minimality condition expressed as $\int_0^T (Y_t - L_t) dK_t = 0$ which is obviously entirely equivalent.

Note that we do the identification between a continous increasing path on $[0, T]$ starting from 0 (an accumulator) and a positive measure dK on $[0, T]$.

Skorohod decomposition problem.

Just like BSDEs are the backward stochastic version of ODEs, reflected BSDEs are the backward stochastic version of ODEs with an obstacle constraint. And the underlying problem of the latter deterministic problem is the Skorohod decomposition problem.

Given a time interval $[0, T]$, a path x and an obstacle l, both $[0, T] \longrightarrow \mathbb{R}$, the Skorohod decomposition problem consists in finding a pair of paths (y, k) (all the paths we consider are continuous), with k increasing starting from 0, such that x decomposes into $x = y - k$, y satisfying the constraint $y \geq l$. There would of course be many such decompositions. Given one, (y, k), one can add any increasing path k' to both y and k

and still have a solution. The problem requires that one finds a solution with minimal k, which makes it unique. If k has to be as small possible, the idea is to take $dk_t = 0$ as much as possible, and allowing $dk_t > 0$ only when the contraint $y \geq l$ would otherwise be broken, which can only happen when the constraint is already "saturated" ($y_t = l_t$). So the problem requires that we find (y, k) satisfying $1_{y_t > l_t} dk_t = 0$.

Note that since $k_0 = 0$, in order for the problem to have a solution the condition $l_0 \leq x_0 = y_0$ must be satisfied. One can rewrite things in differential form. Instead of decomposing $x = y - k$ let us rather write $y = x + k$, so

$$\begin{cases} dy = dx + dk \\ y_0 = x_0 \\ y \geq l \text{ and } 1_{y>l} dk = 0 \end{cases}$$

We see that the problem consists in finding a path y that tracks x as much as possible, mirroring its moves so long as this does not break the constraint $y \geq l$.

The path x could be given by $dx_t = f_t dt$. Now, when $dx_t = f(y_t)dt$, the problem would be an ODE with an (obstacle) constraint.

Let us make the following remark. Whether dx_t drives y down toward the obstacle or whether dl_t drives the obstacle up toward y is really the same from a certain point of view (from the point of view of whether the constraint is being broken whether k should be active). In fact, one can always do a change of frame of reference (translation on the space of paths, $\overrightarrow{y} := y + h$, $\overrightarrow{l} := l + h$, $\overrightarrow{x} := x + h$, k being unchanged). So on pictures or in some computations, one can always assume that one deal with the obstacle $l = 0$ (take $h = -l$) or alternatively have a path x constant (take $h = -x$).

There are several ways to solve the Skorohod problem One way to describe the solution is by giving its dynamics. One sets $y_0 = x_0$ and $k_0 = 0$. Then

$$dy_t = \begin{cases} \text{if } y_t > l_t \text{ then } dx_t \text{ (the constraint can't be broken here, set } dk_t = 0) \\ \text{if } y_t = l_t \text{ then} \\ \quad \text{if } dx_t \geq dl_t \text{ then } dx_t \text{ (contact but no passing below, set } dk_t = 0) \\ \quad \text{if } dx_t < dl_t \text{ then } dl_t \text{ (to avoid passing below, and set } dk_t = dl_t - dx_t) \end{cases}$$

Another way to solve it is to first to a change of frame of reference so that $dx_t = 0$. So consider $\overrightarrow{x_t} = x_t - (x_t - x_0)$, $\overrightarrow{l_t} = l_t - (x_t - x_0)$, $\overrightarrow{y_t} = y_t - (x_t - x_0)$. Then it becomes clear that

$$\overrightarrow{y_t} = \sup \left\{x_0, (\overrightarrow{l}_s)_{s \leq t}\right\}$$

and then just define $k = \overrightarrow{y} - \overrightarrow{x}$. Undoing the transform leads to $y_t = x_t + \sup\{(l_s - x_s)^+, s \leq t\} = x_t + k_t$. Note that $\overrightarrow{y}_t = y_t - \int_0^t dx_s$ is the smallest increasing path (starting from x_0) dominating $\overrightarrow{l}_t = l_t - \int_0^t dx_s$.

Having seen the Skorohod decomposition problem, and more generally ODEs with an obstacle constraint, a reflected BSDE is the backward stochastic version of the ODE with obstacle (when the drift is a fixed process $f_t dt$ which does not depend on the solution, it is the backward stochastic version of the Skorohod problem).

Let us note that there is not really reflection here. In the simple case where $x = l = 0$, then $y = k = 0$ is the solution and one cannot say that $y = 0$ is in any way reflected on the obstacle $l = 0$. But the backward stochastic version is nonetheless known as BSDEs with reflection. As mentioned above, it is more a problem of finding a path y that follows as much as possibly the dynamics prescribed by x (by $f_t dt$), while satisfying a constraint.

Snell envelopes and optimal stopping.

Just like the underlying problem for a BSDE was the semimartingale decomposition, the underlying problem for a reflected BSDE is the Snell envelope (by underlying problem here we mean the base case where the drift is a fixed process, not dependending on the solution).

Let us consider a reflected BSDE with drift $f_t dt$:

$$\begin{cases} dY_t = -f_t dt - dK_t + Z_t dW_t \\ Y_T = \xi \ , \ Y_t \geq L_t \\ \quad 1_{Y_t > L_t} dK_t = 0 \end{cases}$$

Doing as above, we can define $\overrightarrow{Y}_t = Y_t + \int_0^t f_s ds$ and see that

$$d(Y_t + \int_0^t f_s ds) = -dK_t + Z_t dW_t.$$

So $\overrightarrow{Y}$ is a supermartingale, which dominates $L+V$ where $V_t = \int_0^t f_s ds$. And it is the smallest such supermartingale, that is to say the Snell envelope of $L+V$: $Y+V = Snell(L+V)$.

One can see that Y is given by

$$Y_t = \sup_\tau E\bigg(\int_t^\tau f_u du + L_\tau 1_{\tau<T} + \xi 1_{\tau=T} \bigg| \mathcal{F}_t \bigg)$$

where τ runs through the stopping times $\mathcal{T}_t^T$ such that $t \leq \tau \leq T$. This is the optimal stopping representation of Y : each Y_t is the value of an optimal stopping problem.

In the same way as the solutions of BSDEs can be thought of as nonlinear martingales, solutions of reflected BSDEs can be though of as nonlinear Snell envelopes.

Link with BSDEs : supersolutions.

In view of the above idea, that of a smallest process dominating another one, we can give an alternative description of the solution to a reflected BSDE.

Given a BSDE

$$\begin{cases} dY_t = -f(Y_t, Z_t)dt + Z_t dW_t \\ Y_T = \xi \ , \end{cases}$$

a supersolution is any (Y', Z') such that

$$\begin{cases} dY'_t \leq -f(Y'_t, Z'_t)dt + Z'_t dW_t \\ Y'_T = \xi \ . \end{cases}$$

Setting dK'_t to be the difference $-dY'_t - f(Y'_t, Z'_t)dt + Z'_t dW_t \geq 0$, we see that, by very definition of dK', $dY'_t = -f(Y'_t, Z'_t)dt - dK'_t + Z'_t dW_t$. Conversely, for any (Y', Z', K') satisfying the latter equation, with $dK'_t \geq 0$, (Y', Z') is a supersolution to the BSDE.

The solution to a reflected BSDE is therefore a supersolution to the corresponding BSDE. It is one that dominates the obtacle. And the minimality condition (the

Skorohod condition) says that it is the smallest supersolution dominating the obstacle.

Link with BSDEs : penalization.

So far, except maybe in the latest subsubsection, we have described reflected BSDEs in an intrinsic way. Whether when saying that reflected BSDEs are the backward stochastic version of ODEs with obstacle, or when saying that reflected BSDEs produce nonlinear Snell envelopes, we did not rely on BSDEs to understand reflected BSDEs. These ideas are the ones at work in the research presented in this chapter. However, it will be good to have also in mind the following approach to reflected BSDEs, which has been used in several existing works.

Reflected BSDEs can be seen at the limit of penalized BSDEs,

$$\begin{cases} dY_t^n = -\big[f(Y_t^n, Z_t^n) + n(Y_t^n - L_t)^-\big]dt + Z_t^n dW_t \\ Y_T^n = \xi \end{cases}$$

In the above dynamics equation, Y^n is subject to an upward drift (or force) $n(Y_t^n - L_t)^- dt$, in addition to $f(Y_t^n, Z_t^n)dt$. This force is proportional to how far below the obstacle L_t the solution Y_t^n is, but the proportionality coefficient, n, becomes increasingly big. It acts somehow like an elastic force, and at the limit $n \longrightarrow +\infty$ becomes rigid, preventing the solution from passing under L at all. Set $dK_t^n = n(Y_t^n - L_t)^- dt$. If one can show that the sequence converge (Y^n, Z^n, K^n), and converges in some sense strong enough that one can take the limit on the above equation (this can be proven under standard assumptions), then the limit $(Y^\infty, Z^\infty, K^\infty)$ is a solution to the reflected BSDE.

3.1.3 Review of the literature

Reflected BSDEs were introduced by El Karoui, Kapoudjian, Pardoux, Peng and Quenez in [30]. These authors considered the case where f is Lipschitz, the terminal condition is square-integrable and the lower obstacle a continuous square-integrable semimartingale, the natural extension of Pardoux and Peng [65].

The first results for quadratic BSDEs (that is, when f is allowed to have a quadratic growth in z) were obtained by Kobylanski in [50], under the assumption that ξ is bounded and f Lipschitz in y. Lepeltier and San Martin [55] allowed f to have slightly superlinear growth in y. Kobylanski, Lepeltier, Quenez and Torrès [51] were then able to prove the analogue results for RBSDEs by extending the techniques of [50] and [55].

The coefficient f can have any growth in the y variable if it satisfies the monotonicity condition (see for instance Pardoux [64]), which is encountered in reaction-diffusion equations. Briand, Lepeltier and San Martin [15] studied the case of quadratic growth BSDEs with such an assumption, and this was then extended to reflected BSDEs by Xu [79].

Briand and Hu [13, 14] extended Kobylanski's results [50] to the case of an unbounded terminal condition, and this case was further studied in the recent works of Delbaen, Hu and Richou [25, 26] and Barrieu and El Karoui [4]. Lepeltier and Xu [56] could then treat the case of RBSDEs with unbounded ξ, while Bayraktar and Yao [5] removed the condition that L be bounded.

However, the case of quadratic BSDEs remains significantly more difficult than that of Lipschitz BSDEs, and the methods used initially are often quite involved. Recently, Tevzadze [77] and Briand and Elie [12] gave simpler approaches for the case when ξ is bounded.

The above works ([30], [51], [79], [56] and [5]) concerning the well-posedness of reflected BSDEs considered a Brownian setting. However, BSDEs have been studied in a general martingale setting (see El Karoui and Huang [29], Tevzadze [77], Morlais [62], Barrieu and El Karoui [4]), and in a general filtered probability space in Cohen and Elliott [20].

In this chapter, we obtain the well-posedness of a general class of quadratic RBSDEs driven by a continuous martingale and with a bounded terminal condition in a simple, self-contained way. We show the existence of solutions in the cases where the dependence of f in y is Lipschitz, slightly-superlinear or monotone with arbitrary growth. We also obtain the special comparison theorem for the increasing processes under minimal assumptions. Finally, we obtain a local Lipschitz estimate in BMO for the martingale part of the solution.

3.1.4 Overview of the content of this chapter.

We now describe the organization of this chapter, explaning the results we obtained and the techniques used for this, thereby specifying our own contribution.

In section 3.3, we first obtain the standard comparison theorem in our setting, using a linearization and the BMO argument from Hu, Imkeller and Müller [40], as opposed to via an optimal stopping representation and comparison for BSDEs (see [51]). We note that this result, which guarantees uniqueness, holds naturally for f only locally Lipschitz in y, instead of globally Lipschitz as often assumed ([12], [77]). We then prove the special comparison theorem for reflected BSDEs, which allows one to compare the increasing processes when one RBSDE solution dominates another. This theorem was first proved in Hamadène, Lepeltier and Matoussi [37], and reused in Peng and Xu [71]. In the papers ([37], [71], [54], [51]) where it appears, the proof always relies on the penalization approach to reflected BSDEs and the comparison theorem for standard BSDEs, comparing quantities which, at the limit, become the increasing processes. The statement and the new proof we provide here are more intimately related to the nonlinear Snell envelope approach reflected BSDEs and hold under minimal assumptions. In particular, because we do not rely on a comparison theorem, they hold without the regularity assumptions usually made on f.

In section 3.4 we prove the existence of solutions to the reflected BSDEs when f is quadratic in z and Lipchitz in y. To this end, we generalize the technique introduced by Tevzadze [77] for BSDEs. The idea there is to first use the fixed point theorem to obtain a solution to a quadratic BSDEs when $f(\cdot, 0, 0)$ and ξ are sufficiently small ($f(s, 0, 0)ds$ is the residual drift, that drives the solution even if $(Y_s, Z_s) = (0, 0)$), and then to build upon this partial result to obtain a solution for general $f(\cdot, 0, 0)$ and ξ.

This technique can be understood as a type of "vertical" splitting and recombination, and is in that sense an analogue to what is done for Lipschitz BSDEs. In that classical case, if one works with the natural norm on the space where one looks for solutions, which in that context is the space of square-integrable processes, one finds that the fixed point theorem applies if the time interval is small enough. A natural way to use this is then to split ("horizontally") a general time interval into pieces small enough that one can obtain a solution on each interval, and patch them together to obtain a solution on the whole interval. For quadratic BSDEs, since one can apparently solve the BSDE only for small data, the idea is to split a general set of data into pieces small enough that one can obtain a solution for each piece, and then combine them to obtain a solution to the initial problem (this is the spirit in which the technique is decribed in [78] and [48]). In our view, one can also understand this method as a series of perturbations. One first solves a BSDE with microscopic data, then successively solves pertubation equations and adds the associated solutions, allowing the size of the data to grow at each step. At the end, one has built a solution to the initial BSDE with macroscopic data.

In order to use the fixed point theorem, one mainly needs to understand the underlying backward stochastic problem. For BSDEs, this underlying problem is the semimartingale decomposition. For reflected BSDEs, it is a Snell envelope problem. However, for the perturbation procedure to work well, the underlying problem should be a linear problem (for instance, it has been applied recently in Kazi-Tani, Possamai and Zhou [48] to BSDEs with jumps). This way, the equations satisfied by the perturbations are of the same nature as the equations satisfied by the solutions. This is not the case for reflected BSDEs. It is however possible to identify the equation that a perturbation should satisfy. The obstacle cannot be perturbed during the procedure, but this can be dealt with by assuming from the start that it is negative, a case which covers all the others by a simple translation. In particular, unlike in [51], we do not need L to be bounded but only require it to be upper bounded.

We then study the stability of the solution with respect to changes in the terminal condition ξ and in the residual drift $f(\cdot, 0, 0)$. We obtain for the martingale part of the solution a local Lipschitz estimate in the space BMO. Global Lipschitz bounds in $\mathcal{H}^p$ were obtained already in Briand, Delyon, Hu, Pardoux and Stoica [11] (see also Briand and Confortola [9], Ankirchner, Imkeller and Dos Reis [2]). Kazi-Tani, Possamai and Zhou [48] provide a global $\frac{1}{2}$-Hölder estimate in the smaller space BMO. Here, we can obtain a stronger regularity for small perturbations, essentially by bootstrapping a weaker regularity result.

Finally, in section 3.5, we extend the scope of the existence theorem of section 3.4. In that latter case, f is Lipschitz in y and the sequence of perturbations described above can be performed uniformly without problem. However, when the first derivative f_y is not a bounded function of y, the maximal allowed size for a perturbation depends on the size of the solution to the reflected BSDE that one wants to perturb, so it is not clear a priori that the procedure would terminate after finitely many perturbations. We show, however, that this is the case as soon as one can obtain an a priori bound for Y in $\mathcal{S}^\infty$. We can therefore extend the existence theorem of section 3.4 to the case where f is slightly-superlinear and to the case where f is monotone with arbitrary growth (as studied respectively by [51] and [79] in a Brownian setting), using the same perturbation technique.

In the following section, we specify the notation and the framework that will be used throughout this chapter, and we give the precise assumptions under which we work.

3.2 Setting

We will study the following general reflected BSDE :

$$\begin{cases} dY_s = -dV(Y,N)_s - dK_s + dN_s, \\ Y_T = \xi, \\ Y_t \geq L_t \text{ for all } t \leq T, \text{ and} \\ K \text{ increasing, continuous, starting from 0 and such that } 1_{\{Y_s > L_s\}} dK_s = 0 \end{cases} \tag{3.2.1}$$

where the drift is given by

$$dV(Y,N)_s = f(s, Y_s, Z_s\sigma_s)dC_s + d\langle \nu, N^\perp \rangle_s + g_s d\langle N^\perp \rangle_s.$$

This is referred to as the reflected BSDE of data $(V, \xi, L) = (f, \nu, g, \xi, L)$.

The framework is a filtered probability space $\big(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\in[0,\mathbb{T}]}, P\big)$ satisfying the usual conditions, where $\mathbb{T} > 0$ is a finite time horizon. T is an $\mathcal{F}$-stopping time valued in $[0, \mathbb{T}]$ (bounded stopping time). The continuous square-integrable martingale M is assumed to be BMO (see below). All the processes considered are continuous.

C is a continuous and progressively measurable increasing process (starting from 0) such that, roughly, all the finite variational processes which are related to the data (not depending on the solution) are absolutely continuous with respect to it. In particular, $d\langle M\rangle_s = a_s dC_s = \sigma_s\sigma_s^* dC_s$. It is assumed that the positive symmetric matrix a (or equivalently σ) is bounded away from 0 and infinity (i.e. bounded and uniformly elliptic).

The data of the BSDE (coefficients f, ν, g of the drift V, terminal condition ξ, obstacle L) are as follows :

- $f : \Omega \times [0,T] \times \mathbb{R} \times M_{1,d}(\mathbb{R}) \to \mathbb{R}$ is $Prog \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(M_{1,d}(\mathbb{R}))$-measurable, where $M_{n,d}(\mathbb{R})$ is the space of $n \times d$ matrices with entries in $\mathbb{R}$. $Prog = Prog(\mathcal{F}_T)$ is the progressively measurable sigma-field on the interval $[|0,T|]$ (the set of pairs (ω, t) such that $t \leq T(\omega)$).

- ν is a BMO martingale orthogonal to M (that is $\langle \nu, M \rangle = 0$)

- g is a progressively measurable and bounded scalar process.

- ξ is an $\mathcal{F}_T$-measurable, bounded random variable.

- L is a continuous semimartingale bounded above.

Throughout the chapter, we assume that f has at most quadratic growth in the variable z, in the following sense :

($\mathbf{A_{qg}}$) There exists a growth function $\lambda(\cdot)$ (i.e. $\lambda : \mathbb{R} \to \mathbb{R}_+$ symmetric, increasing on $\mathbb{R}_+$, bounded below by 1) and a positive process $h \in L^2_{BMO}$ (i.e. $\int hdM \in BMO$, see below) such that :

$$|f(t,y,z)| \leq \lambda(y)\big(h_t^2 + |z|^2\big).$$

The assumption as written above allows for any growth in y, although more specific assumptions on this are made in sections 3.3, 3.4 and 3.5.

Solutions to the reflected BSDE.

A solution to the reflected BSDE is generally understood as a triple $S = (Y, N, K)$ where Y is a semimartingale, N a square-integrable martingale ($\in \mathcal{H}^2$) and K an increasing process (starting from 0), such that (3.2.1) is satisfied, with $N = \int ZdM + N^\perp$.

Note that a solution can also be understood as a pair $S = (Y, N)$ such that, definining K from $K_0 = 0$ and the dynamics equation in (3.2.1), K is indeed found to be increasing and satisfies the Skorohod condition. This will often be what is meant by solution in the rest of the chapter.

Under the assumption of quadratic growth and bounded terminal condition, we consider only bounded solutions : $Y \in \mathcal{S}^\infty$. For those, N is found to be a BMO martingale (see the a priori estimate of proposition 3.2.1 below). So a solution will always be understood as being in $\mathcal{S}^\infty \times BMO$ $(\times \mathcal{A})$, where BMO and $\mathcal{A}$ are described below.

Spaces of processes, notation.
We will make use of the following notation for particular spaces of processes, which we define or recall the definition of.

- $BMO(P)$ is the space of all the BMO P-martingales, that is those for which the norm

$$\|N\|^2_{BMO(P)} = \sup_{t\in\mathcal{T}_0^T} \|E_P(\langle N\rangle_T - \langle N\rangle_t|\mathcal{F}_t)\|_\infty$$

  is finite, where $\mathcal{T}_0^T$ is the set of stopping times t such that $0 \leq t \leq T$. The mention of the measure P will be omitted whenever no confusion is possible. When $X \in BMO$, then $\mathcal{E}(X)$ is a UI martingale, so one can define a measure Q by stating that on $\mathcal{F}_t$, $\frac{dQ}{dP} = \mathcal{E}(X)_t$. Also, we will use frequently the fact that for any $N \in BMO(P)$, $\widetilde{N} = N - \langle X, N\rangle$ is in $BMO(Q)$ (cf Kazamaki [47], theorems 2.3 and 3.3).

- L^2_{BMO} is the space of processes h such that $\int hdM \in BMO$. We equip it with the norm

$$\|h\|^2_{L^2_{BMO}} = \sup_t \left\| E\left(\int_t^T h_s^2 dC_s \middle| \mathcal{F}_t \right) \right\|_\infty$$

- $\mathcal{A}$ is the space of accumulators, that is : progressively measurable, continuous, increasing processes starting from 0.

- $L^{\infty,2}$ is the space of processes x such that $\int_0^T |x_s|^2 dC_s \in L^\infty$, and $L^{\infty,1}$ is that of processes such that $\int_0^T |x_s| dC_s \in L^\infty$, with norms $\|x\|^2_{\infty,2} = \|\int_0^T |x_s|^2 dC_s\|_\infty$ and $\|x\|_{\infty,1} = \|\int_0^T |x_s| dC_s\|_\infty$.

In the growth assumptions on y made later on, we use a fixed positive process $r \in L^{\infty,2}$, which is part of the framework (like T, M, C). It is there to take into account the fact that C_T might not be bounded (in facts, if C incorporates a local time of a diffusion on a regular boundary, it has exponential moments but is not bounded). In the case where $dC_s = ds$ and T is a constant, $r = 1$ and $\|r\|_{\infty,2} = \sqrt{T}$.

For processes $S = (Y, N) \in \mathcal{S}^\infty \times BMO(Q)$, we use the norm $\|S\|_Q^2 := \|Y\|_{\mathcal{S}^\infty}^2 + \|N\|_{BMO(Q)}^2$. As usual the mention of the measure Q is omitted whenever no confusion arises. By L^∞ we denote the space of bounded random variables, and we use the norm $\| \ \|_\infty$, whether those random variables are $\mathbb{R}$-valued (like ξ) or path-valued (like g).

BMO property for N.

Proposition 3.2.1. *Let f satisfy $(\mathbf{A_{qg}})$, $\nu \in BMO$ and g be bounded. Let Y be a continuous semimartingale, N be a square-integrable martingale and K be an increasing process such that Y has the decomposition :*

$$dY = -dV(Y, N) - dK + dN,$$

where $dV(Y, N)_s = f(s, Y_s, Z_s\sigma_s)dC_s + d\langle \nu, N^\perp\rangle_s + g(s)d\langle N^\perp\rangle_s$. If Y is bounded (i.e. $Y \in \mathcal{S}^\infty$), then $N \in BMO$ and $K \in \mathcal{A}_{BMO}$.

Here, $\mathcal{A}_{BMO}$ refers to the increasing processes $K \in \mathcal{A}$ such that the norm $\|K\|_{\mathcal{A}_{BMO}} = \sup_t \|E\big(K_T - K_t|\mathcal{F}_t\big)\|_\infty$ is finite. Note that this statement is, to some extent, not so much about solutions to a (possibly reflected) BSDE but about quadratic semimartingales (see Barrieu and El Karoui [4]), and quadratic semimartingales are considered here up to a monotonous process.

Remark 3.2.2. The result implies in particular the following : if Y is a bounded semimartingale with decomposition $dY = -dV - dK + dN$, with K monotonous (which boils down to increasing, up to considering $-Y$) and if the process V is in L^1_{BMO} (i.e. $\sup_t \|E\big(\int_t^T |dV_s|\big|\mathcal{F}_t\big)\|_\infty < +\infty$), then $N \in BMO$ and $K \in \mathcal{A}_{BMO}$.

Proof. The proof uses the usual exponential transform. Let $\mu \in \mathbb{R}$, whose sign and value will be chosen later. By Itô's formula for the process $\exp(\mu Y)$ between a stopping time $t \in \mathcal{T}_0^T$ and T one has

$$\begin{aligned} e^{\mu Y_t} - \mu \int_t^T e^{\mu Y_s} dK_s + \frac{\mu^2}{2} \int_t^T e^{\mu Y_s} d\langle N\rangle_s = e^{\mu Y_T} + \mu \int_t^T e^{\mu Y_s} dV_s \\ - \mu \int_t^T e^{\mu Y_s} dN_s \ . \end{aligned} \tag{3.2.2}$$

Since $Y \in \mathcal{S}^\infty$, the process $e^{\mu Y}$ is bounded, and since N is a square-integrable martingale, $\int e^{\mu Y} dN$ is a martingale. We have

$$|dV_s| \leq |f(s, Y_s, Z_s\sigma_s)|dC_s + |d\langle \nu, N^\perp\rangle_s| + |g_s||d\langle N^\perp\rangle_s|,$$

and using the quadratic growth assumption on f we have

$$|f(s, Y_s, Z_s\sigma_s)| \leq \lambda(Y_s)\big(h_s^2 + |Z_s\sigma_s|^2\big) \leq \Lambda\big(h_s^2 + |Z_s\sigma_s|^2\big),$$

where $\Lambda = \lambda(\|Y\|_{\mathcal{S}^\infty})$. Using the Kunita-Watanabe inequality and $ab \leq a^2 + b^2$, we see that

$$\begin{aligned} E\bigg(\int_t^T e^{\mu Y_s}|dV_s|\bigg|\mathcal{F}_t\bigg) \leq \Lambda E\bigg(\int_t^T e^{\mu Y_s}\big(h_s^2 + |Z_s\sigma_s|^2\big)dC_s\bigg|\mathcal{F}_t\bigg) + E\bigg(\int_t^T e^{\mu Y_s} d\left\langle \nu\right\rangle_s \bigg|\mathcal{F}_t\bigg) \\ E\bigg(\int_t^T e^{\mu Y_s} d\left\langle N^\perp\right\rangle_s \bigg|\mathcal{F}_t\bigg) + \|g\|_\infty E\bigg(\int_t^T e^{\mu Y_s} d\left\langle N^\perp\right\rangle_s \bigg|\mathcal{F}_t\bigg). \end{aligned}$$

Recall that by the orthogonality of M and $N^\perp$, $d\left\langle N\right\rangle = |Z\sigma|^2 dC + d\left\langle N^\perp\right\rangle$, and therefore both $|Z\sigma|^2 dC$ and $d\left\langle N^\perp\right\rangle$ are less than or equal to $d\left\langle N\right\rangle$. Therefore,

$$\begin{aligned} E\bigg(\int_t^T e^{\mu Y_s}|dV_s|\bigg|\mathcal{F}_t\bigg) \leq \Lambda E\bigg(\int_t^T e^{\mu Y_s} h_s^2 dC_s\bigg|\mathcal{F}_t\bigg) + E\bigg(\int_t^T e^{\mu Y_s} d\left\langle \nu\right\rangle_s \bigg|\mathcal{F}_t\bigg) \\ \big(\Lambda + 1 + \|g\|_\infty\big) E\bigg(\int_t^T e^{\mu Y_s} d\left\langle N\right\rangle_s \bigg|\mathcal{F}_t\bigg). \end{aligned}$$

So, setting $b = (\Lambda + 1 + \|g\|_\infty)$, and taking the conditional expectation of (3.2.2) with respect to $\mathcal{F}_t$, one has

$$\begin{aligned} 0 - \mu E\bigg(\int_t^T e^{\mu Y_s} dK_s\bigg|\mathcal{F}_t\bigg) + \Big\{\frac{\mu^2}{2} - |\mu| b\Big\} E\bigg(\int_t^T e^{\mu Y_s} d\left\langle N\right\rangle_s \bigg|\mathcal{F}_t\bigg) \\ \leq e^{|\mu|\|Y\|_{\mathcal{S}^\infty}} + |\mu|\bigg[\Lambda E\bigg(\int_t^T e^{\mu Y_s} h_s^2 dC_s\bigg|\mathcal{F}_t\bigg) + E\bigg(\int_t^T e^{\mu Y_s} d\left\langle \nu\right\rangle_s \bigg|\mathcal{F}_t\bigg)\bigg] - 0. \end{aligned}$$

We now choose $\mu = -4b$, so $\frac{\mu^2}{2} - |\mu| b = 4b^2$. Since $b \geq 1$, $b^2 \geq b$. We now use the fact that $e^{-|\mu|\|Y\|_{\mathcal{S}^\infty}} \leq e^{\mu Y_s} \leq e^{|\mu|\|Y\|_{\mathcal{S}^\infty}}$ and take the the $\sup_t$, so we obtain finally

$$\|K\|_{\mathcal{A}_{BMO}} + \|N\|_{BMO}^2 \leq \frac{e^{8b\|Y\|_{\mathcal{S}^\infty}}}{2b}\Big[1 + 4b\big(\Lambda\|h\|_{L^2_{BMO}}^2 + \|\nu\|_{BMO}^2\big)\Big] < +\infty.$$

$\square$

Note however that while it indeed gives a bound for $N \in BMO$, this estimate does not guarantee that if $\|Y\|_{\mathcal{S}^\infty} \longrightarrow 0$ then $\|N\|_{BMO} \longrightarrow 0$.

3.3 Comparison theorems and uniqueness.

3.3.1 Comparison theorem.

We prove below the comparison theorem in our setting, which guarantees uniqueness in the existence theorems of sections 3.4 and 3.5. The regularity assumption that we require for the theorem to hold is, for notational simplicity, the following :

($\mathbf{A_{Df}}$) The function f is of class $\mathcal{C}^1$ (in the variable (y,z), for all ω, t) with

$$|f_y(t,y,z)| \le \rho(y) r_t^2 \qquad \text{and} \qquad |f_z(t,y,z)| \le \rho'(y)(h_t + |z|),$$

for some growth functions ρ and ρ', and some positive process $h \in L^2_{BMO}$.

Theorem 3.3.1. *Consider two sets of data (f, ν, g, ξ, L) and (f', ν, g', ξ', L'), and assume that :*

1. *there exist solutions (Y, N, K) and (Y', N', K') to the corresponding reflected BSDEs,*
2. *the parameters are ordered : $f' \le f$, $g' \le g$, $\xi' \le \xi$ and $L' \le L$,*
3. *f is regular enough : it satisfies ($\mathbf{A_{Df}}$).*

Then one has $Y' \le Y$.

While the proof given in [51] in a Brownian setting uses an optimal stopping representation and the comparison theorem for BSDEs, we rely here on a classical linearization argument and the properties of solutions to a linear BSDE. More precisely, we study the positive part $(\Delta Y)^+$, where $\Delta X = X' - X$ for a generic quantity X, and show that $(\Delta Y)^+ \le 0$.

Proof. Denoting by l the local time of ΔY in 0, the Itô-Tanaka formula gives

$$\begin{aligned} d(\Delta Y)^+_s &= 1_{\{\Delta Y_s > 0\}} d\Delta Y_s + \frac{1}{2} dl_s \\ &= 1_{\{\Delta Y_s > 0\}} \Big[- d\Delta V_s - d\Delta K_s + d\Delta N_s \Big] + \frac{1}{2} dl_s \ . \end{aligned} \tag{3.3.1}$$

Now, gathering terms, rewriting differences, and linearizing some,

$$\begin{aligned} d\Delta V_s &= \Big[f'(s, Y'_s, Z'_s\sigma_s) - f(s, Y_s, Z_s\sigma_s) \Big] dC_s \\ &\quad + d\langle \nu, (N')^\perp - N^\perp \rangle + g'_s d\langle (N')^\perp \rangle_s - g_s d\langle N^\perp \rangle_s \\ \\ &= \Big[(\Delta f)(s, Y'_s, Z'_s\sigma_s) + f(s, Y'_s, Z'_s\sigma_s) - f(s, Y_s, Z_s\sigma_s) \Big] dC_s \\ &\quad + d\langle \nu, \Delta N^\perp \rangle_s + (\Delta g)_s \, d\langle (N')^\perp \rangle_s + g_s \Big[d\langle (N')^\perp \rangle - d\langle N^\perp \rangle_s \Big] \\ \\ &= \Big[(\Delta f)(s, Y'_s, Z'_s\sigma_s) + f(s, Y'_s, Z'_s\sigma_s) - f(s, Y_s, Z'_s\sigma_s) \Big] dC_s \\ &\quad + \Big[f(s, Y_s, Z'_s\sigma_s) - f(s, Y_s, Z_s\sigma_s) \Big] dC_s \\ &\quad + d\langle \nu, \Delta N^\perp \rangle_s + (\Delta g)_s \, d\langle (N')^\perp \rangle_s + g_s d\langle (N')^\perp + N^\perp, \Delta N^\perp \rangle \\ \\ &= \Big[\Delta f + F_y \Delta Y + F_z \Delta Z \sigma \Big] dC \\ &\quad + d\langle \nu', \Delta N^\perp \rangle + (\Delta g) \, d\langle (N')^\perp \rangle, \end{aligned}$$

where

$$\begin{aligned} F_y(s) &= F_y(s, Y_s, Y'_s, Z'_s\sigma_s) = \int_0^1 f_y(s, Y_s + u\Delta Y_s, Z'_s\sigma_s) du \ , \\ F_z(s) &= F_z(s, Y_s, Z_s\sigma_s, Z'_s\sigma_s) = \int_0^1 f_z(s, Y_s, Z_s\sigma_s + u\Delta Z_s\sigma_s) du \text{ and} \\ \nu' &= \nu + \int g d(N^\perp + (N')^\perp) \ . \end{aligned}$$

So we can rewrite (3.3.1) as

$$d(\Delta Y)^+ = -dD - F_y(\Delta Y)^+ dC + 1_{\{\Delta Y>0\}}\Big[d\Delta N - F_z \Delta Z \sigma dC - d\langle \nu', \Delta N^\perp\rangle\Big] \ , \tag{3.3.2}$$

where

$$dD = 1_{\{\Delta Y>0\}}\Big[\Delta f dC_s + \Delta g \, d\langle (N')^\perp\rangle + d\Delta K\Big] - \frac{1}{2}dl$$

is a decreasing process. Indeed $\Delta f \leq 0$, $\Delta g \leq 0$, $dl \geq 0$ and

$$1_{\{\Delta Y>0\}} d\Delta K = \underbrace{1_{\{\Delta Y>0\}} dK'}_{=0} - \underbrace{1_{\{\Delta Y>0\}} dK}_{\geq 0} \leq 0$$

because $dK \geq 0$ and on $\{\Delta Y > 0\}$ we have $Y' > Y \geq L \geq L'$, hence $dK' = 0$.

$(\Delta Y)^+$ is therefore seen as the solution to a linear equation (3.3.2). Define the integrating factor $B_t = e^{\int_0^t F_y(u) dC_u}$ and the measure Q by $\frac{dQ}{dP} = \mathcal{E}(\int F_z \sigma^{-1} dM + \nu')_t$. By the assumption on f_z and the fact that $h \in L^2_{BMO}$ and ν, N, and N' are in BMO (recall proposition 3.2.1 and the definition of a solution), $\int F_z \sigma^{-1} dM + \nu'$ is in BMO, and therefore Q is indeed well defined. Then, $\widetilde{\Delta N} = \Delta N - \int F_z \Delta Z \sigma dC_s - \langle \nu', \Delta N^\perp\rangle$ is a $BMO(Q)$-martingale. By the assumption on f_y, the process $B.$ is bounded so $\int B_u 1_{\Delta Y_u > 0} d\widetilde{\Delta N_u}$ is again a Q-martingale. Therefore, looking at the dynamic of $\widehat{Y} = BY$ under Q we finally find that

$$0 \leq (\Delta Y_t)^+ = E_Q\bigg(e^{\int_t^T F_y(v)dC_v}(\Delta\xi)^+ + \int_t^T e^{\int_t^u F_y(v) dC_v} dD_u \Big| \mathcal{F}_t\bigg) \leq 0 \ .$$

□

Remark 3.3.2. The previous theorem is stated, for convenience, for a function f which is C^1. Typically, for the comparison theorem one only requires f to be locally Lipschitz, in which case the processes F_y, F_z have to be replaced by the differential quotients : $\delta_y f(Y, Y', Z'\sigma) = \frac{f(Y',Z') - f(Y,Z')}{Y'-Y} 1_{Y \neq Y'}$, etc, and the above proof works as long as $F_y \in L^{\infty,1}$ and $F_z \in L^2_{BMO}$. These criteria are satisfied as soon as

($\mathbf{A}_{\text{lLip}}$) There exist growth functions ρ and ρ', and a process $h \in L^2_{BMO}$ such that

$$|f(t,y',z') - f(t,y,z)| \leq \rho(y,y') r_t^2 |\Delta y| + \rho'(y,y')\big(h_t + |z| + |z'|\big)|\Delta z| \ .$$

Note : when ρ, ρ' are constants this is the standard assumption of local Lipchitz regularity made in the quadratic BSDE literature (see for instance Briand and Elie [12] and Tevzadze [77]). However, since we are dealing with bounded solutions, the assumption can be weakened to the case ($\mathbf{A}_{\text{lLip}}$) where ρ, ρ' are growth functions.

3.3.2 Special comparison theorem.

When the two sets of data are in a comparison configuration and when the lower obstacles are the same, one can say more than $Y' \leq Y$ and also compare the increasing processes of the two solutions, K' and K.

Proposition 3.3.3. *Let (f, g, ν, ξ, L) and (f', g', ν, ξ', L) be some data, and assume that :*

1. *there exist solutions $S = (Y, N, K)$ and $S' = (Y', N', K')$ to the corresponding RBSDEs,*
2. *the drift coefficients are ordered : $f' \leq f$, $g' \leq g$,*
3. *Y' is dominated by Y : $Y' \leq Y$.*

Then it is the case that $dK_t \leq dK'_t$.

The intuition is quite clear. First, since one has $Y'_t \leq Y_t$, if Y doesn't touch the barrier ($Y_t > L_t$), then $dK_t = 0$ and whether $Y'_t > L_t$ or $Y'_t = L_t$, one has $dK'_t \geq 0 = dK_t$. So the only non-trivial case is when Y touches the barrier, and therefore Y' as well. In that case, since the extra forces dK' and dK are minimal, they only prevent the drifts dV' and dV from driving the solutions Y' and Y under the obstacle. But since $dV'_t \leq dV_t$ in that case, the correction that could be needed for Y will be less than that needed for Y'. The proof makes this heuristics rigorous.

Unlike in [37], [71], [54], [51], the proof we give here works under minimal assumptions and in particular does not require a regularity assumptions on f, since it does not rely on the comparison theorem for BSDEs.

Proof. In this proof, contrary to the rest of the chapter, ΔX denotes $X - X'$ for a generic quantity X. In order to deal with what happens locally when the process ΔY touches 0, we proceed as in El Karoui *et al.* [30] : write down the structure of ΔY and ΔY^+, argue that these two processes are equal (since by assumption $\Delta Y \geq 0$), identify their finite variational and martingale parts, and then extract the relevant information. Our goal is to prove that $d\Delta K \leq 0$.

We have

$$\begin{aligned} d\Delta Y &= -d\Delta V - d\Delta K + d\Delta N \qquad \text{and} \\ d(\Delta Y)^+ &= 1_{\{\Delta Y > 0\}} d\Delta Y + \frac{1}{2} dl \ , \end{aligned}$$

where l is the local time of ΔY at 0. Identifying the finite variational and martingale parts, we see that

$$\begin{aligned} -d\Delta V - d\Delta K &= 1_{\{\Delta Y > 0\}} \Big(-d\Delta V - d\Delta K \Big) + \frac{1}{2} dl \qquad \text{and} \\ d\Delta N &= 1_{\{\Delta Y > 0\}} d\Delta N \ , \end{aligned}$$

that is to say

$$\begin{aligned} 1_{\{\Delta Y = 0\}} \Big(-d\Delta V - d\Delta K \Big) &= \frac{1}{2} dl \qquad \text{and} \\ 1_{\{\Delta Y = 0\}} d\Delta N &= 0 \ . \end{aligned}$$

The second equation implies, by Itô's isometry and the orthogonality between M and $\Delta N^\perp$, that $1_{\Delta Y = 0} \Big(|\Delta Z \sigma|^2 dC + d\langle \Delta N^\perp \rangle \Big) = 0$. So we know that on the set $\{Y' = Y\}$ (i.e. against $1_{\{\Delta Y = 0\}}$) we have $Y = Y'$ and $Z = Z'$. We also notice that by the Kunita-Watanabe inequality, $1_{\{\Delta Y = 0\}} d\langle \nu', \Delta N^\perp \rangle = 0$ for any continuous semimartingale ν'.

The drift term can be rewritten, using $\Delta \nu = \nu - \nu = 0$,

$$\begin{aligned} d\Delta V_t &= \big(f(S) - f'(S') \big) dC + d\left\langle \nu, N^\perp \right\rangle - d\left\langle \nu, (N')^\perp \right\rangle + g d\left\langle N^\perp \right\rangle - g' d\left\langle (N')^\perp \right\rangle \\ &= \big(f(S) - f(S') + (\Delta f)(S') \big) dC + d\left\langle \nu, \Delta N^\perp \right\rangle + d\left\langle \Delta \nu, (N')^\perp \right\rangle \\ &\qquad + g\Big[d\left\langle N^\perp \right\rangle - d\left\langle (N')^\perp \right\rangle \Big] + (\Delta g) d\left\langle (N')^\perp \right\rangle \\ &= \Big[\big(f(S) - f(S') \big) dC + d\left\langle \nu, \Delta N^\perp \right\rangle + g d\left\langle N^\perp + (N')^\perp, \Delta N^\perp \right\rangle \Big] \\ &\qquad + \Big[(\Delta f)(S') dC + d\left\langle (\Delta \nu), (N')^\perp \right\rangle + (\Delta g) d\left\langle (N')^\perp \right\rangle \Big] \\ &= \Big[\big(f(S) - f(S') \big) dC + d\left\langle \nu', \Delta N^\perp \right\rangle \Big] + \Big[d(\Delta V)(S') \Big] \ , \end{aligned}$$

where $\nu' = \nu + \int g d(N^\perp + (N')^\perp)$. By the assumptions on the coefficients, we know that $d(\Delta V)(S')_t =: dI_t \geq 0$. So we find that against $1_{\{\Delta Y_t=0\}}$ we have

$$1_{\{\Delta Y_t=0\}} d\Delta V_t = 0 + 1_{\{\Delta Y_t=0\}} dI_t \ .$$

In the end,

$$1_{\{\Delta Y=0\}} \big(- dI - d\Delta K \big) = \frac{1}{2} dl \ ,$$

so

$$1_{\{\Delta Y=0\}} d\Delta K = -1_{\{\Delta Y=0\}} \underbrace{dI}_{\geq 0} - \frac{1}{2} \underbrace{dl}_{\geq 0} \leq 0 \ ,$$

and so we have proven that $1_{\{\Delta Y=0\}} d\Delta K \leq 0$. And when $\Delta Y > 0$, one has $Y > Y' \geq L' = L$ so $dK = 0 \leq dK'$, and therefore $1_{\{\Delta Y>0\}} d\Delta K \leq 0$, which completes the proof. □

3.4 Existence and stability.

In this section we work under the assumption that the derivatives of f are controlled in the following way :

($\mathbf{A_{der}}$) f is twice continuously differentiable in the variables (y, z) and there exists $\rho, \rho', \lambda > 0$, and $h \in L^2_{BMO}$ such that

$$|f_y(t, y, z)| \leq \rho r_t^2 \qquad \text{and} \qquad |f_z(t, y, z)| \leq \rho'(h_t + |z|) \ ,$$
$$|f_{yy}(t, y, z)| \leq \lambda r_t^2, \quad |f_{yz}(t, y, z)| \leq \lambda r_t \quad \text{and} \quad |f_{zz}(t, y, z)| \leq \lambda \ .$$

Rather than aiming to construct a solution to (3.2.1) by an approximation procedure on the data, as was done in the Brownian setting (see [51]), we work in a more direct way, as in section 5 of [30], and for this we adapt the pertubation procedure introduced in [77] for BSDEs. We then analyze the dependence of the solution on the data.

3.4.1 Principle.

As said in the introduction, the strategy is to first apply the fixed point theorem. To perform this, one will use only the following assumption on f :

($\mathbf{A_{locLip}}$) The function f is differentiable at $(0,0)$ (in (y,z), for all (ω,s)), and there exist $\lambda>0$ such that, writing $\beta_s=f_y(s,0,0)$ and $\gamma_s=f_z(s,0,0)$, one has

- for all ω,s,y_1,y_2,z_1,z_2 :

$$\begin{aligned}&|f(s,y_1,z_1)-f(s,y_2,z_2)-\beta_s(y_1-y_2)-\gamma_s(z_1-z_2)|\\&\qquad\le\lambda\Big(r_s|y_1|+r_s|y_2|+|z_1|+|z_2|\Big)\big(r_s|y_1-y_2|+|z_1-z_2|\big)\ ,\end{aligned}$$

- $\gamma\in L^2_{BMO}$ and $\beta\in L^{\infty,1}$ (that is : $\int_0^T|\beta_s|dC_s\in L^\infty$),

which follows naturally from the assumption on the second derivative of f in ($\mathbf{A_{der}}$). In all generality, it allows also for quadratic growth in y. So what one actually proves first is that when f satisfies this assumption (with possibly quadratric growth in y and z), and when the data are small enough (in a sense to specify), there exists a solution.

The perturbations procedure is then carried as follows for a reflected BSDE with obstacle $L\le 0$. Split the initial data in n pieces : $(\xi^i)_{i=1\dots n}$ and $(\alpha^i)_{i=1\dots n}$ such that $\sum_{i=1}^n\xi^i=\xi$ and $\sum_{i=1}^n\alpha^i=\alpha$, where $\alpha=f(\cdot,0,0)$, and such that for each i, (ξ^i,α^i) is small enough. For the sake of the proof we take the particular decomposition given by $\xi^i:=\xi^{(n)}=\frac{1}{n}\xi$ and $\alpha^i:=\alpha^{(n)}=\frac{1}{n}\alpha$, for n big enough, though other decompositions would do.

First, there is a solution $S^1=(Y^1,N^1,K^1)$ to the reflected BSDE (3.2.1) with small data $(f-\alpha+\alpha^1,\nu,g,\xi^1,L)$.

Now, unless otherwise specified, we denote by $\overline{x}^k$ the sum $\Sigma_{j=1}^k x^j$, for a general quantity x indexed by $\{1,\dots,n\}$. For $i=2$ to n, having obtained a solution $\overline{S}^{i-1}=(\overline{Y}^{i-1},\overline{N}^{i-1},\overline{K}^{i-1})$ to the reflected BSDE (3.2.1) with parameters $(f-\alpha+\overline{\alpha}^{i-1},\nu,g,\overline{\xi}^{i-1},L)$, one incorporates one more (α^i,ξ^i) in the system. One first constructs the perturbation $S^i=(Y^i,N^i,K^i)$ solving the pertubation equation

$$\begin{cases}dY^i=-dV^i(Y^i,N^i)-dK^i+dN^i\ ,\\ Y^i_T=\xi^i\ ,\\ \overline{Y}^{i-1}+Y^i\ge\overline{L}^{i-1}+L^i\ ,\\ d\overline{K}^{i-1}+dK^i\ge 0\text{ and }(d\overline{K}^{i-1}+dK^i)(\overline{Y}^{i-1}+Y^i>\overline{L}^{i-1}+L^i)=0\end{cases}\tag{3.4.1}$$

with drift given by

$$\begin{aligned} dV^i(Y^i,N^i)_s &= \big[f(\overline{S}^{i-1}+S^i)-f(\overline{S}^{i-1})+\alpha^i_s\big]dC_s \\ &\quad + d\langle \nu + \int 2gd(\overline{N}^{i-1})^\perp,(N^i)^\perp\rangle_s + gd\langle (N^i)^\perp\rangle_s \\ &= \big[\overline{f}^{i-1}(S^i)+\alpha^i_s\big]dC_s + d\langle \overline{\nu}^{i-1},(N^i)^\perp\rangle_s + gd\langle (N^i)^\perp\rangle_s\ , \end{aligned}$$

where $\overline{\nu}^{i-1} = \nu + \int 2gd(\overline{N}^{i-1})^\perp$, and $\overline{f}^{i-1}$ is the function f recentered around $\overline{S}^{i-1}$. It satisfies $\overline{f}^{i-1}(0)=0$ so the residual-drift (constant part) in this equation is given by α^i. So the parameters $(\overline{f}^{i-1}+\alpha^i,\nu^{i-1},g,\xi^i,L)$ here are small in the required sense. Finally, one sums $\overline{S}^i := \overline{S}^{i-1}+S^i$ to obtain a solution to the reflected BSDE (3.2.1) of parameters $(f-\alpha+\overline{\alpha}^i,\nu,g,\overline{\xi}^i,L)$. For $i=n$ this provides a solution to the reflected BSDE of interest.

This allows us to conclude to existence for those reflected BSDEs with negative obtacles. But then we can show that up to translation, this covers all the cases where the obstacle is upper-bounded.

Note already that the above perturbation equation (3.4.1) is not a RBSDE in the variable $S^i=(Y^i,N^i,K^i)$ because K^i is not necessarily increasing. It could be viewed as a reflected BSDE in the variable $(Y^i,N^i,\overline{K}^i)$ but this point of view will not be used (see the remark after proposition 3.4.4 and remark 3.4.5 after its proof). Also, note that the solution S^1 to the initial, small RBSDE can be viewed as a perturbation : $\overline{S}^1 = 0+S^1$, 0 being the solution to the RBSDE of parameters $(f-f(\cdot,0,0),\nu,g,0,L)$. So it would be enough to study only the pertubation equations, but it seemed clearer to treat first the small reflected BSDEs and then deal with what changes for the perturbation equations.

3.4.2 Existence for small reflected BSDEs.

Underlying problem.

In order to use the fixed point theorem, we need to check that the underlying problem, that is to say the backward stochastic problem that one sees when the drift dV_t is a fixed process and doesn't depend on the solution, defines indeed a map from $\mathcal{S}^\infty \times BMO$ to itself. For reflected BSDEs, as was explained in El Karoui *et al.* [30], the solution is the Snell envelope of a certain process (more precisely, $Y+\int_0^\cdot dV_s$ is the Snell envelope of $L+\int_0^\cdot dV_s$).

Proposition 3.4.1. *Let* $V \in L^1_{BMO}$ *(in the sense that* $\sup_t \| E\big(\int_t^T |dV_s| \big| \mathcal{F}_t\big)\|_\infty < +\infty$*),* $\xi \in L^\infty$*, and* L *be upper bounded. There exist a unique* $(Y, N, K) \in \mathcal{S}^\infty \times BMO \times \mathcal{A}$ *solution to the reflected BSDE :*

$$\begin{cases} dY = -dV - dK + dN \ , \\ Y_T = \xi \ , \\ Y \geq L \ \text{and} \ 1_{\{Y > L\}} dK = 0 \ . \end{cases} \tag{3.4.2}$$

In particular, this applies when $dV_s = dV(y,n)_s = f(s, y_s, z_s\sigma_s)dC_s + d\langle \nu, n^\perp\rangle_s + g_s d\langle n^\perp\rangle_s$*, for* f *satisfying the quadratic growth condition* $(\mathbf{A_{qg}})$*,* $\nu \in BMO$*,* $g \in L^\infty$ *and* $(y, n) \in \mathcal{S}^\infty \times BMO$*.*

Proof. We know from proposition 5.1 in El Karoui *et al.* [30] that Y_t is given by

$$Y_t = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_t^T} E\bigg(\int_t^\tau dV_s + L_\tau 1_{\tau < T} + \xi 1_{\tau = T} \bigg| \mathcal{F}_t \bigg) \ , \tag{3.4.3}$$

where $\mathcal{T}_t^T$ are the stopping times τ such that $t \leq \tau \leq T$, and that the square integrable martingale N and the increasing process K are the Doob-Meyer decomposition of the supermartingale $Y + V$. Our goal is to check that (Y, N) is indeed in $\mathcal{S}^\infty \times BMO$. For an upper bound on Y_t, we have

$$\begin{aligned} E\bigg(\int_t^\tau dV_s + L_\tau 1_{\tau < T} + \xi 1_{\tau = T} \bigg| \mathcal{F}_t \bigg) &\leq E\bigg(\int_t^T |dV_s| \bigg| \mathcal{F}_t \bigg) + E\Big(L_\tau^+ \Big| \mathcal{F}_t \Big) + E\Big(\xi^+ \Big| \mathcal{F}_t \Big) \\ &\leq \|V\|_{L^1_{BMO}} + \|L^+\|_\infty + \|\xi^+\| \end{aligned}$$

for any stopping time τ, so $Y_t \leq \|V\|_{L^1_{BMO}} + \|L^+\|_\infty + \|\xi^+\|$. For a lower bound, since Y solves (3.4.2), and using the fact that K is increasing, we have

$$\begin{aligned} Y_t &= E\bigg(\xi + \int_t^T dV_s + \big(K_T - K_t\big) \bigg| \mathcal{F}_t \bigg) \\ &\geq E\bigg(\xi + \int_t^T dV_s \bigg| \mathcal{F}_t \bigg) \\ &\geq -\|\xi^-\|_\infty - \|V\|_{L^1_{BMO}} \ , \end{aligned}$$

so Y is indeed in $\mathcal{S}^\infty$. One can then invoke remark 3.2.2 after proposition 3.2.1 to conclude that $N \in BMO$.

We now prove the second assertion. For a drift process V of the form described above,

$$|dV_s| \leq |f(s, y_s, z_s\sigma_s)|dC_s + |d\left\langle \nu, n^\perp \right\rangle_s| + |g_s|\,|d\left\langle n^\perp \right\rangle_s|.$$

Using the assumption ($\mathbf{A_{qg}}$) on f and the Kunita-Watanabe inequality, we have, similarly as in proposition 3.2.1,

$$\begin{aligned} E\left(\int_t^T |dV_s|\Big|\mathcal{F}_t\right) &\leq \lambda(\|y\|_{\mathcal{S}^\infty})E\left(\int_t^T h_s^2 + |z_s\sigma_s|^2 dC_s\Big|\mathcal{F}_t\right) \\ &\quad + E\left(\int_t^T |d\left\langle \nu, n^\perp\right\rangle_s|\Big|\mathcal{F}_t\right) + \|g\|_\infty E\left(\int_t^T d\left\langle n^\perp\right\rangle_s \Big|\mathcal{F}_t\right) \\ &\leq \Lambda\|h\|^2_{L^2_{BMO}} + \Big(\Lambda + 1 + \|g\|_\infty\Big)\|n\|^2_{BMO} + \|\nu\|^2_{BMO}, \end{aligned}$$

where $\Lambda = \lambda(\|y\|_{\mathcal{S}^\infty})$. Hence $V \in L^1_{BMO}$ as wanted. □

Existence for RBSDEs with small data.

First one proves that there is a solution when the data are small and when, essentially, the drift is purely quadratic in the solution.

Proposition 3.4.2. *Let $\lambda > 0$. Let f satisfy assumption ($\mathbf{A_{locLipz}}$), with parameters ($\beta = 0, \gamma = 0, \lambda, r$) and be such $\alpha = f(\cdot, 0, 0) \in L^{\infty,1}$ (i.e. : $\int_0^T |\alpha_s| dC_S \in L^\infty$). Let $\nu = 0 \in BMO$ and g be bounded by λ. There exists $\epsilon_0 = \epsilon_0(\lambda, r) > 0$ such that if the size of the data*

$$\boldsymbol{D} = \|\xi\|_\infty + \|f(\cdot, 0, 0)\|_{\infty,1} + \|L^+\|_\infty \leq \epsilon_0\ ,$$

then there exists a solution $S = (Y, N, K) \in \mathcal{S}^\infty \times BMO(P) \times \mathcal{A}$ to the reflected BSDE (3.2.1) with data (V, ξ, L), where $dV(Y,N)_s = f(s, Y_s, Z_s\sigma_s)dC_s + g_s d\langle N^\perp\rangle_s$.

More precisely,

$$\epsilon_0(\lambda, r) = \frac{1}{2^{10}\lambda\Big(\|r\|^2_{\infty,2} + 2\Big)}\ .$$

Also, for any $R \leq R_0(\lambda, r) = \frac{1}{2^5\lambda\left(\|r\|^2_{\infty,2}+2\right)}$, if $\boldsymbol{D} \leq \frac{R}{2^5}$, then this solution is known *to satisfy*

$$\|S\|^2 = \|Y\|^2_{\mathcal{S}^\infty} + \|N\|^2_{BMO(P)} \leq R^2 \ .$$

Proof. We study the map $Sol : \mathcal{S}^\infty \times BMO \to \mathcal{S}^\infty \times BMO$ which sends (y, n) on the solution (Y, N) to the reflected BSDE

$$\begin{cases} dY = -dV(y,n) - dK + dN \ , \\ Y_T = \xi \ , \\ Y \geq 0 \text{ and } 1_{\{Y>0\}} dK = 0 \ , \end{cases} \tag{3.4.4}$$

where $dV(y,n)_s = f(s, y_s, z_s\sigma_s)dC_s + g_s d\langle n^\perp\rangle_s$. This map is well defined according to proposition 3.4.1, and $(Y, N) \in \mathcal{S}^\infty \times BMO$ is a solution of (3.2.1) if and only if it is a fixed point of Sol. It will be seen that Sol is not a contraction on the whole space, but it is on a small ball, and it stabilizes such a small ball if the data are small enough. Therefore there exists at least one fixed point in the space.

We study first the regularity of Sol. Take $s = (y, n)$ and $s' = (y', n')$ in $\mathcal{S}^\infty \times BMO$, write $S = Sol(s)$, $S' = Sol(s')$, and $\Delta x = x' - x$ for a generic quantity x. The semimartingale decomposition of ΔY is $d\Delta Y = -d\Delta V - d\Delta K + d\Delta N$, and the terminal value is 0. Therefore, applying Itô's formula to $(\Delta Y)^2$ between $t \in \mathcal{T}_0^T$ and T, and taking the expectation conditional to $\mathcal{F}_t$ one has, since $\int_0^\cdot \Delta Y d\Delta N$ is a martingale,

$$\begin{aligned} (\Delta Y_t)^2 + E\Big(\int_t^T d\langle \Delta N\rangle_s|\mathcal{F}_t\Big) &= 0^2 + 2E\Big(\int_t^T \Delta Y_s d\Delta V_s|\mathcal{F}_t\Big) \\ &\quad + 2E\Big(\int_t^T \Delta Y_s d\Delta K_s|\mathcal{F}_t\Big) - 0 \ . \end{aligned} \tag{3.4.5}$$

Let us now look at the third term on the right-hand side. Using the fact that $YdK = LdK$ and $Y'dK' = LdK'$ one has

$$\begin{aligned} \Delta Y d\Delta K &= (Y' - Y)dK' - (Y' - Y)dK \\ &= \underbrace{(L - Y)}_{\leq 0}\underbrace{dK'}_{\geq 0} - \underbrace{(Y' - L)}_{\geq 0}\underbrace{dK}_{\geq 0} \leq 0 \ . \end{aligned}$$

Let us now deal with the second term :

$$E\Big(\int_t^T \Delta Y_s d\Delta V_s|\mathcal{F}_t\Big) \leq \|\Delta Y\|_\infty E\Big(\int_t^T |d\Delta V_s||\mathcal{F}_t\Big)\ .$$

The assumption on f gives

$$\begin{aligned}|d\Delta V_s| \leq \lambda\Big(r_s|y_s| + r_s|y'_s| + |z_s\sigma_s| + |z'_s\sigma_s|\Big)\big(r_s|\Delta y_s| + |\Delta z_s\sigma_s|\big)dC_s \\ + |g_s||d\langle\Delta n^\perp, n^\perp + (n')^\perp\rangle_s|.\end{aligned}$$

Consequently, using the Cauchy-Schwartz and the Kunita-Watanabe inequalities, and the elementary inequality $(\sum_{i=1}^n a_i)^2 \leq n\sum a_i^2$, we have

$$\begin{aligned}
&E\bigg(\int_t^T |d\Delta V|\Big|\mathcal{F}_t\bigg)\\
&\quad\leq 2^{\frac{3}{2}}\lambda E\bigg(\int_t^T r_s^2|y_s|^2 + r_s^2|y'_s|^2 + |z_s\sigma_s|^2 + |z'_s\sigma_s|^2\ dC_s\Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\\
&\qquad\qquad\times E\bigg(\int_t^T r_s^2|\Delta y_s|^2 + |\Delta z_s\sigma_s|^2\ dC_s\Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\\
&\qquad + \|g\|_\infty E\bigg(\int_t^T d\left\langle\Delta n^\perp\right\rangle_s\Big|\mathcal{F}_t\bigg)^{\frac{1}{2}} E\bigg(\int_t^T d\left\langle n^\perp + (n')^\perp\right\rangle_s\Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\ .
\end{aligned}$$

Now, by orthogonality, one has $|z_s\sigma_s|^2 dC_s + d\left\langle n^\perp\right\rangle_s = d\left\langle n\right\rangle_s$ so in particular each term on the left-hand side of this equation is smaller than or equal to the right-hand side.

So

$$
\begin{aligned}
&E\Big(\int_t^T |d\Delta V|\Big|\mathcal{F}_t\Big)\\
&\le 2^{\frac{3}{2}}\lambda E\bigg(\int_t^T \Big(r_s^2|y_s|^2+r_s^2|y_s'|^2\Big)dC_s + d\left\langle n\right\rangle_s + d\left\langle n'\right\rangle_s \Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\\
&\qquad\qquad \times E\bigg(\int_t^T r_s^2|\Delta y_s|^2 dC_s + d\left\langle \Delta n\right\rangle_s \Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\\
&\qquad + \|g\|_\infty E\bigg(\int_t^T d\left\langle \Delta n\right\rangle_s \Big|\mathcal{F}_t\bigg)^{\frac{1}{2}} E\bigg(\int_t^T d\left\langle n+(n')\right\rangle_s \Big|\mathcal{F}_t\bigg)^{\frac{1}{2}}\\
&\le 2^{\frac{3}{2}}\lambda\Big(\|r\|_{\infty,2}^2\|y\|_{\mathcal{S}^\infty}^2+\|r\|_{\infty,2}^2\|y'\|_{\mathcal{S}^\infty}^2+\|n\|_{BMO}^2+\|n'\|_{BMO}^2\Big)^{\frac{1}{2}}\\
&\qquad\qquad \times\Big(\|r\|_{\infty,2}^2\|\Delta y\|_{\mathcal{S}^\infty}^2+\|\Delta n\|_{BMO}^2\Big)^{\frac{1}{2}}\\
&\qquad + \|g\|_\infty\|\Delta n\|_{BMO}\ \|n+n'\|_{BMO}\\
&\le 2^{\frac{3}{2}}\lambda\ \big(\|r\|_{\infty,2}^2+1\big)\Big(\|y\|_{\mathcal{S}^\infty}^2+\|n\|_{BMO}^2+\|y'\|_{\mathcal{S}^\infty}^2+\|n'\|_{BMO}^2\Big)^{\frac{1}{2}}\\
&\qquad\qquad \times\Big(\|\Delta y\|_{\mathcal{S}^\infty}^2+\|\Delta n\|_{BMO}^2\Big)^{\frac{1}{2}}\\
&\qquad + \|g\|_\infty\Big(\|n\|_{BMO}+\|n'\|_{BMO}\Big)\|\Delta n\|_{BMO}.
\end{aligned}
$$

Now, by definition of the norm on $\mathcal{S}^\infty\times BMO$, $\|y\|_{\mathcal{S}^\infty}^2+\|n\|_{BMO}^2=\|s\|^2$. Again, this implies in particular that $\|n\|_{BMO}^2\le\|s\|^2$. So, recalling that $\|g\|_\infty\le\lambda$, using $(a^2+b^2)^{\frac{1}{2}}\le a+b$ and majorizing $1\le 2^{\frac{3}{2}}$ (for the 2nd inequality), we have

$$
\begin{aligned}
E\Big(\int_t^T |d\Delta V|\Big|\mathcal{F}_t\Big) &\le 2^{\frac{3}{2}}\lambda\ \big(\|r\|_{\infty,2}^2+1\big)\Big(\|s\|^2+\|s'\|^2\Big)^{\frac{1}{2}}\Big(\|\Delta s\|^2\Big)^{\frac{1}{2}}+\lambda\Big(\|s\|+\|s'\|\Big)\|\Delta s\|\\
&\le 2^{\frac{3}{2}}\lambda\ \big(\|r\|_{\infty,2}^2+1\big)\big(\|s\|+\|s'\|\big)\|\Delta s\|+2^{\frac{3}{2}}\lambda\ \big(\|s\|+\|s'\|\big)\|\Delta s\|\\
&\le 2^{\frac{3}{2}}\lambda\big(\|r\|_{\infty,2}^2+1+1\big)\big(\|s'\|+\|s\|\big)\|\Delta s\|\ .
\end{aligned}
$$

Equation (3.4.5) then yields, using $2ab\le\frac{1}{4}a^2+4b^2$ and $(a+b)^2\le 2(a^2+b^2)$,

$$
\begin{aligned}
(\Delta Y_t)^2+E\Big(\int_t^T d\langle\Delta N\rangle_s\Big|\mathcal{F}_t\Big) \le \frac{1}{4}\|\Delta Y\|_{\mathcal{S}^\infty}^2&\\
+4\times 2^3\lambda^2\big(\|r\|_{\infty,2}^2+2\big)^2\times 2\big(\|s'\|^2+\|s\|^2\big)\|\Delta s\|^2\ ,&
\end{aligned}
$$

and by taking the sup, we finally find, since $\|\Delta Y\|_{\mathcal{S}^\infty} \leq \|\Delta S\|$, that

$$\|\Delta S\|^2 \leq 2^8 \lambda^2 \big(\|r\|^2_{\infty,2} + 2\big)^2 \big(\|s\|^2 + \|s'\|^2\big) \|\Delta s\|^2 \ . \tag{3.4.6}$$

Let us now study the size of $S = Sol(s)$. Following the very same computations and arguments as for ΔS we have first

$$(Y_t)^2 + E\Big(\int_t^T d\langle N\rangle_s|\mathcal{F}_t\Big) \leq \|\xi\|^2_\infty + 2E\bigg(\int_t^T Y_s dV_s\Big|\mathcal{F}_t\bigg) + 2E\bigg(\int_t^T Y_s dK_s\Big|\mathcal{F}_t\bigg). \tag{3.4.7}$$

Since $YdK = LdK$ and, importantly, since K is increasing, one can write

$$\begin{aligned}\int_t^T Y_s dK_s = \int_t^T L_s dK_s &\leq \|L^+\|_{\mathcal{S}^\infty}\big(K_T - K_t\big)\\ &= \|L^+\|_{\mathcal{S}^\infty}\bigg(Y_t - \xi - \int_t^T dV + (N_T - N_t)\bigg) \ ,\end{aligned}$$

so that

$$E\bigg(\int_t^T YdK\Big|\mathcal{F}_t\bigg) \leq \|L^+\|_{\mathcal{S}^\infty}|Y_t| + \|L^+\|_{\mathcal{S}^\infty}\|\xi\|_\infty + \|L^+\|_{\mathcal{S}^\infty} E\bigg(\int_t^T |dV|\Big|\mathcal{F}_t\bigg) + 0 \ .$$

Reinjecting this into (3.4.7), then using the Young inequality, in particular the case $2ab \leq 8a^2 + \frac{1}{8}b^2$, leads to

$$(Y_t)^2 + E\bigg(\int_t^T d\langle N\rangle_s|\mathcal{F}_t\bigg) \leq \Big(2\|\xi\|^2_\infty + 10\|L^+\|^2_{\mathcal{S}^\infty}\Big) + \frac{1}{4}\|Y\|^2_{\mathcal{S}^\infty} + 9E\bigg(\int_t^T |dV|\Big|\mathcal{F}_t\bigg)^2 \ .$$

Now, by the assumption on f,

$$\begin{aligned}|dV_s| &\leq \Big[f(s,0,0) + \lambda\big(r_s|y_s| + |z_s\sigma_s|\big)^2\Big]dC_s + |g_s|d\langle n^\perp\rangle_s\\ &\leq \Big[f(s,0,0) + 2\lambda\big(r_s^2|y_s|^2 + |z_s\sigma_s|^2\big)\Big]dC_s + |g_s|d\langle n^\perp\rangle_s\end{aligned}$$

so, by the same argumentation as for ΔV above,

$$\begin{aligned}E\Big(\int_t^T |dV|\Big|\mathcal{F}_t\Big) &\le \|f(\cdot,0,0)\|_{\infty,1} + 2\lambda\big(\|r\|_{\infty,2}^2+1\big)\|s\|^2 + \lambda\|s\|^2 \\ &\le \|f(\cdot,0,0)\|_{\infty,1} + 2\lambda\big(\|r\|_{\infty,2}^2+2\big)\|s\|^2.\end{aligned}$$

Consequently, after taking $\sup_t$ and using $\|Y\|_{\mathcal{S}^\infty} \le \|S\|$, one has

$$\begin{aligned}\|S\|^2 \le &\Big(4\|\xi\|_\infty^2 + 20\|L^+\|_{\mathcal{S}^\infty}^2\Big) + \frac{1}{2}\|S\|^2 \\ &+ 18\times 2\times\Big[\|f(\cdot,0,0)\|_{\infty,1}^2 + 2^2\lambda^2\big(\|r\|_{\infty,2}^2+2\big)^2\|s\|^4\Big]\ .\end{aligned}$$

Collecting the terms in $\|S\|^2$ and majorizing largely one has finally

$$\|S\|^2 \le 2^9\boldsymbol{D}^2 + 2^9\lambda^2\big(\|r\|_{\infty,2}^2+2\big)^2\|s\|^4\ , \tag{3.4.8}$$

where $\boldsymbol{D} = \|\xi\|_\infty + \|L^+\|_{\mathcal{S}^\infty} + \|f(\cdot,0,0)\|_{\infty,1}$ and we used $a^2+b^2+c^2 \le (a+b+c)^2$.

To have Sol be a contraction on a closed (and therefore complete) ball $\overline{B}(0,R)$ of $\mathcal{S}^\infty \times BMO$, we see from (3.4.6) and (3.4.8) that we would like the radius R and the size $\boldsymbol{D}$ of the data to be sufficiently small so that $2^9\lambda^2\big(\|r\|_{\infty,2}^2+2\big)^2R^2 \le \frac{1}{2}(<1)$ and $2^9\boldsymbol{D}^2 + 2^9\lambda^2\big(\|r\|^2+2\big)^2R^4 \le R^2$. This is the case as soon as

$$\begin{aligned}R &\le R_0(\lambda,r) := \frac{1}{2^5\lambda\big(\|r\|_{\infty,2}^2+2\big)} \\ \boldsymbol{D} &\le \frac{R}{2^5} \le \frac{R_0(\lambda,r)}{2^5} =: \epsilon_0(\lambda,r)\ .\end{aligned}$$

□

We now remove the assumption that the linear terms in the drift are null.

Proposition 3.4.3. *Let $\lambda > 0$. Let f satisfy assumption ($\mathbf{A_{locLipz}}$), with parameters (β,γ,λ,r) and be such that $\alpha = f(\cdot,0,0) \in L^{\infty,1}$ (i.e. : $\int_0^T |f(s,0,0)|dC_s \in L^\infty$). Let $\nu \in BMO$ and g be bounded by λ. There exists $\epsilon_0 = \epsilon_0(\beta,\lambda,r) > 0$ such that if the size of the data*

$$\boldsymbol{D} = \|\xi\|_\infty + \|f(\cdot,0,0)\|_{\infty,1} + \|L^+\|_\infty \le \epsilon_0\ ,$$

then there exists a solution $S = (Y,N,K) \in \mathcal{S}^\infty \times BMO(P) \times \mathcal{A}$ to the reflected BSDE

(3.2.1) with data (f, ν, g, ξ, L).

More precisely,

$$\epsilon_0(\beta, \lambda, r) = \frac{e^{-2\|\beta\|_{\infty,1}}}{2^{10}\lambda\big(\|r\|^2_{\infty,2} + 2\big)} \ .$$

Also, for any $R \leq R_0(\widehat{\lambda}, r) = \frac{1}{2^5\widehat{\lambda}\big(\|r\|^2_{\infty,2}+2\big)}$, where $\widehat{\lambda} = \exp\big(\|\beta\|_{\infty,1}\big)\lambda$, if $\boldsymbol{D} \leq \exp(-\|\beta\|_{\infty,1})\frac{R}{2^5}$, then this solution is known to satisfy

$$\|\widehat{S}\|^2_Q = \|\widehat{Y}\|^2_{\mathcal{S}^\infty} + \|\widehat{\widetilde{N}}\|^2_{BMO(Q)} \leq R^2 \ ,$$

where $\widehat{Y}_t = e^{\int_0^t \beta_u dC_u} Y_t$ and $\widehat{\widetilde{N}}$ is the martingale part of $\widehat{Y}$ under $Q : \frac{dQ}{dP} = \mathcal{E}(\int \gamma\sigma^{-1}dM + \nu)$.

Proof. Write $f(t,y,z) = \beta_t y + \gamma_t z + h(t,y,z)$, where $\beta_t = f_y(t,0,0)$ and $\gamma_t = f_z(t,0,0)$ (so that $h(t,0,0) = f(t,0,0) = \alpha_t$). Note that h satisfies ($\mathbf{A_{locLipz}}$) with parameters $(\beta = 0, \gamma = 0, \lambda, r)$.

The idea is that if (Y, N, K) is a solution to the reflected BSDE (3.2.1), one can eliminate the linear terms $(\beta_t Y_t + \gamma_t Z_t \sigma_t)dC_t + d\left\langle \nu, N^\perp \right\rangle_t$ in the drift $dV(Y,N)_t$ by a pair of transforms and obtain a reflected BSDE with purely quadratic drift. Proposition 3.4.2 guarantees the existence of a solution to such a RBSDE, so undoing the transforms yields a solution to (3.2.1).

In view of this, let us define the measure Q by $\frac{dQ}{dP} = \mathcal{E}(L)$ where $L = \int \gamma\sigma^{-1}dM + \nu$. Then $\widetilde{M} := M - \langle L, M\rangle = M - \int \gamma\sigma^* dC$ is a $BMO(Q)$-martingale. Define also $B = \exp\Big(\int_0^\cdot \beta_u dC_u\Big)$, which is a bounded process. Define the transformed data

$$\begin{aligned}
&\widehat{h}(s,y,z) = B_s h(s, B_s^{-1}y, B_s^{-1}z) \ , \\
&\widehat{g}_s = B_s^{-1} g_s \ , \\
&\widehat{\xi} = B_T \xi \ , \\
&\widehat{L} = BL \ .
\end{aligned}$$

Note that $\widehat{h}$ satisfies ($\mathbf{A_{locLipz}}$) with parameters $(\beta = 0, \gamma = 0, \widehat{\lambda}, r)$ where $\widehat{\lambda} = \lambda\exp(\|\beta\|_{\infty,1})$. Proposition 3.4.2 ensures the existence of a solution $(\widehat{Y}, \widehat{\widetilde{N}}, \widehat{K}) \in \mathcal{S}^\infty \times BMO(Q) \times \mathcal{A}$ under Q to the reflected BSDE (3.2.1) with transformed data $(\widehat{h}, \nu =$

$0, \widehat{g}, \widehat{\xi}, \widehat{L})$. Indeed,

$$\begin{aligned}
&\|\widehat{g}\|_\infty \leq \exp(\|\beta\|_{\infty,1})\|g\|_\infty \leq \exp(\|\beta\|_{\infty,1})\lambda = \widehat{\lambda} < +\infty \ , \\
&\|\widehat{h}(\cdot,0,0)\|_{\infty,1} \leq \exp(\|\beta\|_{\infty,1})\|f(\cdot,0,0)\|_{\infty,1} < +\infty \ , \\
&\|\widehat{\xi}\|_\infty \leq \exp(\|\beta\|_{\infty,1})\|\xi\|_\infty < +\infty \ , \\
&\|\widehat{L^+}\|_{\mathcal{S}^\infty} \leq \exp(\|\beta\|_{\infty,1})\|L^+\|_{\mathcal{S}^\infty} < +\infty \ ,
\end{aligned}$$

so if $\boldsymbol{D} \leq \exp(-\|\beta\|_{\infty,1})\epsilon_0(\widehat{\lambda}, r) = \exp(-2\|\beta\|_{\infty,1})\epsilon_0(\lambda, r)$, proposition 3.4.2 applies.

Now, define $Y = B^{-1}\widehat{Y}$, $\widetilde{N} = \int_0^\cdot B^{-1} d\widehat{\widetilde{N}} = \int \widetilde{Z} d\widetilde{M} + \widetilde{N}^\perp$ and $K = \int_0^\cdot B^{-1} d\widehat{K}$. The Girsanov $(Q \to P)$-transform of $\widetilde{N}$,

$$\begin{aligned}
N &= \widetilde{N} + \left\langle L, \widetilde{N} \right\rangle \\
&= \int \widetilde{Z} d\widetilde{M} + \widetilde{N}^\perp + \int \gamma \widetilde{Z} \sigma dC + \left\langle \nu, \widetilde{N}^\perp \right\rangle \\
&= \int \widetilde{Z} dM + N^\perp \ ,
\end{aligned}$$

is a $BMO(P)$-martingale. Y is a bounded semimartingale, since B^{-1} is bounded, and differentiating $Y = B^{-1}\widehat{Y}$ shows that (Y, N, K) is a solution to the reflected BSDE (3.2.1) with data (f, ν, g, ξ, L), as we wanted. □

3.4.3 Perturbation of a reflected BSDE.

We now deal with existence for perturbation equations like (3.4.1). We assume we have a solution $S^1 = (Y^1, N^1, K^1)$ to a reflected BSDE with data (f, ν, g, ξ^1, L^1), and want to construct a solution $\overline{S}^2$ to a reflected BSDE with slightly different data $(f + \alpha^2, \nu, g, \xi^1 + \xi^2, L^1 + L^2)$. The idea is to construct the difference $S^2 = (Y^2, N^2, K^2) = \overline{S}^2 - S^1$. The next proposition shows how this can be done despite the fact that K^2 is not an increasing process, so long as one does not change the obstacle ($L^2 = 0$).

Proposition 3.4.4. *Let f satisfy $(\mathbf{A_{der}})$ with parameters $(\rho, \rho', \lambda, r, h)$ and be such that $\alpha = f(\cdot, 0, 0) \in L^{\infty,1}$, let $g \in L^\infty$ be bounded by λ and $\nu \in BMO$. Let also $\xi^1 \in L^\infty$ and L^1 be upper bounded. Assume that there exists a solution $S^1 = (Y^1, N^1, K^1)$ to the RBSDE (3.2.1) with data (f, g, ν, ξ^1, L^1). Now let $\xi^2 \in L^\infty$ and $\alpha^2 \in L^{\infty,1}$ (and*

$L^2 = 0$). If

$$\delta \boldsymbol{D} = \|\xi^2\|_\infty + \|\alpha^2\|_{\infty,1} \leq \epsilon_0(\rho, 2\lambda, r) = \frac{e^{-2\rho\|r\|^2_{\infty,2}}}{2^{10}(2\lambda)\big(\|r\|^2_{\infty,2} + 2\big)},$$

then there exist $S^2 = (Y^2, N^2, K^2)$ where $Y^2 \in \mathcal{S}^\infty$, $N^2 \in BMO(P)$ and K^2 has finite variation, solving the perturbation equation

$$\begin{cases} dY^2 = -dV^2(Y^2, N^2) - dK^2 + dN^2 \ , \\ \ Y^2_T = \xi^2 \ , \\ \ Y^1 + Y^2 \geq L^1 + L^2 \ , \\ dK^1 + dK^2 \geq 0 \ \textit{and} \ 1_{\{Y^1+Y^2>L^1+L^2\}}(dK^1 + dK^2) = 0 \end{cases} \tag{3.4.9}$$

with drift given by

$$\begin{aligned} dV^2(Y^2, N^2)_s = &\big[f(s, Y^2_s + Y^1_s, Z^2_s\sigma_s + Z^1_s\sigma_s) - f(s, Y^1_s, Z^1_s\sigma_s) + \alpha^2_s\big]dC_s \\ &+ d\Big\langle \nu + \int 2gd(N^1)^\perp, (N^2)^\perp \Big\rangle_s + g_s d\Big\langle (N^2)^\perp \Big\rangle_s . \end{aligned}$$

So $\overline{S}^2 := S^1 + S^2$ is a solution to the RBSDE (3.2.1) with data $(f + \alpha^2, g, \nu, \xi^1 + \xi^2, L^1)$.

We further know that for any $R \leq R_0(\widehat{2\lambda}, r) = \frac{1}{2^5\widehat{2\lambda}\left(\|r\|^2_{\infty,2}+2\right)}$, where $\widehat{2\lambda} = 2\lambda \exp\left(\rho\|r\|^2_{\infty,2}\right)$, if $\delta\boldsymbol{D} \leq \exp(-\rho\|r\|^2_{\infty,2})\frac{R}{2^5}$, then this solution satisfies

$$\|\widehat{S^2}\|^2_Q = \|\widehat{Y^2}\|^2_{\mathcal{S}^\infty} + \|\widehat{\widehat{N^2}}\|^2_{BMO(Q)} \leq R^2 \ .$$

Note that while $S^2 = (Y^2, N^2, K^2)$ is not the solution to a reflected BSDE, $(Y^2, N^2, \overline{K}^2)$ is. However the drift there would be $dV^2(Y^2, N^2)_s - dK^1_s$, whose residual action (when $(Y^2, N^2) = (0, 0)$) is $\alpha^2_s dC_s - dK^1_s$, and this has no reason to be small. We can therefore not simply invoke proposition 3.4.3 to construct $(Y^2, N^2, \overline{K}^2)$ and we need to argue further.

Proof. The majority of computations that would need to be done here, related to the dynamics of Y^2, are very similar to those in the proposition 3.4.3 about small RBSDEs, so we only do the part which is different.

Define $\overline{f}(s, y, z) = f(s, y+Y^1_s, z+Z^1_s\sigma_s) - f(s, Y^1_s, Z^1_s\sigma_s) + \alpha^2_s$. Note that since f satisfies ($\mathbf{A_{der}}$), $\overline{f}$ satisfies ($\mathbf{A_{locLip}}$) with parameters $(\overline{\beta}, \overline{\gamma}, 2\lambda, r)$, where $\overline{\beta} = f_y(\cdot, Y^1, Z^1\sigma)$ and $\overline{\gamma} = f_z(\cdot, Y^1, Z^1\sigma)$. We have $\|\overline{\beta}\|_{\infty,1} \leq \rho \|r\|^2_{\infty,2} < +\infty$ and $\overline{\gamma} \in L^2_{BMO}$.

Following the same approach as for RBSDEs, we first look at the underlying problem of finding $S^2 = (Y^2, N^2, K^2)$ solving the perturbation equation (3.4.9) when the drift process is

$$\begin{aligned} dV^2_s = dV^2(y^2, n^2)_s = \big[f(s, y^2_s + Y^1_s, z^2_s\sigma_s + Z^1_s\sigma_s) - f(s, Y^1_s, Z^1_s\sigma_s) + \alpha^2_s\big]dC_s \\ +d\Big\langle \nu + \int 2gd(N^1)^\perp, (n^2)^\perp \Big\rangle_s + g_s d\Big\langle (n^2)^\perp \Big\rangle_s . \end{aligned}$$

If S^2 is a solution, $\overline{S}^2 = S^1 + S^2$ is then solution to the reflected BSDE (3.4.2) with drift process given by $d\overline{V}^2_s = dV^1(Y^1, N^1)_S + dV^2(y^2, n^2)_s = \big[f(s, y^2_s + Y^1_s, z^2_s\sigma_s + Z^1_s\sigma_s) + \alpha^2_s\big]dC_s + d\langle \nu, (N^1)^\perp + (n^2)^\perp \rangle_s + g_s d\langle (N^1)^\perp + (n^2)^\perp \rangle_s$. But proposition 3.4.1 guarantees the existence and uniqueness of such an $\overline{S}^2$, hence that of the sought S^2. This allows to define a map Sol' from $\mathcal{S}^\infty \times BMO$ to itself.

Now, to find a solution S^2 to the perturbation equation (3.4.9), we proceed like in propositions 3.4.2 and 3.4.3, the difference being in dealing with dK^2 which is not monotonous anymore here. Up to doing the usual transformations (proposition 3.4.3), let us assume that the drift is purely quadratic as in proposition 3.4.2. Then, Itô's formula first leads to the estimates

$$\begin{aligned} |\Delta Y^2_t|^2 + E\bigg(\int_t^T d\left\langle \Delta N^2 \right\rangle_s \Big| \mathcal{F}_t \bigg) \leq 2E\bigg(\int_t^T \Delta Y^2_s d\Delta V^2_s \Big| \mathcal{F}_t \bigg) + 2E\bigg(\int_t^T \Delta Y^2_s d\Delta K^2_s \Big| \mathcal{F}_t \bigg), \\ |Y^2_t|^2 + E\bigg(\int_t^T d\left\langle N^2 \right\rangle_s \Big| \mathcal{F}_t \bigg) \leq \|\xi^2\|^2_\infty + 2E\bigg(\int_t^T Y^2_s dV^2_s \Big| \mathcal{F}_t \bigg) + 2E\bigg(\int_t^T Y^2_s dK^2_s \Big| \mathcal{F}_t \bigg) . \end{aligned}$$

For the term in $\Delta Y^2 d\Delta K^2$ one has (even if $L^2 \neq 0$)

$$\begin{aligned} \Delta Y^2 d\Delta K^2 &= \big((Y^2)' - Y^2\big)d(K^2)' - \big((Y^2)' - Y^2\big)dK^2 \\ &= \big((\overline{Y}^2)' - \overline{Y}^2\big)\big(d(\overline{K}^2)' - dK^1\big) - \big((\overline{Y}^2)' - \overline{Y}^2\big)\big(d\overline{K}^2 - dK^1\big) \\ &= \big((\overline{Y}^2)' - \overline{Y}^2\big)d(\overline{K}^2)' - \big((\overline{Y}^2)' - \overline{Y}^2\big)d\overline{K}^2 \\ &= \underbrace{\big(\overline{L}^2 - \overline{Y}^2\big)}_{\leq 0} \underbrace{d(\overline{K}^2)'}_{\geq 0} - \underbrace{\big((\overline{Y}^2)' - \overline{L}^2\big)}_{\geq 0} \underbrace{d\overline{K}^2}_{\geq 0} \leq 0. \end{aligned}$$

For the term in $Y^2 dK^2$ one has however, since $L^2 = 0$,

$$\begin{aligned}
Y^2 dK^2 &= (Y^2 - L^2) dK^2 + L^2 dK^2 \\
&= \big((\overline{Y}^2 - Y^1) - (\overline{L}^2 - L^1)\big) dK^2 + L^2 dK^2 \\
&= \big((\overline{Y}^2 - \overline{L}^2) - (Y^1 - L^1)\big)(d\overline{K}^2 - dK^1) + L^2 dK^2 \\
&= \underbrace{(\overline{Y}^2 - \overline{L}^2) d\overline{K}^2}_{=0} - \underbrace{(\overline{Y}^2 - \overline{L}^2) dK^1}_{\geq 0} - \underbrace{(Y^1 - L^1) d\overline{K}^2}_{\geq 0} + \underbrace{(Y^1 - L^1) dK^1}_{=0} + L^2 dK^2 \\
&\leq L^2 dK^2 = 0.
\end{aligned}$$

Having observed this, the rest is like the analysis of the map Sol and the ϵ_0 is the same. So in the end, provided that that

$$\delta \boldsymbol{D} = \|\xi^2\|_\infty + \|\alpha^2\|_{\infty,1} \leq \epsilon_0(\rho, 2\lambda, r) = \frac{e^{-2\rho|r^2|}}{2^{10}(2\lambda)\big(\|r\|^2_{\infty,2} + 2\big)},$$

there exists a solution (Y^2, N^2, K^2) to the perturbation equation (3.4.9). $\square$

Remark 3.4.5. Note that uniqueness holds for the perturbation equations. First, under $(\mathbf{A_{der}})$, $(\mathbf{A_{lLip}})$ holds and so does uniqueness for reflected BSDEs. Then, one can argue that if Y^2 and $(Y^2)'$ are two solutions to (3.4.9), then $\overline{Y}^2 = Y^1 + Y^2$ and $(\overline{Y}^2)' = Y^1 + (Y^2)'$ are two solutions to the same reflected BSDE, so $\overline{Y}^2 = (\overline{Y}^2)'$, and therefore $Y^2 = (Y^2)'$. Alternatively we can also argue that if (Y^2, N^2, K^2) is a solution to (3.4.9), then $(Y^2, N^2, \overline{K}^2)$ solves a reflected BSDE (3.2.1) for which uniqueness holds.

3.4.4 Existence theorem.

We can now prove the existence theorem of this section.

Theorem 3.4.6. *Let f satisfy $(\mathbf{A_{der}})$ with parameters $(\rho, \rho', \lambda, r, h)$ and be such that $f(\cdot, 0, 0) \in L^{\infty,1}$. Let $\nu \in BMO$, $g \in L^\infty$ be bounded by λ, $\xi \in L^\infty$, and L be upper bounded. There exists a solution $(Y, N, K) \in \mathcal{S}^\infty \times BMO \times \mathcal{A}$ to the RBSDE (3.2.1) with data (f, g, ν, ξ, L).*

Proof. The proof is done in two steps. First, we show that one can indeed reduce the problem to the case $L \leq 0$, by translation. Existence for the RBSDE with $L \leq 0$

is then proved by repeatedly perturbing a solution to a similar RBSDE with smaller data.

Step 1. If (Y, N, K) is a solution to the RBSDE, and U is an upper bound for L, set $\overrightarrow{Y} = Y - U$. We see that

$$\begin{cases} d\overrightarrow{Y} = dY - dU = -dV - dK + dN - 0 = -d\overrightarrow{V} - dK + dN \ , \\ \overrightarrow{Y}_T = \xi - U =: \overrightarrow{\xi} \ , \\ \overrightarrow{Y} = Y - U \geq L - U =: \overrightarrow{L} \ , \\ 1_{\{\overrightarrow{Y} > L-U\}} dK = 1_{\{Y>L\}} dK = 0 \ . \end{cases}$$

Here we defined

$$\begin{aligned} d\overrightarrow{V} &= dV(Y, N) \\ &= dV(\overrightarrow{Y} + U, N) \\ &= f(s, \overrightarrow{Y} + U, Z\sigma) dC + d\langle \nu, N^\perp \rangle + g d\langle N^\perp \rangle \\ &= \overrightarrow{f}(s, \overrightarrow{Y}, Z\sigma) dC + d\langle \nu, N^\perp \rangle + g d\langle N^\perp \rangle \ . \end{aligned}$$

It is clear that $\overrightarrow{f}$ still satisfies $(\mathbf{A_{der}})$ with parameters $(\rho, \rho', \lambda, r, h)$. And from the assumption on f_y one has

$$|\overrightarrow{f}(s, 0, 0)| = |f(s, U, 0)| \leq |f(s, 0, 0)| + \rho r_s^2 U \ ,$$

so $\overrightarrow{\alpha} = \overrightarrow{f}(\cdot, 0, 0) \in L^{\infty,1}$.

In the end, $(\overrightarrow{Y}, N, K) \in \mathcal{S}^\infty \times BMO \times \mathcal{A}$ is a solution to the reflected BSDE of parameters $(\overrightarrow{f}, \nu, g, \overrightarrow{\xi}, \overrightarrow{L})$ satisfying the same assumptions, but with $\overrightarrow{L} \leq 0$.

Step 2. We now focus on the case $L \leq 0$. Consider ϵ_0 given by proposition 3.4.4. For $n \in \mathbb{N}^*$, we define $\xi^{(n)} = \frac{1}{n}\xi$ and $\alpha^{(n)} = \frac{1}{n}\alpha$ (where $\alpha = f(\cdot, 0, 0)$). We split the data uniformly, that is we consider $\xi^i = \xi^{(n)}$ and $\alpha^i = \alpha^{(n)}$ for all $i \in \{1, \ldots, n\}$. We choose n big enough so that one has $\boldsymbol{D}^{(n)} := \|\xi^{(n)}\|_\infty + \|\alpha^{(n)}\|_{\infty,1} = \frac{1}{n}\boldsymbol{D} \leq \epsilon_0$.

First, by proposition 3.4.3, there exists a solution (Y^1, N^1, K^1) to the RBSDE (3.2.1) with data $(f - \alpha + \alpha^1, \nu, g, \xi^1, L)$.

Next, for $i = 2$ to n, having obtained a solution $(\overline{Y}^{i-1}, \overline{N}^{i-1}, \overline{K}^{i-1})$ to the RBSDE

(3.2.1) with data $(f-\alpha+\overline{\alpha}^{i-1},\nu,g,\overline{\xi}^{i-1},L)$, proposition 3.4.4 provides a solution (Y^i,N^i,K^i) to the perturbation equation (3.4.1) and therefore a solution $(\overline{Y}^i,\overline{N}^i,\overline{K}^i)$ to the RBSDE (3.2.1) with parameters $(f-\alpha+\overline{\alpha}^i,\nu,g,\overline{\xi}^i,L)$. For $i=n$, since $\overline{\xi}^n=\xi$ and $\overline{\alpha}^n=\alpha$, $(\overline{Y}^n,\overline{Z}^n,\overline{K}^n)$ is a solution to the RBSDE of interest, which ends the proof. $\square$

3.4.5 Stability in $\mathcal{S}^\infty\times BMO$.

Given that uniqueness holds, the a posteriori bounds that come with the construction of a perturbation $\delta S=S'-S$ to a solution S in proposition 3.4.4 readily shows the continuity of the map $(\xi,\alpha)\mapsto(Y,N)$, from $L^\infty\times L^{\infty,1}$ to $\mathcal{S}^\infty\times BMO$.

We now derive an estimate which shows that it is locally Lipchitz, by a sort of bootstrap argument on the above stability result, as well as a BMO-norm equivalence. In the proposition below, we consider a fixed set of data (f,ν,g,ξ,L) and the associated solution $S=(Y,N,K)$, and we define $\alpha=f(\cdot,0,0)$. Now, for close data $(f+\delta\alpha',\nu,g,\xi',L)$ and $(f+\delta\alpha'',\nu,g,\xi'',L)$, we consider the solutions S' and S''. Set $\delta\xi'=\xi'-\xi$ and $\delta\xi''=\xi''-\xi$. We use the notation $\delta S'=S'-S$, $\delta S''=S''-S$ for the perturbations around S and $\Delta S=S''-S'=\delta S''-\delta S'$. What we show is that if $(\delta\xi',\delta\alpha')$ and $(\delta\xi'',\delta\alpha'')$ are sufficiently small, the distance $\|\Delta S\|$ is linearly controlled by the distance $\Delta\boldsymbol{D}=\|\Delta\xi\|_\infty+\|\Delta\alpha\|_{\infty,1}=\|\xi''-\xi'\|_\infty+\|\delta\alpha''-\delta\alpha'\|_{\infty,1}$. That is, $(\xi',\alpha')\mapsto(Y',N')$ is locally Lipschitz at the point (ξ,α).

Proposition 3.4.7. *Suppose that f satisfies $(\mathbf{A_{der}})$ with parameters (ρ,ρ',λ,r,h), that $\alpha=f(\cdot,0,0)\in L^{\infty,1}$, that $\nu\in BMO$, that g is bounded by λ and that L is upper bounded. We consider $\xi\in L^\infty$ and the solution (Y,N,K) to the reflected BSDE of parameters (f,ν,g,ξ,L).*

Now, for any $(\xi',\delta\alpha')$ and $(\xi'',\delta\alpha'')\in L^\infty\times L^{\infty,1}$, let $S'=(Y',N',K')$ and $S''=(Y'',N'',K'')$ be the solutions to the reflected BSDEs of parameters $(f+\delta\alpha',\nu,g,\xi',L)$ and $(f+\delta\alpha'',\nu,g,\xi'',L)$ respectively.

If $\delta\boldsymbol{D}'=\|\delta\xi^{'}\|_\infty+\|\delta\alpha^{'}\|_{\infty,1}$ and $\delta\boldsymbol{D}''=\|\delta\xi^{''}\|_\infty+\|\delta\alpha^{''}\|_{\infty,1}$ satisfy

$$\delta\boldsymbol{D}' \text{ and } \delta\boldsymbol{D}''\leq\frac{1}{\sqrt{2}}\epsilon_0(\overline{\beta},2\lambda,r)=\frac{1}{\sqrt{2}}\frac{e^{-2\|\overline{\beta}\|_{\infty,1}}}{2^{10}(2\lambda)\big(\|r\|^2_{\infty,2}+2\big)}\ ,$$

where $\overline{\beta} = f_y(\cdot, Y, Z\sigma)$*, then we have*

$$\|Y'' - Y'\|_{\mathcal{S}^\infty} \leq 2^5 e^{2\|\overline{\beta}\|_{\infty,1}} \big(\|\xi'' - \xi'\|_\infty + \|\alpha'' - \alpha'\|_{\infty,1}\big) \qquad \textit{and}$$
$$\|N'' - N'\|_{BMO(P)} \leq 2^5 C(Y,N)\, e^{2\|\overline{\beta}\|_{\infty,1}} \big(\|\xi'' - \xi'\|_\infty + \|\alpha'' - \alpha'\|_{\infty,1}\big) \ ,$$

where $C(Y,N)$ *is a constant depending on* (Y,N)*.*

Proof. We know that $\overline{f}(s,\delta y,\delta z) := f(s, Y_s + \delta y, Z_s\sigma_s + \delta z) - f(s, Y_s, Z_s\sigma_s)$ satisfies $(\mathbf{A_{locLip}})$ with parameters $(\overline{\beta}, \overline{\gamma}, 2\lambda, r)$, where $\overline{\beta} = f_y(\cdot, Y, Z\sigma)$ and $\overline{\gamma} = f_z(\cdot, Y, Z\sigma)$ satisfy $\|\overline{\beta}\|_{\infty,1} \leq \rho \|r\|^2_{\infty,2}$ and $\overline{\gamma} \in L^2_{BMO}$; and $\overline{\nu} = \nu + \int 2g dN \in BMO$. We linearize $\overline{f}$ like in proposition 3.4.3 : $\overline{f}(s,\delta y,\delta z) = \overline{\beta}_s \delta y + \overline{\gamma}_s \delta z + \overline{h}(s,\delta y,\delta z)$.

Since the difference $\Delta Y = Y'' - Y'$ has the dynamics

$$\begin{aligned} d\Delta Y_s = &-\big[\Delta\alpha_s + \overline{\beta}_s \Delta Y_s + \overline{\gamma}_s \Delta Z_s \sigma_s + \{\overline{h}(s, \delta Y''_s, \delta Z''_s\sigma_s) - \overline{h}(s, \delta Y', \delta Z'_s\sigma_s)\}\big] dC_s \\ &- d\left\langle \overline{\nu}, (\Delta N)^\perp \right\rangle_s - g_s d\left\langle (\delta N'')^\perp + (\delta N')^\perp, (\Delta N)^\perp \right\rangle_s - d\Delta K_s + d\Delta N_s \ , \end{aligned}$$

doing the usual transformations, with $\frac{dQ}{dP} = \mathcal{E}(\int \overline{\gamma}\sigma^{-1} dM + \overline{\nu})$ and $\overline{B} = e^{\int_0^\cdot \overline{\beta}_u dC_u}$, the standard computations give, like in (3.4.6),

$$\|\widehat{\Delta S}\|^2_Q \leq 2^9 \widehat{\Delta E} + 2^9 \widehat{2\lambda}^2 (\|r\|^2_{\infty,2} + 2)^2 \big(\|\widehat{\delta S''}\|^2_Q + \|\widehat{\delta S'}\|^2_Q\big) \|\widehat{\Delta S}\|^2_Q$$

But $\delta S''$ and $\delta S'$ are the unique solutions to the perturbation equations (3.4.9) with data $(\overline{f} + \delta\alpha'', \overline{\nu}, g, \delta\xi'', L)$ and $(\overline{f} + \delta\alpha', \overline{\nu}, g, \delta\xi', L)$, and by the way they were constructed in proposition 3.4.4 (recall that $\delta\boldsymbol{D}', \delta\boldsymbol{D}'' \leq \exp(-\|\overline{\beta}\|_{\infty,1}) \frac{1}{2^5} \frac{R_0(\widehat{2\lambda}, r)}{\sqrt{2}}$) we know that they satisfy

$$\|\widehat{\delta S''}\|^2_Q,\ \|\widehat{\delta S'}\|^2_Q \leq \frac{R_0(\widehat{2\lambda}, r)^2}{2} \ , \qquad \text{so}$$
$$\|\widehat{\delta S''}\|^2_Q + \|\widehat{\delta S'}\|^2_Q \leq R_0(\widehat{2\lambda}, r)^2 = \frac{1}{2^{10} \widehat{2\lambda}^2 (\|r\|^2 + 2)^2} \ .$$

Reinjecting this in the previous estimate we have $\|\widehat{\Delta S}\|^2_Q \leq 2^9 \widehat{\Delta E} + \frac{1}{2} \|\widehat{\Delta S}\|^2_Q$ and therefore

$$\|\widehat{\Delta S}\|^2_Q = \|\widehat{\Delta Y}\|^2_{\mathcal{S}^\infty} + \|\widehat{\widehat{\Delta N}}\|^2_{BMO(Q)} \leq 2^{10} \widehat{\Delta E} \ .$$

Then this implies that $\|\widehat{\Delta Y}\|_{\mathcal{S}^\infty} \leq 2^5 \widehat{\Delta \boldsymbol{D}}$ and so $\|\Delta Y\|_{\mathcal{S}^\infty} \leq 2^5 e^{2\|\overline{\beta}\|_{\infty,1}} \Delta\boldsymbol{D}$. For the same reason, $\|\widetilde{\Delta N}\|_{BMO(Q)} \leq 2^5 e^{2\|\overline{\beta}\|_{\infty,1}} \Delta \boldsymbol{D}$. By theorem 3.6 in Kazamaki, $\|\Delta N\|_{BMO(P)} \leq C(Q)\|\widetilde{\Delta N}\|_{BMO(Q)}$ where the constant depends only on Q, or equivalently on the martingale $\int \overline{\gamma}\sigma^{-1}dM + \overline{\nu}$, and *in fine* on (Y, N). □

Note that the interesting part of the above result is the martingale estimate. Indeed, the estimate for $Y'' - Y'$ in $\mathcal{S}^\infty$ actually holds for any size of data (as can be seen by linearizing the drift, doing a change of measure to get rid of all the terms in N and solving for Y). As mentionned in the introduction, we know that $(\xi, \alpha) \mapsto N$ is global Lipschitz in $\mathcal{H}^p$, and $\frac{1}{2}$-Hölder in BMO. The above estimate shows it is in fact locally Lipschitz in BMO.

3.5 Existence under more general assumptions.

In theorem 3.4.6, the existence of a solution was proved under $(\mathbf{A_{der}})$, so in particular under the assumption that f is a Lipschitz function of y, and therefore at most linear in y. In this section, we extend this result to more general assumptions on f.

To some extent, we would like to replace ρ, ρ', λ which are constants in $(\mathbf{A_{der}})$ by arbitrary growth functions (while of course still assuming that f ends up with a growth in y compatible with existence of solutions). Looking back at proposition 3.4.4, we see that when ρ is a growth function, the maximal size ϵ allowed for a perturbation (ξ^2, α^2) of the parameters would depend on the size $\|Y^1\|_{\mathcal{S}^\infty}$ of the solution. It is therefore not clear that one can choose ϵ_0 and the decomposition $\xi = \sum_{i=1}^n \xi^i$, $\alpha = \sum_{i=1}^n \alpha^i$ uniformly for the perturbation procedure in the proof of theorem 3.4.6, or to put things differently, that a series of pertubations could terminate in finitely many steps. This however can be guaranteed if one can obtain an *a priori* bound for the solutions to reflected BSDEs with drift (f, ν, g).

Case of a superlinear growth in y.
In the following theorem, we extend theorem 3.4.6 to the case where f can have slightly-superlinear growth in y.

Theorem 3.5.1. *Consider a set of data (f, ν, g, ξ, L) satisfying the assumptions of theorem 3.4.6, but with ρ, ρ', λ in $(\mathbf{A_{der}})$ being growth function instead of constants. Further assume that $|f(t, y, 0)| \leq |f(t, 0, 0)| + \varphi(y)$ for a growth function φ such that $\int_1^{+\infty} \frac{1}{\varphi(y)} dy = +\infty$. Then there exists a solution (Y, N, K) to the reflected BSDE (3.2.1) with data (f, ν, g, ξ, L).*

Proof. We will apply the perturbation procedure as was done previously when ρ, ρ', λ were constants.

First, by the estimate in theorem 1 in Kobylanski *et al.* [51], we know that there exists a function F increasing (a growth function) such that for any set of data (f, ν, g, ξ, L) satisfying the assumptions and for any solution (Y, N, K) we have $\|Y\|_{\mathcal{S}^\infty} \leq F(\|\xi\|_\infty, \|\alpha\|_{\infty,1})$. Now, for a fixed set of data, we define $\rho_{\max} = \rho\big(F(\|\xi\|_\infty, \|\alpha\|_{\infty,1})\big)$. We fix n big enough that

$$\boldsymbol{D}^{(n)} = \frac{\boldsymbol{D}}{n} \leq \frac{e^{-2\rho_{\max}\|r\|^2_{\infty,2}}}{2^{10}\big(2\lambda(1)\big)\big(\|r\|^2_{\infty,2}+2\big)} = \epsilon_0(\rho_{\max}, 2\lambda(1), r) \ .$$

We will construct n solutions S^i of reflected BSDEs or perturbation equations such that for each equation, the size of the data is $\boldsymbol{D^i} = \boldsymbol{D^{(n)}}$ and the size of the solution is such that $\|\widehat{Y^i}\|_{\mathcal{S}^\infty} \leq 1$. Note that the $\widehat{\ }$ here indicates the multiplication by $B^i = \exp\big(\int_0^\cdot f_y(\overline{S}^{i-1}_u)dC_u\big)$.

We know that we can do a transation to be reduced to the case $L \leq 0$ so we assume from now on that $L \leq 0$. Define, for $i = 1 \dots n$, $\xi^i := \xi^{(n)} = \frac{1}{n}\xi$ and $\alpha^i := \alpha^{(n)} = \frac{1}{n}\alpha$ (uniform decomposition of ξ and α).

For $i = 1$, we first build a solution $S^1 = (Y^1, N^1, K^1)$ to the reflected BSDE (3.2.1) with parameters $(f - \alpha + \alpha^1, \nu, g, \xi^1, L)$. Proposition 3.4.3 as it is stated doesn't strictly apply, but we can adapt the proof. We define the integrating factor $B = e^{\int \beta dC}$ with $\beta = \overline{\beta}^0 = f_y(\cdot, 0, 0) \in L^{\infty,1}$ and the new measure Q by $\frac{dQ}{dP} = \mathcal{E}(\int \gamma\sigma^{-1}dM + \nu)$ where $\gamma = \overline{\gamma}^0 = f_z(\cdot, 0, 0) \in L^2_{BMO}$. Then, like in proposition 3.4.2, we look for a solution $(\widehat{Y^1}, \widehat{\widehat{N^1}}, \widehat{K^1})$ to the reflected BSDE with no linear term, via the fixed point theorem. We look for a solution in a ball of radius R and now further demand that $R \leq 1$, so that the conditions to be met are that

$$R \leq R_0(\widehat{2\lambda(1)}, r) = \frac{1}{2^5\widehat{2\lambda(1)}\big(\|r\|^2_{\infty,2}+2\big)} \text{ and } \widehat{\boldsymbol{D}^1} = \frac{\widehat{\boldsymbol{D}}}{n} \leq \frac{R_0(\widehat{2\lambda(1)}, r)}{2^5} = \epsilon_0(\widehat{2\lambda(1)}, r) \ ,$$

where $\widehat{2\lambda(1)} = e^{\|\overline{\beta}^0\|_{\infty,1}}(2\lambda(1))$. Now, since we have chosen n such that $\frac{\boldsymbol{D}}{n} \leq \frac{e^{-2\rho_{\max}\|r\|^2_{\infty,2}}}{2^{10}(2\lambda(1))\left(\|r\|^2_{\infty,2}+2\right)}$,

$$
\begin{aligned}
\frac{\widehat{\boldsymbol{D}}}{n} \leq e^{\|\overline{\beta}^0\|_{\infty,1}}\frac{\boldsymbol{D}}{n} \leq \frac{e^{\|\overline{\beta}^0\|_{\infty,1}}e^{-2\rho_{\max}\|r\|^2_{\infty,2}}}{2^{10}\big(2\lambda(1)\big)\big(\|r\|^2_{\infty,2}+2\big)} &= \frac{e^{2\|\overline{\beta}^0\|_{\infty,1}}e^{-2\rho_{\max}\|r\|^2_{\infty,2}}}{2^{10}\widehat{2\lambda(1)}\big(\|r\|^2_{\infty,2}+2\big)} \\
&\leq \frac{1}{2^{10}\widehat{2\lambda(1)}\big(\|r\|^2_{\infty,2}+2\big)} = \epsilon_0(\widehat{2\lambda(1)}, r)
\end{aligned}
$$

because $\|\overline{\beta}^0_{\infty,1}\| \leq \rho(0)\|r\|^2_{\infty,2}$ and $\rho(0) \leq \rho_{\max}$ by construction. So we indeed get a solution $(\widehat{Y^1}, \widehat{\widetilde{N^1}}, \widehat{K^1})$ and doing the reverse transforms gives a solution $S^1 = (Y^1, N^1, K^1)$ to the reflected BSDE with the linear terms.

For $i = 2 \dots n$, we have a solution $\overline{S}^{i-1}$ to the reflected BSDE (3.2.1) with parameters $(f - \alpha + \overline{\alpha}^{i-1}, \nu, g, \overline{\xi}^{i-1}, L)$ and want to construct the appropriate perturbation S^i. We simply do the same computations as in proposition 3.4.4, using the integrating factor $\overline{B}^{i-1} = e^{\int \overline{\beta}^{i-1}dC}$ where $\overline{\beta}^{i-1} = f_y(\cdot, \overline{Y}^{i-1}, \overline{Z}^{i-1}\sigma)$, and the change of measure $\frac{dQ}{dP} = \mathcal{E}(\int \overline{\gamma}^{i-1}\sigma^{-1}dC + \overline{\nu}^{i-1})$ where $\overline{\gamma}^{i-1} = f_z(\cdot, \overline{Y}^{i-1}, \overline{Z}^{i-1}\sigma)$ and $\overline{\nu}^{i-1} = \nu + \int 2gd(\overline{N}^{i-1})^{\perp}$.

Because we know, by the a priori estimate on solutions of the reflected BSDE, that

$$
\|\overline{Y}^{i-1}\|_{\mathcal{S}^\infty} \leq F\left(\left\|\frac{(i-1)}{n}\xi\right\|_\infty, \left\|\frac{(i-1)}{n}\alpha\right\|_{\infty,1}\right) \leq F\big(\|\xi\|_\infty, \|\alpha\|_{\infty,1}\big),
$$

we know that $\|\overline{\beta}^{i-1}\|_{\infty,1} \leq \rho(\|\overline{Y}^{i-1}\|_{\mathcal{S}^\infty})\|r\|^2_{\infty,2} \leq \rho_{\max}\|r\|^2_{\infty,2}$. Therefore, just as above, we find that the size $\widehat{\boldsymbol{D}^i}$ of the data is indeed small enough, and so we can construct the perturbation S^i. $\square$

As can be seen from the proof above, the key to the generalization is to have an a priori estimate $\|Y\|_{\mathcal{S}^\infty} \leq F(\|\xi\|_\infty, \|\alpha\|_{\infty,1})$ for some growth function F.

Case of f monotone and with arbitraty growth in y.

We can also generalize the result of theorem 3.4.6 to the case where f is so-called monotonous (or 1-sided Lipschitz) in y, with arbitrary growth.

Theorem 3.5.2. *Consider a set of parameters (f, ν, g, ξ, L) satisfying the assumptions of theorem 3.4.6, but with ρ, ρ', λ in $(\mathbf{A_{der}})$ being growth function instead of constants.*

Further assume that $|f(t,y,0)| \leq |f(t,0,0)| + \varphi(y)$ *for a growth function* φ *and that there exists a constant* μ *such that for all* y, y', z, s, ω*,*

$$(y'-y)\big(f(s,y',z) - f(s,y,z)\big) \leq \mu\, r_s^2\, |y'-y|^2$$

Then there exists a solution (Y,N,K) *to the reflected BSDE (3.2.1) with parameters* (f,ν,g,ξ,L)*.*

As remarked above, it is enough to have an a priori estimate for $\|Y\|_{\mathcal{S}^\infty}$. One can use the one obtained in the proof of theorem 3.1 in [79]. Alternatively, having argued that it is enough to study the case where the obstacle is negative, one can linearize the driver in the N variable, and do a measure change. Then, using Itô with $|\cdot|^2$ to take advantage of the monotonicity condition, one could conclude via standard estimations that

$$\|Y\|^2_{\mathcal{S}^\infty} \leq 2e^{4\mu\|r\|^2_{\infty,2}}\big(\|\xi\|^2_\infty + 2\|\alpha\|^2_{\infty,1}\big) =: F\big(\|\xi\|_\infty, \|\alpha\|_{\infty,1}\big)^2.$$

3.6 Technical details : differential calculus lemmas.

In the above presentation of our results (sections 3.2 to 3.5), the focus was kept strictly on reflected BSDEs. A number of claims were not proven, which involve only elementary differential calculus. This is because we wanted to keep the proofs clear, and not make them longer with elementary verifications. We provide the associated proofs in this section.

3.6.1 Recall of the assumptions

Let us first recall the assumptions that were used. Recall that r is a positive process in $L^{\infty,2}$.

($\mathbf{A_{qg}}$) There exists a growth function $\lambda(\cdot)$ (i.e. $\lambda : \mathbb{R} \to \mathbb{R}_+$ symmetric, increasing on $\mathbb{R}_+$, bounded below by 1) and a positive process $h \in L^2_{BMO}$ (i.e. $\int h dM \in BMO$, see below) such that :

$$|f(t,y,z)| \leq \lambda(y)\big(h_t^2 + |z|^2\big).$$

($\mathbf{A_{locLip}}$) The function f is differentiable at $(0,0)$ (in (y,z), for all (ω,s)), and there exist $\lambda>0$ such that, writing $\beta_s=f_y(s,0,0)$ and $\gamma_s=f_z(s,0,0)$, one has

- for all $\omega, s, y_1, y_2, z_1, z_2$:

$$\begin{aligned}|f(s,y_1,z_1)-f(s,y_2,z_2)-\beta_s(y_1-y_2)-\gamma_s(z_1-z_2)|\\ \leq \lambda\Big(r_s|y_1|+r_s|y_2|+|z_1|+|z_2|\Big)\big(r_s|y_1-y_2|+|z_1-z_2|\big)\ ,\end{aligned}$$

- $\gamma\in L^2_{BMO}$ and $\beta\in L^{\infty,1}$ (that is : $\int_0^T|\beta_s|dC_s\in L^\infty$),

($\mathbf{A_{der}}$) f is twice continuously differentiable in the variables (y,z) and there exists $\rho,\rho',\lambda>0$, and $h\in L^2_{BMO}$ such that

$$\begin{gathered}|f_y(t,y,z)|\leq \rho r_t^2 \qquad \text{and} \qquad |f_z(t,y,z)|\leq \rho'(h_t+|z|)\ ,\\ |f_{yy}(t,y,z)|\leq \lambda r_t^2,\quad |f_{yz}(t,y,z)|\leq \lambda r_t \quad \text{and} \quad |f_{zz}(t,y,z)|\leq \lambda\ .\end{gathered}$$

We will refer to these assumptions as ($\mathbf{A_{qg}}$)$\{\lambda,h\}$, ($\mathbf{A_{locLip}}$)$\{\beta,\gamma,\lambda,r\}$ and finally ($\mathbf{A_{der}}$)$\{\rho,\rho',\lambda,r,h\}$. More specifically, the assumptions in ($\mathbf{A_{der}}$) related to the first derivative Df and the second derivative D^2f are denoted by (ctrldD^2)$\{\lambda,r\}$ and (ctrldD^1)$\{\rho,\rho',r,h\}$ respectively. In ($\mathbf{A_{locLip}}$), the regularity estimate on its own will be referred to as the $locLipz\{r,\lambda\}$ estimate.

3.6.2 Lemmas related to those assumptions

In the lemmas below, we gather some facts about the interplay between the different assumptions. In particular, how the different transforms which we use affect the assumptions.

Translations and going from ($\mathbf{A_{der}}$) to ($\mathbf{A_{locLip}}$)

We need to consider some translated of f : given $Y\in\mathcal{S}^\infty$ and $Z\in L^2_{BMO}$, we define

$$\overline{f}(t,y,z)=f(t,y+Y,z+Z).$$

The lemmas below are used in two places. First, when we perturbate an equation and consider that a difference in drifts is a drift in the delta variables. There, we need to consider $\overline{f}(t,y,z) = f(t,y+Y,z+Z) - f(t,y,z)$ which is as above, plus a constant. Second, when we do a change of framework of reference, doing a translation by the constant U.

Lemma 3.6.1 (no change in the growth of the second derivative after translation).
f satisfies (ctrldD^2)$\{\lambda, r\}$ $\Longrightarrow$ $\overline{f}$ satisfies (ctrldD^2)$\{\lambda, r\}$

Proof. Quite obvious.

$$
\begin{aligned}
|\overline{f}_{yy}(t,y,z)| &= |f_{yy}(t,y+Y_t,z+Z_t)| \le \lambda(y+\|Y\|_\infty) r_t^2 \\
|\overline{f}_{yz}(t,y,z)| &= |f_{yz}(t,y+Y_t,z+Z_t)| \le \lambda(y+\|Y\|_\infty) r_t \\
|\overline{f}_{zz}(t,y,z)| &= |f_{zz}(t,y+Y_t,z+Z_t)| \le \lambda(y+\|Y\|_\infty)
\end{aligned}
$$

So in case the parameter λ is a growth function, we have the new growth function $\lambda'(y) = \lambda(y+\|Y\|)$. Otherwise, we just have the same parameter $\lambda > 0$ (and in both cases, the same $r \in L^{2,\infty}$). □

Lemma 3.6.2 (change in the growth of the first derivatives after translation).
f satisfies (ctrldD^1)$\{\rho, \rho', r, h\}$ $\Longrightarrow$ $\overline{f}$ satisfies (ctrldD^1)$\{\rho, \rho', r, h+Z\}$

Proof. Same kind of obviousness.

$$
\begin{aligned}
&|\overline{f}_y(t,y,z)| = |f_y(t,y+Y_t,z+Z_t)| \le \rho(y+\|Y\|_\infty) r_t^2 \\
&|\overline{f}_z(t,y,z)| = |f_z(t,y+Y_t,z+Z_t)| \le \rho'(y+\|Y\|_\infty)\big(h_t+|z+Z_t|\big) \le \rho'(..)\Big(\big[h_t+|Z_t|\big]+|z|\Big)
\end{aligned}
$$

and $h+Z \in L^2_{BMO}$ by the assumption on Z. □

Combining the two we obtain the following useful lemma.

Lemma 3.6.3 (change in the assumption ($\mathbf{A_{der}}$) after translation).
f satisfies($\mathbf{A_{der}}$)$\{\rho, \rho', \lambda, r, h\}$ $\Longrightarrow$ $\overline{f}$ satisfies ($\mathbf{A_{der}}$)$\{\rho, \rho', \lambda, r, h+Z\}$.

Lemma 3.6.4 (how the assumption on the growth of $D^2 f$ implies the locLipz regularity).
f satisfies (ctrldD^2)$\{\lambda, r\}$ $\Longrightarrow$ f satisfies the (locLipz)$\{2\lambda, r\}$ estimate.

Proof. This proof requires computations.

Take $w_1 = (y_1, z_1)$ and $w_2 = (y_2, z_2)$. We want to prove that $f(w_2) - f(w_1) - Df(0).(w_2 - w_1)$ is bounded by the right quantitity. Using Taylor's formula (with integral remainder),

$$
\begin{aligned}
&f(w_2) - f(w_1) - Df(0).(w_2 - w_1) \\
&\quad = f(w_2) - f(w_1) - Df(w_1).(w_2 - w_1) + \Big(Df(w_1) - Df(0)\Big).(w_2 - w_1) \\
&\quad = \int_0^1 (1-u) D^2 f(w_1 + u\Delta w) du.(\Delta w)^{\otimes 2} + \int_0^1 D^2 f(uw_1) du.w_1.\Delta w
\end{aligned}
$$

so taking the modulus, using the assumptions on $D^2 f$ (with possibly a growth function λ), then recognizing the development of $(a+b)^2$ and $(a+b)(c+d)$, and noting that $|\Delta y| \leq |y_1| + |y_2|$ and $|\Delta z| \leq |z_1| + |z_2|$, one sees that

$$
\begin{aligned}
&|f(w_2) - f(w_1) - Df(0).(w_2 - w_1)| \\
&\quad \leq \int_0^1 (1-u)|D^2 f(w_1 + u\Delta w).(\Delta w)^{\otimes 2}| du + \int_0^1 |D^2 f(uw_1).w_1.\Delta w| du \\
&\quad \leq \int_0^1 (1-u)\lambda(|y_1 + u\Delta y|)\Big[r_t^2|\Delta y|^2 + 2r_t|\Delta y||\Delta z| + |\Delta z|^2\Big] du \\
&\qquad + \int_0^1 \lambda(|y_1 + u\Delta y|)\Big[r_t^2|y_1||\Delta y| + r_t|y_1||\Delta z| + r_t|z_1||\Delta y| + |z_1||\Delta z|\Big] du \\
&\quad \leq \lambda(|y_1| + |y_2|)\Big[r_t^2|\Delta y|^2 + 2r_t|\Delta y||\Delta z| + |\Delta z|^2\Big] \\
&\qquad + \lambda(|y_1| + |y_2|)\Big[r_t^2|y_1||\Delta y| + r_t|y_1||\Delta z| + r_t|z_1||\Delta y| + |z_1||\Delta z|\Big] \\
&\quad \leq \lambda(..)\Big[r_t|\Delta y| + |\Delta z|\Big]^2 \\
&\qquad \lambda(..)\Big(r_t|\Delta y| + |\Delta z|\Big)\Big(r_t|y_1| + |z_1|\Big) \\
&\quad \leq \lambda(..)\Big(r_t|\Delta y| + |\Delta z|\Big)\Big(r_t|y_1| + r_t|y_2| + |z_1| + |z_2|\Big) \\
&\qquad \lambda(..)\Big(r_t|\Delta y| + |\Delta z|\Big)\Big(r_t|y_1| + |z_1|\Big) \\
&\quad \leq 2\lambda(|y_1| + |y_2|)\Big(r_t|y_1| + r_t|y_2| + |z_1| + |z_2|\Big)\Big(r_t|\Delta y| + |\Delta z|\Big) \;,
\end{aligned}
$$

which is the *locLipz* estimate required, with $\lambda' = 2\lambda$ if λ is a constant (and r is unchanged, as usual). □

Lemma 3.6.5 (how the assumption on the growth of $D^1 f$ gives the integrability of β

and γ).

$$f \text{ satisfies (ctrldD}^1\text{)}\{\rho,\rho',r,h\} \Longrightarrow \begin{cases} \beta = f_y(\cdot,0,0) \in L^{\infty,1} \text{ (with } \|\beta\|_{\infty,1} \le \rho\|r\|^2_{\infty,2}\text{)} \\ \gamma = f_z(\cdot,0,0) \in L^2_{BMO} \text{ (with } \|\gamma\|_{BMO} \le \rho'\|h\|_{BMO}\text{)} \end{cases}$$

Proof. Quite obvious.

$\beta = f_y(\cdot,0,0)$ so $|\beta_t| \le \rho r_t^2$, and since $r \in L^{\infty,2}$, $\beta \in L^{\infty,1}$ with $\|\beta\| \le \rho\|r\|^2$

$\gamma = f_z(\cdot,0,0)$ so $|\gamma_t| \le \rho'(h_t + |0|)$, and since $h \in L^2_{BMO}$, $\gamma \in L^2_{BMO}$ with $\|\gamma\| \le \rho'\|h\|$ □

Combining those facts, we obtain the following lemma.

Lemma 3.6.6. *f satisfies $(\mathbf{A_{der}})\{\rho,\rho',\lambda,r,h\} \Longrightarrow \overline{f}$ satisfies $(\mathbf{A_{locLip}})\{\overline{\beta},\overline{\gamma},2\lambda,r\}$. Here, of course, $\overline{\beta}_t = \overline{f}_y(t,0,0) = f_y(t,Y_t,Z_t)$ and $\overline{\gamma}_t = \overline{f}_z(t,0,0) = f_z(t,Y_t,Z_t)$, so that*

$$\|\overline{\beta}\|_{\infty,1} \le \rho\|r\|^2_{\infty,2}$$
$$\|\overline{\gamma}\|_{L^2_{BMO}} \le \rho'\|h+Z\|_{L^2_{BMO}}$$

Proof. First, note that if a given f satisfies $(\mathbf{A_{der}})\{\rho,\rho',\lambda,r,h\}$, then by lemma 3.6.4 and 3.6.5, one can say that f satisfies $(\mathbf{A_{locLip}})\{\beta,\gamma,2\lambda,r\}$.

Now, by lemma 3.6.3, since f satisfies $(\mathbf{A_{der}})\{\rho,\rho',\lambda,r,h\}$, so does $\overline{f}$ but with parameters $\{\rho,\rho',\lambda,r,h+Z\}$. And by the point above, $\overline{f}$ satisfies $(\mathbf{A_{locLip}})\{\beta,\gamma,2\lambda,r\}$, and $\|\overline{\beta}\|_{\infty,1} \le \rho\|r\|^2_{\infty,2}$ and $\|\overline{\gamma}\|_{L^2_{BMO}} \le \rho'\|h+Z\|_{L^2_{BMO}}$. □

Linearization : removal of linear terms

One can always remove the linear terms in f by looking at h defined by

$$f(t,y,z) = \beta_t.y + \gamma_t.z + h(t,y,z).$$

Here, h includes the residual drift $\alpha = f(\cdot,0,0)$.

Lemma 3.6.7. *f satisfies $(\mathbf{A_{locLip}})\{\beta,\gamma,\lambda,r\} \Longrightarrow h$ satisfies $(\mathbf{A_{locLip}})\{\beta=0,\gamma=0,\lambda,r\}$*

Proof.

$$\begin{aligned}|h(y',z')-h(y,z)-0-0| &= |f(y',z')-f(y,z)-\beta(y'-y)-\gamma(z'-z)|\\ &\le \lambda\Big(r|y|+r|y'|+|z|+|z'|\Big)\big(r|\Delta y|+|\Delta z|\big)\end{aligned}$$

□

$\widehat{f}$: scalings

Let B be a strictly positive process, bounded away from 0 and $+\infty$. We define

$$\widehat{f}(t,y,z) = B_t f(t, B_t^{-1}y, B_t^{-1}z)$$

Lemma 3.6.8. *f satisfies $(\mathbf{A_{locLip}})\{\beta,\gamma,\lambda,r\} \Longrightarrow \widehat{f}$ satisfies $(\mathbf{A_{locLip}})\{\beta,\gamma,\widehat{\lambda},r\}$ Here,*

$$\widehat{\lambda} = \|B^{-1}\|_\infty \lambda$$

Note that in practice, $B = \exp(\int_0^{\cdot} \beta du)$ so if $\beta \in L^{\infty,1}$ then $\|B\|_\infty, \|B^{-1}\|_\infty \le e^{\|\beta\|_{\infty,1}}$.

Proof. One computes

$$\begin{aligned}&|\widehat{f}(y',z')-\widehat{f}(y,z)-\beta(y'-y)-\gamma(z'-z)|\\ &\quad= B|f(B^{-1}y',B^{-1}z')-f(B^{-1}y,B^{-1}z)-\beta(B^{-1}y'-B^{-1}y)-\gamma(B^{-1}z'-B^{-1}z)|\\ &\quad\le B\lambda\Big(r|B^{-1}y|+r|B^{-1}y'|+|B^{-1}z|+|B^{-1}z'|\Big)\big(r|\Delta B^{-1}y|+|\Delta B^{-1}z|\big)\\ &\quad= BB^{-1}B^{-1}\lambda\Big(r|y|+r|y'|+|z|+|z'|\Big)\big(r|\Delta y|+|\Delta z|\big)\\ &\quad\le \|B^{-1}\|\lambda\Big(r|y|+r|y'|+|z|+|z'|\Big)\big(r|\Delta y|+|\Delta z|\big)\end{aligned}$$

So setting $\widehat{\lambda} = \|B^{-1}\|_\infty\lambda$ indeed shows that $\widehat{f}$ is (locLipz)$\{\beta,\gamma,\widehat{\lambda},r\}$ □

Note : we need to have this lemma, and then as a rigour-protection it needs to have a proof, which is not much informative. What really happens behind this is quite clear though. When we do a scaling, the term $f(t,y,z)$ is multiplied by B (and becomes $Bf(t,y,z) = Bf(t, B^{-1}\widehat{y}, B^{-1}\widehat{z}))$. The coefficients β, γ of the linear terms are unchanged in the scaling, while the coefficient λ of the quadratic term is inversely scaled.

Verifying that we are always under $(\mathbf{A_{qg}})$

It is clear that if f satisfies $(\mathbf{A_{der}})$, the bounds on the first derivative imply that f has at most quadratic growth in z (and at most linear growth in y, if ρ and ρ' are constants).

Lemma 3.6.9 (from (ctrldD^1) to $(\mathbf{A_{qg}})$)**.**

$$\begin{cases} f \textit{ satisfies (ctrld}D^1\textit{)}\{\rho, \rho', r, h\} \\ \alpha = f(\cdot, 0, 0) \textit{ is in } L^1_{BMO} \end{cases} \Longrightarrow f \textit{ satisfies } (\mathbf{A_{qg}})\{\lambda', h'\}$$

Proof. Again a straight-forward estimation. Using Taylor's formula we get

$$\begin{aligned} f(y,z) &= f(0,0) + \int_0^1 Df(uy,uz)du.(y,z) \\ &= f(0,0) + \int_0^1 f_y(uy,uz)du.y + \int_0^1 f_z(uy,uz)du.z \end{aligned}$$

so

$$\begin{aligned} |f(t,y,z)| &\leq |f(t,0,0)| + \int_0^1 \rho(u|y|)r_t^2 du|y| + \int_0^1 \rho'(u|y|)(h_t + |uz|)du|z| \\ &\leq |\alpha| + \rho(y)r_t^2|y| + \rho'(y)(h_t + |z|)|z| \\ &\leq |\alpha| + \rho(y)r_t^2|y| + \frac{1}{2}\rho'(y)h_t^2 + \frac{3}{2}\rho'(y)|z|^2 \\ &\leq \Big(1 + \rho(y)|y| + \frac{1}{2}\rho'(y) + \frac{3}{2}\rho'(y)\Big)\big(|\alpha| + r_t^2 + h_t^2 + |z|^2\big) \end{aligned}$$

Now, $\alpha \in L^1_{BMO}$, $r^2 \in L^{\infty,1} \subseteq L^1_{BMO}$ and $h^2 \in L^1_{BMO}$, so we see that f satisfies (atmostquad)$\{\lambda', h'\}$ where

$$\begin{aligned} (h'_t)^2 &= |\alpha_t| + r_t^2 + h_t^2 \\ \lambda'(y) &= 1 + \rho(y)|y| + \frac{1}{2}\rho'(y) + \frac{3}{2}\rho'(y) \end{aligned}$$

$\square$

Now, we also need to check that the ($\mathbf{A_{locLip}}$) assumption, which is a by-product of ($\mathbf{A_{der}}$) but used on its own from time to time, also implies ($\mathbf{A_{qg}}$). This ultimately means that we are just using the second derivative to check that f satisfies ($\mathbf{A_{qg}}$).

Lemma 3.6.10 (from ($\mathbf{A_{locLip}}$) to ($\mathbf{A_{qg}}$))**.**
$$\begin{cases} f \textit{ satisfies } (\mathbf{A_{locLip}})\{\beta, \gamma, \lambda, r\} \\ \alpha = f(\cdot, 0, 0) \textit{ is } L^1_{BMO} \end{cases} \Longrightarrow f \textit{ satisfies } (\mathbf{A_{qg}})\{\lambda'(y), h\}$$

Proof. Simply estimate, using the *locLipz* regularity :

$$\begin{aligned} |f(t,y,z)| &\le |f(t,0,0)| + |\beta||y| + |\gamma||z| + \lambda\big(r_t|y| + |z|\big)^2 \\ &\le |\alpha| + |\beta||y| + \frac{1}{2}|\gamma|^2 + \frac{1}{2}|z|^2 + 2\lambda r^2|y|^2 + 2\lambda|z|^2 \\ &\le \big(1 + |y| + 2\lambda|y|^2 + 2\lambda\big)\big(|\alpha| + |\beta| + |\gamma|^2 + r_t^2 + |z|^2\big) \ . \end{aligned}$$

Now, note that $|\alpha| \in L^1_{BMO}$, $|\beta| \in L^{\infty,1} \subseteq L^1_{BMO}$, $|\gamma|^2 \in L^1_{BMO}$ and $r^2 \in L^{\infty,1} \subseteq L^1_{BMO}$. So defining

$$\begin{aligned} h_t^2 &= |\alpha_t| + |\beta_t| + |\gamma_t|^2 + r_t^2 \\ \lambda'(y) &= 1 + |y| + 2\lambda|y|^2 + 2\lambda \end{aligned}$$

we see that $h \in L^2_{BMO}$ and λ' is a growth function (λ can equally well be a growth function). $\square$

Chapter 4

Analysis of the time-discretization for FBSDEs with monotone drivers with polynomial growth.

4.1 Introduction.

4.1.1 Motivation.

We now return to our core subject, recentering the discussion on the relationship between parabolic PDEs and standard BSDEs. However we investigate this relationship under a different angle, and this chapter will be concerned with the numerical aspects of BSDEs.

In the previous chapter, aside from the fact that we considered reflected BSDEs and not merely standard BSDEs, we discussed backward stochastic problems set in a very general setting. The reason for this was that such a generality, at least in its formalism, allows to cover at once several types of PDE problems. We now come back to the standard set-up of Markovian (F)BSDEs in a Brownian setting. On the analytical aspects however, we will retain the monotonicity condition encountered in the previous chapter.

It is an important task to develop methods for solving BSDEs numerically.

The connection between PDEs and BSDEs always holds at least formally. For this reason, FBSDEs can be thought of as a different language to study the same thing

: PDEs. There is then a one-to-one correspondence between the analytical results on PDEs on one side and FBSDEs on the other side. Every result proved on the BSDE side might, in principle, be proved equally on the PDE side with pure PDE methods. There is however a more important justification for continuing the intense effort currently directed toward BSDEs.

Through BSDEs, one obtains new numerical methods for solving PDEs. On the long term, even if they eventually do not turn out to be much faster than the numerical methods for PDEs, BSDE-based methods are usually of Monte-Carlo type and as such they are immune to, or less subject to the curse of dimensionality. They would therefore be the only tool available for some problems. This make it important to understand better and develop further the numerical methods for BSDEs, a research sub-area which is still relatively young.

So far, the major part of the research on numerical methods for BSDEs has focused on the analysis of time discretization error, and did so for the cases of nonlinearities that are either Lipschitz or quadratic in the secondary variable Z. The aim of the research we present in this chapter is to understand how to handle the time discretization for BSDEs in the case where the nonlinearity coefficient is monotone and can have polynomial growth in the primary variable Y. More precisely, we aim primarily at understanding what happens to the usual backward Euler scheme, proving that it converges if indeed it does, and proposing an alternative scheme when it does not.

This monotonicity assumption is relevant for PDEs describing reaction-diffusion phenomena. Indeed, in many such equations, the function f is a polynomial (in v), for example the Allen-Cahn equation, the FitzHugh-Nagumo equations (with or without recovery) or the standard non-linear heat and Schrödinger equation (see [38], [74], [32], [52] and references).

The applicability of the results we develop here is not restricted to the modeling of physical phenomena. It is also possible to extend the work we develop to the Brownian-Lévy setting and apply it for instance to problems of contingent claim hedging in defaultable markets, see e.g. instance [36].

4.1.2 Introduction to the numerical methods for BSDEs.

Before moving on to presenting the research done for this project, let us take the time to present briefly the numerical methods for BSDEs.

Consider a Markovian forward-backward SDE

$$X_t^{r,x} = x + \int_r^t b(s, X_s^{r,x})ds + \int_r^t \sigma(s, X_s^{r,x})dW_s \ , \tag{4.1.1}$$

$$Y_t^{r,x} = g(X_T^{r,x}) + \int_t^T f(u, X_u^{r,x}, Y_u^{r,x}, Z_t^{r,x})du - \int_t^T Z_u^{r,x}dW_u \ , \tag{4.1.2}$$

for a given $(r, x) \in [0, T] \times \mathbb{R}^d$, and $r \leq t \leq T$. We know that this is connected to the PDE

$$\begin{aligned} &v_t + v_{xx} \cdot a + v_x\, b + f(t, x, u, u_x\sigma) = 0 \qquad \text{on } [0, T[\times\mathbb{R}^d, \\ &v(T, \cdot) = g \ , \end{aligned}$$

where $a = \sigma\sigma^*$, $v_{xx} \cdot a = \sum_{i,j=1}^d v_{ij}a_{ij}$ and $v_x\, b = \sum_{i=1}^d v_i b_i$. The link is that $Y_t^{r,x} = v(t, X_t^{r,x})$ and $Z_t^{r,x} = (v_x\sigma)(t, X_t^{r,x})$.

Not knowing how to solve explicity the FBSDE (4.1.1)-(4.1.2) for a general nonlinear f, we would like to obtain a numerically computable approximation of the solution.

Time-discretization.

Let us drop the superscript r, x for the moment, and focus on the approximation of the solution $(X_t, Y_t, Z_t)_{t\in[0,T]}$. The forward equation (4.1.1) is not coupled to the backward equation (4.1.2), so it can be solved first.

Let us a consider a partition $\pi : 0 = t_0 < t_1 < \ldots < t_N = T$ of $[0, T]$. One can easily obtain an approximation $(X_i)_{i=0\ldots N}$ for the SDE solution $(X_t)_{t\in[0,T]}$. For instance, with the simple Euler scheme (known as Euler-Maruyama scheme in the context of SDEs), we would define $X_0 = x$ and then, for $i = 1 \ldots N$,

$$X_{i+1} = X_i + b(t_i, X_i)h_{i+1} + \sigma(t_i, X_i)\Delta W_{i+1} \ .$$

Here $h_{i+1} = t_{i+1} - t_i$ and $\Delta W_{i+1} = W_{t_{i+1}} - W_{t_i}$. Existing results for the numerical approximation of SDE solutions ensure, under various conditions, that a whole range of schemes converge, that is to say, (X_i) become a good approximation for (X_t), in some sense, when the modulus $|\pi| = \max_i |h_i|$ of the partition goes to 0 (see Kloeden and Platen [49]).

So one can consider this part of the problem solved, assume we have already computed an approximation (X_i) of the solution to the SDE, and focus on the approximation of the BSDE.

How do we discretize the BSDE ? Let us look at the BSDE (4.1.2) between t_i and t_{i+1} :

$$\begin{aligned} Y_{t_i} &= Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt - \int_{t_i}^{t_{i+1}} Z_t dW_t \\ &= E\Big(Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt \Big| \mathcal{F}_{t_i} \Big) \end{aligned}$$

To define an approximation Y_i of Y_{t_i}, one may want to simply approximate the integrals by increments. For instance, if we used a right-end rectangle rule for the integrals, we would define

$$Y_i = Y_{i+1} + f(t_i, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1} - Z_{i+1}\Delta W_{i+1} \ ,$$

Having already computed (Y_{i+1}, Z_{i+1}), this Y_i is readily computable. However, just like for solutions to BSDEs, we want Y_i to be adapted. Thinking of the connection with PDEs, this Y_i that we define would be an approximation of $v(t_i, X_{t_i})$, so it should not depend on the paths of the Brownian motion after t_i. Taking the conditional expection $E\big(\cdot \big| \mathcal{F}_i \big)$, where $\mathcal{F}_i = \sigma(X_j, j \leq i)$, we obtain

$$Y_i := E\Big(Y_{i+1} + f(t_i, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1} \Big| \mathcal{F}_i \Big) \ ,$$

and this will be our Y_i. Now, we also need to approximate Z_{t_i}. Various heuristics can be used to obtain such an approximation. For instance, one could think that $Z = \langle Y, W \rangle$ in continuous time. So, taking the discrete time equivalent, we could think of multiplying the discrete dynamics equation

$$Y_{t_i} \approx Y_{i+1} + f(t_i, X_{t_{i+1}}, Y_{t_{i+1}}, Z_{t_{i+1}})h_{i+1} - Z_{t_i}\Delta W_{i+1} \ ,$$

(note that we use Z_{t_i} here and not $Z_{t_{i+1}}$) by ΔW_{i+1}, and take the condition expectation. Using the martingale property, we end up obtaining the following candidate for Z_i :

$$Z_i := E\bigg(\Big(Y_{i+1} + f(t_i, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1}\Big)\frac{\Delta W_{i+1}}{h_{i+1}}\bigg|\mathcal{F}_i\bigg) \ ,$$

In fact, interpreting more strictly $\int_{t_i}^{t_{i+1}} Z_t dt = \int_{t_i}^{t_{i+1}} d\left\langle Y, W\right\rangle_t$, we could in fact propose

$$Z_i := E\bigg(Y_{i+1}\frac{\Delta W_{i+1}}{h_{i+1}}\bigg|\mathcal{F}_i\bigg) \ ,$$

which also works.

Summarizing, the above heuristics lead to considering, for instance, the following scheme. Define $Y_N = g(X_N)$, $Z_N = 0$ and then, for $i = N-1 \dots 0$,

$$\begin{aligned} Y_i &= E\Big(Y_{i+1} + f(t_i, X_{i+1}, Y_{i+1}, Z_{i+1})h_{i+1}\Big|\mathcal{F}_i\Big) \\ Z_i &= E\bigg(Y_{i+1}\frac{\Delta W_{i+1}}{h_{i+1}}\bigg|\mathcal{F}_i\bigg). \end{aligned}$$

This is the (explicit) backward Euler scheme for BSDEs. It can be called in the BSDE context the Bouchard-Touzi-Zhang scheme (see Crisan and Manolarakis [21]). A number of results ensure that, under standard conditions on f and g, this scheme converges, that is to say (Y_i, Z_i) becomes a good approximation for (Y, Z) in some sense, as the modulus (or mesh) of the subdivision π goes to 0.

Approximation of the conditional expectations.

However, these random variables Y_i and Z_i are not yet computable, since they are defined using a conditional expectation. One needs to also approximate these in order to obtain a fully implementable scheme.

While the research presented in this chapter is not concerned with this question, let us mention briefly what those methods are. The quantization method consists in discretizing the space of values that X and Y can take, so that the conditional expectations can be computed with finitely many conditional probabilities. The latter are in general computed by Monte Carlo methods. The projection on a basis of functions is based on the vision of a conditional expectation as a projection on a certain subspace of L^2. One can also make use of the Malliavin derivatives. Finally, the cubature method was recently introduced in this context of BSDEs. This method approximate the law of the forward process and then compute expectations as integrals against this approximated law. We refer the reader to the review paper [21] by Crisan and Manolarakis for more details about these four methods.

4.1.3 Review of the literature.

While some attempts have been made before (see Crisan and Manolarakis [21] for those early references), the study of the convergence of numerical schemes for BSDEs really took off after the fundamental works of Zhang and Ma ([81, 58, 57]). They obtained in particular the path-regularity theorem for Lipschitz BSDEs, which says that the trajectories of Z are continuous in some L^2 sense. This result is crucial in order to prove the convergence of the scheme if one wants to allow f to depend also on Z.

Following this, Zhang [82] and then Bouchard and Touzi [7] proposed the backward Euler scheme (time-discretization) for BSDEs and were able to prove its convergence for Lipschitz BSDEs. Higher-order discretization schemes have also been proposed (Crisan and Manolarakis [22], Chassagneux and Crisan [18], Chassagneux [17]).

In addition to the study of the convergence of the time-discretization, some works have studied the error created by the approximation of the conditional expectations (see Gobet and Turkedjiev [35] and references therein).

The above mentioned works all study the case of Lipschitz coefficient f. For f quadratic in z, the path-regularity theorem was obtained by Imkeller and dos Reis [46] and Richou [72]. In these papers, the essential idea to approximate numerically the solution was to truncate the quadratic BSDEs to a Lipschitz one, controlling the distance between the true solution and the solution to the truncated BSDE, and then apply a standard time-discretization to approximate the solution to the truncated (Lipschitz) BSDE. Very recently, Chassagneux and Richou [19] have proposed a more direct procedure.

The work presented in this chapter tackles the time-discretization of BSDEs with a monotone driver which can have polynomial growth in the y variable.

4.1.4 Overview of the content of this chapter.

We look further at the connection between parabolic PDEs and FBSDEs with monotone drivers f of polynomial growth, which has been studied in Pardoux [64], Briand and Carmona [8] and Briand, Delyon, Hu, Pardoux and Stoica [11]. By monotonicity we mean (see section 4.2) that $\langle v'-v, f(v')-f(v)\rangle \leq \mu|v'-v|^2$, for some $\mu \geq 0$, and any v, v' (one can also find the terminology that f is 1-sided Lipschitz). We extend the three works mentioned above on the theoretical side by providing further results for the FBSDE in question (classical and Malliavin differentiability, representation formula, path-regularity theorem).

Then, we proceed to a thorough analysis of various numerical methods that open the door to Monte Carlo methods for solving numerically the corresponding PDEs.

Before presenting our results on the time-discretization, let us illustrate with an example why the explicit Euler scheme can explode under this type of assumption on f.

Consider the following simple example (for further details and notational setup see Section 4.2 and Section 4.7.1)

$$Y_t = \xi - \int_t^1 Y_s^3 \mathrm{d}s - \int_t^1 Z_s \mathrm{d}W_s, \quad t \in [0,1] \tag{4.1.3}$$

with the terminal condition $\xi \in \mathcal{F}_1$. For any $\xi \in L^p$ for $p \geq 2$ there exists a (square-integrable) solution (Y, Z) to the above BSDE.

Fix the number of time-discretization points to be $N+1 > 0$. The explicit Euler scheme for the above equation with uniform time step $h = 1/N$ is, with the notation $Y_i := Y_{i/N}$, given by

$$Y_i = \mathbb{E}[Y_{i+1} - Y_{i+1}^3 h|\mathcal{F}_i] = \mathbb{E}[Y_{i+1}(1 - hY_{i+1}^2)|\mathcal{F}_i], \quad i = 0, \ldots, N-1, \tag{4.1.4}$$

where $Y_N = \xi$.

It is a simple calculation (see Section 4.7.1 for the details) to show that if

$$\xi \geq 2\sqrt{N} \quad \text{then} \quad |Y_i| \geq 2^{2^{N-i}}\sqrt{N} \quad \text{for} \quad i = 0, \ldots, N-1. \tag{4.1.5}$$

With this simple computation in mind it is possible to show that there exists a random variable ξ whose moments of any order are finite and for which the explicit Euler scheme diverges. The proof of the following result can be found in Section 4.7.1.

Lemma 4.1.1. *Let π^N be the uniform grid over the interval $[0,1]$ with $N+1$ points, N an even number ($t = 1/2$ is common to all grids π^N). For any $\xi \in L^p(\mathcal{F}_1)$, for $p \geq 1$, let (Y, Z) denote the solution to (4.1.3).*

Then there exists a random variable ξ such that

$$\lim_{N\to\infty} \mathbb{E}[\,|Y^{(N)}_{\frac{1}{2}}|\,] = +\infty,$$

where $Y^{(N)}_{\frac{1}{2}}$ is the Euler approximation of Y on the time point $t = 1/2$ via (4.1.4) over the grids π^N.

The special random variable ξ we work with is normally distributed and it is known that $\mathbb{P}[|\xi| > 2\sqrt{N}]$ is exponentially small (see Lemma 4.7.1). What our counter-example shows is that although ξ may take very large values on an event with exponentially small probability, the impact of these very large values when propagated through the Euler explicit scheme is doubly-exponential (see (4.1.5)).

This double-exponential impact is precisely a consequence of the superlinearity of the driver. In general, the terminal condition ξ is an unbounded random variable (RV) so there is a positive probability of the scenario where $\xi \geq 2\sqrt{N}$ no matter how small a time-step we choose. This indicates that, in general, the explicit Euler scheme may diverge, as it happens in SDE context [43]. Therefore one needs to seek alternative (for example implicit) approximations for BSDE with polynomial drivers that are also monotone and/or find conditions under which it is possible for the explicit scheme to work, as explicit schemes have certain computational advantages over implicit ones.

Our contribution on the practical side, concerning the convergence or not of the time-discretization is the following.

We extend the canonical Zhang path regularity theorem (see Ma and Zhang [57], [46]), originally proved under Lipschitz assumptions, to our polynomial growth monotone driver setting

For our non-Lipschitz setting we provide a thorough analysis of the family of θ-schemes, where $\theta \in [0,1]$ characterizes the degree of implicitness of the scheme. Contrary to the FBSDEs with Lipschitz drivers, we show that choosing $\theta \geq 1/2$ is essential to ensure the stability of the scheme, in a similar way to the SDE context (see Mao and Szpruch [59]). This is to our knowledge the first result in the numerical BSDEs literature that shows a superior stability of the implicit scheme over the standard explicit one. We also generalize the concept of stability for discretization schemes (see that in Chassagneux [16, 17]). This, among others things, paves a way for deriving higher order approximations schemes for FBSDEs with non-Lipschitz drivers. As an example, we prove a higher order of convergence for the trapezoidal scheme (the case $\theta = 1/2$).

We construct an appropriately tamed version of the explicit Euler scheme for which the required stability property can be recovered. This allows to obtain convergence of the scheme. Interestingly enough in the special case where the driver of the FBSDEs does not depend on the SDE solution it is enough to appropriately tame the terminal condition, leaving the rest of the Euler approximation unchanged.

As a rule of thumb, implicit schemes tend to be more robust than explicit ones. Unfortunately implicit schemes involve solving an implicit equation, which creates an extra layer of complexity when compared to explicit schemes. A secondary aim of this work is to distinguish under which conditions explicit and implicit schemes can be used.

As standard in numerical analysis, we derive the global error estimates of various numerical schemes by analyzing their one-step errors and stability properties (which allows to study how errors propagate with time). We formulate the *Fundamental Lemma* (following the nomenclature from Milstein [61]) that states how to estimate the global error of a stable approximation scheme in terms of its local errors. The lemma is proved under minimal assumptions. We stress that a similar approach has been used in Chassagneux and Crisan [18], Chassagneux [16, 17], however their results are not sufficiently general to deal with non-Lipschitz drivers.

The structure of the global error estimate given by the Fundamental Lemma allows to study in a very easy and transparent way the special case of the θ-scheme with $\theta = 1/2$ (trapezoidal rule) which has a higher order of convergence. In this context we also conjecture a candidate for the 2nd order scheme.

Concerning the implementation of the presented schemes we propose an alternative estimator of the component Z whose standard deviation, contrary to usual estimator, does not explode as the time step vanishes.

Finally, in proving convergence for the mostly-implicit schemes, we prove L^p-type uniform bounds for the scheme, extending the classical L^2-bound obtained previously for the discretization of Lipschitz FBSDEs (see Bouchard and Touzi [7] or Gobet and Turkedjiev [35]).

In Section 4.2 we define notation and recall standard results from the literature that we use in this chapter. In Section 4.3 we establish first order variational results for the solution of the FBSDEs as well as stating the path regularity results required for the study of numerical schemes within the FBSDE framework. In the remaining sections we discuss several numerical schemes: in Section 4.4 we define the numerical discretization procedure and state general estimates for integrability and on the local errors. In Section 4.5 we establish the convergence of the implicit dominating schemes and in Section 4.6 the convergence of the tamed fully explicit scheme.

4.2 Preliminaries

4.2.1 Notation

Throughout let us fix $T > 0$. We work on a canonical Wiener space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a d-dimensional Wiener process $W = (W^1, \cdots, W^d)$ restricted to the time interval $[0, T]$. We denote by $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ its natural filtration enlarged in the usual way by the $\mathbb{P}$-zero sets and by $\mathbb{E}$ and $\mathbb{E}[\cdot|\mathcal{F}_t] = \mathbb{E}_t[\cdot]$ the usual expectation and conditional expectation operator respectively.

For vectors $x = (x^1, \cdots, x^d)$ in the Euclidean space $\mathbb{R}^d$ we denote by $|\cdot|$ and $\langle, \rangle$ the canonical Euclidean norm and inner product (respectively) while $\|\cdot\|$ is the matrix norm in $\mathbb{R}^{k\times d}$ (when no ambiguity arises we use $|\cdot|$ as $\|\cdot\|$); for $A \in \mathbb{R}^{k\times d}$ A^* denotes the transpose of A; I_d denotes the d-dimensional identity matrix. For a map $b : \mathbb{R}^m \to \mathbb{R}^d$, we denote by ∇b its $\mathbb{R}^{d\times m}$-valued Jacobian matrix (gradient in case $d = 1$) whenever it exists. To denote the j-th first derivative of $b(x)$ for $x \in \mathbb{R}^m$ we write $\nabla_{x_j} b$ (valued in $\mathbb{R}^{d\times 1}$). For $b(x, y) : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^k$ we write $\nabla_x h$ or $\nabla_y h$ to refer to its Jacobian matrix (gradient if $k = 1$) with relation to x and y respectively. Δ denotes the canonical Laplacian operator.

We define the following spaces for $p > 1$, $q \geq 1$, $n, m, d, k \in \mathbb{N}$

- $C^{0,n}([0, T] \times \mathbb{R}^d, \mathbb{R}^k)$ is the space of continuous functions endowed with the $\|\cdot\|_\infty$-norm that are n-times continuously differentiable in the spatial variable ; $C_b^{0,n}$

contains all bounded functions of $C^{0,n}$; the first superscript 0 is dropped for functions independent of time;

- $L^p(\mathcal{F}_t, \mathbb{R}^d)$, $t \in [0,T]$, is the space of d-dimensional $\mathcal{F}_t$-measurable RVs X with norm $\|X\|_{L^p} = \mathbb{E}[\,|X|^p]^{1/p} < \infty$; L^∞ refers to the subset of essentially bounded RVs;

- $\mathcal{S}^p([0,T]\times\mathbb{R}^d)$ is the space of d-dimensional measurable $\mathcal{F}$-adapted processes Y satisfying $\|Y\|_{\mathcal{S}^p} = \mathbb{E}[\sup_{t\in[0,T]}|Y_t|^p]^{1/p} < \infty$; $\mathcal{S}^\infty$ refers to the subset of $\mathcal{S}^p(\mathbb{R}^d)$ of absolutely uniformly bounded processes;

- $\mathcal{H}^p([0,T]\times\mathbb{R}^{n\times d})$ is the space of d-dimensional measurable $\mathcal{F}$-adapted processes Z satisfying $\|Z\|_{\mathcal{H}^p} = \mathbb{E}[\left(\int_0^T |Z_s|^2 \mathrm{d}s\right)^{p/2}]^{1/p} < \infty$;

- $\mathbb{D}^{k,p}(\mathbb{R}^d)$ and $\mathbb{L}_{k,d}(\mathbb{R}^d)$ are the spaces of Malliavin differentiable RVs and processes, see Section 4.7.2.

4.2.2 Setting

We want to study the forward-backward SDE system with dynamics (4.1.1)-(4.1.2), for $(t,x)\in[0,T]\times\mathbb{R}^d$ and $\Theta^{t,x} := (X^{t,x}, Y^{t,x}, Z^{t,x})$. Here we work, for $s\in[t,T]$, with the filtration $\mathcal{F}^t_s := \sigma\left(W_r - W_t : r\in[t,s]\right)$, completed with the $\mathbb{P}$-Null measure sets of $\mathcal{F}$. Concerning the functions appearing in (4.1.1) and (4.1.2) we will work with the following assumptions.

(HX0) $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ are 1/2-Hölder continuous in their time variable, are Lipschitz continuous in their spatial variables, satisfy $\|b(\cdot,0)\|_\infty + \|\sigma(\cdot,0)\|_\infty < \infty$ and hence satisfy $|b(\cdot,x)| + |\sigma(\cdot,x)| \le K(1+|x|)$ for some $K>0$.

(HY0) $g : \mathbb{R}^d \to \mathbb{R}^k$ is a Lipschitz function of linear growth; $f : [0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d} \to \mathbb{R}^k$ is a continuous function such that for some $L, L_x, L_y, L_z > 0$ for all $t, t', x, x', y, y', z, z'$

$$
\begin{aligned}
|f(t,x,y,z)| &\le L + L_x|x| + L_y|y|^m + L_z\|z\|, \quad m\ge 1,\\
\langle y'-y, f(t,x,y',z) - f(t,x,y,z)\rangle &\le L_y|y'-y|^2, \qquad (4.2.1)\\
|f(t,x,y,z) - f(t',x',y,z')| &\le L_t|t-t'|^{\frac{1}{2}} + L_x|x-x'| + L_z\|z-z'\|.
\end{aligned}
$$

(HY0$_{\text{loc}}$) (HY0) holds and, given L_y it holds for all t, x, y, y', z that

$$|f(t,x,y,z) - f(t,x,y',z)| \leq L_y(1 + |y|^{m-1} + |y'|^{m-1})|y - y'|, \quad m \geq 1. \quad (4.2.2)$$

(HXY1) (HX0), (HY0$_{\text{loc}}$) hold; $g \in C^1$ and $b, \sigma, f \in C^{0,1}$.

We state in the next remark some useful consequences of the monotonicity condition (4.2.1).

Remark 4.2.1. Under Assumption (HY0), for all t, x, y, y', z, z' and any $\alpha > 0$ we have

$$\begin{aligned}
&\langle y' - y, f(t,x,y',z') - f(t,x,y,z)\rangle \\
&\quad = \langle y' - y, f(t,x,y',z') - f(t,x,y,z')\rangle + \langle y' - y, f(t,x,y,z') - f(t,x,y,z)\rangle \\
&\quad \leq L_y|y' - y|^2 + L_z|y' - y||z' - z| \\
&\quad \leq (L_y + \alpha)|y' - y|^2 + \frac{L_z^2}{4\alpha}|z' - z|^2.
\end{aligned}$$

Moreover

$$\begin{aligned}
\langle y, f(t,x,y,z)\rangle &= \langle y - 0, f(t,x,y,z) - f(t,x,0,z)\rangle + \langle y, f(t,x,0,z)\rangle \\
&\leq L_y|y|^2 + |y|\big(L + L_x|x| + L_z|z|\big) \\
&\leq (L_y + \alpha)|y|^2 + \frac{3L^2}{4\alpha} + \frac{3L_x^2}{4\alpha}|x|^2 + \frac{3L_z^2}{4\alpha}|z|^2. \qquad (4.2.3)
\end{aligned}$$

4.2.3 Basic results

In this subsection we recall several auxiliary results concerning the solution of (4.1.1)-(4.1.2) that will become useful later. These results follows from [64] and [8].

Theorem 4.2.2 (Existence and uniqueness)**.** *Let (HX0) and (HY0) hold. Then FBSDE* (4.1.1)-(4.1.2) *has a unique solution* $(X, Y, Z) \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ *for any* $p \geq 2$*. Moreover, it holds for some constant* $C_p > 0$ *that*

$$\|Y\|_{\mathcal{S}^p}^p + \|Z\|_{\mathcal{H}^p}^p \leq C_p\big\{\|g(X_T)\|_{L^p}^p + \|f(\cdot, X_\cdot, 0, 0)\|_{\mathcal{H}^p}^p\big\} \leq C_p(1 + |x|^p). \qquad (4.2.4)$$

Proof. The existence and uniqueness results for SDE (4.1.1) follow from standard SDE literature. The existence and uniqueness result for the BSDE follows from Proposition

2.2 in [64], since the SDE results imply that $X \in \mathcal{S}^p$ for any $p \geq 2$, along with linear growth in x of g and f. The estimates for $Y \in \mathcal{S}^p$ for any $p \geq 2$ and $Z \in \mathcal{H}^p$ follow from the pathwise inequality

$$|Y_t|^2 + \big(1 - \frac{3L_z^2}{2\alpha}\big)\mathbb{E}_t\left[\int_t^T |Z_u|^2 \mathrm{d}u\right] \leq C_{\alpha,T,t}\mathbb{E}_t\Big[|g(X_T)|^2 + \int_t^T \frac{3}{4\alpha}|f(u, X_u, 0, 0)|^2 \mathrm{d}u\Big], \tag{4.2.5}$$

where $C_{\alpha,T,t} = \exp\{2(L_y + \alpha)(T - t)\}$, for any $\alpha > 0$ and $t \in [0, T]$. This last inequality follows from the proof of Proposition 2.2 and Exercise 2.3 in [64], (see also Theorem 3.6 in [8]). □

We now state a result concerning a priori estimates for BSDEs.

Theorem 4.2.3 (A priori estimate)**.** *Let $p \geq 2$ and for $i \in \{1, 2\}$, let $\Theta^i = (X^i, Y^i, Z^i)$ be the solution of FBSDE* (4.1.1)-(4.1.2) *with functions b^i, σ^i, g^i, f^i satisfying (HX0)-(HY0). Then there exists $C_p > 0$ depending only on p and the constants in the assumptions such that for $i \in \{1, 2\}$*

$$\begin{aligned} &\|Y^1 - Y^2\|_{\mathcal{S}^p}^p + \|Z^1 - Z^2\|_{\mathcal{H}^p}^p \\ &\quad \leq C_p\Big\{\mathbb{E}\Big[|g^1(X_T^1) - g^2(X_T^2)|^p + \Big(\int_0^T |f^1(s, X_s^1, Y_s^i, Z_s^i) - f^2(s, X_s^2, Y_s^i, Z_s^i)|\mathrm{d}s\Big)^p\Big]\Big\}. \end{aligned} \tag{4.2.6}$$

Proof. See Proposition 3.2 and Corollary 3.3 in [8]. □

Corollary 4.2.4 (Markov property and sample path continuity)**.** *The mapping $(t, x) \mapsto Y_t^{t,x}(\omega)$ is continuous. There exist two $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^k)$ and $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^{k\times d})$ measurable deterministic functions u and v (respectively) s.th.*

$$\begin{aligned} Y_s^{t,x} &= u(s, X_s^{t,x}) \quad s \in [t, T],\ \mathrm{d}\mathbb{P} - a.s. \\ Z_s^{t,x} &= v(s, X_s^{t,x})\sigma(s, X_s^{t,x}) \quad s \in [t, T],\ \mathrm{d}\mathbb{P} \times \mathrm{d}s - a.s. \end{aligned} \tag{4.2.7}$$

Moreover, the Markov property holds $Y_{t+h}^{t,x} = Y_{t+h}^{t+h, X_{t+h}^{t,x}}$ for any $h \geq 0$ and $u \in C^{0,0}([0, T] \times \mathbb{R}^k)$.

Proof. See Section 3 in [64]. The sample path continuity of $Y_t^{t,x}$ follows from the mean-square continuity of $(Y_s^{t,x})_{s\in[t,T]}$ for $x \in \mathbb{R}^k$, $0 \leq t \leq s \leq T$, which in turn follows from inequality (4.2.6). combined with the Lipschitz property of $x \mapsto g(x)$ and

$(t,x) \mapsto f(t,x,\cdot,\cdot)$ along with the continuity properties of $(t,x) \mapsto X_\cdot^{t,x}$ solution to (4.1.1).

The Markov property follows from Remark 3.1 [64] and the continuity of $u(t,x)$ is implied by that of $Y_t^{t,x}$. $\square$

4.2.4 Non-linear Feynman–Kac formula

As pointed out in the introduction, our aim is to deepen the connection between FBSDEs and PDEs via the so called non-linear Feynman–Kac formula, i.e. we study the probabilistic representation of the solution to a class of parabolic PDEs on $\mathbb{R}^k$ with polynomial growth coefficients that are associated with FBSDE (4.1.1)-(4.1.2). For $(t,x) \in [0,T] \times \mathbb{R}^d$, denote by $\mathcal{L}$ the infinitesimal generator of the Markov process $X^{t,x}$ solution to (4.1.1)

$$\mathcal{L} := \frac{1}{2} \sum_{i,j=1}^{d} ([\sigma\sigma^*]_{ij})(t,x)\partial^2_{x_i x_j} + \sum_{i=1}^{d} b_i(t,x)\partial_{x_i}, \tag{4.2.8}$$

and consider for a function $v = (v_1, \cdots, v_k)$ the following system of backward semi-linear parabolic PDEs for $i \in \{1, \cdots, k\}$

$$-\partial_t v_i(t,x) - \mathcal{L}v_i(t,x) - f_i\big(t,x,v(t,x),(\nabla v\sigma)(t,x)\big) = 0, \qquad v(T,x) = g(x). \tag{4.2.9}$$

In rough it can be easily proved using Itô's formula that if $v \in C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R}^k)$ solves the above PDE then $Y_t := v(t,X_t)$ and $Z_t := (\nabla v\, \sigma)(t,X_t)$ solves BSDE (4.1.2) (see Proposition 3.1 in [64]). But the more interesting result is the converse one, i.e. that $u(t,x) := Y_t^{t,x}$ is the solution of the PDE (in some sense). It was established in Theorem 3.2 of [64] (recalled next) that indeed $(t,x) \mapsto Y_t^{t,x}$ is the viscosity solution of the PDE.

Theorem 4.2.5. *Let (HX0), (HY0) hold and take $(t,x) \in [0,T] \times \mathbb{R}^d$. Furthermore, assume that the i-th component of the driver function f depends only on the i-th row of the matrix $z \in \mathbb{R}^{k\times d}$, i.e. $f_i(t,x,y,z) = f_i(t,x,y,z^i)$.*

Then $u(t,x) := Y_t^{t,x}$ is a continuous function of (t,x) that grows at most polynomially at infinity and is a viscosity solution of (4.2.9) *(in the sense of Definition 3.2 in [64]).*

Remark 4.2.6 (Multi-dimensional case)**.** The proof of Theorem 4.2.5 relies on a BSDE comparison theorem that holds only in the case $k = 1$ (i.e. when Y is one-dimensional). Nonetheless, with the restriction imposed by (HY0), it is still possible to use the said comparison theorem to prove Theorem 4.2.5, we point the reader to Theorem 2.4 and Remark 2.5 in [64].

It is possible to show that $(t,x) \mapsto Y_t^{t,x}$ is the solution to (4.2.9) not only in the viscosity sense, but also in weak sense (in weighted Sobolev spaces), this has been done in [60] and [83].

One equation covered by our setting is the FitzHugh-Nagumo PDE with recovery, usually used in biology to model the electrical distribution of the heart.

Example 4.2.7 (The FH-N equation with recovery)**.** Let $(t,x) \in [0,T] \times \mathbb{R}^d$, $g = (g_u, g_v)$, $f = (f_u, f_v)$ and $g, f, (u,v) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^2$. The FH-N PDE has a dynamics of the type

$$
\begin{aligned}
&-\partial_t u - \Delta u - f_u(u,v) = 0, \quad -\partial_t v - \Delta v - f_v(u,v) = 0, \\
&\quad \text{with } \ u(T,\cdot) = g_u(\cdot), \ v(T,\cdot) = g_v(\cdot).
\end{aligned}
$$

where $f_u(u,v) = u - u^3 + v$ and $f_v(u,v) = u - v$. f clearly satisfies, (HY0) and (HY0$_{\text{loc}}$).

4.3 Representation results, path regularity and other properties

As seen before $u(t,x) := Y_t^{t,x}$ solves the PDE (4.2.9) in viscosity sense. If $u \in C^{1,2}$ we would also obtain the representation of the process Z as $Z_t^{t,x} = (\nabla_x u \sigma)(t,x)$, but in view of Theorem 4.2.5 we have not given meaning to $\nabla_x u$. The main aim of this section is to first prove some representation formulas, that express Z as a function of Y and X, then use these representation formulas to obtain the so called L^2- (and L^p-) path regularity results needed to prove the convergence of the numerical discretization of FBSDE (4.1.1)-(4.1.2) in the later sections.

4.3.1 Differentiability in the spatial parameter

Take the system (4.1.1)-(4.1.2) into account. We now show that the smoothness of the FBSDE parameters b, σ, g, f carries over to the solution process $\Theta = (X, Y, Z)$.

Theorem 4.3.1. *Let (HXY1) hold and $(t,x) \in [0,T] \times \mathbb{R}^d$.*

Then u (from (4.2.7)) is continuously differentiable in its spatial variable. Moreover, the triple $\nabla_x \Theta^{t,x} = (\nabla_x X^{t,x}, \nabla_x Y^{t,x}, \nabla_x Z^{t,x}) \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ for any $p \geq 2$ and solves for $0 \leq t \leq s \leq T$

$$\begin{cases} \nabla_x X_s^{t,x} & = I_d + \int_t^s (\nabla_x b)(r, X_r^{t,x}) \nabla_x X_r^{t,x} \mathrm{d}r + \int_t^s (\nabla_x \sigma)(r, X_r^{t,x}) \nabla_x X_r^{t,x} \mathrm{d}W_r, \\ \nabla_{x_i} Y_s^{t,x} & = (\nabla_x g)(X_T^{t,x}) \nabla_{x_i} X_T^{t,x} - \int_s^T \nabla_{x_i} Z_r^{t,x} \mathrm{d}W_r + \int_t^T F(r, \nabla_{x_i} \Theta_r^{t,x}) \mathrm{d}r \end{cases} \tag{4.3.1}$$

for $i \in \{1, \cdots, d\}$ and with[1]

$$F : (\omega, r, x, \chi, \Upsilon, \Gamma) \mapsto \big(\nabla_x f\big)(r, \Theta_r^{t,x}) \cdot \chi + \big(\nabla_y f\big)(r, \Theta_r^{t,x}) \cdot \Upsilon + \big(\nabla_z f\big)(r, \Theta_r^{t,x}) \cdot \Gamma.$$

There exists a positive constant C_p independent of x such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \|(\nabla_x Y^{t,x}, \nabla_x Z^{t,x})\|_{\mathcal{S}^p \times \mathcal{H}^p} \leq C_p. \tag{4.3.2}$$

Furthermore, for u as in (4.2.7) we have for $x \in \mathbb{R}^d$ and $0 \leq t \leq s \leq T$

$$\nabla_x Y_s^{t,x} = (\nabla_x u)(s, X_s^{t,x}) \nabla_x X_s^{t,x} \quad \mathbb{P}-a.s. \quad \text{and} \quad \|\nabla_x u\|_\infty < \infty. \tag{4.3.3}$$

We recall that $\nabla_x Y^{t,x}$ is $\mathbb{R}^{k \times d}$-valued and $\nabla_{x_i} Y^{t,x}$ denotes its i-th column. Similar notation follows for $\nabla_x X$ and $\nabla_x Z$.

Proof. Throughout fix $(t,x) \in [0,T] \times \mathbb{R}^d$ and let $\{e_i\}_{i \in \{1, \cdots, d\}}$ be the canonical unit vectors of $\mathbb{R}^d$. Let $i \in \{1, \cdots, d\}$.

The results concerning SDE (4.1.1) follow from those in Subsection 2.5 in [46]. We start by showing that the partial derivatives $(\nabla_{x_i} Y^{t,x}, \nabla_{x_i} Z^{t,x})$ for any i exist, then we will show the full differentiability. We start by proving that (4.3.1) has indeed a solution for every i. Unfortunately, the driver of (4.3.1) does not satisfy (HY0) and hence we cannot quote Theorem 4.2.2 directly; we use a more general result from [11]. We remark though, that the techniques used to obtain moment estimates of the form of (4.2.4) and (4.2.6) are the same in both [11] and [64].

FBSDE (4.3.1) has a unique solution $\Xi^{t,x,i} := (\nabla_{x_i} X^{t,x}, U^{t,x,i}, V^{t,x,i}) \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$

[1] The term $(\nabla_z f)(\cdot, \Theta) \cdot \Gamma$ can be better understood if one interprets z in f not as in $\mathbb{R}^{k \times d}$ but as $(\mathbb{R}^d)^k$, i.e. f recieves not a matrix but its $\mathbb{R}^d$-valued k lines.

for any $p \geq 2$, where (U^i, V^i) replaces $(\nabla_{x_i} Y, \nabla_{x_i} Z)$. This follows by a direct application of Theorem 4.2 in [11]. It is easy to see that under (HXY1) the conditions (H1)-(H5) in [11] (p118-119) are satisfied. First, under (HXY1), standard SDE theory (see e.g. Theorem 2.4 in [46]) ensures that $\nabla_x X \in \mathcal{S}^p$ for all $p \geq 2$, which along with $\nabla_x g, \nabla_x f \in C_b^{0,0}$, implies in turn that the terminal condition $(\nabla_x g)(X_T^{t,x})\nabla_{x_i} X_T^{t,x} \in L^p_{\mathcal{F}_T}$ and the term $\big(\nabla_x f\big)(\cdot, \Theta_\cdot^{t,x})\nabla_{x_i} X_\cdot^{t,x} = F(\cdot, \nabla_{x_i} X_\cdot^{t,x}, 0, 0) \in \mathcal{S}^p$ for any $p \geq 2$. Given the linearity of F and the Lipschitz property of f in its z-variable it follows that F is uniformly Lipschitz in Γ. Moreover, since f satisfies (4.2.1) it implies that F is monotone[2] in Υ, i.e.

$$\big\langle \Upsilon - \Upsilon', \big(\nabla_y f\big)(\cdot, \Theta_\cdot^{t,x}) \cdot (\Upsilon - \Upsilon')\big\rangle \leq L_y |\Upsilon - \Upsilon'|^2, \quad \text{for any } \Upsilon, \Upsilon' \in \mathbb{R}^k. \tag{4.3.4}$$

The continuity of $\Upsilon \mapsto F(r, x, \chi, \Upsilon, \Gamma)$ is also clear. Lastly, the linearity of F, the fact that $\Theta \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ for any $p \geq 2$ and (4.2.2) implies that condition (H5) in [11] is also satisfied, i.e. that for any $R > 0$, $\sup_{|\Upsilon| \leq R} |F(r, x, \nabla_{x_i} X_r^{t,x}, \Upsilon, 0) - F(r, x, \nabla_{x_i} X_r^{t,x}, 0, 0)| \in L^1([t, T] \times \Omega)$. We are therefore under the conditions of Theorem 4.2 in [11], as claimed.

In view of (4.2.3) and the linearity of F one can obtain moment estimates in the style of (4.2.4) by following arguments similar to those in the proof of Theorem 4.2.2 (recall that (4.2.3) takes in this case a very simple form). In view of (4.2.4), we have (recall that $\nabla X \in \mathcal{S}^p$ for all $p \geq 2$)

$$\begin{aligned} \|U^i\|_{\mathcal{S}^p}^p + \|V^i\|_{\mathcal{H}^p}^p &\leq C_p \left\{ \|(\nabla_x g)(X_T^{t,x})\nabla_{x_i} X_T^{t,x}\|_{L^p}^p + \|\big(\nabla_x f\big)(\cdot, \Theta_\cdot^{t,x})\nabla_{x_i} X_\cdot^{t,x}\|_{\mathcal{H}^p}^p \right\} \\ &\leq C_p \|\nabla_{x_i} X^{t,x}\|_{\mathcal{S}^p}^p \leq C_p, \end{aligned} \tag{4.3.5}$$

where C_p does *not* depend on x, t or i.

In order to obtain results on the first order variation of the solution, we follow standard BSDE techniques used already in [46], [10] or [28]; we start by studying the behaviour of $\Theta^{t,x+\varepsilon e_i} - \Theta^{t,x}$ for any $\varepsilon > 0$. Take $h \in \mathbb{R}^d$. Via the stability of SDEs and inequality (4.2.6) (and (HY0)), it is clear that a constant $C_p > 0$ independent of x exists such that

$$\lim_{h \to 0} \|\Theta^{t,x+h} - \Theta^{t,x}\|_{\mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p} \leq \lim_{h \to 0} C_p \|X^{x+h} - X^x\|_{\mathcal{S}^p} \leq \lim_{h \to 0} C_p |h| = 0. \tag{4.3.6}$$

[2] This follows easily from the differentiability of f, its monotonicity in y and the definition of directional derivative.

Define $\delta\Theta^{\varepsilon,i} := (\delta X^{\varepsilon,i}, \delta Y^{\varepsilon,i}, \delta Z^{\varepsilon,i}) := (\Theta^{t,x+\varepsilon e_i} - \Theta^{t,x})/\varepsilon - (\nabla_{x_i} X^{t,x}, U^{t,x,i}, V^{t,x,i})$ for which

$$\begin{aligned}\delta Y_s^{\varepsilon,i} = &\big[\frac{1}{\varepsilon}\big(g(X_T^{t,x+\varepsilon e_i}) - g(X_T^{t,x})\big) - (\nabla_x g)(X_T^{t,x})\nabla_{x_i} X_T^{t,x}\big] - \int_s^T \delta Z_r^{\varepsilon,i} \mathrm{d}W_r \\ &+ \int_s^T \big[\frac{1}{\varepsilon}\big(f(r,\Theta_r^{t,x+\varepsilon e_i}) - f(r,\Theta_r^{t,x})\big) - F(r,x,\nabla_{x_i} X_r^{t,x}, U_r^{t,x,i}, V_r^{t,x,i})\big]\mathrm{d}r.\end{aligned} \tag{4.3.7}$$

Using the differentiability of the involved functions we can re-write (4.3.7) as a linear FBSDE with random coefficients satisfying in its essence a (HY0) type assumption: for $s \in [t,T]$, $j \in \{1,\cdots,d\}$

$$\begin{cases} \delta X_s^{\varepsilon,j} &= 0 + \int_t^s \big[b_x^{\varepsilon,j}(r)\delta X_r^{\varepsilon,j} + \delta\nabla b_r^{\varepsilon}\ \nabla_{x_j} X_r^{t,x}\big]\mathrm{d}r + \int_t^s \big[\sigma_x^{\varepsilon,j}(r)\delta X_r^{\varepsilon,j} + \delta\nabla\sigma_r^{\varepsilon}\ \nabla_{x_j} X_r^{t,x}\big]\mathrm{d}W_r, \\ \delta Y_s^{\varepsilon,i} &= \big[g_x^{\varepsilon,i}(T)\delta X_T^{\varepsilon,i} + \delta\nabla g_T^{\varepsilon}\ \nabla_{x_i} X_T^{t,x}\big] - \int_s^T \delta Z_r^{\varepsilon,i}\mathrm{d}W_r \\ &\quad + \int_s^T \big[f_x^{\varepsilon,i}(r)\delta X_r^{\varepsilon,i} + f_y^{\varepsilon,i}(r)\delta Y_r^{\varepsilon,i} + f_z^{\varepsilon,i}(r)\delta Z_r^{\varepsilon,i} + \delta\nabla f_r^{\varepsilon}\cdot(\nabla_{x_i} X_r^{t,x}, U_r^{t,x,i}, V_r^{t,x,i})\big]\mathrm{d}r, \end{cases} \tag{4.3.8}$$

where $\delta\nabla f$ and $\delta\nabla\varphi$ denote the differences

$$\delta\nabla f_\cdot^{\varepsilon} := \big(f_x^{\varepsilon,i}, f_y^{\varepsilon,i}, f_z^{\varepsilon,i}\big)(\cdot) - \big(\nabla_x f, \nabla_y f, \nabla_z f\big)(\cdot, \Theta_\cdot^{t,x}),$$

and

$$\delta\nabla\varphi_\cdot^{\varepsilon} := \varphi_x^{\varepsilon,i}(\cdot) - \nabla_x\varphi(\cdot, \Theta_\cdot^{t,x}),$$

for $\varphi \in \{b,\sigma,g\}$ (with some abuse of notation) and $r \in [t,T]$, and where we defined

$$\begin{aligned}\varphi_x^{\varepsilon,i}(r) :&= \int_0^1 (\nabla_x\varphi)\left(r, (1-\lambda)X_r^{t,x} + \lambda X_r^{t,x+\varepsilon e_i}\right)\mathrm{d}\lambda \\ &= \int_0^1 (\nabla_x\varphi)\left(r, X_r^{t,x} + \lambda\left(X_r^{t,x+\varepsilon e_i} - X_r^{t,x}\right)\right)\mathrm{d}\lambda,\end{aligned}$$

and $f_*^{\varepsilon,i}$ for $* \in \{x, y, z\}$ in the following way:

$$\begin{aligned}
f_z^{\varepsilon,i}(r) &:= \int_0^1 (\nabla_z f)\big(r, X_r^{t,x+\varepsilon e_i}, Y_r^{t,x+\varepsilon e_i}, Z_r^{t,x} + \lambda(Z_r^{t,x+\varepsilon e_i} - Z_r^{t,x})\big)\mathrm{d}\lambda,\\
f_y^{\varepsilon,i}(r) &:= \int_0^1 (\nabla_y f)(r, X_r^{t,x+\varepsilon e_i}, Y_r^{t,x} + \lambda(Y_r^{t,x+\varepsilon e_i} - Y_r^{t,x}), Z_r^{t,x})\mathrm{d}\lambda,\\
f_x^{\varepsilon,i}(r) &:= \int_0^1 (\nabla_x f)(r, X_r^{t,x} + \lambda(X_r^{t,x+\varepsilon e_i} - X_r^{t,x}), Y_r^{t,x}, Z_r^{t,x})\mathrm{d}\lambda.
\end{aligned}$$

The assumptions imply immediately that $b_x^{\varepsilon,i}, \sigma_x^{\varepsilon,i}, f_x^{\varepsilon,i}, f_z^{\varepsilon,i}$ are uniformly bounded, while $f_y^{\varepsilon,i} \in \mathcal{S}^p$, $p \geq 2$ (thanks to $HY0_{loc}$). Furthermore, using estimate (4.2.4) (along with $\|X^{t,x}\|_{\mathcal{S}^p}^p \leq C_p(1 + |x|^p)$), (4.3.5), (4.3.6), the continuity of $\varphi \in \{b, \sigma, g\}$ and its derivative it is easy to see that, in combination with the dominated convergence theorem, one has

$$\lim_{\varepsilon \to 0} \big\{ \|\varphi_x^{\varepsilon,i}(\cdot) - \nabla_x \varphi(\cdot, \Theta_\cdot^{t,x})\|_{\mathcal{S}^p} + \|\big(f_x^{\varepsilon,i}, f_y^{\varepsilon,i}, f_z^{\varepsilon,i}\big)(\cdot) - (\nabla_x f, \nabla_y f, \nabla_z f)(\cdot, \Theta_\cdot^{t,x})\|_{\mathcal{H}^p} \big\} = 0. \tag{4.3.9}$$

We remark that in the above limit a localization argument for the convergence of $f_y^{\varepsilon,i}(\cdot)$ to $\nabla_y f(\cdot, \Theta_\cdot)$ is required, namely that we work inside a ball (of any given radius) centered around x in which all points $x + \varepsilon e_i \in \mathbb{R}^d$ as ε vanishes are contained. We do not detail the argumentation since it is similar to that given in e.g. [46], [9] or [28].

With this in mind we return to (4.3.7), written in the form of (4.3.8), and since it is a linear FBSDE satisfying the monotonicity condition (4.2.1) we have via Corollary 3.3 in [8] (essentially our moment estimate (4.2.4) for FBSDE (4.3.8)) in combination with (4.3.5), (4.3.6) and (4.3.9), that for any i

$$\lim_{\varepsilon \to 0} \|\frac{1}{\varepsilon}(\Theta^{t,x+\varepsilon e_i} - \Theta^{t,x}) - (\nabla_{x_i} X^{t,x}, U^{t,x,i}, V^{t,x,i})\|_{\mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p} = 0, \qquad \text{for any } p \geq 2.$$

Since the limit exists we identify $(\nabla_{x_i} Y^{t,x}, \nabla_{x_i} Z^{t,x})$ with $(U^{t,x,i}, V^{t,x,i})$ and, moreover, estimate (4.3.5) implies estimate (4.3.2). Furthermore, the above limit implies in particular that (take $s = t$)

$$\nabla_{x_i} u(t,x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[u(t, x + \varepsilon e_i) - u(t,x)] = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[Y_t^{t,x+\varepsilon e_i} - Y_t^{t,x}] = \nabla_{x_i} Y_t^{t,x}.$$

Observing that the RHS of (4.3.5) is a constant independent of $t \in [0, T]$, $x \in \mathbb{R}^d$ and

$i \in \{1, \cdots, d\}$ we can conclude that $\|\nabla_{x_i} u\|_\infty = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla_{x_i} Y_t^{t,x}| < \infty$.

It is clear that $(\nabla_{x_i} Y_s^{t,x})_{s\in[t,T]}$ is continuous in its time parameter as it is a solution to a BSDE; we now focus on the continuity of $x \mapsto \nabla_{x_i} Y_t^{t,x}$. Let $x, x' \in \mathbb{R}^d$. The difference $\nabla_{x_i} Y^{t,x} - \nabla_{x_i} Y^{t,x'}$ is the solution to a *linear* FBSDE following from (4.3.1). As before, it is easy to adapt the computations and apply Corollary 3.3 in [8] (essentially our moment estimate (4.2.6) for FBSDEs (4.3.1)) to the difference $\nabla_{x_i} Y_s^{t,x} - \nabla_{x_i} Y_s^{t,x'}$ yielding

$$
\begin{aligned}
&\|\nabla_{x_i} Y^{t,x} - \nabla_{x_i} Y^{t,x'}\|_{\mathcal{S}^2}^2 \\
&\leq C_p \Big\{ \big\| (\nabla_x g)(X_T^{t,x}) \nabla_{x_i} X_T^{t,x} - (\nabla_x g)(X_T^{t,x'}) \nabla_{x_i} X_T^{t,x'} \big\|_{L^2}^2 \\
&\qquad + \mathbb{E}\Big[\Big(\int_0^T |F(r, x, \nabla_{x_i} X_r^{t,x}, \nabla_{x_i} Y_r^{t,x}, \nabla_{x_i} Z_r^{t,x}) - F(r, x', \nabla_{x_i} X_r^{t,x'}, \nabla_{x_i} Y_r^{t,x}, \nabla_{x_i} Z_r^{t,x}) | \mathrm{d}s \Big)^p \Big] \Big\}.
\end{aligned}
$$

Given the known results on SDEs, the linearity of F, (4.3.5), the continuity of the derivatives of f and (4.3.6), dominated convergence theorem yields that $\|\nabla_{x_i} Y^{t,x} - \nabla_{x_i} Y^{t,x'}\|_{\mathcal{S}^2}^2 \to 0$ as $x' \to x$ uniformly on compact sets. This mean-square continuity of $\nabla_{x_i} Y^{t,x}$ implies in particular that $\nabla_{x_i} Y_t^{t,x} = \nabla_{x_i} u(t,x)$ is continuous. In conclusion, we just proved that for any $i \in \{1, \cdots, d\}$ the partial derivatives $\nabla_{x_i} u$ exist and are continuous, hence, standard multi-dimensional real analysis implies that u is continuously differentiable in its spatial variables. This argumentation is similar to that in the proof of Corollary 4.2.4.

We are left to prove (4.3.3). Note that for any $\varepsilon > 0$ we have $(Y_s^{t,x+\varepsilon e_i} - Y_s^{t,x})/\varepsilon = (u(s, X_s^{t,x+\varepsilon e_i}) - u(s, X_s^{t,x}))/\varepsilon$. By sending $\varepsilon \to 0$ and using the (continuous) differentiability of u, we have $\nabla_x Y_s^{t,x} = (\nabla_x u)(s, X_s^{t,x}) \nabla_x X_s^{t,x}$. Hence, as the RHS of (4.3.5) is a constant independent of $t \in [0,T]$, $x \in \mathbb{R}^d$ and i we can conclude (let $s \searrow t$) that $\|\nabla_x u\|_\infty = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla_x Y_t^{t,x}| < \infty$. □

4.3.2 Malliavin differentiability

As in the previous section we show a form of regularity of the solution Θ to (4.1.1)-(4.1.2), namely the stochastic variation of Θ in the sense of Malliavin's calculus.

Theorem 4.3.2 (Malliavin differentiability)**.** *Let (HXY1) hold. Then the solution* $\Theta = (X, Y, Z)$ *of* (4.1.1)-(4.1.2) *verifies*

- *$X \in \mathbb{L}_{1,2}$ and DX admits a version $(u,t) \mapsto D_u X_t$ satisfying for $0 \le u \le t \le T$*

$$D_u X_t = \sigma(u, X_u) + \int_u^t (\nabla_x b)(s, X_s) D_u X_s \mathrm{d}s + \int_u^t (\nabla_x \sigma)(s, X_s) D_u X_s \mathrm{d}W_s.$$

Moreover, for any $p \ge 2$ there exists $C_p > 0$ such that

$$\sup_{u \in [0,T]} \|D_u X\|_{\mathcal{S}^p}^p \le C_p(1 + |x|^p). \tag{4.3.10}$$

- *for any $0 \le t \le T$, $x \in \mathbb{R}^m$ we have $(Y, Z) \in \mathbb{L}_{1,2} \times \big(\mathbb{L}_{1,2}\big)^d$. A version of $(DY, DZ)_{0 \le u,t \le T}$ satisfies : for $t < u \le T$, $D_u Y_t = 0$ and $D_u Z_t = 0$, and for $0 \le u \le t$,*

$$D_u Y_t = (\nabla_x g)(X_T) D_u X_T + \int_t^T \langle (\nabla f)(s, \Theta_s), D_u \Theta_s \rangle \mathrm{d}s - \int_t^T D_u Z_s \mathrm{d}W_s. \tag{4.3.11}$$

Moreover, $(D_t Y_t)_{0 \le t \le T}$ defined by the above equation is a version of $(Z_t)_{0 \le t \le T}$.

- *the following representation holds for any $0 \le u \le t \le T$ and $x \in \mathbb{R}^m$*

$$D_u X_t = \nabla_x X_t (\nabla_x X_u)^{-1} \sigma(u, X_u) 1_{[0,u]}(t), \tag{4.3.12}$$
$$D_u Y_t = \nabla_x Y_t (\nabla_x X_u)^{-1} \sigma(u, X_u), \quad a.s., \tag{4.3.13}$$
$$Z_t = \nabla_x Y_t (\nabla_x X_t)^{-1} \sigma(s, X_t), \quad a.s.. \tag{4.3.14}$$

Remark 4.3.3 (Y is already in $\mathbb{L}^{1,2}$)**.** Via Theorem 4.3.1 we know that $u \in C^{0,1}$. Under (HXY1) it is known that $X \in \mathbb{L}^{1,2}$ (see [63]) hence using the chain rule (for Malliavin calculus, see Proposition 1.2.3 in [63]) we obtain $Y_\cdot = u(\cdot, X_\cdot) \in \mathbb{L}_{1,2}$. A careful analysis of Theorem 4.3.1 and the results about $\nabla_x u$ show that indeed $X, Y \in \mathbb{L}^{1,p}$ for all $p \ge 2$ (just combine (4.3.10) with (4.7.1) as described in Subsection 4.7.2).

Using the fact that $X, Y \in \mathbb{L}^{1,2}$, the statement of Theorem 4.3.2 follows easily if the driver f in (4.1.2) does not depend on z. One would argue in the following way: for any $t \in [0,T]$

$$\left(g(X_T) - Y_t + \int_t^T f(r, X_r, Y_r) \mathrm{d}r \right)_{t \in [0,T]} \in \mathbb{L}^{1,2} \Rightarrow \left(\int_t^T Z_r \mathrm{d}W_r \right)_{t \in [0,T]} \in \mathbb{L}^{1,2} \Leftrightarrow Z \in \mathbb{L}^{1,2},$$

this follows from the definition of the BSDE (4.1.2) itself and Theorem 4.7.3. The dynamics of (4.3.11) and the representation formulas (4.3.13), (4.3.14) follow by arguments similar to those given below.

Proof of Theorem 4.3.2. The first part of the statement is trivial as it follows from standard SDE theory, see e.g. [63] or Theorem 2.5 in [46]. To prove the other statements of the theorem, we will use an identification trick by taking advantage of the fact we already know that $Y \in \mathbb{L}^{1,2}$ (see Remark 4.3.3).

Let (X, Y, Z) be the solution of (4.1.1)-(4.1.2) and define the following BSDE:

$$U_t = g(X_T) + \int_t^T \widehat{f}(r, V_r)\mathrm{d}r - \int_t^T V_r \mathrm{d}W_r, \tag{4.3.15}$$

where the driver $\widehat{f} : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is defined as

$$\widehat{f}(t, v) := f(t, X_t, Y_t, v) = f\left(t, X_t, u(t, X_t), v\right). \tag{4.3.16}$$

It is clear that: $g(X_T) \in \mathbb{D}^{1,2}$, $f(\cdot, X_\cdot, Y_\cdot, 0) \in \mathbb{L}^{1,p}$ for all $p \geq 2$ (see Remark 4.3.3) and that $v \mapsto \widehat{f}(\cdot, v)$ is a Lipschitz continuous function, all these imply in particular via Lipschitz BSDE theory (see Theorem 2.1, Proposition 2.1 in [31]) that there exists a pair $(U, V) \in \mathcal{S}^2 \times \mathcal{H}^2$ solving (4.3.15). Furthermore, Theorem 4.2.2 in [31] states that the solution to (4.1.2) is unique and hence the solution of (4.3.15) verifies $(U, V) = (Y, Z)$.

Proposition 5.3 in [31], yields the existence of the Malliavin derivatives (DU, DV) of (U, V) with the following dynamics. Set $\Xi := (X, Y, V)$, then for $t < u \leq T$ we have $D_u U_t = 0$, $D_u V_t = 0$ and

$$D_u U_t = (\nabla_x g)(X_T) D_u X_T + \int_t^T \big\langle (\nabla f)(s, \Xi_s), (D_u \Xi_s) \big\rangle \mathrm{d}s - \int_t^T D_u V_s \mathrm{d}W_s, \qquad 0 \leq u \leq t.$$

Since $(U, V) = (Y, Z)$ then from the above BSDE for (DU, DV) follows BSDE (4.3.11). Moreover, Proposition 5.9 in [31] yields (4.3.13) and (4.3.14) for (U, V) which carry out for (Y, Z). ∎

4.3.3 Representation results

Here we combine the results of the two previous subsections to obtain representation formulas that will allow us to establish the path regularity properties of Y and Z required for the convergence proof of the numerical discretization.

Theorem 4.3.4. *Let the assumptions of Theorem 4.3.1 and 4.3.2 hold.*

Then the following representation holds for all $0 \le t \le s \le T$ $\mathrm{d}\mathbb{P} - a.s.$

$$Z_s^{t,x} = (\nabla_x u\sigma)\,(s, X_s^{t,x}) \tag{4.3.17}$$

$$= \nabla_x Y_s^{t,x} \left(\nabla_x X_s^{t,x}\right)^{-1} \sigma(s, X_s^{t,x}). \tag{4.3.18}$$

Moreover $Z \in \mathcal{S}^q$ *for any* $q \ge 2$ *with*

$$\|Z\|_{\mathcal{S}^q} \le C_q(1 + |x|^q), \quad q \ge 2. \tag{4.3.19}$$

Proof. The representation $Z = \nabla Y(\nabla X)^{-1}\sigma(\cdot, X)$ follows from Theorem 4.3.2, while from Theorem 4.3.1 we have

$$\begin{aligned} Z_s^{t,x} &= \nabla_x Y_s^{t,x}(\nabla_x X_s^{t,x})^{-1}\sigma(s, X_s^{t,x}) = (\nabla_x u)(s, X_s^{t,x}) \left(\nabla_x X_s^{t,x} \left(\nabla_x X_s^{t,x}\right)^{-1}\right) \sigma(s, X_s^{t,x}) \\ &= (\nabla_x u)(s, X_s^{t,x})\sigma(s, X_s^{t,x}). \end{aligned}$$

Since all the involved processes (in the RHS) are continuous we can identify Z with its continuous version. Moreover, as all the processes in the RHS belong to $\mathcal{S}^p$ for all $p \ge 2$ it follows that $Z \in \mathcal{S}^p$ for all $p \ge 2$. Combining Hölder's inequality with the fact that $X, \nabla X \in \mathcal{S}^p$ for all $p \ge 2$ and estimate (4.3.2), leads to (4.3.19), i.e.

$$\begin{aligned} \|Z\|_{\mathcal{S}^p} &= \|\nabla_x Y_\cdot^{t,x}(\nabla_x X_\cdot^{t,x})^{-1}\sigma(\cdot, X_\cdot^{t,x})\|_{\mathcal{S}^p} \\ &\le C_p\|\nabla_x Y^{t,x}\|_{\mathcal{S}^{3p}}\|(\nabla_x X)^{-1}\|_{\mathcal{S}^{3p}}\|1 + X^{t,x}\|_{\mathcal{S}^{3p}} \le C_p(1 + |x|). \end{aligned}$$

□

4.3.4 Path regularity results

Now let π be a partition of the interval $[0, T]$, say $0 = t_0 < \cdots < t_i < \cdots < T_N = T$, and mesh size $|\pi| = \max_{i=0,\cdots,N-1}(t_{i+1} - t_i)$. Given π, we also consider $r_\pi = |\pi|/\big(\min_{i=0,\cdots,N-1}(t_{i+1} - t_i)\big)$.

Let Z be the control process in the solution to BSDE (4.1.2), under (HX0)-(HY0). We define a set of random variables $\{\bar{Z}_{t_i}\}_{t_i\in\pi}$ term wise given by

$$\bar{Z}_{t_i} = \frac{1}{t_{i+1}-t_i}\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s \big| \mathcal{F}_{t_i}\Big], \quad 0\leq i\leq N-1, \quad \text{and} \quad \bar{Z}_{t_N} = Z_T. \qquad (4.3.20)$$

The RV Z_T can be obtained using (4.3.17), namely $Z_T = (\nabla_x g)(X_T)\sigma(T, X_T)$ when $g\in C^1$. If g is only Lipschitz continuous then one easily sees that a RV $G\in L^\infty(\mathcal{F}_T)$ exists such that $Z_T = G\sigma(T,X_T)$. In any case, under (HX0) and (HY0) it easily follows that

$$\bar{Z}_{t_N} = Z_T \in L^p(\mathcal{F}_T) \quad \text{for any } p\geq 2 \qquad \text{and} \qquad \bar{Z}_{t_i}\in L^2 \quad \text{for any } t_i\in\pi. \quad (4.3.21)$$

It is not difficult to show that $\bar{Z}_{t_i}$ is the best $\mathcal{F}_{t_i}$-measurable square integrable RV approximating Z in $\mathcal{H}^2([t_i,t_{i+1}])$, i.e.

$$\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_i}|^2 \mathrm{d}s\Big] = \inf_{\xi\in L^2(\Omega,\mathcal{F}_{t_i})} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s-\xi|^2\mathrm{d}s\Big]. \qquad (4.3.22)$$

Let now $\bar{Z}_t := \bar{Z}_{t_i}$ for $t\in[t_i,t_{i+1})$, $0\leq i\leq N-1$. It is equally easy to see that $\bar{Z}$ converges to Z in $\mathcal{H}^2$ as $|\pi|$ vanishes: since Z is adapted, the family of processes Z^π indexed by our partition defined by $Z^\pi_t = Z_{t_i}$ for $t\in[t_i,t_{i+1})$ converges to Z in $\mathcal{H}^2$ as $|\pi|$ goes to zero. Since $\{\bar{Z}\}$ is the best $\mathcal{H}^2$-approximation of Z, we obtain

$$\|Z-\bar{Z}\|_{\mathcal{H}^2} \leq \|Z-Z^\pi\|_{\mathcal{H}^2}\to 0, \quad \text{as} \quad |\pi|\to 0,$$

although without knowing the rate of this convergence.

The next result expresses the modulus of continuity (in the time variable) for Y and Z.

Theorem 4.3.5 (Path regularity). *Let (HX0), (HY0$_{loc}$) hold. Then the unique solution (X,Y,Z) to (4.1.1)-(4.1.2) satisfies $(X,Y,Z)\in\mathcal{S}^p\times\mathcal{S}^p\times\mathcal{S}^p$ for all $p\geq 2$. Moreover,*

(i) for any $p\geq 2$ there exists a constant $C_p>0$ such that for $0\leq s\leq t\leq T$ we have

$$\mathbb{E}\left[\sup_{s\leq u\leq t}|Y_u-Y_s|^p\right] \leq C_p(1+|x|^p)|t-s|^{\frac{p}{2}}; \qquad (4.3.23)$$

(ii) for any $p \geq 2$ there exists a constant $C_p > 0$ such that for any partition π of $[0,T]$ with mesh size $|\pi|$

$$\sum_{i=0}^{N-1} \mathbb{E}\Big[\Big(\int_{t_i}^{t_{i+1}} |Z_t - Z_{t_i}|^2 \mathrm{d}t\Big)^{\frac{p}{2}} + \Big(\int_{t_i}^{t_{i+1}} |Z_t - Z_{t_{i+1}}|^2 \mathrm{d}t\Big)^{\frac{p}{2}}\Big] \leq C_p(1+|x|^p)|\pi|^{\frac{p}{2}}, \tag{4.3.24}$$

(iii) in particular, there exists a constant C such that for any partition $\pi = \{0 = t_0 < \cdots < t_N = T\}$ of the interval $[0,T]$ with mesh size $|\pi|$ we have

$$\max_{0\leq i\leq N-1} \sup_{t\in[t_i,t_{i+1}]} \Big\{\mathbb{E}\Big[|Y_t - Y_{t_i}|^2\Big] + \mathbb{E}\Big[|Y_t - Y_{t_{i+1}}|^2\Big]\Big\} + \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_i}|^2 \mathrm{d}s\Big] \leq C|\pi|.$$

If, as $|\pi| \to 0$, r_π remains bounded [3] *then*

$$\sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_{i+1}}|^2 \mathrm{d}s\Big] \leq C|\pi|.$$

Proof. Fix $(t,x) \in [0,T] \times \mathbb{R}^d$, take $s \in [t,T]$ and throughout this proof we work with $\Theta^{t,x}$ and $\nabla_x \Theta^{t,x}$; to avoid a notational overload we omit the super- and subscript and write Θ and $\nabla\Theta$.

We first prove point (i) and (ii) under Assumption (HXY1), then we use a mollification argument to recover the case (HX0)-(HY0$_{\text{loc}}$). We then explain how (iii) is obtained.

Proof of (i) under (HXY1): $Z \in \mathcal{S}^q$ for any $q \geq 2$ follows from (4.3.19). Hence writing the BSDE for the difference $Y_u - Y_s$ we have

$$Y_u - Y_s = \int_s^u f(r,\Theta_r)\mathrm{d}r - \int_s^u Z_r \mathrm{d}W_r \leq \int_s^u K\left(1 + |X_r| + |Y_r|^m + |Z_r|\right)\mathrm{d}r - \int_s^u Z_r \mathrm{d}W_r$$

hence taking absolute values, the sup over $u \in [s,t]$, power p, expectations and Jensen's inequality leads, for some constant $C_p > 0$, to

$$\mathbb{E}\Big[\sup_{u\in[s,t]} |Y_u - Y_s|^p\Big] \leq C_p\Big\{|t-s|^p\big(1 + \|(X,Y,Z)\|^p_{\mathcal{S}^p\times\mathcal{S}^p\times\mathcal{S}^p}\big) + \mathbb{E}\Big[\sup_{u\in[s,t]}\Big|\int_s^u Z_r \mathrm{d}W_r\Big|^p\Big]\Big\}.$$

[3] This is trivially satisfied for the uniform grid for which $r_\pi = 1$.

Applying Burkholder-Davis-Gundy's inequality (BDG) to the last term in the RHS yields

$$\mathbb{E}\Big[\sup_{u\in[s,t]}\Big|\int_s^u Z_r \mathrm{d}W_r\Big|^p\Big] \leq C_p \mathbb{E}\Big[\Big(\int_s^t |Z_r|^2 \mathrm{d}r\Big)^{\frac{p}{2}}\Big] \leq C_p |t-s|^{\frac{p}{2}} \|Z\|^p_{\mathcal{S}^p}.$$

It then follows that

$$\mathbb{E}\Big[\sup_{u\in[s,t]}|Y_u - Y_s|^p\Big] \leq C_p\left\{|t-s|^p + |t-s|^{\frac{p}{2}}\right\} \leq C_p(1+|x|^p)|t-s|^{\frac{p}{2}}.$$

Proof of (ii) under (HXY1): To prove the desired inequality we use the representation (4.3.14) (alternatively (4.3.18)). We first estimate the difference $\mathbb{E}[\big(\int_{t_i}^{t_{i+1}}|Z_s - Z_{t_i}|^2\mathrm{d}s\big)^{p/2}]$. The difference $Z_s - Z_{t_i}$ can be written as $Z_s - Z_{t_i} = I_1 + I_2$ with $I_2 := (\nabla Y_s - \nabla Y_{t_i})\,(\nabla X_{t_i})^{-1}\,\sigma(t_i, X_{t_i})$ and

$$I_1 := \nabla Y_s\left\{\left((\nabla X_s)^{-1} - (\nabla X_{t_i})^{-1}\right)\sigma(s,X_s) + (\nabla X_{t_i})^{-1}\left[\sigma(s,X_s) - \sigma(t_i,X_{t_i})\right]\right\}.$$

The estimation of I_1 is rather easy as it relies on Hölder's inequality combined with (4.3.2), (HX0), Theorems 2.3 and 2.4 in [46] (see proof of Theorem 5.5(i) in [46]), in short we have

$$\mathbb{E}[|I_1|^p] \leq C_p(1+|x|^p)|\pi|^{\frac{p}{2}}.$$

Concerning the second part, the estimation of I_2, it follows from an adaptation of the proof of Theorem 5.5(ii) in [45]. We reformulate the main argument and skip the obvious details. Let us start with a simple trick, as $s \in [t_i, t_{i+1}]$,

$$\begin{aligned}\mathbb{E}\Big[&|(\nabla Y_s - \nabla Y_{t_i})(\nabla X_{t_i})^{-1}\sigma(t_i,X_{t_i})|^p\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\big[|\nabla Y_s - \nabla Y_{t_i}|^p\big|\mathcal{F}_{t_i}\big]|(\nabla X_{t_i})^{-1}\sigma(t_i,X_{t_i})|^p\Big].\end{aligned} \tag{4.3.25}$$

Writing the BSDE for the difference $\nabla Y_s - \nabla Y_{t_i}$ for $t_i \leq s \leq t_{i+1}$ we have for some constant $C > 0$

$$\begin{aligned}&\mathbb{E}\Big[|\nabla Y_s - \nabla Y_{t_i}|^p\Big|\mathcal{F}_{t_i}\Big] \leq C\,\mathbb{E}\left[\widehat{I}_{[t_i,t_{i+1}]}\big|\mathcal{F}_{t_i}\right]\\ &\qquad\text{where}\quad \widehat{I}_{[t_i,t_{i+1}]} := \Big(\int_{t_i}^{t_{i+1}}|(\nabla f)(r,\Theta_r)|\,|\nabla\Theta_r|\mathrm{d}r\Big)^p + \Big(\int_{t_i}^{t_{i+1}}|\nabla Z_r|^2\mathrm{d}r\Big)^{p/2},\end{aligned}$$

where we used the conditional BDG inequality and maximized over the time interval $[t_i, t_{i+1}]$.

Combining these last two inequalities and observing that since ∇X_{t_i} and $\sigma(X_{t_i})$ are $\mathcal{F}_{t_i}$-adapted we can drop the conditional expectation from (4.3.25). Hence, for some $C > 0$,

$$
\begin{aligned}
\sum_{i=0}^{N-1} \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} |I_2|^2 \mathrm{d}s\right)^{\frac{p}{2}}\right] &\leq C|\pi|^{\frac{p}{2}-1} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[\,|I_2|^p\right] \mathrm{d}s \\
&\leq C\,|\pi|^{\frac{p}{2}-1} \sum_{i=0}^{N-1} |\pi| \mathbb{E}\big[\,|(\nabla X_{t_i})^{-1}\sigma(t_i, X_{t_i})|^p\, \widehat{I}_{[t_i,t_{i+1}]}\,\big] \\
&\leq C\,|\pi|^{\frac{p}{2}}\, \mathbb{E}\left[\sup_{0\leq t\leq T} |(\nabla X_t)^{-1}\sigma(t, X_t)|^p \sum_{i=0}^{N-1} \widehat{I}_{[t_i,t_{i+1}]}\right] \\
&\leq C\,|\pi|^{\frac{p}{2}}\, \|(\nabla X)^{-1}\|_{\mathcal{S}^{3p}}^{\frac{1}{3}} \|1 + X\|_{\mathcal{S}^{3p}}^{\frac{1}{3}} \|\widehat{I}_{[0,T]}\|_{L^1} \\
&\leq C\,(1 + |x|^p)|\pi|^{\frac{p}{2}}.
\end{aligned}
$$

The last line follows from standard inequalities (sum of powers is less than the power of the sum), the growth conditions on ∇f and the fact that for any $q \geq 2$ we have: $X, \nabla X, (\nabla X)^{-1} \in \mathcal{S}^q$, $Y, Z, \nabla Y \in \mathcal{S}^q$ and $\nabla Z \in \mathcal{H}^q$.

Collecting now the estimates we obtain the desired result for the difference $Z_s - Z_{t_i}$. To have the same estimate for the difference $Z_s - Z_{t_{i+1}}$ we need only to repeat the above calculations with a minor change in order to incorporate the $Z_{t_{i+1}}$: one writes $Z_s - Z_{t_{i+1}}$ with the help of I_1^{i+1} and I_2^{i+1}, which are I_1 and I_2 respectively but with t_{i+1} instead of t_i. The estimate for I_1^{i+1} follows from SDE theory in the same fashion as for I_1 above; concerning I_2^{i+1} one just needs another small trick,

$$
\begin{aligned}
I_2^{i+1} &= \big(\nabla Y_s - \nabla Y_{t_{i+1}}\big) \big(\nabla X_{t_{i+1}}\big)^{-1} \sigma(t_{i+1}, X_{t_{i+1}}) \\
&\leq \big(|\nabla Y_s| + |\nabla Y_{t_{i+1}}|\big) \big[\, \big(\nabla X_{t_{i+1}}\big)^{-1} \sigma(t_{i+1}, X_{t_{i+1}}) - (\nabla X_{t_i})^{-1} \sigma(t_i, X_{t_i})\big] && (4.3.26) \\
&\quad + \big(\nabla Y_s - \nabla Y_{t_{i+1}}\big) (\nabla X_{t_i})^{-1} \sigma(t_i, X_{t_i}) && (4.3.27)
\end{aligned}
$$

The rest of the proof follows just like before, like I_1 for (4.3.26) and like I_2 for (4.3.27).

Final step - (i) and (ii) under (HX0)-(HY0$_{loc}$) - arguing via mollification: In this step we rely on a standard mollification argument. Note that the mollified drivers are still monotone.

Take b^n, σ^n, g^n, f^n as mollified versions of b, σ, g, f in their spatial variables such that the mollified functions satisfy uniformly (in n) (HX0) and (HY0$_{\text{loc}}$), with uniform Lipschitz and monotonicity constant. Moreover, Theorem 4.2.2 ensures that $\Theta = (X^n, Y^n, Z^n) \in \mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ for any $p \geq 2$ and solves (4.1.1)-(4.1.2) with b^n, σ^n, g^n, f^n replacing b, σ, g, f. Furthermore, in view of (4.2.6) and the standard theory of SDEs it is rather simple to deduce that $\Theta^n \to \Theta$ as $n \to \infty$ in $\mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$ for all $p \geq 2$.

For each $n \in \mathbb{N}$ estimates (4.3.23) and (4.3.24) hold for Θ^n. Since b^n, σ^n, g^n, f^n satisfy (HX0) and (HY0$_{\text{loc}}$) uniformly in n then it is easy to check that the constants appearing on the RHS of (4.3.23) and (4.3.24) are independent of n. Hence, by taking the limit of $n \to \infty$ in (4.3.23) and (4.3.24) and given the convergence $\Theta^n \to \Theta$ as $n \to \infty$ (and the continuity of the involved functions) the statement follows.

Proof of (iii) under (HX0)-(HY0$_{loc}$): The estimates concerning Y and $\bar{Z}_{t_i}$ follow trivially from (4.3.23) on the one hand, and (4.3.24) combined with (4.3.22) on the other hand. For the difference $Z_s - \bar{Z}_{t_{i+1}}$ more care is required,

$$\begin{aligned}\sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_{i+1}}|^2 \mathrm{d}s\Big] &\leq 2 \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - Z_{t_{i+1}}|^2 + |Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2 \mathrm{d}s\Big] \\ &\leq C|\pi| + 2 \sum_{i=0}^{N-1} (t_{i+1} - t_i) \mathbb{E}\Big[|Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2\Big],\end{aligned}$$

where the last inequality follows from the proof of *ii*). We next estimate the last term in the RHS, since $\bar{Z}_{t_N} = Z_T$ by construction

$$\begin{aligned}&\sum_{i=0}^{N-1} (t_{i+1} - t_i) \mathbb{E}\Big[|Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2\Big] = \sum_{i=0}^{N-2} (t_{i+1} - t_i) \mathbb{E}\Big[|Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2\Big] \\ &\quad \leq r_\pi \sum_{i=0}^{N-2} (t_{i+2} - t_{i+1}) \mathbb{E}\Big[|Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2\Big] \leq r_\pi \sum_{i=0}^{N-2} \int_{t_{i+1}}^{t_{i+2}} \mathbb{E}\Big[|Z_{t_{i+1}} - \bar{Z}_{t_{i+1}}|^2\Big] \mathrm{d}s \\ &\quad \leq r_\pi \sum_{j=1}^{N-1} \int_{t_j}^{t_{j+1}} \mathbb{E}\Big[|Z_{t_j} - \bar{Z}_{t_j}|^2\Big] \mathrm{d}s \leq 2r_\pi \sum_{i=0}^{N-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - Z_{t_i}|^2 + |Z_s - \bar{Z}_{t_i}|^2 \mathrm{d}s\Big],\end{aligned}$$

where we made use of the assumption on the grid. The result now follows by combining (iii) with the above estimates and having in mind that r_π is uniform over the partition. □

Corollary 4.3.6. *Let (HX0), (HY0) hold and take the family $\{\bar{Z}_{t_i}\}_{t_i \in \pi}$. For any $p \geq 1$*

there exists constant C_p independent of $|\pi|$ such that

$$\mathbb{E}\Big[\sum_{i=0}^{N-1}\big(|\bar{Z}_{t_i}|^2(t_{i+1}-t_i)\big)^p\Big] \leq C_p < \infty.$$

If, moreover, (HY0$_{loc}$) holds then $\max_{t_i\in\pi}\mathbb{E}[\,|\bar{Z}_{t_i}|^{2p}] \leq C_p < \infty$.

Proof. The second sattement follows easily from the definition of $\bar{Z}_{t_i}$ (see (4.3.20)) and the fact that $Z \in \mathcal{S}^p$ for any $p \geq 2$ (Theorem 4.3.5). And, under (HY0$_{\text{loc}}$) the 2nd statement implies the first.

We leave the proof of the first statement for the interested reader. The proof is based on standard integral manipulations combining the definition of $\bar{Z}$, Jensen's inequality and the tower property of the conditional expectation. $\square$

4.3.5 Some finer properties

Here we discuss properties of the solution to (4.1.1)-(4.1.2) in more specific settings. The first lemma concerns a set-up where Z belongs to $\mathcal{S}^\infty$ (rather than $\mathcal{H}^2$ or $\mathcal{S}^2$).

Proposition 4.3.7 (The additive noise case)**.** *Let (HX0)-(HY0$_{loc}$) hold. Assume additionally that $\sigma(t,x)=\sigma(t)$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$. Then $Z\in\mathcal{S}^\infty$.*

Proof. Assume first that (HXY1) also hold. Then the result follows easily by combining the representation formula (4.3.17) with the 2nd part of (4.3.3) and injecting that σ is uniformly bounded.

Now using a standard mollification argument, as was used in the last step of the proof of Theorem 4.3.5, one easily concludes that the result also holds under (HX0)-(HY0$_{\text{loc}}$). $\square$

If the initial data g and $f(\cdot,\cdot,0,0)$ are bounded then so will be the Y process; the second component, Z will also satisfy a type of boundedness condition (see (4.3.28) below).

Lemma 4.3.8 (The bounded setting)**.** *Let (HX0), (HY0) hold and further that g and $(t,x)\mapsto f(t,x,0,0)$ are uniformly bounded then $(Y,Z)\in\mathcal{S}^\infty\times\mathcal{H}^2$.*

Denoting $\mathcal{T}_{[0,T]}$ the set of all stopping times $\tau\in[0,T]$, then Z satisfies further[4] *for*

[4]This means Z belongs to the so called $\mathcal{H}_{BMO}$-spaces, see Subsection 2.3 in [46] or Section 10.1 in [78].

some constant $K_{BMO} > 0$

$$\sup_{\tau\in\mathcal{T}_{[0,T]}} \big\|\mathbb{E}\big[\int_\tau^T |Z_s|^2 \mathrm{d}s \big| \mathcal{F}_\tau\big]\big\|_\infty \le K_{BMO} < \infty. \tag{4.3.28}$$

The constant K_{BMO} depends only on $\|Y\|_{\mathcal{S}^\infty}$, the bounds for g, $f(\cdot,\cdot,0,0)$ and the constants appearing in (HY0).

Proof. The boundedness of Y follows from (4.2.5) by using that $g(X_\cdot), f(\cdot, X_\cdot, 0, 0) \in \mathcal{S}^\infty$. Knowing that $Y \in \mathcal{S}^\infty$ we can easily adapt the proof of Lemma 10.2 in [78] to our setting, where we make use of the inequality $|z| \le 1 + |z|^2$, to obtain (4.3.28); an alternative proof would be to use (4.2.5). □

The first of the above results implies that Z is bounded. Such a setting also includes the case of $\sigma(t, x) = 1$ which is common in many applications in reaction-diffusion equations. The next result provides another type of control for the growth of the process Z without the boundedness assumption on σ.

Proposition 4.3.9. *Let the assumptions of Lemma 4.3.8 hold. Assume further that $|Z|^2$ is a sub-martingale then $|Z_t| \le K_{BMO}/\sqrt{T-t}$, $\forall\, t \in [0, T]$ $\mathbb{P} - a.s.$.*

In particular, if σ is uniformly elliptic and (HXY1) holds then there exists $C > 0$ such that $|\nabla_x u(t, x)| \le C/\sqrt{T-t}$, $\forall\, (t, x) \in [0, T) \times \mathbb{R}^n$.

Proof. The first statement follows by a careful but rather clean analysis of the fact that Z satisfies (4.3.28), which in particular means any $t \in [0, T]$ $\mathbb{P}$-a.s.

$$K_{BMO} \ge \mathbb{E}\big[\int_t^T |Z_s|^2 \mathrm{d}s \big| \mathcal{F}_t\big] = \int_t^T \mathbb{E}\left[|Z_s|^2 \big| \mathcal{F}_t\right] \mathrm{d}s \ge \int_t^T |Z_t|^2 \mathrm{d}s = |Z_t|^2 (T-t),$$

where we applied Fubini then used the sub-martingale property of Z^2. The sought statement now follows by a direct rewriting of the above inequality. The second statement in the proposition follows from the first by using the representation $Z_t^{t,x} = (\nabla_x u\sigma)(t, x)$ and the ellipticity of σ. □

4.4 Numerical discretization and general estimates.

In this section and the following ones, we discuss the numerical approximation of (4.1.1)-(4.1.2). We consider a regular partition[5] π of $[0,T]$ with $N+1$ points $t_i = ih$ for $i = 0, \cdots, N$ with $h := T/N$.

Remark 4.4.1 (On constants)**.** Throughout the rest of this work we introduce a generic constant $c > 0$, that will always be independent of h or N, though it may depend on the problem's data, namely the constants appearing in the assumptions, and may change from line to line.

4.4.1 Discretization of the SDE and further setup

Numerical methods for SDEs with Lipschitz continuous coefficients are well understood, see Section 10 in [49]. Therefore, we postulate that there exists a family of RVs $\{X_i\}_{i=0,\dots,N}$ that approximates the solution X to (4.1.1) over the grid π. More exactly, for any $p \geq 2$ there exists a constant $c = c(T,p,x)$ such that

$$\sup_{N\in\mathbb{N}} \max_{i=0,\dots,N} \mathbb{E}\big[\,|X_i|^p\big] \leq c \quad \text{and} \quad \max_{i=0,\dots,N} \mathbb{E}\big[\,|X_{t_i} - X_i|^p\big]^{\frac{1}{p}} \leq ch^\gamma, \quad \gamma \geq \frac{1}{2}, \tag{4.4.1}$$

where γ is called the rate of the strong convergence and the RVs $\{X_{t_i}\}_{t_i\in\pi}$ are the solution to (4.1.1) on the grid points π. Under (HX0) the Euler scheme give an approximation with $\gamma = 1/2$. For conditions required for the higher order schemes we refer to [49].

Throughout the rest of this work we assume that the family $\{X_i\}_{i=0,\dots,N}$ has been computed; we denote by $\{\mathcal{F}_i\}_{i=0,\dots,N}$ the associated discrete-time filtration $\mathcal{F}_i := \sigma(X_j, j = 0, \cdots, i)$ and with respect to this filtration we define the operator $\mathbb{E}_i[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_i]$.

For the analysis of the time-discretization error, we also make use of the following standard path-regularity estimate for X, which holds under (HX0): there exists a constant $c > 0$ such that

$$\max_{i=0,\dots,N-1} \sup_{t_i\leq s\leq t_{i+1}} \Big\{\mathbb{E}\big[\,|X_s - X_{t_i}|^2\big]^{\frac{1}{2}} + \mathbb{E}\big[\,|X_s - X_{t_{i+1}}|^2\big]^{\frac{1}{2}}\Big\} \leq c\,h^{\frac{1}{2}}. \tag{4.4.2}$$

[5] We point out that the results we state would hold for non-uniform time-steps, but we work with a regular partition for notational clarity and to keep the focus on the main issues.

4.4.2 Fundamental lemma for convergence

The goal of this section is to present a general result on the numerical approximations of the BSDE (4.1.2). To estimate the global error, we follow a path, classical in numerical analysis for SDEs and ODEs, where the local error is decomposed into two parts: one-step error and propagation of error with time (controlled by the stability of the scheme). Although this type of analysis has already been used in the context of Lipschitz BSDEs (see e.g. [23], [16] or [17]), we generalize it to the non-Lipschitz framework we are working with. More precisely, the Fundamental Lemma we present below allows us to cope with schemes which lack stability in the sense of [17][6].

Theorem 4.3.5 implies that to approximate (Y, Z) solution to (4.1.2) one needs only to approximate the family $\{(Y_{t_i}, \bar{Z}_{t_i})\}_{t_i\in\pi}$ (recall (4.3.20)) on π via a family of RVs $\{(Y_i, Z_i)\}_{i=0,\cdots,N}$, the said numerical approximation.

The error criteria that we consider for the numerical approximation is standard and given by

$$\mathrm{ERR}_\pi(Y,Z) := \Big(\max_{i=0,\dots,N} \mathbb{E}\big[\,|Y_{t_i} - Y_i|^2\big] + \sum_{i=0}^{N-1} \mathbb{E}\big[|\bar{Z}_{t_i} - Z_i|^2\big] h\Big)^{\frac{1}{2}}. \qquad (4.4.3)$$

In abstract terms, a discretization scheme generates recursively a family of RVs $\{(Y_i, Z_i)\}_{i=0,\cdots,N}$ approximating $\{(Y_{t_i}, \bar{Z}_{t_i})\}_{t_i\in\pi}$ via Markovian operators Φ_i

$$\Phi_i : L^2(\mathcal{F}_{i+1}) \times L^2(\mathcal{F}_{i+1}) \to L^2(\mathcal{F}_i) \times L^2(\mathcal{F}_i)$$

in the following way. Start with an initial approximation (Y_N, Z_N) and define for $i \in \{N-1, \cdots, 0\}$ $(Y_i, Z_i) = \Phi_i(Y_{i+1}, Z_{i+1})$. The purpose of the Fundamental Lemma below is to formulate in a transparent way the ingredients required to show convergence of $\{(Y_i, Z_i)\}_{i=0,\cdots,N}$ to $\{(Y_{t_i}, \bar{Z}_{t_i})\}_{t_i\in\pi}$ in the norm (4.4.3).

As usual the estimation of the global error requires controls on how the error propagates at each step. Since (Y_i, Z_i) is obtained via Φ_i from the input (Y_{i+1}, Z_{i+1}) we introduce the following notation: given a $\mathcal{F}_{t_{i+1}}$-measurable input $(\mathcal{Y}, \mathcal{Z})$, the pair $(Y_{i,(\mathcal{Y},\mathcal{Z})}, Z_{i,(\mathcal{Y},\mathcal{Z})})$ denotes the associated output of $\Phi_i(\mathcal{Y}, \mathcal{Z})$; writing (Y_i, Z_i) without specifying the input denotes the canonical output of $\Phi_i(Y_{i+1}, Z_{i+1})$, that is, we refer to the family of RV's $\{(Y_i, Z_i)\}_{i=0,\dots,N}$.

[6] See definition 2.1 in [17] with $\zeta_i^Y = \zeta_i^Z = 0$ for $i = 0, \dots, N-1$.

As we already indicated, we decompose the error into two parts: the one-step discretization error and the propagation to time t_i of the error from time t_{i+1}. Given $i \in \{0, \cdots, N-1\}$ we write

$$Y_{t_i} - Y_i := Y_{t_i} - Y_{i,(Y_{i+1}, Z_{i+1})} = \underbrace{\Big(Y_{t_i} - Y_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})}\Big)}_{\text{one-step error}} + \underbrace{\Big(Y_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})} - Y_{i,(Y_{i+1}, Z_{i+1})}\Big)}_{\text{stability of the scheme}},$$

and similarly for Z

$$\bar{Z}_{t_i} - Z_i := \bar{Z}_{t_i} - Z_{i,(Y_{i+1}, Z_{i+1})} = \underbrace{\Big(\bar{Z}_{t_i} - Z_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})}\Big)}_{\text{one-step error}} + \underbrace{\Big(Z_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})} - Z_{i,(Y_{i+1}, Z_{i+1})}\Big)}_{\text{stability of the scheme}}.$$

We now give meaning to our concept of stability, generalizing that in [16] and [17].

Definition 4.4.2 (Scheme stability)**.** *We say that the numerical scheme* $\{(Y_i, Z_i)\}_{i=0,\cdots,N}$ *is* stable *if for some* $\rho > 0$ *there exists a constant* $c > 0$ *such that*

$$\begin{aligned}\mathbb{E}[|Y_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})} - Y_{i,(Y_{i+1}, Z_{i+1})}|^2] + \rho \mathbb{E}[|Z_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})} - Z_{i,(Y_{i+1}, Z_{i+1})}|^2]h \\ \leq (1 + ch)\left(\mathbb{E}[|Y_{t_{i+1}} - Y_{i+1}|^2] + \frac{\rho}{4}\mathbb{E}[|\bar{Z}_{t_{i+1}} - Z_{i+1}|^2]h\right) + \mathbb{E}[H_i], \quad (4.4.4)\end{aligned}$$

where $H_i \in L^1(\mathcal{F}_i)$ *and moreover* $\{H_i\}_{i=0,\cdots,N-1}$ *satisfies*

$$\mathcal{R}^{\mathcal{S}}(H) := \max_{i=0,\ldots,N-1} \sum_{j=i}^{N-1} e^{c(j-i)h}\mathbb{E}[H_j] \to 0, \quad as \quad h \to 0.$$

Remark 4.4.3. In the case where f is a globally Lipschitz function, it can be shown for both implicit and explicit schemes that $H_i = 0$ (see [23] or [17]). In our setting, where f is a monotone function with polynomial growth y, it may not always not be possible to dominate the term H_i by zero. Nevertheless, our definition of stability is sufficient for our numerical analysis, as we can control the stability remainder term $\mathcal{R}^{\mathcal{S}}(H)$.

We also point out that it is important that in (4.4.4), we have $\rho > \frac{\rho}{4}$ (compare LHS with RHS). This later allows the use of Lemma 4.7.4.

We now state the Fundamental Lemma which is the basis of the error analysis throughout.

Lemma 4.4.4 (Fundamental Lemma). *Assume that the numerical scheme* $\{(Y_i, Z_i)\}_{i=0,\cdots,N}$ *is stable. Denoting the one-step discretization errors for* $i = 0, \cdots, N-1$ *by*

$$\tau_i(Y) := \mathbb{E}[|Y_{t_i} - Y_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})}|^2] \quad \textit{and} \quad \tau_i(Z) := \mathbb{E}[|\bar{Z}_{t_i} - Z_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})}|^2 h], \tag{4.4.5}$$

there exists a constant $C = C(\rho, T, c)$ *such that*

$$\begin{aligned}\big(\mathrm{ERR}_\pi(Y,Z)\big)^2 \\ \leq C\Big\{\mathbb{E}[\,|Y_{t_N} - Y_N|^2] + \mathbb{E}[|\bar{Z}_{t_N} - Z_N|^2]h + \sum_{i=0}^{N-1}\Big(\frac{\tau_i(Y)}{h} + \tau_i(Z)\Big)\Big\} + (1+h)\mathcal{R}^{\mathcal{S}}(H)\end{aligned} \tag{4.4.6}$$

This result states in a rather clear fashion (although $\mathcal{R}^{\mathcal{S}}(H)$ is unknown at this point) what is required in order to have convergence of the numerical scheme. One needs a control on the approximation of the terminal conditions, a control on the sum of the local discretization errors and a control on the stability remainder $\mathcal{R}^{\mathcal{S}}(H)$.

Proof. We decompose the error as explained above and use Young's inequality to get

$$|Y_{t_i} - Y_i|^2 \leq (1 + \frac{1}{h})|Y_{t_i} - Y_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})}|^2 + (1+h)|Y_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})} - Y_{i,(Y_{i+1},Z_{i+1})}|^2$$

and

$$|\bar{Z}_{t_i} - Z_i|^2 h \leq 2|\bar{Z}_{t_i} - Z_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})}|^2 h + 2|Z_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})} - Z_{i,(Y_{i+1},Z_{i+1})}|^2 h.$$

Take $\rho > 0$ from (4.4.4) and combine (4.4.5) with the above, it then follows that

$$\begin{aligned}&\mathbb{E}[|Y_{t_i} - Y_i|^2] + \frac{\rho}{2}\mathbb{E}[|\bar{Z}_{t_i} - Z_i|^2]h \\ &\quad\leq (1+h)\mathbb{E}[|Y_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})} - Y_{i,(Y_{i+1},Z_{i+1})}|^2] \\ &\qquad + \rho\mathbb{E}[|Z_{i,(Y_{t_{i+1}},\bar{Z}_{t_{i+1}})} - Z_{i,(Y_{i+1},Z_{i+1})}|^2]h + \Big((1+\frac{1}{h})\tau_i(Y) + \rho\tau_i(Z)\Big).\end{aligned}$$

Since $\rho \leq (1+h)\rho$, by the stability of the scheme (see (4.4.4)) it follows that

$$\begin{aligned}
&\mathbb{E}[|Y_{t_i} - Y_i|^2] + \frac{\rho}{2}\mathbb{E}[|\bar{Z}_{t_i} - Z_i|^2]h \qquad (4.4.7)\\
&\qquad \leq (1+h)(1+ch)\Big(\mathbb{E}[|Y_{t_{i+1}} - Y_{i+1}|^2] + \frac{\rho}{4}\mathbb{E}[|\bar{Z}_{t_{i+1}} - Z_{i+1}|^2]h\Big)\\
&\qquad\qquad + \Big((1+\frac{1}{h})\tau_i(Y) + \rho\tau_i(Z) + (1+h)\mathbb{E}[H_i]\Big).
\end{aligned}$$

Taking $I_i := |Y_{t_i} - Y_i|^2 + \frac{\rho}{4}|\bar{Z}_{t_i} - Z_i|^2 h$ we have

$$I_i + \frac{\rho}{4}\mathbb{E}[|\bar{Z}_{t_i} - Z_i|^2]h \leq (1+h)(1+ch)I_{i+1} + \Big((1+\frac{1}{h})\tau_i(Y) + \rho\tau_i(Z) + (1+h)\mathbb{E}[H_i]\Big),$$

and we conclude the proof using Lemma 4.7.4. □

4.4.3 Discretization of the backward component

Let throughout $t_i, t_{i+1} \in \pi$. To approximate the solution (Y, Z) to (4.1.2) we need two approximations, one for the Y component and one for the Z component. Write (4.1.2) over the interval $[t_i, t_{i+1}]$ and take $\mathcal{F}_{t_i}$-conditional expectations to obtain

$$Y_{t_i} = \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s, \Theta_s)\mathrm{d}s\big]. \tag{4.4.8}$$

For the Z component, one multiplies (4.1.2) (written over the interval $[t_i, t_{i+1}]$) by ΔW_{i+1} and takes $\mathcal{F}_{t_i}$-conditional expectations to obtain (using Itô's Isometry) the implicit formula

$$0 = \mathbb{E}_{t_i}\big[\Delta W_{i+1}\big(Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s, \Theta_s)\mathrm{d}s\big)\big] - \mathbb{E}_{t_i}\big[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s\big]. \tag{4.4.9}$$

One now obtains a scheme by approximating the Lebesgue integral via the θ-integration rule (indexed by a parameter $\theta \in [0,1]$), i.e. for some function ψ

$$\int_{t_i}^{t_{i+1}} \psi(s)\mathrm{d}s \approx \big[\theta\psi(t_i) + (1-\theta)\psi(t_{i+1})\big](t_{i+1} - t_i), \quad \theta \in [0,1].$$

This type of approximation is known to be of first order for $\theta \neq 1/2$ and of higher order for $\theta = 1/2$, see [76] at the end of this section. Unfortunately, with the results obtained so far (see Section 4.3) we are not able to prove the convergence of a general higher order approximation in its full generality; roughly, the issue boils down to obtaining controls on $|\partial^2_{xx}v|$ where v is solution to (4.2.9). However, under the results of Section 4.3 we do not even know if $\partial^2_{xx}v$ exists. Under the assumption that f is independent of z we can prove that the scheme is indeed of higher order (in the y component); the general case is left for future research.

From (4.4.9) above we have (compare with (4.3.20))

$$\bar{Z}_{t_i} := \frac{1}{h}\mathbb{E}_{t_i}\Big[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s\Big] = \frac{1}{h}\mathbb{E}_{t_i}\Big[\Delta W_{t_{i+1}}\Big(Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s,\Theta_s)\mathrm{d}s\Big)\Big],$$

and we approximate $(Z_s)_{s\in[t_i,t_{i+1}]}$ via $\bar{Z}_{t_i}$ and $\bar{Z}_{t_{i+1}}$ rather than Z_{t_i} or $Z_{t_{i+1}}$. Following the notation for Θ we denote $\bar{\Theta}_{t_i} := (X_{t_i}, Y_{t_i}, \bar{Z}_{t_i})$ and using the θ-integration rule it follows

$$Y_{t_i} = \mathbb{E}_{t_i}\left[Y_{t_{i+1}} + h\left[\theta f(t_i,\bar{\Theta}_{t_i}) + (1-\theta) f(t_{i+1},\bar{\Theta}_{t_{i+1}})\right] + \int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\right], \quad (4.4.10)$$

$$\bar{Z}_{t_i} = \mathbb{E}_{t_i}\left[\frac{\Delta W_{t_{i+1}}}{h}\Big(Y_{t_{i+1}} + (1-\theta) f(t_{i+1},\bar{\Theta}_{t_{i+1}})h + \int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)\right], \quad (4.4.11)$$

where the error term is, for $s \in [t_i, t_{i+1}]$, defined as $R(s) := \theta R^I(s) + (1-\theta)R^E(s)$ where

$$R^I(s) := f(s,\Theta_s) - f(t_i,\bar{\Theta}_{t_i}) \quad \text{and} \quad R^E(s) := f(s,\Theta_s) - f(t_{i+1},\bar{\Theta}_{t_{i+1}}). \quad (4.4.12)$$

Remark 4.4.5. For the error analysis here and in the following section we always understand the set of RVs $\{(Y_{t_i}, \bar{Z}_{t_i})\}_{t_i\in\pi}$ as the true solution of the BSDE on the partition points $t_i \in \pi$ but in the set-up of (4.4.10) and (4.4.11). We emphasize that our numerical scheme does not aim at approximating Z itself over π but the family $\{\bar{Z}_{t_i}\}_{t_i\in\pi}$.

The order of the approximation depends on the smoothness of the driver f and the properties of the other coefficients. Ignoring the error term R we arrive at the following discretization scheme: define $Y_N := g(X_N)$ and $Z_N := 0$ and for $i = 0, 1, \ldots, N-1$:

$$Y_i := \mathbb{E}_i\Big[Y_{i+1} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})h\Big] + \theta f(t_i, X_i, Y_i, Z_i)h, \quad (4.4.13)$$

$$Z_i := \mathbb{E}_i\Big[\frac{\Delta W_{t_{i+1}}}{h}\Big(Y_{i+1} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})h\Big)\Big] \quad (4.4.14)$$

We point out that we aim at 1st order schemes, so setting $Z_N = 0$ is not an issue. For a higher order schemes, Z_T needs to be approximated in a more robust fashion, e.g. following (4.3.21), $Z_T = (\nabla_x g)(X_T)\sigma(T, X_T) \approx (\nabla_x g)(X_N)\sigma(T, X_N) = Z_N$ (under the extra assumption that ∇g is Lipschitz).

We can already estimate the error on the terminal conditions, which is the first term in the global error estimate of Lemma 4.4.4.

Lemma 4.4.6. *Let (HX0), (HY0) hold. Then there exists a constant c such that (recall* (4.3.20)*)*

$$\mathbb{E}[|Y_{t_N} - Y_N|^p]^{\frac{1}{p}} \le ch^\gamma \quad \text{for any } p \ge 2 \qquad \text{and} \qquad \mathbb{E}[|\bar{Z}_{t_N} - Z_N|^2 h] \le c\,h, \quad (4.4.15)$$

where γ is the order of the approximation $\{X_i\}_{i=0,\dots,N}$ of X (according to (4.4.1)*).*

Assume that $g \in C_b^1$ and that ∇g is Lipschitz continuous. Define $Z_N := (\nabla_x g)(X_N)\sigma(T, X_N)$ then $\mathbb{E}[|\bar{Z}_{t_N} - Z_N|^2 h] \le c\,h^2$.

Proof. The error estimate on Y_{t_N} results from the Lipschitz regularity of g and the estimate on $\mathbb{E}[|X_{t_N} - X_N|^2]$ given by (4.4.1). For the error estimate on Z, we have $Z_N = 0$, and $\bar{Z}_{t_N} = Z_T$, which in turn implies $\mathbb{E}[|\bar{Z}_{t_N} - Z_N|^2 h] = \mathbb{E}[|Z_T|^2]h \le c\,h$ where we have used (4.3.21).

The second estimate follows easily using that $\bar{Z}_T = Z_T = \nabla g(X_T)\sigma(T, X_T)$ and using the Lipschitz property of ∇g and σ, Cauchy-Schwartz' inequality and (4.4.1). □

4.4.4 Existence and local estimates for the general θ-scheme

In this subsection we start the study of the θ-scheme (4.4.13)-(4.4.14) by analyzing *one step* of it, i.e. going from time t_{i+1} to t_i. To simplify notation, we define $f_{i+1} := f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1})$ and $A_{i+1} := Y_{i+1} + (1-\theta)f_{i+1}h$.

Along with (HX0) and (HY0) we make the temporary assumption that $Y_{i+1}, Z_{i+1}, f_{i+1} \in L^2$ and analyze how this integrability carries on to the next time step; such integrability assumption is clearly satisfied by Y_N, Z_N and f_N.

For $\theta = 0$ (i.e. the explicit case) the scheme step is well defined as Y_i and Z_i can be easily computed. For $\theta > 0$, there is no issue in defining Z_i from (4.4.14), but unlike the Lipschitz case, it is not immediate that the solution Y_i to the implicit equation (4.4.13) exists. We need to show that there exists a unique Y_i solving $Y_i = \mathbb{E}_i[A_{i+1}] + \theta f(t_i, X_i, Y_i, Z_i)h$, where $\mathbb{E}_i[A_{i+1}]$, X_i and Z_i are already known. This follows from Theorem 26.A in [80] (p557). Define (almost surely) the map $F : y \mapsto y - \theta f\big(t_i, X_i(\omega), y, Z_i(\omega)\big)h$. This map is strongly monotone (increasing) in the sense of Definition 25.2 in [80], i.e. there exists a $\mu > 0$ such that for all y, y',

$$\langle y' - y, F(y') - F(y)\rangle \geq \mu |y' - y|^2.$$

Indeed, from (HY0) and Remark 4.2.1 we have

$$\langle y' - y, F(y') - F(y)\rangle \geq (1 - \theta L_y h)|y' - y|^2,$$

so if $h < 1/(\theta L_y)$ we can take $\mu = (1 - \theta L_y h) > 0$. This (almost surely) guarantees the existence of a unique $Y_i(\omega) = F^{-1}\big(\mathbb{E}_i[(A_{i+1})](\omega)\big)$, as needed. By the monotonicity of F, Y_i can be quickly computed using, for example, Newton-Raphson type methods.

Now, Y_i so defined is only a $\mathcal{F}_i$-measurable RV. [7] The following proposition guarantees that if $\theta > 0$, the pair (Y_i, Z_i) and the term f_i are square integrable provided the corresponding RVs at t_{i+1} also are. So for every N, by iteration, (Y_i, Z_i) is well-defined for $i = N-1, \cdots, 0$. For $\theta \geq 1/2$, this estimate also leads to a uniform bound, as will become clear in the next section.

Proposition 4.4.7. *Let (HX0), (HY0) hold, $\theta \in [0,1]$ and set $h \leq \min\{1, [4\theta\big(L_y + 3d\theta L_z^2\big)]^{-1}\}$. Then there exists a constant c such that for any $i \in \{0, \cdots, N-1\}$*

$$\begin{aligned} |Y_i|^2 + \frac{1}{2d}|Z_i|^2 h + 2\theta^2 |f_i|^2 h^2 \leq (1 + ch)\,\mathbb{E}_i\big[|Y_{i+1}|^2 + \frac{1}{8d}|Z_{i+1}|^2 h\big] + ch \\ + c\,(|X_i|^2 + \mathbb{E}_i\big[|X_{i+1}|^2\big])h + 2\big(1-\theta\big)^2 \mathbb{E}_i\big[|f_{i+1}|^2\big]h^2. \end{aligned} \tag{4.4.16}$$

Proof of Proposition 4.4.7 . Let $i \in \{0, \cdots, N-1\}$. First we estimate Z_i. The martingale property of ΔW_{i+1} yields

$$Z_i h = \mathbb{E}_i\big[\Delta W_{i+1} A_{i+1}\big] = \mathbb{E}_i\big[\Delta W_{i+1}(A_{i+1} - \mathbb{E}_i[A_{i+1}])\big]. \tag{4.4.17}$$

[7] The previous explanation only justified the existence of Y_i as a function from Ω to $\mathbb{R}^k$. To obtain that it is measurable, one should rather consider the map $G : (a, y) \mapsto \big(a, y - \theta f(t_i, a, y)h\big)$, where $a = (x, z) \in \mathbb{R}^{d\times k\times d}$ and $f(t, a, y) = f(t, x, y, z)$. It is again seen to be strongly monotonous, so it is invertible and Theorem 26.A asserts that G^{-1} is continuous (Lipschitz in fact), hence measurable.

By the Cauchy-Schwartz inequality,

$$|Z_i|^2 h \leq d\left\{\mathbb{E}_i[A_{i+1}^2] - \mathbb{E}_i[A_{i+1}]^2\right\}. \tag{4.4.18}$$

We now proceed with the estimation of Y_i. We first rewrite

$$Y_i = \mathbb{E}_i[A_{i+1}] + \theta f_i h \Leftrightarrow Y_i - \theta f_i h = \mathbb{E}_i[A_{i+1}]$$

and then square both sides of the RHS of the above equivalence to obtain

$$|Y_i|^2 = \mathbb{E}_i[A_{i+1}]^2 + 2\theta\langle Y_i, f_i\rangle h - \theta^2|f_i|^2h^2.$$

This simple manipulation allows us to take advantage of the monotonicity of f (see (4.2.1)) and will be reused frequently. By the estimate of Remark 4.2.1, with an $\alpha > 0$ to be chosen later, the previous equality leads to

$$|Y_i|^2 \leq \mathbb{E}_i[A_{i+1}]^2 + 2\theta(L_y + \alpha)|Y_i|^2 h + \theta B(i,\alpha) + \frac{3\theta L_z^2}{2\alpha}|Z_i|^2 h - \theta^2|f_i|^2h^2,$$

where $B(i,\alpha) := (3L^2h + 3L_x^2|X_i|^2h)/(2\alpha)$. Now, for $\epsilon = 1/d$, we combine the above estimate with (4.4.18) to obtain

$$\begin{aligned}|Y_i|^2 + \epsilon|Z_i|^2 h \leq{}& (1-\epsilon d)\,\mathbb{E}_i[A_{i+1}]^2 + \epsilon d\mathbb{E}_i[A_{i+1}^2] \\ &+ 2\theta(L_y+\alpha)|Y_i|^2h + \frac{3\theta L_z^2}{2\alpha}|Z_i|^2h + \theta B(i,\alpha) - \theta^2|f_i|^2h^2.\end{aligned}$$

Reorganizing the terms leads to

$$\Big(1 - 2\theta(L_y+\alpha)h\Big)|Y_i|^2 + \Big(\epsilon - \frac{3\theta L_z^2}{2\alpha}\Big)|Z_i|^2h \leq \mathbb{E}_i[A_{i+1}^2] + \theta B(i,\alpha) - \theta^2|f_i|^2h^2. \tag{4.4.19}$$

Using again Remark 4.2.1 with $\alpha' > 0$ we obtain

$$\begin{aligned}A_{i+1}^2 \leq{}& |Y_{i+1}|^2 + (1-\theta)2(L_y+\alpha')|Y_{i+1}|^2h \\ &+ (1-\theta)\frac{3L_z^2}{2\alpha'}|Z_{i+1}|^2h + (1-\theta)B(i+1,\alpha') + (1-\theta)^2|f_{i+1}|^2h^2,\end{aligned}$$

which in turns leads to

$$
\begin{aligned}
&\Big(1-2\theta(L_y{+}\alpha)h\Big)|Y_i|^2+\Big(\epsilon-\frac{3\theta L_z^2}{2\alpha}\Big)|Z_i|^2h\\
&\qquad\le\Big(1+(1-\theta)2(L_y+\alpha')h\Big)\mathbb{E}_i\Big[|Y_{i+1}|^2\Big]+(1-\theta)\frac{3L_z^2}{2\alpha'}\mathbb{E}_i\Big[|Z_{i+1}|^2\Big]h+\ H_i^\theta \qquad (4.4.20)\\
&\qquad\quad+\theta B(i,\alpha)+(1-\theta)\mathbb{E}_i[B(i+1,\alpha')],
\end{aligned}
$$

where

$$
H_i^\theta:=\big(1-\theta\big)^2\mathbb{E}_i\big[|f_{i+1}|^2\big]h^2-\theta^2|f_i|^2h^2 \qquad (4.4.21)
$$

Now, we choose $\alpha=3d\theta L_z^2$ (so that $\epsilon-\frac{3\theta L_z^2}{2\alpha}=\frac{1}{2d}$) and $\alpha'=24d(1-\theta)L_z^2$ (so that $(1-\theta)\frac{3L_z^2}{2\alpha'}\le\frac{1}{16d}$). Since $h\le\min\{1,[4\theta(L_y+3d\theta L_z^2)]^{-1}\}$ it is true that $2\theta(L_y+\alpha)h\le 1/2$. We also observe that for $x\in[0,1/2]$, $1\le 1/(1-x)\le 1+2x\le 2$ and as a consequence

$$
\begin{aligned}
|Y_i|^2+\frac{1}{2d}|Z_i|^2h\le&\big(1+4\theta(L_y+\alpha)h\big)\big(1+2(1-\theta)(L_y+\alpha')h\big)\mathbb{E}_i\big[|Y_{i+1}|^2\big]\\
&+\frac{1}{8d}\mathbb{E}_i\big[|Z_{i+1}|^2\big]h+2\theta B(i,\alpha)+2(1-\theta)\mathbb{E}_i\big[B(i+1,\alpha')\big]+2H_i^\theta.
\end{aligned}
$$

Defining $c:=4\theta(L_y+\alpha)+2(1-\theta)(L_y+\alpha')+8\theta(L_y+\alpha)(1-\theta)(L_y+\alpha')$ we clearly have

$$
\big(1+4\theta(L_y+\alpha)h\big)\big(1+2(1-\theta)(L_y+\alpha')h\big)\le 1+ch.
$$

We can now conclude to the announced estimate

$$
\begin{aligned}
|Y_i|^2+\frac{1}{2d}|Z_i|^2h\le&\big(1+ch\big)\big(\mathbb{E}_i\big[|Y_{i+1}|^2\big]+\frac{1}{8d}\mathbb{E}_i\big[|Z_{i+1}|^2\big]h\big)\\
&+2\theta B(i,\alpha)+2(1-\theta)\mathbb{E}_i[B(i+1,\alpha')]+2H_i^\theta, \qquad (4.4.22)
\end{aligned}
$$

provided one passes the term $2\theta^2|f_i|^2h^2$ in $2H_i^\theta$ to the LHS. This concludes the proof. □

4.4.5 Local discretization error

The schemes we propose in this work (tamed explicit and implicit-dominant) require a different argumentation with regard to the stability property (see (4.4.4)). However, the analysis of the local discretization errors $\tau_i(Y)$ and $\tau_i(Z)$ (see (4.4.5)) is the same for both types of schemes, therefore in this section we carry it out for a general $\theta \in [0,1]$. Later on, we give particular attention to the higher order approximation case $\theta = 1/2$ which corresponds to the trapezoidal rule.

We follow the notation of Subsection 4.4.2 and set, for $i = 0, 1, \ldots, N-1$, $\widehat{Y}_i = Y_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})}$ and $\widehat{Z}_i = Z_{i,(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})}$; that is $(\widehat{Y}_i, \widehat{Z}_i)$ is the solution to

$$\widehat{Y}_i = \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})h\big] + \theta f(t_i, X_i, \widehat{Y}_i, \widehat{Z}_i)h \qquad (4.4.23)$$

$$\widehat{Z}_i = \mathbb{E}_{t_i}\Big[\frac{\Delta W_{t_{i+1}}}{h}\Big(Y_{t_{i+1}} + (1-\theta) f(t_{i+1}, X_{i+1}, Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})h\Big)\Big]. \qquad (4.4.24)$$

Remark 4.4.8. We know from Proposition 4.4.7 that, for $h \le \min\{1, [4\theta\big(L_y + 3d\theta L_z^2\big)]^{-1})\}$, the RV's $\{(\widehat{Y}_i, \widehat{Z}_i)\}_{i=0,\cdots,N}$ are well defined and square integrable. Furthermore, estimate (4.4.16), together with the growth assumption on f in (HY0), (4.4.1) for X_{i+1}, Theorem 4.2.2 for $Y_{t_{i+1}}$ and Corollary 4.3.6 for $\bar{Z}_{t_{i+1}}$, guarantee immediately that for any $p \ge 2$, there exists a constant c such that

$$\max_{i=0,\ldots,N} \mathbb{E}[|\widehat{Y}_i|^p] \le c \qquad (4.4.25)$$

This fact will be needed later in Section 4.5.

The next result estimates the one-step discretization errors $\tau_i(Y)$ and $\tau_i(Z)$ of the approximation in terms of the error process R (as defined in (4.4.12)). Afterward we discuss the behavior of R itself. Following the notation above we have

Lemma 4.4.9. *Let (HX0) and (HY0) hold. Then for any $\theta \in [0,1]$ there exists a constant c such that for any $i \in \{0, \cdots, N-1\}$*

$$|Y_{t_i} - \widehat{Y}_i|^2 + |\bar{Z}_{t_i} - \widehat{Z}_i|^2 h \le c\,\mathbb{E}_i\Big[\Big(\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)^2\Big] + c\,(1-\theta)^2 \mathbb{E}_i[|X_{t_{i+1}} - X_{i+1}|^2]h^2 + c\,\theta^2 |X_{t_i} - X_i|^2 h^2.$$

Proof. Let $i \in \{0, \cdots, N-1\}$. Recalling (4.4.11), (4.4.24) and that $\bar{\Theta}_{t_i} := (X_{t_i}, Y_{t_i}, \bar{Z}_{t_i})$

we have

$$\bar{Z}_{t_i} - \widehat{Z}_i = \mathbb{E}_i\left[\frac{\Delta W_{t_{i+1}}}{h}\Big((1-\theta)\big[f(t_{i+1},\bar{\Theta}_{t_{i+1}}) - f(t_{i+1},X_{i+1},Y_{t_{i+1}},\bar{Z}_{t_{i+1}})\big]h + \int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)\right],$$

which by the Cauchy-Schwartz's inequality and the Lipschitz property of $x \mapsto f(\cdot,x,\cdot,\cdot)$ leads to

$$h|\bar{Z}_{t_i} - \widehat{Z}_i|^2 \leq 2d\,\mathbb{E}_i\Big[\Big(\int_{t_i}^{t_{i+1}} R_u \mathrm{d}u\Big)^2\Big] + 2d(1-\theta)^2 L_x^2 \mathbb{E}_i\big[\,|X_{t_{i+1}} - X_{i+1}|^2\big]h^2.$$

For the Y-part, similarly by recalling (4.4.10) and (4.4.23) we have

$$\begin{aligned}
Y_{t_i} - \widehat{Y}_i =& \mathbb{E}_i\Big[\int_{ti}^{t_{i+1}} R(s)\mathrm{d}s + (1-\theta)\Big(f(t_{i+1},\bar{\Theta}_{t_{i+1}}) - f(t_{i+1},X_{i+1},Y_{t_{i+1}},\bar{Z}_{t_{i+1}})\Big)h\Big] \\
&+ \theta\Big(f(t_i,\bar{\Theta}_{t_i}) - f(t_i,X_i,\widehat{Y}_i,\widehat{Z}_i)\Big)h \\
=& \mathbb{E}_i\left[\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s + (1-\theta)\Big(f(t_{i+1},\bar{\Theta}_{t_{i+1}}) - f(t_{i+1},X_{i+1},Y_{t_{i+1}},\bar{Z}_{t_{i+1}})\Big)h\right] \\
&+ \theta\Big(f(t_i,X_{t_i},Y_{t_i},\bar{Z}_{t_i}) - f(t_i,X_i,Y_{t_i},\widehat{Z}_i)\Big)h \\
&\quad + \theta\Big(f(t_i,X_i,Y_{t_i},\widehat{Z}_i) - f(t_i,X_i,\widehat{Y}_i,\widehat{Z}_i)\Big)h.
\end{aligned}$$

To obtain the estimate for $|Y_{t_i} - \widehat{Y}_i|^2$, similarly as in the proof of Theorem 4.4.7, we pass the last term in the RHS to the LHS, square both sides, expand the square on the LHS, pass the cross term to the RHS and dominate it on the RHS using (4.2.1). By collecting only the convenient terms in the LHS and using Assumption (HY0) on the RHS we get

$$\begin{aligned}
|Y_{t_i} - \widehat{Y}_i|^2 \leq& 3\,\mathbb{E}_i\Big[\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big]^2 + 6\,\theta^2 L_z^2 |\bar{Z}_{t_i} - \widehat{Z}_i|^2 h^2 + 2\,\theta L_y |Y_{t_i} - \widehat{Y}_i|^2 h \\
&+ 6\,\theta^2 L_x^2 |X_{t_i} - X_i|^2 h^2 + 3(1-\theta)^2 L_x^2 \mathbb{E}_i[|X_{t_{i+1}} - X_{i+1}|^2]h^2,
\end{aligned}$$

which implies, using the estimate for $|\bar{Z}_{t_i} - \widehat{Z}_i|^2$, that

$$\begin{aligned}
(1-2\theta L_y h)|Y_{t_i} - \widehat{Y}_i|^2 \leq& (3 + 12d\theta^2 L_z^2 h)\mathbb{E}_i\Big[\Big(\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)^2\Big] + 6\,\theta^2 L_x^2 |X_{t_i} - X_i|^2 h^2 \\
&+ 3(1-\theta)^2 L_x^2 (1 + 4d\theta^2 L_z^2 h)\mathbb{E}_i\big[|X_{t_{i+1}} - X_{i+1}|^2\big]h^2.
\end{aligned}$$

Noting that h is such that $2\theta L_y h \leq 1/2$ and by combining the estimates for $|Y_{t_i} - \widehat{Y}_i|^2$ and $|\bar{Z}_{t_i} - \widehat{Z}_i|^2$ the sought result follows. $\square$

We now estimate the integral of the error function R (see (4.4.12)).

Lemma 4.4.10. *Let (HX0), (HY0$_{loc}$) hold. Then there exists $c > 0$, for any $\theta \in [0,1]$ and $i \in \{0, \cdots, N-1\}$ such that*

$$\mathbb{E}\Big[\Big(\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)^2\Big] \leq c\,h^3 + c\,h\,\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_i}|^2 \mathrm{d}s + \int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_{i+1}}|^2 \mathrm{d}s\Big].$$

Proof. Following from (4.4.12) we estimate R via R^I and R^E: using (HY0$_{loc}$), Cauchy-Schwarz's inequality and Fubini's theorems we have (recall that $\Theta = (X, Y, Z)$ and $\bar{\Theta}_{t_i} = (X_{t_i}, Y_{t_i}, \bar{Z}_{t_i})$)

$$\begin{aligned}
\mathbb{E}\left[\Big(\int_{t_i}^{t_{i+1}} R^I(s)\mathrm{d}s\Big)^2\right] &= \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} [f(s,\Theta_s) \pm f(s, X_s, Y_{t_i}, Z_s) - f(t_i, \bar{\Theta}_{t_i})]\mathrm{d}s\right)^2\right] \\
&\leq 2h\mathbb{E}\left[\int_{t_i}^{t_{i+1}} 3L_y^2\big(1 + |Y_s|^{2(m-1)} + |Y_{t_i}|^{2(m-1)}\big)|Y_s - Y_{t_i}|^2 \mathrm{d}s + \ \alpha_i\right] \\
&\leq 2h\left(\int_{t_i}^{t_{i+1}} L_y^2 \mathbb{E}\big[3\big(1 + |Y_s|^{4(m-1)} + |Y_{t_i}|^{4(m-1)}\big)\big]^{1/2} \mathbb{E}\big[|Y_s - Y_{t_i}|^4\big]^{1/2}\mathrm{d}s + \mathbb{E}[\alpha_i]\right),
\end{aligned}$$

where $\alpha_i = 3\int_{t_i}^{t_{i+1}} \big[L_t^2|s - t_i| + L_x^2|X_s - X_{t_i}|^2 + L_z^2|Z_s - \bar{Z}_{t_i}|^2\big]\mathrm{d}s$.

Using Theorem 4.2.2 and (4.3.23) to deal with the Y component, yields the estimate

$$\mathbb{E}\left[\Big(\int_{t_i}^{t_{i+1}} R^I(s)\mathrm{d}s\Big)^2\right] \leq ch^3 + 6hL_z^2\mathbb{E}\big[\int_{t_i}^{t_{i+1}} |Z_s - \bar{Z}_{t_i}|^2\mathrm{d}s\big] + 6h^2L_x^2 \sup_{t_i \leq s \leq t_{i+1}} \mathbb{E}[|X_s - X_{t_i}|^2].$$

Similar arguments allow a similar estimate for R^E but with terms $\bar{Z}_{t_{i+1}}$ and $X_{t_{i+1}}$ instead of $\bar{Z}_{t_i}$ and X_{t_i}. We conclude by recalling (4.4.2). $\square$

The trapezoidal integration case

Here, we refine the analysis of the local discretization error from Lemma 4.4.10 for the case $\theta = 1/2$ in order to obtain better global error estimates. We drop the Z-dependence in f due to lacking regularity results. Approximation (4.4.14) is found by approximating the last integral on the RHS of (4.4.9) by a 1st order approximation and so, it should be clear that at best the overall order of the scheme would be one (in the next section we propose a candidate for higher order approximation of Z). We point out nonetheless that many reaction-diffusion equations have a driver f that only depends on Y. For ease of the presentation we also assume that f does not depend on the forward process X and omit the time dependence (these can be easily extended).

We write, similarly to (4.4.10),

$$\int_{t_i}^{t_{i+1}} f(Y_s)\mathrm{d}s = \frac{h}{2}\big[f(Y_{t_i}) + f(Y_{t_{i+1}})\big] + \int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s,$$

with

$$R(s) := f(Y_s) - \frac{1}{2}\big[f(Y_{t_i}) + f(Y_{t_{i+1}})\big],$$

where, using integration by parts, it can be shown (see [76]) that

$$\mathbb{E}\Big[\Big(\int_{t_i}^{t_{i+1}} R(s)\mathrm{d}s\Big)^2\Big] \leq \frac{h^6}{12^2}\mathbb{E}\big[\sup_{t_i\leq t\leq t_{i+1}} |\partial^2_{yy} f(Y_t)|^2\big]. \tag{4.4.26}$$

Hence, in the special case where the driver of FBSDE under consideration does not depend on the process $(Z_t)_{0\leq t\leq T}$ we can take full advantage of trapezoidal integration rule provided that the second derivatives of f in the y variable has polynomial growth, so that there exists a constant c for which

$$\max_{t_i,t_{i+1}\in\pi} \mathbb{E}\Big[\sup_{t_i\leq t\leq t_{i+1}} |\partial^2_{yy} f(Y_t)|^2\Big] \leq c.$$

The result on the sum of local errors

In view of the above lemmas (as well as the estimate (4.4.1) and the path-regularity Theorem 4.3.5), we can state the following estimates on the sum of the local discretization errors, as appearing in the global error estimate (4.4.6) of Lemma 4.4.4.

Proposition 4.4.11. *Let (HX0), (HY0$_{loc}$) hold and $h \leq \min\{1, [4\theta\big(L_y + 3d\theta L_z^2\big)]^{-1}\}$. For the scheme* (4.4.13)-(4.4.14) *we have the following local error estimates:*

i) For any $\theta \in [0,1]$, there exists a constant $c > 0$ such that $\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} + \tau_i(Z) \leq ch$.

ii) Take $\theta = 1/2$ and scheme (4.4.13). *Assume additionally that $f \in C^2$ does not depend on (t,x,z) and $\partial^2_{yy} f$ has at most polynomial growth, then there exists $c > 0$ such that $\sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h} \leq ch^4$.*

Proof. Recall (4.4.5). The proof of case i) is a direct consequence of Lemma 4.4.9, Lemma 4.4.10, estimate (4.4.1) and the path-regularity Theorem 4.3.5.

For the proof of case ii), remark that (4.4.23) is now independent of Z, and hence using Lemma 4.4.9 in combination with (4.4.26) instead of Lemma 4.4.10 yields the result. $\square$

4.5 Convergence of the implicit-leaning schemes ($1/2 \leq \theta \leq 1$)

In this section, we complete the convergence proof of the theta scheme (4.4.13)-(4.4.14) for $\theta \in [1/2, 1]$. In view of the Fundamental Lemma and Proposition 4.4.11 it boils down to the proof of the stability of the scheme. For the reader's convenience, we state immediately the main result while the rest of the section is devoted to its proof.

Theorem 4.5.1. *Let (HX0), (HY0$_{loc}$) hold and $h \leq \min\{1, [4\theta(L_y + 3d\theta L_z^2)]^{-1}\}$. Let $\gamma \geq 1/2$ be the order of the approximation $\{X_i\}_{i=0,\dots,N}$ of X.*

Then, for the scheme (4.4.13)-(4.4.14) we have:

i) For $\theta \in [1/2, 1]$, there exists a constant c such that $\mathrm{ERR}_\pi(Y, Z) \leq ch^{1/2}$.

ii) Take $\theta = 1/2$ and scheme (4.4.13). Assume that $f \in C^2$, $f(t, x, y, z) = f(y)$ and $\partial^2_{yy} f$ has at most polynomial growth, then there exists $C > 0$ s.th. $\max_{i=0,\dots,N} \mathbb{E}[|Y_{t_i} - Y_i|^2]^{1/2} \leq Ch^{\min\{7/4,\gamma\}}$.

4.5.1 Size estimate for the theta-scheme, for $1/2 \leq \theta \leq 1$

We now show that for $\theta \geq 1/2$ the scheme cannot explode as h vanishes. These L^p estimates will be useful in obtaining the stability of the scheme.

Proposition 4.5.2. *Let (HX0), (HY0) hold, and $h \leq \min\{1, [4\theta(L_y + 3d\theta L_z^2)]^{-1}\}$ and let $\theta \in [1/2, 1]$. Then for any $p \geq 1$, there exists a constant c such that*

$$\max_{i=0,\dots,N} \mathbb{E}\left[|Y_i|^{2p}\right] + \sum_{i=0}^{N-1} \mathbb{E}[\,(|Z_i|^2 h)^p] \leq c\big(1 + \mathbb{E}[\,|X_N|^{2\,m\,p}]\big).$$

Proof. Take $i \in \{0, \cdots, N-1\}$. Let $I_i := |Y_i|^2 + \frac{1}{8d}|Z_i|^2 h + \theta^2 |f(t_i, X_i, Y_i, Z_i)|^2 h^2$. By Proposition 4.4.7 and the fact that $(1-\theta)^2 \le \theta^2$, for $\theta \in [1/2, 1]$, we have

$$I_i + \frac{3}{8d}|Z_i|^2 h \le e^{ch}\mathbb{E}_i\big[I_{i+1}\big] + \mathbb{E}_i\big[\,\beta_i\,\big]h, \quad \text{with } \beta_i := c + c\left(|X_i|^2 + |X_{i+1}|^2\right). \tag{4.5.1}$$

As a consequence of Lemma 4.7.4 we know that, since $\beta_j \ge 0$,

$$I_i + \frac{3}{8d}\mathbb{E}_i\Bigg[\sum_{j=i}^{N-1}|Z_j|^2 h\Bigg] \le e^{cT}\Big(\mathbb{E}_i[I_N] + \sum_{j=i}^{N-1}\mathbb{E}_i[\beta_j]h\Big),$$

in particular, using Jensen's inequality, we obtain further

$$|I_i|^p \le 2^{p-1}e^{c\,p\,T}\Big(\mathbb{E}_i[|I_N|^p] + (N\,h)^{p-1}\sum_{j=0}^{N-1}\mathbb{E}_i[\,|\beta_j|^p]h\Big).$$

This then implies, thanks to (HY0)

$$\max_{i=0,\dots,N}\mathbb{E}[|I_i|^p] \le c(1+|X_N|^{2\,m\,p}) \implies \max_{i=0,\dots,N}\mathbb{E}[|Y_i|^{2p}] \le c(1+|X_N|^{2\,m\,p}).$$

From (4.5.1) we also have

$$\begin{aligned} I_i^p + \Big(\frac{3}{8d}\Big)^p(|Z_i|^2 h)^p &\le \Big(I_i + \frac{3}{8d}|Z_i|^2 h\Big)^p \\ &\le e^{c\,p\,h}\mathbb{E}_i[I_{i+1}^p] + \sum_{j=1}^{p}\binom{p}{j}\Big(e^{ch}\mathbb{E}_i\Big[I_{i+1}\Big]\Big)^{p-j}\Big(\mathbb{E}_i[\beta_i]h\Big)^j, \end{aligned}$$

so that, applying again Lemma 4.7.4 along with Hölder's and Jensen's inequalities we have

$$
\begin{aligned}
&\Big(\frac{3}{8d}\Big)^p \mathbb{E}\Big[\sum_{i=0}^{N-1} \big(|Z_i|^2 h\big)^p\Big] \\
&\quad \le e^{c\,p\,T}\mathbb{E}[|I_N|^p] + \sum_{i=0}^{N-1} e^{c\,i\,h} \sum_{j=1}^{p} \binom{p}{j} \mathbb{E}\Big[\Big(e^{ch}\mathbb{E}_i\Big[I_{i+1}\Big]\Big)^{p-j}\Big(\mathbb{E}_i[\beta_i]h\Big)^j\Big] \\
&\quad \le e^{c\,p\,T}\mathbb{E}[|I_N|^p] + e^{c\,p\,T}\sum_{i=0}^{N-1}\sum_{j=1}^{p} \binom{p}{j} \Big(\mathbb{E}\big[|I_{i+1}|^p\big]\Big)^{\frac{p-j}{p}} \Big(\mathbb{E}[|\beta_i|^p]\Big)^{\frac{j}{p}} h \\
&\quad \le e^{c\,p\,T}\mathbb{E}[|I_N|^p] + e^{c\,p\,T}T\sum_{j=1}^{p} \binom{p}{j} \Big(\max_{i=0,\dots,N} \mathbb{E}\big[|I_{i+1}|^p\big]\Big)^{\frac{p-j}{p}} \Big(\max_{i=0,\dots,N} \mathbb{E}\big[\,|\beta_i|^p\big]\Big)^{\frac{j}{p}}.
\end{aligned}
$$

Due to (HY0) and the previous estimates we arrive, as required, at

$$
\mathbb{E}\Big[\sum_{i=0}^{N-1} \big(|Z_i|^2 h\big)^p\Big] \le c(1+|X_N|^{2\,m\,p}).
$$

$\square$

4.5.2 Stability and convergence of the theta-scheme for $1/2 \le \theta \le 1$

We now study the stability of the scheme in the sense of (4.4.4). We fix $i \in \{0,\dots,N-1\}$ and estimate the distance between $(\widehat{Y}_i, \widehat{Z}_i)$ (see (4.4.23)-(4.4.24)) and (Y_i, Z_i) (see (4.4.13)-(4.4.14)) as a function of the distance between $(Y_{t_{i+1}}, \bar{Z}_{t_{i+1}})$ and (Y_{i+1}, Z_{i+1}).

We use the notation $\delta Y_{i+1} = Y_{t_{i+1}} - Y_{i+1}$, $\delta Z_{i+1} := \bar{Z}_{t_{i+1}} - Z_{i+1}$, as well as

$$
\begin{aligned}
&\delta f_{i+1} = f(t_{i+1}, X_{i+1}, Y_{t_{i+1}}, \bar{Z}_{t_{i+1}}) - f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1}) \qquad \text{and} \\
&\delta A_{i+1} = \delta Y_{i+1} + (1-\theta)\delta f_{i+1} h.
\end{aligned}
$$

Then, denoting by $\widehat{\delta Y_i} = \widehat{Y}_i - Y_i$, $\widehat{\delta Z_i} = \widehat{Z}_i - Z_i$, and $\delta \widehat{f}_i = f(t_i, X_i, \widehat{Y}_i, \widehat{Z}_i) - f(t_i, X_i, Y_i, Z_i)$, we can write that (compare with (4.4.23), (4.4.24), (4.4.13) and (4.4.14))

$$
\widehat{\delta Y_i} = \mathbb{E}_i\big[\delta A_{i+1}\big] + \theta \delta \widehat{f}_i h \qquad \text{and} \qquad \widehat{\delta Z_i} = \mathbb{E}_i\big[\frac{1}{h}\Delta W_{i+1}\delta A_{i+1}\big].
$$

Proposition 4.5.3. *Let (HX0) and (HY0) hold. Then there exists a constant c for any $i \in \{0, \cdots, N-1\}$ and $h \leq \min\{1, [4\theta\big(L_y + d\theta L_z^2\big)]^{-1}\}$ such that*

$$|\widehat{\delta Y_i}|^2 + \frac{1}{2d}|\widehat{\delta Z_i}|^2 h \leq (1 + ch)\, \mathbb{E}_i\Big[|\delta Y_{i+1}|^2 + \frac{1}{8d}|\delta Z_{i+1}|^2 h\Big] + H_i^\theta,$$

where

$$H_i^\theta = \big(1-\theta\big)^2 \mathbb{E}_i\big[|\delta f_{i+1}|^2\big] h^2 - \theta^2 \mathbb{E}_i\big[|\delta \widehat{f}_i|^2\big] h^2. \tag{4.5.2}$$

Proof. This proof is very similar to that of Proposition 4.4.7 therefore we omit it. □

In order to show convergence using Fundamental Lemma 4.4.4, we need to control $\mathcal{R}^{\mathcal{S}}(H)$. For the fully implicit scheme ($\theta = 1$) we have $H_i^\theta = -|\delta\widehat{f}_i|^2 h^2 \leq 0$ and hence the implicit scheme is stable in the classical sense (of [16] or [17]) as we have $\mathcal{R}^{\mathcal{S}}(H) \leq 0$. The next lemma provides, in our setting, a control on $\mathcal{R}^{\mathcal{S}}(H)$ for any $\theta \geq 1/2$.

Lemma 4.5.4. *Let (HX0), (HY0$_{loc}$) hold and take the family $\{H_i\}_{i=0,\cdots,N-1}$ defined in* (4.5.2). *Then for $\theta \geq 1/2$ there exists a constant c such that*

$$\begin{aligned}\mathcal{R}^{\mathcal{S}}(H) = \max_{i=0,\dots,N-1} \mathbb{E}\Big[\sum_{j=i}^{N-1} e^{c(j-i)h} H_j^\theta\Big] \leq& c\, \mathbb{E}[|Y_{t_N} - Y_N|^4]^{\frac{1}{2}} h^2 + c\, \mathbb{E}[\,|\bar{Z}_N - Z_N|^2] h^2 \\ &+ c\Big(\sum_{i=0}^{N-1} \tau_i(Y)\Big)^{\frac{1}{2}} h + c\Big(\sum_{i=0}^{N-1} \tau_i(Z)\Big)^{\frac{1}{2}} h.\end{aligned}$$

Proof. Let $i \in \{0, \cdots, N-1\}$. Since $1/2 \leq \theta \leq 1$, we have $(1-\theta)^2 \leq \theta^2$ and therefore

$$\begin{aligned}\mathbb{E}\Big[\sum_{j=i}^{N-1} e^{c(j-i)h} H_j^\theta\Big] &\leq \theta^2 \mathbb{E}\Big[\sum_{j=i}^{N-1} e^{c(j-i)h}\Big(|\delta f_{j+1}|^2 - |\delta\widehat{f}_j|^2\Big)\, h^2\Big] \\ &= \theta^2 \mathbb{E}\Big[\sum_{j=i}^{N-1} e^{c(j-i)h}\Big(|\delta f_{j+1}|^2 - |\delta f_j + \beta_j|^2\Big) h^2\Big] \\ &\leq \theta^2 \mathbb{E}\Big[\sum_{j=i}^{N-1} e^{c(j-i)h}\Big(e^{ch}|\delta f_{j+1}|^2 - |\delta f_j|^2 - 2\,\langle \delta f_j, \beta_i\rangle - {\beta_j}^2\Big) h^2\Big] \\ &\leq \theta^2 e^{c(N-i)h} \mathbb{E}\big[|\delta f_N|^2\big] h^2 - 2\theta^2 \sum_{j=i}^{N-1} e^{c(j-i)h} \mathbb{E}\big[\,\langle \delta f_j, \beta_j\rangle\,\big] h^2,\end{aligned}$$

where $\beta_i := \delta\widehat{f}_j - \delta f_j = f(t_i, X_j, \widehat{Y}_j, \widehat{Z}_j) - f(t_i, X_j, Y_{t_i}, \bar{Z}_{t_i})$ and we used a telescopic sum. Using now (HY0$_{loc}$) yields

$$\mathbb{E}[|\delta f_N|^2] \le c\,\mathbb{E}[1 + |Y_{t_N}|^{4(m-1)} + |Y_N|^{4(m-1)}]^{\frac{1}{2}}\mathbb{E}[|Y_{t_N} - Y_N|^4]^{\frac{1}{2}} + c\,\mathbb{E}[\,|\bar{Z}_N - Z_N|^2]$$

and

$$\begin{aligned}\mathbb{E}[\langle \delta f_i, \beta_i\rangle]h^2 \le \mathbb{E}[|\delta f_i||\beta_i|]h^2 \le & \mathbb{E}\Big[\Big(|\delta f_i| L_y(1 + |\widehat{Y}_i|^{m-1} + |Y_{t_i}|^{m-1})\Big)^2\Big]^{\frac{1}{2}}\mathbb{E}\big[|\widehat{Y}_i - Y_{t_i}|^2\big]^{\frac{1}{2}}h^2 \\ & + \mathbb{E}\Big[\big(L_z|\delta f_i|\big)^2\Big]^{\frac{1}{2}}\mathbb{E}\big[|\widehat{Z}_i - \bar{Z}_{t_i}|^2\big]^{\frac{1}{2}}h^2 \\ \le & c\,\mathbb{E}\left[B_i^1\right]^{\frac{1}{2}}\mathbb{E}[|\widehat{Y}_i - Y_{t_i}|^2]^{\frac{1}{2}}h + c\,\mathbb{E}\left[B_i^2\right]^{\frac{1}{2}}\mathbb{E}[|\widehat{Z}_i - \bar{Z}_{t_i}|^2 h]^{\frac{1}{2}}h,\end{aligned}$$

where $B_i^2 := |Y_{t_i}|^{2m}h + |Y_i|^{2m}h + |\bar{Z}_{t_i}|^2 h + |Z_i|^2 h$ and

$$B_i^1 := h^2 + |\widehat{Y}_i|^{4m}h^2 + |Y_{t_i}|^{4m}h^2 + |Y_i|^{4m}h^2 + (|\bar{Z}_{t_i}|^2 h)^2 + (|Z_i|^2 h)^2.$$

From Theorem 4.2.2, Corollary 4.3.6, Remark 4.4.8 and Proposition 4.5.2 we have, for the first term of the above inequality

$$\sum_{i=0}^{N-1}\mathbb{E}\big[B_i^1\big]^{\frac{1}{2}}\mathbb{E}\big[\,|\widehat{Y}_i - Y_{t_i}|^2\big]^{\frac{1}{2}}h \le \Big(\sum_{i=0}^{N-1}\mathbb{E}\left[B_i^1\right]\Big)^{\frac{1}{2}}\Big(\sum_{i=0}^{N-1}\tau_i(Y)\Big)^{\frac{1}{2}}h \le c\Big(\sum_{i=0}^{N-1}\tau_i(Y)\Big)^{\frac{1}{2}}h$$

and similarly for the second term

$$\sum_{i=0}^{N-1}\mathbb{E}\left[B_i^2\right]^{\frac{1}{2}}\mathbb{E}[|\widehat{Z}_i - \bar{Z}_{t_i}|^2 h]^{\frac{1}{2}}h \le c\Big(\sum_{i=0}^{N-1}\tau_i(Z)\Big)^{\frac{1}{2}}h.$$

□

The proof of the Theorem 4.5.1

By collecting the above results we can now prove Theorem 4.5.1.

Proof of Theorem 4.5.1. The proof is a combination of the Fundamental Lemma 4.4.4, with Lemma 4.4.6, Proposition 4.4.11 and stability results obtained in this section, namely Proposition 4.5.3 and Lemma 4.5.4.

We move to the proof of part *ii*), the case $\theta = 1/2$. Since in this case f depends only on y, a quick re-run of arguments of the Fundamental Lemma 4.4.4, shows there exists a constant $c > 0$ such that

$$\max_{i=0,\dots,N} \mathbb{E}[|Y_{t_i} - Y_i|^2] \le c\Big\{\mathbb{E}[\,|Y_{t_N} - Y_N|^2] + \sum_{i=0}^{N-1} \frac{\tau_i(Y)}{h}\Big\} + (1+h)\mathcal{R}^{\mathcal{S}}(H).$$

The first two terms on the RHS can be bounded by $ch^{2\gamma} + ch^4$, $c > 0$, using Lemma 4.4.6 and Proposition 4.4.11, respectively. By Lemma 4.5.4 there exists a constant $c > 0$ such that

$$\mathcal{R}^{\mathcal{S}}(H) \le c\,\mathbb{E}[|Y_{t_N} - Y_N|^4]^{\frac{1}{2}} h^2 + c\Big(\sum_{i=0}^{N-1} \tau_i(Y)\Big)^{\frac{1}{2}} h,$$

and using again Lemma 4.4.6 and Proposition 4.4.11 yields $\mathcal{R}^{\mathcal{S}}(H) \le ch^{2\gamma+2} + ch^{7/2}$. By joining these results the theorem's conclusion follows. □

4.5.3 Further remarks

Here, we discuss a true overall 2nd order scheme, namely a 2nd order discretization for Z, and an intuitive variance reduction technique which we have used throughout but not made formally explicit.

The candidate for 2nd order scheme

For the general case were driver depends on Z, the approximation for Z_i, namely (4.4.14), is not enough to obtain a higher order scheme as it is a 1st order approximation. The proper higher order scheme in its full generality follows by applying the trapezoidal rule to *all* integrals present in (4.4.9); as is done for (4.4.8). With some manipulation (left to the reader), we end up with the following approximation for Z_i (compare with (4.4.14)),

$$Z_i = \frac{2}{h}\mathbb{E}_i\big[\Delta W_{t_{i+1}}\,\big(Y_{t_{i+1}} + (1-\theta) f(t_i, X_{i+1}, Y_{i+1}, Z_{i+1})h\big)\,\big] - \mathbb{E}_i\big[Z_{i+1}\big],$$

with $\theta = 1/2$, the terminal condition $Y_N = g(X_N)$, along with (4.4.13) and a suitable approximation for Z_T. An approximation for Z_T is not trivial and can for instance be found via Malliavin calculus. The general treatment of such a scheme is left for future research.

Another type of 2nd order scheme can be found in [22], the approximation there is based in Itô-Taylor expansions.

Controlling the variance of the scheme

If we use the notation set up in Subsection 4.4.4, the approximation (4.4.14) can be written out as $Z_i = \mathbb{E}_i\left[\Delta W_{i+1} A_{i+1}\right]/h$. We point out that implementation wise it is better to use the lower variance approximation (4.4.17) instead of (4.4.14), i.e. to use

$$Z_i = \frac{1}{h}\mathbb{E}_i\left[\Delta W_{i+1}\left(A_{i+1} - \mathbb{E}_i[A_{i+1}]\right)\right], \qquad i = 0, \cdots, N-1.$$

This does not lead to a relevant additional computation effort, as $\mathbb{E}_i[A_{i+1}]$ must be computed for the estimation of the Y_i component. To avoid a long analysis we make some simplifying assumptions in order to better explain the gain: assume $X_t = x + W_t$ and that we are about to compute Z_0 (a standard expectation); assume further (via Doob-Dynkin Lemma) that A_1 can be written as[8] $A_1 = \varphi(X_1) = \varphi(x + \Delta W_1)$ where φ has some regularity so that

$$\varphi(x + \Delta W_1) = \varphi(x) + \varphi'(x)(\Delta W_1) + \frac{1}{2}\varphi''(x^*)(\Delta W_1)^2,$$

where x^* lies between x and $x + \Delta W_1$. Then the Monte-Carlo (MC) estimator for Z_0 from (4.4.14), with M samples of the normal $\mathcal{N}(0,1)$ distribution given by $\{\mathcal{N}^\lambda\}_{\lambda=1,\dots,M}$, and its Standard deviation (Std) are

$$Z_0^{\text{MC},(4.4.14)} = \frac{1}{M}\sum_{\lambda=1}^{M} \frac{\sqrt{h}\mathcal{N}^\lambda}{h}\varphi\big(x + \sqrt{h}\mathcal{N}^\lambda\big) \quad \text{with} \quad \text{Std} \approx \frac{|\varphi(x)|}{\sqrt{h}\sqrt{M}}.$$

Using (4.4.17) instead of (4.4.14) to compute Z_0 would produce the MC estimator and its Std

$$Z_0^{\text{MC},(4.4.17)} = \frac{1}{M}\sum_{\lambda=1}^{M} \frac{\sqrt{h}\mathcal{N}^\lambda}{h}\Big(\varphi\big(x + \sqrt{h}\mathcal{N}^\lambda\big) - \varphi(x)\Big) \quad \text{with} \quad \text{Std} \approx \frac{|\varphi'(x)|}{\sqrt{M}}.$$

[8]If the reader is aware of how conditional expectations in the BSDE framework are calculated, say e.g. via projection over a basis of functions, having a function φ is expected.

Compare now the standard deviation of both estimators. It is crucial for the stability that the denominator of the variance of $Z_0^{\mathrm{MC},(4.4.17)}$ lacks that $\sqrt{h}$ term. If M is kept fixed then as h gets smaller we expect $Z_0^{\mathrm{MC},(4.4.14)}$ to blow up while $Z_0^{\mathrm{MC},(4.4.17)}$ will remain controlled (assuming φ can be controlled[9]). This can be numerically confirmed in [1].

We point out that this simple trick can be adapted to the scheme proposed in the next section as well as to the computation of the 2nd order scheme proposed previously.

4.6 Convergence of a tamed explicit scheme.

Unlike the case $\theta \in [1/2, 1]$, when $\theta < 1/2$, the local estimates of Proposition 4.4.7 cannot be extended to the global ones (as in Proposition 4.5.2). Consequently, we also do not have a control over the stability remainder $\mathcal{R}^{\mathcal{S}}(H)$ (see Definition 4.4.2). To overcome this difficulty, we consider a tamed version of the explicit scheme, which in turn corresponds to a truncation procedure applied to the original BSDE.

Remark 4.6.1 ($m > 1$)**.** In this section we focus exclusively on the case $m > 1$ in Assumptions (HY0). The easier case $m = 1$ does not require taming and stability of the scheme results from a straightforward adaptation of the proof of Proposition 4.6.5.

4.6.1 A tamed explicit scheme.

For any level $L > 0$, we define the thresholding function $T_L : \mathbb{R} \to \mathbb{R}$, $x \mapsto -L \vee x \wedge L$. We denote similarly its extension as a function from $\mathbb{R}^d$ to $\mathbb{R}^d$ (projection on the ball of radius L). We consider the following scheme: define $Y_N := T_{L_h}\big(g(X_N)\big)$, $Z_N := 0$, and for $i = N-1, \dots, 0$,

$$Y_i := \mathbb{E}_i\Big[Y_{i+1} + f\big(t_{i+1}, T_{K_h}(X_{i+1}), Y_{i+1}, Z_{i+1}\big)h\Big], \tag{4.6.1}$$

$$Z_i := \mathbb{E}_i\Big[\frac{\Delta W_{t_{i+1}}}{h}\Big(Y_{i+1} + f\big(t_{i+1}, T_{K_h}(X_{i+1}), Y_{i+1}, Z_{i+1}\big)h\Big)\Big], \tag{4.6.2}$$

[9] In [35] it is shown for the locally Lipschitz driver case that φ is indeed a Lipschitz function of its variables.

where the levels L_h and K_h satisfy $e^{c_1 T}\big(L_h^2 + c_2 T + c_2 T K_h^2\big) \leq h^{-1/(m-1)}$, with

$$c_1 = 2\big(L_y + 12dL_z^2 + 2L_y^2\big) \quad \text{and} \quad c_2 = \max\Big\{\frac{L^2}{4dL_z^2}, \frac{L_x^2}{4dL_z^2}\Big\}.$$

For $h \leq h^*$, where h^* satisfies $e^{c_1 T} c_2 T \leq (h^*)^{-1/(m-1)}/3$ and $h^* \leq 1/(32dL_z^2)$ we can take

$$L_h = \frac{1}{\sqrt{3}} e^{-\frac{1}{2}c_1 T}\Big(\frac{1}{h}\Big)^{\frac{1}{2(m-1)}} \qquad \text{and} \qquad K_h = \frac{1}{\sqrt{3}}\frac{e^{-\frac{1}{2}c_1 T}}{\sqrt{c_2 T}}\Big(\frac{1}{h}\Big)^{\frac{1}{2(m-1)}}.$$

Here we present the main results of this section that states the convergence rate of scheme (4.6.1)-(4.6.2). Its proof is postponed to the end of this section.

Theorem 4.6.2. *Let (HX0), (HY0$_{loc}$) hold and $h \leq h^*$. Assume that the order γ of the approximation $\{X_i\}_{i=0,\cdots,N}$ of X is at least* $1/2$. *Then for the controlled explicit scheme* (4.6.1)-(4.6.2), *there exists a constant c such that $ERR_\pi(Y,Z) \leq c\, h^{1/2}$.*

The idea is that with this control, one can not only obtain uniform bounds for the scheme, but also a nice pathwise bound, ensuring that the output $\{Y_i\}_{i=0,\cdots,N}$ can be controlled. In other words, we show that the bound on the initial condition propagates throughout the whole scheme and hence the scheme is stable in the sense of (4.4.4) where $H_i = 0$.

Note that this controlled scheme is not exactly the scheme (4.4.13)-(4.4.14) with $\theta = 0$. However it can be seen as the case $\theta = 0$ with the functions $T_{L_h} \circ g$ and $f(\cdot, T_{K_h}(\cdot), \cdot, \cdot)$ instead of g and f, we can reuse the results of Section 4.4.

Because the scheme is controlled, we naturally compare first its output $\{(Y_i, Z_i)\}_{i\in\{0,\dots,N\}}$ to $(Y'_{t_i}, \bar{Z}'_{t_i})_{t_i\in\pi}$, where $(Y'_t, Z'_t)_{t\in[0,T]}$ is the solution to the BSDE (4.1.2) with controlled coefficients:

$$Y'_t = T_{L_h}\big(g(X_T)\big) + \int_t^T f\big(u, T_{K_h}(X_u), Y'_u, Z'_u\big)\mathrm{d}u - \int_t^T Z'_u \mathrm{d}W_u, \qquad t \in [0,T]. \tag{4.6.3}$$

In a second step, it will be enough to estimate the distance between the solution (Y', Z') of the truncated BSDE (4.6.3) and the solution (Y, Z) of the original BSDE (4.1.2) in order to conclude to the convergence of the scheme.

In line with Section 4.4 and 4.5 we denote set $\{\bar{Z}'_{t_i}\}_{t_i\in\pi}$ as in (4.3.20), $\widehat{Y}_i =$

$Y_{i,(Y'_{i+1},\bar{Z}'_{i+1})}$ and $\widehat{Z}_i = Z_{i,(Y'_{i+1},\bar{Z}'_{i+1})}$ for $i = 0, \ldots, N-1$, more precisely

$$\widehat{Y}_i := \mathbb{E}_i \left[Y'_{t_{i+1}} + f\big(t_{i+1}, T_{K_h}(X_{i+1}), Y'_{t_{i+1}}, \bar{Z}'_{t_{i+1}}\big)h \right], \tag{4.6.4}$$

$$\widehat{Z}_i := \mathbb{E}_i \left[\frac{\Delta W_{t_{i+1}}}{h} \Big(Y'_{t_{i+1}} + f_h\big(t_{i+1}, X_{i+1}, Y'_{t_{i+1}}, \bar{Z}'_{t_{i+1}}\big)h \Big) \right]. \tag{4.6.5}$$

4.6.2 Size analysis for the scheme

We now show that the tamed Euler scheme has the property that $|Y_i| \le h^{-1/(2m-2)}$ for $i = 0, \cdots, N$. This is of course true for $Y_N = T_{L_h}\big(g(X_N)\big)$ by construction. In the next two propositions we will show this bound propagates through time.

Proposition 4.6.3. *Assume (HX0), (HY0) and that $h \le 1/(32dL_z^2)$. Then there exists $c > 0$ such that if for a given $i \in \{0, \ldots, N-1\}$ one has $|Y_{i+1}| \le h^{-1/(2m-2)}$ then one also has*

$$|Y_i|^2 + \frac{1}{d}|Z_i|^2 h \le (1 + c_1 h)\, \mathbb{E}_i \Big[|Y_{i+1}|^2 + \frac{1}{4d}|Z_{i+1}|^2 h \Big] + c_2 h + c_2 h \mathbb{E}_i \Big[|T_{K_h}(X_{i+1})|^2 \Big].$$

Proof. Take $i \in \{0, \ldots, N-1\}$. We have seen in the proof of Proposition 4.4.7, equation (4.4.20) that, since $\theta = 0$,

$$|Y_i|^2 + \frac{1}{d}|Z_i|^2 h \le \Big(1 + 2(L_y + \alpha')h\Big)\mathbb{E}_i \Big[|Y_{i+1}|^2\Big] + \frac{3L_z^2}{2\alpha'}\mathbb{E}_i\Big[|Z_{i+1}|^2\Big]h + \mathbb{E}_i[B(i+1,\alpha')] + H_i^0,$$

where $B(i+1, \alpha') := (3L^2 h + 3L_x^2 |T_{K_h}(X_{i+1})|^2 h)/2\alpha'$ and

$$H_i^0 = \mathbb{E}_i\big[|f_{i+1}|^2\big]h^2 = \mathbb{E}_i\big[|f(t_{i+1}, T_{K_N}(X_{i+1}), Y_{i+1}, Z_{i+1})|^2\big]h^2.$$

Using (HY0) and the fact that $|Y_{i+1}|^{2(m-1)}h \le 1$, we have

$$\begin{aligned} |f_{i+1}|^2 h^2 &\le 4L^2h^2 + 4L_x^2 |T_{K_h}(X_{i+1})|^2 h^2 + 4L_y^2\big[|Y_{i+1}|^{2(m-1)}h\big]|Y_{i+1}|^2 h + 4L_z^2|Z_{i+1}|^2 h^2 \\ &\le 4L^2h^2 + 4L_x^2 |T_{K_h}(X_{i+1})|^2 h^2 + 4L_y^2 |Y_{i+1}|^2 h + 4L_z^2 h |Z_{i+1}|^2 h, \end{aligned}$$

so we have in the end

$$\begin{aligned} |Y_i|^2 + \frac{1}{d}|Z_i|^2 h \le{}& \Big(1 + 2\big(L_y + \alpha' + 2L_y^2\big)h\Big)\mathbb{E}_i\Big[|Y_{i+1}|^2\Big] + \Big(\frac{3L_z^2}{2\alpha'} + 4L_z^2 h\Big)\mathbb{E}_i\Big[|Z_{i+1}|^2\Big]h \\ &+ \Big(\frac{3L^2}{2\alpha'} + 4L^2 h\Big)h + \Big(\frac{3L_x^2}{2\alpha'} + 4L_x^2 h\Big)\mathbb{E}_i\Big[|T_{K_h}(X_{i+1})|^2\Big]h. \end{aligned}$$

Choose now $\alpha' = 12dL_z^2$ (so that $3L_z^2/(2\alpha') \le 1/(8d)$) and combine with the restriction $h \le 1/(32dL_z^2)$ (so that $4L_z^2 h \le \frac{1}{8d}$). Taking $c_1 = 2\big(L_y + 12dL_z^2 + 2L_y^2\big)$ and

$$c_2 = \max\left\{\frac{3L^2}{24dL_z^2} + \frac{4L^2}{32dL_z^2}, \frac{3L_x^2}{24dL_z^2} + \frac{4L_x^2}{32dL_z^2}\right\} = \max\left\{\frac{L^2}{4dL_z^2}, \frac{L_x^2}{4dL_z^2}\right\},$$

and noting that $1/(4d) \le (1 + c_1 h)/(4d)$, we find the required estimate

$$|Y_i|^2 + \frac{1}{d}|Z_i|^2 h \le \big(1 + c_1 h\big)\mathbb{E}_i\Big[|Y_{i+1}|^2 + \frac{1}{4d}|Z_{i+1}|^2 h\Big] + c_2 h + c_2 h \mathbb{E}_i\Big[|T_{K_h}(X_{i+1})|^2\Big].$$

□

We can then use this local bound to obtain the following pathwise bound.

Proposition 4.6.4. *Let (HX0) and (HY0) hold. For any $i \in \{0, \cdots, N-1\}$,*

$$\begin{aligned}
&|Y_i|^2 + \frac{1}{4d}|Z_i|^2 h + \frac{3}{4d}\mathbb{E}_i\Big[\sum_{j=i}^{N-1}|Z_j|^2 h\Big] \\
&\quad \le e^{c_1(N-i)h}\mathbb{E}_i\big[|Y_N|^2\big] + e^{c_1(N-1-i)h}\left(\sum_{j=i}^{N-1} c_2 h + c_2 h \mathbb{E}_i\Big[|T_{K_h}(X_{i+1})|^2\Big]\right).
\end{aligned}$$

This implies in particular that $|Y_i| \le h^{-1/(2m-2)}$.

Proof. The proof goes by induction. The case $i = N$ is clear. If the estimate is true for $i+1$, noting that $|Y_N| \le L_h$, $|T_{K_h}(x)| \le K_h$ and $e^{c_1 T}\big(L_h^2 + c_2 T + c_2 T K_h^2\big) \le h^{-1/(m-1)}$, we see that $|Y_{i+1}|^2 \le h^{-1/(m-1)}$. Then, combining the estimate of Proposition 4.6.3 and the estimate for $i+1$ (from the induction assumption), as in Lemma 4.7.4, we obtain the desired estimate for i. □

In view of the previous proposition we can derive a similar estimate for the solution (Y', Z') to (4.6.3). Namely, using (4.2.5) with $\alpha = 12dL_z^2$ and combining it further with (HY0), we have

$$\begin{aligned}
|Y_t'|^2 &\le e^{2(L_y + 12dL_z^2)(T-t)}\mathbb{E}_t\Big[|T_{L_h}(g(X_T))|^2 + \int_t^T \frac{1}{16dL_z^2}|f(u, T_{K_h}(X_u), 0, 0)|^2 \mathrm{d}u\Big] \\
&\le e^{c_1(T-t)}\mathbb{E}_t\Big[|T_{L_h}(g(X_T))|^2 + \int_t^T \frac{1}{8dL_z^2}\big(L^2 + L_x^2|T_{K_h}(X_u)|^2\big)\mathrm{d}u\Big] \\
&\le e^{c_1 T}\Big(L_h^2 + c_2 T + c_2 T K_h^2\Big) \le \Big(\frac{1}{h}\Big)^{\frac{1}{m-1}},
\end{aligned}$$

implying in particular that $|Y'_{t_i}| \leq h^{-1/(2m-2)}$ for all i.

These two estimates, ensuring that both Y_i and Y'_{t_i} are bounded by $h^{-1/(2m-2)}$ will be useful in the analysis of the global error, since the explicit scheme is found to be stable (in the sense of (4.4.4)) under this threshold.

4.6.3 Stability analysis for the scheme.

As previously, for any $i \in \{0, \cdots, N-1\}$ we use the notation $\delta Y_{i+1} := Y'_{t_{i+1}} - Y_{i+1}$ and $\delta Z_{i+1} := \bar{Z}'_{t_{i+1}} - Z_{i+1}$, as well as $\delta A_{i+1} := \delta Y_{i+1} + \delta f_{i+1} h$ where δf_{i+1} is given by

$$\delta f_{i+1} := f\big(t_{i+1}, T_{K_h}(X_{i+1}), Y'_{t_{i+1}}, \bar{Z}'_{i+1}\big) - f\big(t_{i+1}, T_{K_h}(X_{i+1}), Y_{i+1}, Z_{i+1}\big).$$

Then, denoting $\widehat{\delta Y}_i = \widehat{Y}_i - Y_i$ and $\widehat{\delta Z}_i = \widehat{Z}_i - Z_i$, we can write

$$\widehat{\delta Y}_i = \mathbb{E}_i\big[\delta A_{i+1}\big] \qquad \text{and} \qquad \widehat{\delta Z}_i = \mathbb{E}_i\big[\frac{1}{h}\Delta W_{t_{i+1}} \delta A_{i+1}\big].$$

We now proceed to show that, because the two inputs satisfy $|Y_{i+1}|, |Y'_{t_{i+1}}| \leq h^{-1/(2m-2)}$, the scheme is stable in the sense that we can obtain the estimate (4.4.4) with $H_i = 0$.

Proposition 4.6.5. *Assume (HX0) and (HY0$_{loc}$). Then there exists a constant c for any $h \leq \min\{1, 1/32dL_z^2\}$, such that*

$$|\widehat{\delta Y}_i|^2 + \frac{1}{d}|\widehat{\delta Z}_i|^2 h \leq (1+ch)\,\mathbb{E}_i\Big[|\delta Y_{i+1}|^2 + \frac{1}{4d}|\delta Z_{i+1}|^2 h\Big], \quad i \in \{0, \cdots, N-1\}.$$

Proof. Let $i \in \{0, \cdots, N-1\}$. Just like for Proposition 4.5.3, the proof mimics the computations of the proof of Proposition 4.4.7 with only a small adjustment for the constants. However, a different argumentation for the term $H_i^0 = |\delta f_{i+1}|^2 h^2$ is required. Using (HY0$_{loc}$), $h \leq 1$ and the bounds $|Y'_{t_{i+1}}|^{2(m-1)}h$, $|Y'_{t_{i+1}}|^{2(m-1)}h \leq 1$, we have

$$\begin{aligned}
|\delta f_{i+1}|^2 h^2 &\leq 2L_y^2\big(1 + |Y'_{t_{i+1}}|^{2(m-1)} + |Y_{i+1}|^{2(m-1)}\big)|Y'_{t_{i+1}} - Y_{i+1}|^2 h^2 + 2L_z^2 |\bar{Z}'_{t_{i+1}} - Z_{i+1}|^2 h^2 \\
&= 2L_y^2\big(h + |Y'_{t_{i+1}}|^{2(m-1)}h + |Y_{i+1}|^{2(m-1)}h\big)h|Y'_{t_{i+1}} - Y_{i+1}|^2 + 2L_z^2 h|\bar{Z}'_{t_{i+1}} - Z_{i+1}|^2 h \\
&\leq 6L_y^2 h|\delta Y_{i+1}|^2 + 2L_z^2 h|\delta Z_{i+1}|^2 h.
\end{aligned}$$

The rest follows as in the proof of Proposition 4.4.7. $\square$

4.6.4 Convergence of the scheme.

The convergence of the scheme is achieved by controlling the error committed by the truncation procedure, $\|Y - Y'\|_{\mathcal{S}^2} + \|Z - Z'\|_{\mathcal{H}^2}$, as a function of the time step, then by controlling the numerical approximation (4.6.4), (4.6.5) of the solution (Y', Z') to (4.6.3).

The Distance between $(Y_i, Z_i)_i$ and $(Y'_{t_i}, \bar{Z}'_{t_i})_i$

We estimate this distance by combining Lemma 4.4.4 and the estimate for the sum of the local discretization errors given by Proposition 4.4.11.

Since the controlled scheme (4.6.1)-(4.6.2) is the $\theta = 0$ scheme with coefficient $f(\cdot, \cdot, T_{K_h}(\cdot), \cdot)$ having the same Lipschitz constants as f, the results of Lemma 4.4.9 and Lemma 4.4.10 are still valid with the same constants. The only difference is that the path-regularity involved is now that of (Y', Z'), but since $T_{L_h} \circ g$ is still Lipschitz, Theorem 4.3.5 indeed applies to (Y', Z'). So we are entitled to use Proposition 4.4.11 and conclude that

$$\max_{i=0,\dots,N} \mathbb{E}[|Y'_{t_i} - Y_i|^2] + \sum_{i=0}^{N-1} \mathbb{E}[|\bar{Z}'_{t_i} - Z_i|^2]h \tag{4.6.6}$$

$$\leq c\Big(\mathbb{E}[|Y'_{t_N} - Y_N|^2] + \mathbb{E}[|\bar{Z}'_{t_N} - Z_N|^2]h\Big) + c\sum_{i=0}^{N-1}\Big(\frac{1}{h}\tau_i(Y) + \tau_i(Z)\Big)$$

$$\leq ch. \tag{4.6.7}$$

We note that the thresholds L_h and K_h have no effect in this estimation.

The Distance between $(Y'_{t_i}, \bar{Z}'_{t_i})_i$ and $(Y_{t_i}, \bar{Z}_{t_i})_i$

Finally, we estimate the distance between $(Y'_{t_i}, \bar{Z}'_{t_i})_i$ and $(Y_{t_i}, \bar{Z}_{t_i})_i$, which gathers all the error induced by the taming. In order to estimate this error, we need to have an estimation of the L^2-distance between X_u and $T_{K_h}(X_u)$ on the one hand, and $g(X_T)$ and $T_{L_h}\big(g(X_T)\big)$ on the other. We give a general estimation for this below.

Proposition 4.6.6. *Let ξ be a random variable in L^q for some $q > 2$, and $L > 0$. Then we have*

$$\mathbb{E}[\,|\xi - T_L(\xi)|^2] \leq 4\mathbb{E}[\,|\xi|^q]\Big(\frac{1}{L}\Big)^{q-2}$$

Proof. Using the facts that $T_L(x) = x$ for $|x| \leq L$ and that $|T_L(\xi)| \leq |\xi|$, together with the Hölder and the Markov inequalities, we have

$$\begin{aligned}\mathbb{E}[|\xi - T_L(\xi)|^2] &= \mathbb{E}[|\xi - T_L(\xi)|^2 1_{\{|\xi|\geq L\}}] \leq 4\mathbb{E}[|\xi|^2 1_{\{|\xi|\geq L\}}] \\ &\leq 4\mathbb{E}[|\xi|^q]^{\frac{2}{q}}\mathbb{P}[|\xi| \geq L]^{1-\frac{2}{q}} \leq 4\mathbb{E}[|\xi|^q]^{\frac{2}{q}}\Big(\frac{\mathbb{E}[|\xi|^q]}{L^q}\Big)^{1-\frac{2}{q}} = 4\mathbb{E}[|\xi|^q]\Big(\frac{1}{L}\Big)^{q(1-\frac{2}{q})}\end{aligned}$$

□

Now, via Jensen's inequality we have

$$|\bar{Z}_{t_i} - \bar{Z}'_{t_i}|^2 h = \left|\frac{1}{h}\mathbb{E}_i\Big[\int_{t_i}^{t_{i+1}} Z_u \mathrm{d}u\Big] - \frac{1}{h}\mathbb{E}_i\Big[\int_{t_i}^{t_{i+1}} Z'_u \mathrm{d}u\Big]\right|^2 h \leq \mathbb{E}_i\Big[\int_{t_i}^{t_{i+1}} |Z_u - Z'_u|^2 \mathrm{d}u\Big],$$

from which it clearly follows that

$$\max_{i=0,\dots,N} \mathbb{E}[|Y_{t_i} - Y'_{t_i}|^2] + \sum_{i=0}^{N-1} \mathbb{E}[|\bar{Z}_{t_i} - \bar{Z}'_{t_i}|^2]h \leq \sup_{t\in[0,T]} \mathbb{E}\Big[|Y_t - Y'_t|^2\Big] + \mathbb{E}\Big[\int_0^T |Z_u - Z'_u|^2 \mathrm{d}u\Big].$$

From the a priori estimate (4.2.6) we have

$$\begin{aligned}&\sup_{t\in[0,T]} \mathbb{E}[|Y_t - Y'_t|^2] + \mathbb{E}\Big[\int_0^T |Z_u - Z'_u|^2 \mathrm{d}u\Big] \\ &\quad\leq c\Big(\mathbb{E}[|g(X_T) - T_{L_N}(g(X_T))|^2] + \mathbb{E}\Big[\int_0^T |f\big(u, X_u, Y'_u, Z'_u\big) - f\big(u, T_{K_N}(X_u), Y'_u, Z'_u\big)|^2 \mathrm{d}u\Big]\Big) \\ &\quad\leq c\Big(\mathbb{E}[|g(X_T) - T_{L_N}(g(X_T))|^2] + L_x^2 \int_0^T \mathbb{E}[|X_u - T_{K_N}(X_u)|^2]\mathrm{d}u\Big) \\ &\quad\leq c\Big(4\Big(\frac{1}{L_h}\Big)^{2m-2}\mathbb{E}[\,|g(X_T)|^{2m}] + \Big(\frac{1}{K_h}\Big)^{2m-2} 4L_x^2 \int_0^T \mathbb{E}[|X_u|^{2m}]\mathrm{d}u\Big),\end{aligned}$$

thanks to Proposition 4.6.6. Now, since $X \in \mathcal{S}^{2m}$ (Theorem 4.2.2), g is of linear growth, and L_h and K_h are of order $h^{-1/(2m-2)}$, we can conclude that

$$\max_{i=0,\dots,N} \mathbb{E}[|Y_{t_i} - Y'_{t_i}|^2] + \sum_{i=0}^{N-1} \mathbb{E}[|\bar{Z}_{t_i} - \bar{Z}'_{t_i}|^2]h \leq ch. \tag{4.6.8}$$

The proof of the Theorem 4.6.2

By collecting the above results we can now prove Theorem 4.6.2.

Proof of Theorem 4.6.2. Estimate (4.6.8) controls the distance between the solution $\{(Y_{t_i}, \bar{Z}_{t_i})\}_{t_i\in\pi}$ to the original BSDE (4.1.2) and the solution $\{(Y'_{t_i}, \bar{Z}'_{t_i})\}_{t_i\in\pi}$ to the truncated one (4.6.3). Then, estimate (4.6.6) controls the error of numerically approximating the truncated BSDE via scheme (4.6.1), (4.6.2). This estimate follows via the Fundamental Lemma 4.4.4 combined with Lemma 4.4.6 for the error of the terminal condition, Proposition 4.4.11 for the control on the sum of local errors $\tau_i(Y)$ and $\tau_i(Z)$ and finally Proposition 4.6.5 implies that H_i is zero and hence that $\mathcal{R}^{\mathcal{S}}(H) = 0$.

Combining all these estimates produces sought conclusion $\mathrm{ERR}_\pi(Y, Z) \le ch^{1/2}$. $\square$

4.7 Technical details and additional results

4.7.1 Motivating example

Before we state the main result we recall a result on the behavior of Gaussian random variables (which we do not prove, but the reader is invited to try, in any case see Lemma 4.1 in [43]). The notation and probability spaces we work with in this section are as stated in Section 4.2.

Lemma 4.7.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $Z : \Omega \to \mathbb{R}$ be an $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable mapping with standard normal distribution. Then for any $x \in [0, \infty)$ it holds that*

$$\mathbb{P}[\,|Z| \ge x\,] \ge \frac{1}{4} x e^{-x^2}.$$

The statement of Lemma 4.1.1 follows from the next lemma.

Lemma 4.7.2. *Let π^N denote the uniform grid of the time interval $[0, 1]$ with $N + 1$ points and step size $h := 1/N$, where $N \in \mathbb{N}$ is an* even *number; $t = 1/2$ is common to all grids π^N. Let (Y, Z) be the unique solution to (4.1.3) with driver $f(y) := -y^3$ and terminal condition $\xi := W_{\frac{1}{2}} \in L^p(\mathcal{F}_1)$ for any $p \ge 1$.*

Denote by $\{Y_i^{(N)}\}_{i\in\{0,\cdots,N\}}$ the Euler approximation of $(Y_t)_{t\in[0,1]}$ defined via (4.1.4) on the grid π^N; denote by $Y_{\frac{1}{2}}^{(N)}$ the approximation at the time point $t = 1/2$ (corresponding to $i = N/2$).

i) For $i \in \{\frac{N}{2}, \cdots, N\}$, on the set $\{\omega : \xi(\omega) \ge 2\sqrt{N}\}$ it holds that $|Y_i(\omega)| \ge 2^{2^{N-i}}\sqrt{N}$,

ii) $\lim_{N\to\infty} \mathbb{E}[\,|Y_{\frac{1}{2}}^{(N)}|\,] = +\infty$.

Proof. The first thing to remark is that no conditional expectation needs to be computed for the scheme (4.1.4) for $i \in \{N/2, \cdots, N\}$ because $\xi = W_{\frac{1}{2}}$ is $\mathcal{F}_t$-adapted for any $t \in [1/2, 1]$. So the scheme's approximations up to $Y^{(N)}_{\frac{1}{2}}$ can be written as

$$Y_N^{(N)} = W_{\frac{1}{2}}, \quad Y_{N-1}^{(N)} = \psi(W_{\frac{1}{2}}), \quad Y_{N-2}^{(N)} = \psi\big(\psi(W_{\frac{1}{2}})\big), \quad \cdots, \quad Y_{\frac{N}{2}}^{(N)} = \psi^{\circ(N/2)}(W_{\frac{1}{2}}),$$

where $\psi(x) := x - hx^3$ and $\psi^{\circ(n)}$ denotes the composition of ψ with itself n-times ($n \in \mathbb{N}$).

Proof of Part i) In this first step we fix N and drop the superscript (N) from $Y^{(N)}$. We work on the event that $\xi = Y_N \geq 2\sqrt{N}$. We have first

$$Y_{N-1} = \mathbb{E}_{N-1}[Y_N - Y_N^3 h] = Y_N(1 - Y_N^2 h).$$

Observe that $Y_N^2 \geq 2^2 N$ which implies $(1 - Y_N^2 h) \leq (1 - 2^2) < 0$. Hence (since $Y_N > 0$)

$$Y_{N-1} = Y_N(1 - Y_N^2 h) \leq -2\sqrt{N}(2^2 - 1) \leq -2^2\sqrt{N} < 0.$$

Next, since $Y_{N-1} < 0$, $Y_{N-1}^2 \geq 2^4 N$ which implies $1 - Y_{N-1}^2 h \leq (1 - 2^4) < 0$. Hence

$$Y_{N-2} = Y_{N-1}(1 - Y_{N-1}^2 h) = -Y_{N-1}(Y_{N-1}^2 h - 1) \geq 2^2\sqrt{N}(2^4 - 1) \geq 2^{2^2}\sqrt{N}.$$

Proceeding by induction we can easily show that

$$|Y_i| \geq 2^{2^{N-i}}\sqrt{N}, \quad i = \frac{N}{2}, \cdots, N.$$

Indeed, assume $Y_{i+1} \geq 2^{2^{N-i-1}}\sqrt{N}$ (note that in the light of the above calculations the negative case is analogous). Then

$$Y_i = Y_{i+1}(1 - Y_{i+1}^2 h) \leq 2^{2^{N-i-1}}\sqrt{N}\Big(1 - \Big(2^{2^{N-i-1}}\Big)^2\Big) \leq -2^{2^{N-i}}\sqrt{N}.$$

Proof of Part ii): It follows easily from Lemma 4.7.1 that

$$\mathbb{P}\big[\,|W_{\frac{1}{2}}| \geq 2\sqrt{N}\,\big] \geq \frac{\sqrt{2}}{2}\sqrt{N}e^{-8N}.$$

Then, using *Part i)* (to go from the 1st to the 2nd line) and the above remark (on the 3rd line) we have

$$\begin{aligned}
\lim_{N\to\infty}\mathbb{E}[\,|Y^{(N)}_{\frac12}|\,] &= \lim_{N\to\infty}\mathbb{E}[1_{\{\xi\geq 2\sqrt{N}\}}|Y^{(N)}_{\frac12}|+1_{\{\xi<2\sqrt{N}\}}|Y^{(N)}_{\frac12}|\,] \geq \lim_{N\to\infty}\mathbb{E}[1_{\{\xi\geq 2\sqrt{N}\}}|Y^{(N)}_{\frac12}|\,]\\
&\geq \lim_{N\to\infty}\mathbb{E}[1_{\{\xi\geq 2\sqrt{N}\}}2^{2^{N-N/2}}\ \sqrt{N}\,]\\
&= \lim_{N\to\infty}2^{2^{N/2}}\ \sqrt{N}\ \mathbb{P}[\,|W_{\frac12}|\geq 2\sqrt{N}\,] \geq \lim_{N\to\infty}2^{(2^{N/2})}\frac{\sqrt{2}}{2}Ne^{-8N}=+\infty.
\end{aligned}$$

$\square$

4.7.2 Basics of Malliavin's calculus

We briefly introduce the main notation of the stochastic calculus of variations also known as Malliavin's calculus. For more details, we refer the reader to [63], for its application to BSDEs we refer to [44]. Let $\mathcal{S}$ be the space of random variables of the form

$$\xi = F\Big(\big(\int_0^T h^{1,i}_s \mathrm{d}W^1_s\big)_{1\leq i\leq n},\cdots,\big(\int_0^T h^{d,i}_s \mathrm{d}W^d_s\big)_{1\leq i\leq n}\Big),$$

where $F\in C^\infty_b(\mathbb{R}^{n\times d})$, $h^1,\cdots,h^n\in L^2([0,T];\mathbb{R}^d)$, $n\in\mathbb{N}$. To simplify notation, assume that all h^j are written as row vectors. For $\xi\in\mathcal{S}$, we define $D=(D^1,\cdots,D^d):\mathcal{S}\to L^2(\Omega\times[0,T])^d$ by

$$D^i_\theta\xi=\sum_{j=1}^n\frac{\partial F}{\partial x_{i,j}}\Big(\int_0^T h^1_t\mathrm{d}W_t,\ldots,\int_0^T h^n_t\mathrm{d}W_t\Big)h^{i,j}_\theta,\quad 0\leq\theta\leq T,\quad 1\leq i\leq d,$$

and for $k\in\mathbb{N}$ its k-fold iteration by $D^{(k)}=(D^{i_1}\cdots D^{i_k})_{1\leq i_1,\cdots,i_k\leq d}$. For $k\in\mathbb{N}$, $p\geq 1$ let $\mathbb{D}^{k,p}$ be the closure of $\mathcal{S}$ with respect to the norm

$$\|\xi\|^p_{k,p}=\mathbb{E}\Big[\|\xi\|^p_{L^p}+\sum_{i=1}^k\||D^{(k)]}\xi|\|^p_{(\mathcal{H}^p)^i}\Big].$$

$D^{(k)}$ is a closed linear operator on the space $\mathbb{D}^{k,p}$. Observe that if $\xi\in\mathbb{D}^{1,2}$ is $\mathcal{F}_t$-measurable then $D_\theta\xi=0$ for $\theta\in(t,T]$. Further denote $\mathbb{D}^{k\infty}=\cap_{p>1}\mathbb{D}^{k,p}$.

We also need Malliavin's calculus for $\mathbb{R}^m$ valued smooth stochastic processes. For $k\in\mathbb{N},p\geq 1$, denote by $\mathbb{L}_{k,p}(\mathbb{R}^m)$ the set of $\mathbb{R}^m$-valued progressively measurable processes $u=(u^1,\cdots,u^m)$ on $[0,T]\times\Omega$ such that

i) For Lebesgue-a.a. $t\in[0,T]$, $u(t,\cdot)\in(\mathbb{D}^{k,p})^m$;

ii) $[0,T] \times \Omega \ni (t,\omega) \mapsto D^{(k)}u(t,\omega) \in (L^2([0,T]^{1+k}))^{d\times n}$ admits a progressively measurable version;

iii) $\|u\|_{k,p}^p = \|u\|_{\mathcal{H}^p}^p + \sum_{i=1}^k \| D^i u \|_{(\mathcal{H}^p)^{1+i}}^p < \infty$.

Note that Jensen's inequality gives[10] for all $p \geq 2$

$$\mathbb{E}\Big[\Big(\int_0^T \int_0^T |D_u X_t|^2 \mathrm{d}u\, \mathrm{d}t\Big)^{\frac{p}{2}}\Big] \leq T^{p/2-1} \int_0^T \|D_u X\|_{\mathcal{H}^p}^p \mathrm{d}u. \tag{4.7.1}$$

We recall a result from [44] concerning the rule for the Malliavin differentiation of Itô integrals which is of use in applications of Malliavin's calculus to stochastic analysis.

Theorem 4.7.3 (Theorem 2.3.4 in [44])**.** *Let $(X_t)_{t\in[0,T]} \in \mathcal{H}^2$ be an adapted process and define $M_t := \int_0^t X_r \mathrm{d}W_r$ for $t \in [0,T]$. Then, $X \in \mathbb{L}^{1,2}$ if and only if $M_t \in \mathbb{D}^{1,2}$ for any $t \in [0,T]$. And moreover for any $0 \leq s,t \leq T$ we have*

$$D_s M_t = X_s 1_{\{s\leq t\}}(s) + 1_{\{s\leq t\}}(s) \int_s^t D_s X_r \mathrm{d}W_r. \tag{4.7.2}$$

4.7.3 A particular Gronwall lemma

We state here a "discrete Gronwall lemma" of some kind, particularly useful for the numerical analysis of BSDEs, and which we use extensively in this work.

Lemma 4.7.4. *Let a_i, b_i, c_i, be such that $a_i, b_i \geq 0$, $c_i \in \mathbb{R}$ for $i = 0,1,\dots,N$. Assume that, for some constant $c > 0$ and $h > 0$, we have*

$$a_i + b_i \leq (1+ch)a_{i+1} + c_i, \quad \text{for} \quad i = 0,1,\dots,N-1. \tag{4.7.3}$$

Then the following inequality holds for every i

$$a_i + \sum_{j=i}^{N-1} b_j \leq e^{c(N-i)h} a_N + \sum_{j=i}^{N-1} e^{c(j-i)h} c_j.$$

Proof. The estimate is clearly true for $i = N-1$ (even for $i = N$ in fact). Then, for

[10] The reason behind this last inequality is that within the BSDE framework the usual tools to obtain a priori estimates yield with much difficulty the LHS while with relative ease the RHS.

any $i \leq N-2$, if it is true for $i+1$, by multiplying both sides by e^{ch} we find that

$$e^{ch}a_{i+1} + e^{ch} \sum_{j=i+1}^{N-1} b_j \leq e^{c(N-i)h} a_N + \sum_{j=i+1}^{N-1} e^{c(j-i)h} c_j$$

Summing this inequality with (4.7.3) and noting that $\sum_{j=i+1}^{N-1} b_j \leq e^{ch} \sum_{j=i+1}^{N-1} b_j$ due to the positivity of the b_j terms gives the sought estimate for any i. □

4.7.4 Preservation of the monotonicity condition by differentiation and mollification.

This subsection brings precisions to section 4.3, theorems 4.3.1 and 4.3.5.

Earlier versions of our results were making use of an assumption called (HY0$^+$) at the time when the research was being carried (not reported here), in order to negotiate two technicalities in the proof of theorems 4.3.1 and 4.3.5. Namely, we needed to guarantee that the driver of the differentiated BSDE is again a monotone driver (in theorem 4.3.1), and that the same goes for mollified drivers (in the proof of theorem 4.3.5). That assumption was eventually removed, as we found it was not necessary. The computations proving our claims without resorting to (HY0$^+$) are in fact relatively simple, so we did not include them in the main text, but we give them below.

The driver of the differentiated BSDE is again monotone.

For the proof of theorem 4.3.1 we need to show that

$$F : (\omega, r, \chi, \Upsilon, \Gamma) \mapsto \big(\nabla_x f\big)(r, \Theta_r^{t,x}) \cdot \chi + \big(\nabla_y f\big)(r, \Theta_r^{t,x}) \cdot \Upsilon + \big(\nabla_z f\big)(r, \Theta_r^{t,x}) \cdot \Gamma$$

satisfies the monotonicity condition. For this, we fix $(\omega, r, \chi, \Gamma)$ and denote for simplicity $F(\Upsilon) = F(\omega, r, \chi, \Upsilon, \Gamma)$ and $\big(\nabla_y f\big)(Y_r) = \big(\nabla_y f\big)(r, \Theta_r^{t,x})$. Using the definition of F, the linearity of $\big(\nabla_y f\big)(Y_r)$, a consequence of the definition of the (Fréchet) derivative, and finally the monotonicity of f, we have

$$
\begin{aligned}
\langle \Upsilon' - \Upsilon | F(\Upsilon') - F(\Upsilon) \rangle &= \langle \Upsilon' - \Upsilon | \nabla_y f(Y_r).\Upsilon' - \nabla_y f(Y_r).\Upsilon \rangle \\
&= \langle \Upsilon' - \Upsilon | \nabla_y f(Y_r).(\Upsilon' - \Upsilon) \rangle \\
&= \left\langle \Upsilon' - \Upsilon \middle| \lim_{\epsilon \to 0} \frac{f\big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - f\big(Y_r\big)}{\epsilon} \right\rangle \\
&= \lim_{\epsilon \to 0} \left\langle \Upsilon' - \Upsilon \middle| \frac{f\big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - f\big(Y_r\big)}{\epsilon} \right\rangle \\
&= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left\langle \frac{\big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - Y_r}{\epsilon} \middle| f\big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - f(Y_r) \right\rangle \\
&= \lim_{\epsilon \to 0} \frac{1}{\epsilon^2} \left\langle \big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - Y_r | f\big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - f(Y_r) \right\rangle \\
&\leq \lim_{\epsilon \to 0} \frac{1}{\epsilon^2} L_y \left| \big(Y_r + \epsilon(\Upsilon' - \Upsilon)\big) - Y_r \right|^2 \\
&= L_y \left| \Upsilon' - \Upsilon \right|^2 .
\end{aligned}
$$

So, as wanted, F is a monotone function of Υ (uniformly in $(\omega, r, \chi, \Gamma)$), and one can take the same constant L_y as for f.

The mollified drivers are again monotone.

In the proof of theorem 4.3.5 we need to show that the mollified drivers f^n are also monotone.

Consider a regularising kernel $(h^n)_{n \in \mathbb{N}}$. That is, we have $h^n \geq 0$, $h^n \in C_b^\infty$, $\int_{\mathbb{R}^k} h^n dx = 1$, and $\int_{\mathbb{R}^k \setminus B(0,\epsilon)} h^n dx \longrightarrow 0$ when $\epsilon \longrightarrow 0$. We define $f^n = f * h^n$, that is $f^n(y) = \int_{\mathbb{R}^k} f(y - x) h^n(x) dx$. Note that for notational simplicity we wrote $f(y)$ for $f(t, x, y, z)$. Using the definition of f^n, the linearity of the integral and the monotonicity of f, we see that

$$
\begin{aligned}
\langle y' - y | f^n(y') - f^n(y) \rangle &= \left\langle y' - y \middle| \int [f(y' - x) - f(y - x)] h^n(x) dx \right\rangle \\
&= \int \left\langle y' - y \middle| f(y' - x) - f(y - x) \right\rangle h^n(x) dx \\
&\leq L_y |y' - y|^2 \int h^n(x) dx \\
&= L_y |y' - y|^2 .
\end{aligned}
$$

So for every n, f^n is monotone with the same constant L_y as f. So (HY0) is satisfied uniformly in n in the mollification argument, as claimed.

4.7.5 Discrete $\mathcal{H}^p$ estimate from the continuous $\mathcal{H}^p$ estimate.

In corollary 4.3.6 we state a discrete $\mathcal{H}^p$-style estimate for $\bar{Z}$ which is then used in section 4.5.

To prove it, we first prove a discrete $\mathcal{S}^p$-style estimate for $\bar{Z}$ under (HY0$_{loc}$), resulting from an $\mathcal{S}^p$ estimate for Z, and then use it to obtain the desired discrete $\mathcal{H}^p$-style estimate for $\bar{Z}$ (as it holds for discrete estimates, just like for the continuous ones, that "$\mathcal{S}^p$ implies $\mathcal{H}^p$"). This therefore relies on an $\mathcal{S}^p$ estimate for Z (theorem 4.3.5), which is an advanced result. However it is a far more basic result that $Z \in \mathcal{H}^p$ (basic in the sense that it does not require the fine differentiability analysis and the results that stem from it, in particular that $Z \in \mathcal{S}^p$), so there ought to be a proof of the discrete $\mathcal{H}^p$-style estimate for $\bar{Z}$ following directly from the $\mathcal{H}^p$ estimate for Z. Such a proof exist, and we give it below.

Start by remarking that it follows from the definition of $\bar{Z}_{t_i}$ that

$$|\bar{Z}_{t_i}|^2 h \leq \mathbb{E}_i\Big(\int_{t_i}^{t_{i+1}} |Z_u|^2 \mathrm{d}u \Big).$$

Now, using this, then Jensen's inequality (on $\mathbb{E}_i(\cdot)$ only) to get the second line, the linearity of $\mathbb{E}[\cdot]$ and $\mathbb{E}[\mathbb{E}_i(\cdot)] = \mathbb{E}[\cdot]$ for the third line, then the elementary inequality

$\sum a_i^p \le (\sum a_i)^p$ to get the fourth, we have

$$\begin{aligned}
\mathbb{E}\Big[\sum_{i=0}^{N-1}\big(|\bar{Z}_{t_i}|^2 h\big)^p\Big] &\le \mathbb{E}\Bigg[\sum_{i=0}^{N-1}\bigg(\mathbb{E}_i\Big(\int_{t_i}^{t_{i+1}}|Z_u|^2\mathrm{d}u\Big)\bigg)^p\Bigg] \\
&\le \mathbb{E}\Bigg[\sum_{i=0}^{N-1}\mathbb{E}_i\bigg(\Big(\int_{t_i}^{t_{i+1}}|Z_u|^2\mathrm{d}u\Big)^p\bigg)\Bigg] \\
&\le \mathbb{E}\Bigg[\sum_{i=0}^{N-1}\Big(\int_{t_i}^{t_{i+1}}|Z_u|^2\mathrm{d}u\Big)^p\Bigg] \\
&\le \mathbb{E}\Bigg[\bigg(\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}}|Z_u|^2\mathrm{d}u\bigg)^p\Bigg] \\
&= \mathbb{E}\Bigg[\bigg(\int_0^T|Z_u|^2\mathrm{d}u\bigg)^p\Bigg] \le C.
\end{aligned}$$

So the discrete $\mathcal{H}^p$-style estimate for $\bar{Z}$ is proved, using only the $\mathcal{H}^p$ estimate for Z.

4.7.6 Well-definedness of the implicit schemes ($\theta > 0$).

In subsection 4.4.4 we justified why, when $\theta > 0$ and the scheme only defines Y_i implicitly, it is indeed true that Y_i is well defined. More precisely, we wanted to ensure that the equation $Y_i = \mathbb{E}_i\left[A_{i+1}\right] + \theta f(t_i, X_i, Y_i, Z_i)h$ defines indeed a unique random variable Y_i in $L^2(\mathcal{F}_i)$.

Define $b_i = \mathbb{E}_i\left[A_{i+1}\right] \in \mathbb{R}^k$ and $a_i = (X_i, Z_i) \in \mathbb{R}^d \times \mathbb{R}^{k,d}$. In subsection 4.4.4, for a fixed ω, we defined the map $F = F_\omega = F_{a_i(\omega)} : y \mapsto y - \theta f\big(t_i, X_i(\omega), y, Z_i(\omega)\big)h$. We showed, using theorem 26.A p557 of [80] that since F is strongly monotone increasing, it is invertible, and in particular we could define $Y_i(\omega) = F^{-1}_{a_i(\omega)}\big(b_i(\omega)\big)$. This defines $Y_i(\omega)$ for each ω but does not guarantee the measurability of Y_i. So we don't know if we define an element of $L^0(\mathcal{F}_i)$. It is a consequence of theorem 26.A from [80] that $F_{a_i(\omega)}$ is continuous (even Lipschitz), hence measurable, but this says nothing about the measurability of $\omega \mapsto F^{-1}_{a_i(\omega)}\big(b_i(\omega)\big)$. What we need is a joint measurability.

For this, the (standard) idea is to define rather the map G from $E = \big(\mathbb{R}^d \times \mathbb{R}^{k,d}\big) \times \mathbb{R}^k$ to itself,

$$G : (a, y) \mapsto \big(a, y - \theta f(t_i, a, y)h\big),$$

where $f(t_i, a, y) = f(t_i, x, y, z)$ for $a = (x, z)$. It is not difficult to check that, like F_ω previously, G is strongly monotone increasing (here the scalar product is that on E). This implies that for every $(a, b) \in E$ we can define $(a', y) = G^{-1}(a, b)$, and G^{-1} is Lipschitz (hence continuous, hence measurable). Actually, since the first component of G maps (a, y) to a, the first component of the inverse maps (a, b) to a, so $a' = a$ above. Now, denoting by P_2 the second projection, from E to $\mathbb{R}^k$, we can define for each ω

$$Y_i(\omega) = P_2\Big(G^{-1}\big(a_i(\omega), b_i(\omega)\big)\Big).$$

Note that it solves $G(a_i(\omega), Y_i(\omega)) = (a_i(\omega), b_i(\omega))$, so $Y_i(\omega) - \theta f\big(t_i, X_i(\omega), Y_i(\omega), Z_i(\omega)\big)h = b_i(\omega)$ as we want. The point is that now $Y_i = P_2 \circ G^{-1} \circ (a_i, b_i)$, so it is indeed $\mathcal{F}_i$ measurable (because $\omega \mapsto (a_i(\omega), b_i(\omega))$ is).

So Y_i is well defined as an element of $L^0(\mathcal{F}_i)$, and the proposition 4.4.7 proves that $Y_i \in L^2$.

Note that it doesn't seem easy to recast the above argumentation directly into the functional space L^2, defining Y_i directly as a point of that space rather than defining it as a function, ω by ω. For Lipschitz coefficients f, both argumentations are possible. But since f can have polynomial growth here, the map

$$F : Y \mapsto Y - \theta f(t_i, X_i, Y, Z_i)h$$

is not necessarily valued in L^2. This is why we wrote the argument ω-wise.

References

[1] S. Alanko and M. Avellaneda. Reducing variance in the numerical solution of BSDEs. *Comptes Rendus Mathématiques*, 351(3–4):135 – 138, 2013.

[2] S. Ankirchner, P. Imkeller, and G. dos Reis. Classical and variational differentiability of BSDEs with quadratic growth. *Electronic Journal of Probability*, 12(53):1418–1453, 2007.

[3] G. Barles and E. Lesigne. SDEs, BSDEs and PDEs. In El Karoui and Mazliak, editors, *Backward stochastic differential equations (Paris, 1995–1996)*, volume 364 of *Pitman Research Notes in Mathematics Series*, pages 47–80. Longman, 1997.

[4] P. Barrieu and N. El Karoui. Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs. *Annals of Probability*, 41:1831–1853, 2013.

[5] E. Bayraktar and S. Yao. Quadratic reflected BSDEs with unbounded obstacles. *Stochastic processes and their applications*, 122(4):1155–1203, 2012.

[6] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44:384–404, 1973.

[7] 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.

[8] P. Briand and R. Carmona. BSDEs with polynomial growth generators. *Journal Of Applied Mathematics And Stochastic Analysis*, 13(3):207–238, 2000.

[9] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochatic Processes and their Applications*, 118(5):818–838, 2008.

[10] P. Briand and F. Confortola. Differentiability of backward stochastic differential equations in Hilbert spaces with monotone generators. *Appl. Math. Optim.*, 57(2):149–176, 2008.

[11] P. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica. L^p solutions of backward stochastic differential equations. *Stochatic Processes and their Applications*, 108(1):109–129, 2003.

[12] P. Briand and R. Elie. A simple constructive approach to quadratic BSDEs with or without delay. *Stochastic processes and their applications*, 123(8):2921–2939, 2013.

[13] P. Briand and Y. Hu. BSDEs with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136(4):604–618, 2006.

[14] P. Briand and Y. Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probability Theory and Related Fields*, 141(3–4):543–567, 2008.

[15] P. Briand, J.-P. Lepeltier, and J. San Martin. One-dimensional backward stochastic differential equations whose coefficient is monotonic in y and non-lipschitz in z. *Bernoulli*, 13(1):80–91, 2007.

[16] J.-F. Chassagneux. *An introduction to the numerical approximation of BSDEs.* March 2012. 2nd Summer School of the Euro-Mediterranean Research Center for Mathematics and its Applications (EMRCMA).

[17] J.-F. Chassagneux. Linear multi-step schemes for BSDEs. *arXiv:1306.5548v1*, 2013.

[18] J.-F. Chassagneux and D. Crisan. Runge-Kutta schemes for BSDEs. *Forthcoming in Annals of Applied Probability*, 2012.

[19] J.-F. Chassagneux and A. Richou. Numerical simulation of quadratic BSDEs. *arXiv:1307.5741*, 2013.

[20] S. Cohen and R. Elliott. Existence uniqueness and comparison for BSDEs in general spaces. *Annals of Probability*, 40(5):2264–2297, 2012.

[21] D. Crisan and K. Manolarakis. Probabilistic methods for semilinear partial differential equations. Applications to finance. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(05):1107–1133, 2010.

[22] D. Crisan and K. Manolarakis. Second order discretization of Backward SDEs. *arXiv:1012.5650v1*, Dec. 2010.

[23] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method: application to nonlinear pricing. *SIAM J. Financial Math.*, 3:534–571, 2012.

[24] J. Cvitanic and I. Karatzas. Backward stochastic differential equations with reflection and Dynkin games. *Annals of Probability*, 24(4):2024–2056, 1996.

[25] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *Annales de l'Institut Henri Poincaré B, Probability and Statistics*, 47(2):559–574, 2011.

[26] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions : the critical case. *arXiv:1303.4859*, 2013.

[27] F. Delbaen and S. Tang. Harmonic analysis of stochastic equations and backward stochastic differential equations. *Probability Theory and Related Fields*, 146(1–2):291–336, 2010.

[28] G. dos Reis, A. Réveillac, and J. Zhang. FBSDE with time delayed generators: L^p-solutions, differentiability, representation formulas and path regularity. *Stochastic Processes and their Applications*, 121:2114–2150, 2011.

[29] N. El Karoui and S. Huang. A general result of existence and uniqueness of backward stochastic differential equations. In El Karoui and Mazliak, editors, *Backward stochastic differential equations (Paris, 1995–1996)*, volume 364 of *Pitman Research Notes in Mathematics Series*, pages 27–36. Longman, 1997.

[30] 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. *Annals of Probability*, 25(2):702–737, 1997.

[31] N. El Karoui, S. Peng, and M.-C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

[32] D. J. Estep, M. G. Larson, and R. D. Williams. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.*, 146(696):viii+109, 2000.

[33] C. Frei. Splitting multidimensional BSDEs and finding local equilibria. *http://www.math.ualberta.ca/ cfrei/*, 2013.

[34] C. Frei and G. dos Reis. A financial market with interacting investors: does an equilibrium exist ? *Mathematics and Financial Economics*, 4(3):161–182, 2011.

[35] E. Gobet and P. Turkedjiev. Approximation of discrete BSDE using least-squares regression. 2011. hal-00642685.

[36] D. Guo, H. Leng, and Q. Zhang. BSDEs with polynomial growth generators in a defaultable market. *Journal of Mathematical Analysis and Applications*, (0):–, 2013.

[37] S. Hamadène, J.-P. Lepeltier, and A. Matoussi. Double barrier reflected backward SDE's with continuous coefficient. In El Karoui and Mazliak, editors, *Backward stochastic differential equations (Paris, 1995–1996)*, volume 364 of *Pitman Research Notes in Mathematics Series*, pages 161–175. Longman, 1997.

[38] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, 1981.

[39] Y. Hu. Probabilistic interpretation of a system of quasilinear elliptic partial differential equations under Neumann boundary conditions. *Stochastic processes and their applications*, 48(1):107–121, 1993.

[40] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *Annals Of Applied Probability*, 15(3):1691–1712, 2005.

[41] Y. Hu and Z. Qian. BMO martingales and positive solutions of heat equations. *arXiv:1201.5454*, 2012.

[42] Y. Hu, Z. Qian, and Z. Zhang. Gradient estimates for porous medium and fast diffusion equations via FBSDE approach. *1206.1394*, 2012.

[43] M. Hutzenthaler, A. Jentzen, and P. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 467(2130):1563–1576, 2011.

[44] P. Imkeller. *Malliavin's calculus and applications in stochastic control and finance*, volume 1 of *IMPAN Lecture Notes*. Polish Academy of Sciences Institute of Mathematics, 2008.

[45] P. Imkeller and G. dos Reis. Corrigendum to "Path regularity and explicit convergence rate for BSDE with truncated quadratic growth". *Stochastic Processes and their Applications*, 120(11):2286–2288, 2010.

[46] P. Imkeller and G. dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Processes and their Applications*, 120(3):348–379, 2010.

[47] N. Kazamaki. *Continuous exponential martingales and BMO*. Springer, 1994.

[48] N. Kazi-Tani, D. Possamai, and C. Zhou. Quadratic BSDEs with jumps and related non-linear expectations: a fixed-point approach. *arXiv:1208.5581*, 2012.

[49] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, 1992.

[50] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28(2):558–602, 2000.

[51] M. Kobylanski, J.-P. Lepeltier, M.-C. Quenez, and S. Torrès. Reflected BSDE with superlinear quadratic coefficient. *Probability and Mathematical Statistics*, 22(1):51–83, 2002.

[52] B. Kovács. Semilinear parabolic problems. Master's thesis, Eötvös Loránd University of Budapest, 2011.

[53] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type (Translated from the Russian by S. Smith.)*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, R.I., 1967.

[54] J.-P. Lepeltier, A. Matoussi, and M. Xu. Reflected backward stochastic differential equations under monotonicity and general increasing growth conditions. *Advances in Applied Probability*, 37(1):134–159, 2005.

[55] J.-P. Lepeltier and J. San Martin. Existence for BSDEs with superlinear-quadratic coefficient. *Stochastics and Stochastic Reports*, 63(3–4):227–240, 1998.

[56] J.-P. Lepeltier and M. Xu. Reflected BSDEs with quadratic growth and unbounded terminal value. *arXiv:0711.0619*, 2007.

[57] J. Ma and J. Zhang. Path regularity for solutions of backward stochastic differential equations. *Probability Theory and Related Fields*, 122(2):163–190, 2002.

[58] J. Ma and J. Zhang. Representation theorems for backward stochastic differential equations. *Annals of Applied Probability*, 12(4):1390–1418, 2002.

[59] X. Mao and L. Szpruch. Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with superlinear diffusion coefficients. *Stochastics, An International Journal of Probability and Stochastic Processes*, 85(1):144–171, 2013.

[60] A. Matoussi and M. Xu. Sobolev solution for semilinear PDE with obstacle under monotonicity condition. *Electronic Journal of Probability*, 13(35):1035–1067, 2008.

[61] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics. Scientific Computation.* Springer-Verlag, Berlin, 2004.

[62] M.-A. Morlais. Quadratic BSDEs driven by continuous martingales and applications to the utility maximization problem. *Finance and Stochastics*, 13(1):121–150, 2009.

[63] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, 2006.

[64] E. Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. In F. Clarke and R. Stern, editors, *Nonlinear analysis, differential equations and control (Montreal, QC, 1998)*, volume 528 of *NATO Sci. Ser. C Math. Phys. Sci.*, pages 505–549. Kluwer Acad. Publ., 1999.

[65] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[66] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In B. Rozovskii and R. Sowers, editors, *Stochastic partial differential equations and their applications*, volume 176 of *Lecture notes in CIS*, pages 200–217. Springer-Verlag, 1992.

[67] E. Pardoux and S. Peng. Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probability Theory and Related Fields*, 98(2):209–227, 1994.

[68] E. Pardoux and S. Zhang. Generalized BSDEs and nonlinear neumann boundary value problems. *Probability Theory And Related Fields*, 110(4):535–558, 1998.

[69] S. Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports*, 37:61–74, 1991.

[70] S. Peng. A generalized dynamic programming principle and Hamilton–Jacobi-Bellman equation. *Stochastics and Stochastic Reports*, 38(2):119–134, 1992.

[71] S. Peng and M. Xu. The smallest g-supermartingale and reflected BSDEs with single and double L^2 obstacles. *Annales de l'Institut Henri Poincaré B, Probability and Statistics*, 41(3):605–630, 2005.

[72] A. Richou. Numerical simulation of BSDEs with drivers of quadratic growth. *Annals of Applied Probability*, 21(5):1933–1964, 2011.

[73] A. Richou. Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition. *Stochastic Processes and their Applications*, 122(9):3173–3208, 2012.

[74] F. Rothe. *Global solutions of reaction-diffusion systems*, volume 1072 of *Lecture Notes in Mathematics*. Springer-Verlag, 1984.

[75] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.

[76] E. Süli and D. Mayers. *An introduction to numerical analysis.* Cambridge University Press, 2003.

[77] R. Tevzadze. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Processes and their Applications*, 118(3):503–515, 2008.

[78] N. Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDEs*. Fields Institute Monograph, Springer, 2013.

[79] M. Xu. Backward stochastic differential equations with reflection and weak assumptions on the coefficients. *Stochastic processes and their applications*, 118(6):968–980, 2008.

[80] E. Zeidler. *Nonlinear functional analysis and its applications. II/B*. Springer-Verlag, New York, 1990.

[81] J. Zhang. *Some fine properties of backward stochastic differential equations.* PhD thesis, Purdue University (USA), 2001.

[82] J. Zhang. A numerical scheme for BSDEs. *Annals of Applied Probability*, 14(1):459–488, 2004.

[83] Q. Zhang and H. Zhao. Probabilistic representation of weak solutions of partial differential equations with polynomial growth coefficients. *J. Theoret. Probab.*, 25(2):396–423, 2012.