Radon Series Comp. Appl. Math **8**, 1–34

Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs

Bruno Bouchard, Romuald Elie, and Nizar Touzi

Abstract. The aim of this paper is to provide a survey on recent advances on probabilistic numerical methods for nonlinear PDEs, which serve as an alternative to classical deterministic schemes and allow to handle a large class of multidimensional nonlinear problems. These probabilistic schemes are based on the stochastic representation of semilinear PDEs by means of backward SDEs, which can be viewed as an extension of the well-known Feynman–Kac formula to the semilinear case. In this context, we first explain how smoothness properties can be obtained for non-reflected BSDEs set in the whole domain by using purely probabilistic techniques introduced in Ma and Zhang [59], and how they can be exploited to provide convergence rates of discrete time Euler-type approximation schemes. We then present some recent extensions to BSDEs with jumps, reflected BSDEs, BSDEs with horizon given by a finite stopping time, and second order BSDEs. The extension to fully nonlinear PDEs requires to enlarge the definition of backward SDEs. However, a natural probabilistic numerical scheme can be introduced formally by evaluating the nonlinear PDE along the trajectory of some chosen underlying diffusion. This point of view shows an intimate connection between the probabilistic numerical schemes reviewed in this paper and the standard finite differences methods. In particular, convergence and error estimates are provided by using the monotone schemes methods from the theory of viscosity solutions.

Key words. Backward stochastic differential equations, numerical resolution, conditional expectation, monte carlo methods.

AMS classification. 65C05, 60H35, 60G20

1 Introduction

The application field offered by the resolution of BSDEs is wide and diverse. As noticed by Bismut [12], Backward Stochastic Differential Equations (hereafter BSDEs) cover in particular the field of stochastic optimal control problems and the corresponding Hamilton–Jacobi–Bellman equations. See also [67], [72] for optimal control of diffusions with jumps, and [15] and [51] for optimal switching problems. Consequently, these techniques are of great interest for applications in economics. This is particularly true in finance, where their use was popularised by El Karoui, Peng and Quenez [48]. In incomplete markets, it can be used to solve optimal management problems by working along the lines presented in, e.g., El Kaouri, Peng and Quenez [49] for recursive utilities, El Karoui and Rouge [50] for indifference pricing, see also the more recent

Third author: This research is part of the Chair *Financial Risks* of the *Risk Foundation* sponsored by Société Générale, the Chair *Derivatives of the Future* sponsored by the Fédération Bancaire Française, the Chair *Finance and Sustainable Development* sponsored by EDF and Calyon.

paper [46] and the references therein, and Peng [68] for risk measures. Adding jumps to financial assets dynamics can be done as in [10] or [36]. BSDEs with jumps are also used for the valuation of financial derivatives with default risk. The valuation of American or game options follows from the resolution of reflected BSDEs, see [47]. The list of potential applications is long.

More generally, it is well known that the so-called Feynman Kac representation extends to second order semilinear PDEs which can be interpreted in terms of decoupled Forward-Backward SDEs, see Pardoux and Peng [63, 64] and Pardoux [62]. The term “decoupled” refers to the fact that the dynamic of the forward process does not depend on the solution of the Backward SDE. Hence, solving a BSDE or the related semilinear PDE is essentially the same. It was actually first proposed by Ma, Protter and Yong [58], see also [34], to use an estimation of the solution of the semilinear PDE to provide an approximation of the solution of Forward-Backward SDEs. However, solving PDEs in high dimension by finite-difference or finite-element schemes is a challenging task and any deterministic procedure suffers from the so-called “curse of dimensionality” which, very quickly, prevents from a good approximation of the solution of the PDE.

As a powerful alternative to classical purely deterministic techniques, more probabilistic approaches, directly based on the resolution of BSDEs, have been introduced these last years. We can essentially classify them in two categories. The first one concerns the approximation of the driving Brownian motion by a random walk on a finite tree. Such techniques have been widely studied, see e.g. [1], [28], [29] or [57], however they can be efficient only in low dimension since the size of the tree explodes exponentially with the dimension. The second one is based on Monte-Carlo type techniques. The main idea consists in simulating the Forward process by using standard Euler schemes and then computing backward the solution of the BSDE associated to the simulated forward paths by: 1. Discretising in time the BSDE and writing an Euler-type backward scheme which involves conditional expectations, 2. Approximating the conditional expectations by using non-parametric regression methods, see [22], [56], [27], [43] or integration by parts based estimators, see [55], [17] and [20]. The main advantage of such methods is that, like other Monte-Carlo type methods, the convergence rate does not depend a-priori on the dimension of the problem and they should therefore less suffer from the curse of dimensionality. In practice, the variance and/or the convergence rate of the estimators of the conditional expectations usually depends on the complexity/dimension of the problem, and the above comment has to be tempered. There is actually a third class of methods which lies in between the two previous ones and which were developed in a series of papers by Bally, Pagès and Printems, see e.g. [3], [4] and the references therein. In this approach, the original forward process is replaced by a quantised version in discrete time, i.e. a projection on a finite grid, whose associated Backward SDE can be computed by backward induction. The difference with the random walk technique is that the grid is computed in some optimal way, which allows to consider high dimension problems with a limited number of points, see the above quoted papers.

We shall not detail the pure numerical parts in this paper but rather insists on the first step of all these algorithms: the discrete time approximation of the BSDE which leads to a Backward Euler-type scheme.

Not surprisingly, the discrete time approximation error depends heavily on the regularity of the solution itself. The good news is that it essentially only depends on it. Therefore, all the analysis breaks down to that of the regularity of the solution of the BSDE. This has been first observed by Zhang [73, 74] and Bouchard and Touzi [20], who exploited the regularity results of Ma and Zhang [59] to provide a $|\pi|^{1/2}$ converge rate for decoupled Forward-Backward SDEs with Lipschitz coefficients, for a time-grid modulus $|\pi|$. Various extensions have been then developed in the literature and are reviewed in this paper: reflected BSDEs, the presence of a jump component, and the corresponding Dirichlet problem.

More recently, the backward schemes suggested by the discretisation of BSDEs have also been extended to the case of fully nonlinear PDEs by Fahim, Touzi and Warin [37]. The connection with an extended class of BSDEs, namely second order BSDEs, was outlined in [25]. The last section of this paper provides a presentation of the probabilistic numerical schemes which avoids BSDEs and only requires to evaluate the nonlinear PDE along the trajectory of some chosen underlying diffusion. In particular, this point of view highlights the relationship of the probabilistic schemes studied in this paper with the corresponding standard deterministic schemes. The convergence results and the corresponding error estimates in the fully nonlinear case are reviewed in the final section of this paper, and are based on the monotone schemes method initiated within the theory of viscosity solutions by Barles and Souganidis [9], Krylov [52, 53, 54], and Barles and Jakobsen [6, 7, 8].

To conclude, we should note that the case of coupled Forward-Backward SDEs has been considered in e.g. [31], [32], see also the references therein. The nature of the analysis being very different from the ones carried out here, we shall not discuss this case here (although it is important). See also [11] for Picard type iterations techniques.

Notations: Any element of $\mathbb{R}^d$ is viewed as a column vector and we denote by $'$ the transposition. $\mathbb{M}^d$ (resp. $\mathcal{S}_d$) denotes the set of $d \times d$ (resp. symmetric) matrices. For a smooth function $(t,x) \mapsto \varphi(t,x)$, we denote by $\partial_t \varphi$ its first derivative with respect to t, and by $D\varphi$ and $D^2\varphi$ its Jacobian and Hessian matrix with respect to x. It the function depends on more than one space variable, we use D_x, D_y, etc ... to denote the partial Jacobian with respect to a specific argument.

2 Numerical approximation of decoupled forward backward SDE

Let us first consider the simplest case of decoupled Forward-Backward SDE on the whole domain:

$$X_t = x + \int_0^t b(X_s)ds - \int_0^t \sigma(X_s)dW_s \tag{2.1}$$

$$Y_t = g(X_1) + \int_t^1 h(X_s, Y_s, Z_s)ds - \int_t^1 Z_s' dW_s, \quad t \le 1 \tag{2.2}$$

where W is d-dimensional Brownian Motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose associated complete and right-continuous filtration is denoted by $\mathbb{F} = (\mathcal{F}_t)_{t\leq 1}$. Here, $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{M}^d$, $g : \mathbb{R}^d \to \mathbb{R}$ and $h : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are some Lipschitz continuous functions.

2.1 Discrete time approximation

The discrete-time approximation is constructed as follows.

Euler scheme for the Forward process. Given a time-grid $\pi := \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ of $[0, 1]$ with modulus $|\pi|$, we first approximate the Forward process X by its Euler scheme X^π defined as

$$X_0^\pi := x , \quad \text{and} \quad X_{t_i+1}^\pi := b(X_{t_i}^\pi)\Delta t_i + \sigma(X_{t_i}^\pi)\Delta W_{t_i} \;\; \text{for} \;\; i < n , \tag{2.3}$$

where $\Delta t_i := t_{i+1} - t_i$ and $\Delta W_{t_i} := W_{t_{i+1}} - W_{t_i}$.

Euler scheme for the Backward process. As for the solution (Y, Z) of the BSDE, we first approximate the terminal condition $Y_1 = g(X_1)$ by simply replacing X by its Euler scheme : $Y_1 \simeq g(X_1^\pi)$. Then, motivated by the formal Euler discretisation:

$$\begin{aligned} Y_{t_i} &= Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} h(X_s, Y_s, Z_s)ds - \int_{t_i}^{t_{i+1}} Z_s' dW_s \\ {}''\simeq'' &\;\; Y_{t_{i+1}} + h(X_{t_i}^\pi, Y_{t_i}, Z_{t_i})\Delta t_i - Z_{t_i}'\Delta W_{t_i} , \end{aligned}$$

we define the discrete-time approximation of the BSDE as follows:
1. Taking expectation conditionally to $\mathcal{F}_{t_i}$ on both sides leads to

$$Y_{t_i} \simeq \mathbb{E}\left[Y_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + h(X_{t_i}^\pi, Y_{t_i}, Z_{t_i})\Delta t_i .$$

2. Pre-multiplying by ΔW_{t_i} and then taking conditional provides

$$0 \simeq \mathbb{E}\left[Y_{t_{i+1}}\Delta W_{t_i} \mid \mathcal{F}_{t_i}\right] - Z_{t_i}\Delta t_i .$$

This formal approximation argument leads to a Backward Euler scheme $(Y_{t_i}^\pi, \bar{Z}_{t_i}^\pi)$ of the form

$$\begin{cases} \bar{Z}_{t_i}^\pi &= \frac{1}{\Delta t_i}\,\mathbb{E}\left[Y_{t_{i+1}}^\pi \Delta W_{t_i} \mid \mathcal{F}_{t_i}\right] \\ Y_{t_i}^\pi &= \mathbb{E}\left[Y_{t_{i+1}}^\pi \mid \mathcal{F}_{t_i}\right] + h\left(X_{t_i}^\pi, Y_{t_i}^\pi, \bar{Z}_{t_i}^\pi\right)\Delta t_i \end{cases} \tag{2.4}$$

with terminal condition given by $Y_1^\pi = g(X_1^\pi)$. The reason why we denote the Z-component by $\bar{Z}^\pi$, and not simply Z^π, will become clear in Remark 2.1.2 below.

Remark 2.1.1 (Explicit or implicit schemes). Note that the above scheme is implicit as $Y_{t_i}^\pi$ appears in both sides of the equation. Since h is assumed to be Lipschitz and since it is multiplied by Δt_i, which is intended to be small, the equation can however be solved numerically very quickly by standard fixed point methods. Nevertheless, we could also consider an explicit version of this scheme by replacing the second equation in (2.4) by

$$Y^\pi_{t_i} = \mathbb{E}\left[Y^\pi_{t_{i+1}} + h\left(X^\pi_{t_i}, Y^\pi_{t_{i+1}}, \bar{Z}^\pi_{t_i}\right)\Delta t_i \mid \mathcal{F}_{t_i}\right] .$$

This will not change the convergence rate, see [18]. Both can thus be used indifferently depending on the exact nature of the coefficients.

Remark 2.1.2 (Continuous time version of the Euler scheme). Note that the martingale representation theorem implies that we can find a square integrable process Z^π such that

$$Y^\pi_{t_i} = Y^\pi_{t_{i+1}} + h\left(X^\pi_{t_i}, Y^\pi_{t_i}, \bar{Z}^\pi_{t_i}\right)\Delta t_i - \int_{t_i}^{t_{i+1}} (Z^\pi_s)' dW_s .$$

This allows to consider a continuous version of the Backward Euler scheme by setting

$$Y^\pi_t := Y^\pi_{t_{i+1}} + \int_t^{t_{i+1}} h\left(X^\pi_{t_i}, Y^\pi_{t_i}, \bar{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} (Z^\pi_s)' dW_s \quad \text{for } t \in [t_i, t_{i+1}) . \tag{2.5}$$

It then follows from the Itô isometry that

$$\bar{Z}^\pi_{t_i} = \frac{1}{\Delta t_i}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z^\pi_s ds \mid \mathcal{F}_{t_i}\right] . \tag{2.6}$$

In other words, the step constant process $\bar{Z}^\pi$, defined as $\bar{Z}^\pi_t := \bar{Z}^\pi_{t_i}$ for $t \in [t_i, t_{i+1})$, can be seen as the best $L^2(\Omega \times [0,1], d\mathbb{P} \otimes dt)$ approximation of Z^π in the class of adapted processes which are constant on each interval $[t_i, t_{i+1})$.

2.2 Error decomposition

In order to compare (Y, Z) and (Y^π, Z^π), let us introduce

$$\bar{Z}_{t_i} := \frac{1}{\Delta t_i}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z_s ds \mid \mathcal{F}_{t_i}\right]$$

and observe that, since h is Lipschitz continuous,

$$\begin{aligned} -dY_t &= h(X_t, Y_t, Z_t)dt - Z_t' dW_t \\ &\simeq h(X^\pi_t, Y_{t_i}, \bar{Z}_{t_i})dt - Z_t' dW_t \quad \text{for } t \in [t_i, t_{i+1}) \end{aligned} \tag{2.7}$$

up to an error term of order $C\left(|X_t - X^\pi_{t_i}| + |Y_t - Y_{t_i}| + |Z_t - \bar{Z}_{t_i}|\right) dt$ for some constant $C > 0$. Next observe from the definition of $\bar{Z}_{t_i}$ and $\bar{Z}^\pi_{t_i}$ and Jensen's inequality that

$$\mathbb{E}\left[|\bar{Z}_{t_i} - \bar{Z}^\pi_{t_i}|^2\right] \le \mathbb{E}\left[\frac{1}{\Delta t_i}\int_{t_i}^{t_{i+1}} |Z_t - Z^\pi_t|^2 dt\right] . \tag{2.8}$$

It then follows from (2.5)–(2.7), recalling the involved error terms, and standard techniques for BSDEs with Lipschitz coefficients that

$$\max_{0\le i<n} \mathbb{E}\left[|Y_{t_i}-Y^\pi_{t_i}|^2+\int_{t_i}^{t_{i+1}}|Z_t-Z^\pi_t|^2dt\right]$$
$$\le\ C\,\mathbb{E}\left[\sup_{t_i\le t\le t_{i+1}}|X_t-X^\pi_{t_i}|^2+\sup_{t_i\le t\le t_{i+1}}|Y_t-Y_{t_i}|^2+\int_{t_i}^{t_{i+1}}|Z_t-\bar Z_{t_i}|^2dt\right]$$

for some $C>0$. Using (2.8) again and standard estimates for Forward Euler schemes, it follows that

$$\begin{aligned}
\mathrm{Err}(\pi)^2 &:= \max_{0\le i<n}\mathbb{E}\left[\sup_{t_i\le t\le t_{i+1}}|Y_t-Y^\pi_{t_i}|^2\right]+\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}|Z_t-\bar Z^\pi_{t_i}|^2dt\right]\\
&\le 2\max_{0\le i<n}\mathbb{E}\left[\sup_{t_i\le t\le t_{i+1}}\left(|Y_{t_i}-Y^\pi_{t_i}|^2+|Y_t-Y_{t_i}|^2\right)\right]\\
&\quad +2\left(\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}|Z_t-\bar Z_{t_i}|^2+|\bar Z_{t_i}-\bar Z^\pi_{t_i}|^2dt\right]\right)\\
&\le C\ \left(|\pi|+\mathcal{R}^2_Y(\pi)+\mathcal{R}^2_Z(\pi)\right)
\end{aligned}$$

for some $C>0$, where

$$\begin{aligned}
\mathcal{R}^2_Y(\pi) &:= \max_{0\le i<n}\mathbb{E}\left[\sup_{t_i\le t\le t_{i+1}}|Y_t-Y_{t_i}|^2\right]\quad\text{and}\\
\mathcal{R}^2_Z(\pi) &:= \mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}|Z_t-\bar Z_{t_i}|^2dt\right]
\end{aligned}$$

can be seen as (squared) modulus of regularity of Y and Z.

Proposition 2.2.1 ([20], [74]). *There is a constant $C>0$, which is independent on π such that:*

$$\mathrm{Err}(\pi)\quad\le\quad C\ \left(|\pi|+\mathcal{R}^2_Y(\pi)+\mathcal{R}^2_Z(\pi)\right)^{\frac12}\ .$$

Remark 2.2.2. 1. Note that, by continuity of the process Y, the quantity $\sup_{t_i\le t\le t_{i+1}}|Y_t-Y_{t_i}|^2$ converges $\mathbb{P}-$a.s. to 0 as $|\pi|\to 0$. Since the coefficients are Lipschitz, we also have $\mathbb{E}\left[\sup_{t\le 1}|Y_t|^2\right]<\infty$. It thus follows from dominated convergence that

$$\mathcal{R}^2_Y(\pi)\longrightarrow 0\ \text{ as }\ |\pi|\to 0\ .$$

2. Also note that the step constant process $\bar Z$, defined as $\bar Z_t:=\bar Z_{t_i}$ for $t\in[t_i,t_{i+1})$, can be seen as the best $L^2(\Omega\times[0,1],d\mathbb{P}\otimes dt)$ approximation of Z in the class of adapted processes which are constant on each interval $[t_i,t_{i+1})$. Since such processes can be used to approximate Z in $L^2(\Omega\times[0,1],d\mathbb{P}\otimes dt)$, it follows that

$$\mathcal{R}^2_Z(\pi)\longrightarrow 0\ \text{ as }\ |\pi|\to 0\ .$$

Combining Proposition 2.2.1 and the last remark, implies the first converge result.

Corollary 2.2.3. *We have:* $\text{Err}(\pi)^2 \longrightarrow 0$ *as* $|\pi| \to 0$.

Remark 2.2.4. It follows from the interpretation of $\bar{Z}$ in terms of the best approximation of Z by step constant processes that

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}^\pi_{t_i}|^2 dt\right] \geq \mathcal{R}^2_Z(\pi) \ .$$

In the case where $h \equiv 0$, then Y is a martingale and Y_{t_i} is therefore the best $L^2(\Omega, d\mathbb{P})$ approximation of Y_t for $t \in [t_i, t_{i+1}]$ by a $\mathcal{F}_{t_i}$-measurable random variable. The Burckolder–Davis–Gundy's inequality then implies that

$$\mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - Y^\pi_{t_i}|^2\right] \geq \ c\, \mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - Y_{t_i}|^2\right] \ ,$$

for some $c > 0$. In view of Proposition 2.2.1, this shows that

$$c\ \left(\mathcal{R}^2_Y(\pi) + \mathcal{R}^2_Z(\pi)\right) \leq \text{Err}(\pi)^2 \leq C\ \left(|\pi| + \mathcal{R}^2_Y(\pi) + \mathcal{R}^2_Z(\pi)\right) \ .$$

Up to a $|\pi|$ term this can actually be shown even if $h \not\equiv 0$. This result is not surprising. As usual, the convergence rate should depend on the regularity of the functions/processes to be approximated.

2.3 Path regularity and convergence rate

In view of Proposition 2.2.1, the whole analysis of the convergence rate of the discrete-time approximation breaks down to the analysis of the (squared) modulus of continuity $\mathcal{R}^2_Y(\pi)$ and $\mathcal{R}^2_Z(\pi)$. The regularity of Y is standard in our context where the coefficients are Lipschitz continuous.

Proposition 2.3.1. *There exists $C > 0$ independent on π such that $\mathcal{R}^2_Y(\pi) \leq C\ |\pi|$.*

Note that this implies some regularity on the (unique viscosity) solution v of the semilinear PDE

$$\begin{aligned} -\mathcal{L}v - h(\cdot, v, \sigma' Dv) &= 0 \ \text{ on } [0,1) \times \mathbb{R}^d \\ v(T-, \cdot) &= g \ \text{ on } \mathbb{R}^d \ , \end{aligned}$$

where $\mathcal{L}$ is the Dynkin operator associated to X,

$$\mathcal{L}v := \partial_t v + \langle b, Dv\rangle + \frac{1}{2}\text{Tr}\left[\sigma\sigma' D^2 v\right] \ ,$$

since it is related to (Y, Z) by the mapping $Y = v(\cdot, X)$ and (at least formally) $Z = \sigma(X)' Dv(\cdot, X)$, see e.g. [64]. As a matter of fact, it follows from our Lipschitz continuity assumption that this solution is Lipschitz continuous in space and $1/2$-Hölder in time. Since X is itself $1/2$-Hölder in time in L^2, this actually implies the regularity statement of Proposition 2.3.1.

It is however less standard to study the regularity of the Z-component. If Dv, whenever it is well-defined, were Lipschitz continuous in space and $1/2$-Hölder in time, we would easily deduce that $\mathcal{R}_Z^2(\pi)$ is of order $|\pi|$. But such a result is not true in general, and indeed much stronger than what we need.

A pure probabilistic analysis of the modulus of regularity of Z has been carried out by Ma and Zhang [59] and Zhang [74], actually in a more general setting. It is based on the following analysis ($d = 1$ to avoid notational complexities):

Step 1. The first step consists in assuming that g and σ are C_b^1 and using the fact that (Y, Z) are differentiable in the Malliavin sense and that there exists a version of their Malliavin derivative (DY, DZ) which solves

$$D_s Y_t = Dg(X_1) D_s X_1 + \int_t^1 Dh(\Theta_r) D_s(X_r, Y_r, Z_r) dr - \int_t^1 D_s Z_r dW_r \ , \ s \leq 1 \ ,$$

where $\Theta := (X, Y, Z)$, see e.g. [61]. Moreover, Itô's formula and uniqueness of solutions of BSDEs with Lipschitz coefficients imply that it is related to the solution $(\nabla Y, \nabla Z)$ of

$$\nabla Y_t = Dg(X_1) \nabla X_1 + \int_t^1 Dh(\Theta_r) \nabla(X_r, Y_r, Z_r) dr - \int_t^1 \nabla Z_r dW_r \ , \ s \leq 1 \ ,$$

where ∇X denotes the first variation process of X, by the relation $D_s(X, Y, Z) = \mathbf{1}_{[s,1]} \nabla(X, Y, Z)(\nabla X_s)^{-1} \sigma(X_s)$.

Step 2. On the other hand, the identity

$$Y_t = Y_0 - \int_0^t h(\Theta_r) dr + \int_0^t Z_r dW_r$$

implies that $D_t Y_t = Z_t$, up to passing to a suitable version of these processes. It thus follows that

$$Z_t = \left(Dg(X_1) \nabla X_1 + \int_t^1 Dh(\Theta_r) \nabla \Theta_r dr - \int_t^1 \nabla Z_r dW_r \right) (\nabla X_t)^{-1} \sigma(X_t) \ .$$

Setting

$$\eta := \nabla X \exp \left\{ \int_0^\cdot D_z f(\Theta_r) dW_r - \int_0^\cdot \left(\frac{1}{2} |D_z f(\Theta_r)|^2 + D_y f(\Theta_r) \right) dr \right\} \ , \qquad (2.9)$$

we then deduce from Itô's formula that

$$Z_t = \mathbb{E} \left[Dg(X_1) \eta_1 + \int_t^1 D_x h(\Theta_r) \eta_r dr \mid \mathcal{F}_t \right] \eta_t^{-1} \sigma(X_t) \ . \qquad (2.10)$$

Step 3. The rest of the proof is based on a martingale argument. For sake of simplicity, let us consider the case $Dh \equiv 0$, $\nabla X \equiv 1$ and σ is constant. Then

$$Z_t = \mathbb{E}\left[Dg(X_1) \mid \mathcal{F}_t\right] \sigma \ ,$$

and, for $t \in [t_i, t_{i+1}]$,

$$\mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] \le \mathbb{E}\left[|Z_{t_{i+1}}|^2 - |Z_{t_i}|^2\right]$$

so that, by 2. of Remark 2.2.2,

$$\mathcal{R}_Z^2(\pi) \le \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] dt \le |\pi| \mathbb{E}\left[|Z_1|^2\right] = |\pi| \mathbb{E}\left[|Dg(X_1)|^2\right] \ .$$

Also more technical, the general case is essentially treated similarly, by using the Lipschitz continuity of the coefficients. The above proof is performed for C_b^1 coefficients, but the final bound depends only on the Lipschitz constants. It can then be extended by standard approximation techniques.

Theorem 2.3.2 ([74]). *If the coefficients are Lipschitz, then* $\mathcal{R}_Z^2(\pi) = O(|\pi|)$.

Combining the last theorem with Propositions 2.2.1 and 2.3.1 provides a $|\pi|^{\frac{1}{2}}$ convergence rate for the Backward Euler scheme.

Corollary 2.3.3. *If the coefficients are Lipschitz, then* $\mathrm{Err}(\pi) \le O(|\pi|^{\frac{1}{2}})$.

Remark 2.3.4. Importantly, all these results are obtained under the single assumption that the coefficients are Lipschitz continuous. In particular, no ellipticity condition is imposed on σ.

Remark 2.3.5. The $|\pi|^{\frac{1}{2}}$ convergence rate is clearly optimal and is similar to the one obtained for Forward SDEs.

Remark 2.3.6. Since no ellipticity assumption is made on σ, one component of X can be chosen so as to coincide with time. It is therefore not a restriction to consider a-priori time homogeneous dynamics. Note however, that a similar analysis could be performed under the assumption that the b, σ and f are only $1/2$-Hölder in the time component, as opposed to Lipschitz continuous.

Remark 2.3.7. Similar results are obtained in [74] in the case where the terminal condition is of the form $f(X)$ where f is a function of the whole path of X on $[0, 1]$, under suitable Lipschitz continuity assumptions on the space of continuous paths. The case of non-regular terminal conditions has been discussed in [44]. The grid has then to be refined near the terminal time $t = 1$ in order to compensate for the explosion of the gradient Dv. However, their proof relies on PDE arguments and requires strong smoothness assumptions on the coefficients as well as a uniform ellipticity condition on σ, a condition that was not required so far.

2.4 Weak expansion in the smooth case

So far, we have only discussed the strong convergence rate. But, as for linear problems, which correspond to the case where h does not depend on Y and Z, we can also study the weak approximation error. Such an analysis as been performed recently by Gobet and Labart [42], thus extending previous results of Chevance [26] who considered the case where h does not depend on Z.

Their result is of the form:

$$
\begin{aligned}
Y_t - Y_t^\pi &= Dv(t,X_t)(X_t - X_t^\pi) + O(|\pi|) + O\left(|X_t - X_t^\pi|^2\right)\\
Z_t - \bar{Z}_t^\pi &= D\left[Dv(t,X_t)\sigma\right]'(X_t - X_t^\pi) + O(|\pi|) + O\left(|X_t - X_t^\pi|^2\right) \ ,
\end{aligned}
$$

where $(X_t^\pi)_{t\le 1}$ is the continuous version of the Euler scheme of X and (Y^π, Z^π) are defined consistently with the explicit Backward Euler scheme by

$$
\begin{aligned}
Y_t^\pi &= \mathbb{E}\left[Y_{t_{i+1}}^\pi + h\left(X_{t_i}^\pi, Y_{t_{i+1}}^\pi, \bar{Z}_{t_i}^\pi\right)(t_{i+1}-t) \mid \mathcal{F}_t\right]\\
\bar{Z}_t^\pi &= \frac{1}{t_{i+1}-t}\,\mathbb{E}\left[Y_{t_{i+1}}^\pi (W_{t_{i+1}} - W_t) \mid \mathcal{F}_t\right] \ , \ t \in [t_i, t_{i+1}) \ .
\end{aligned}
$$

In particular, this implies that $v(0,X_0) - Y_0^\pi = Y_0 - Y_0^\pi = O(|\pi|)$. Moreover, when the grid is uniform, this allows to deduce from the weak convergence of the process $|\pi|^{-\frac{1}{2}}(X - X^\pi)$ that the process $|\pi|^{-\frac{1}{2}}(X - X^\pi, Y - Y^\pi, Z - \bar{Z}^\pi)$ weakly converges too.

As usual, this requires much stronger assumptions than only our previous Lipschitz continuity conditions, in particular they need to assume that σ is uniformly elliptic.

3 Extensions to other semilinear cases

The analysis of the previous Sections 2.2 and 2.3 consists in two parts:
1. Control the convergence rate of the discrete time approximation in terms of the modulus of regularity of Y and Z.
2. Study these modulus of regularity. Here, the difficult task consists in controlling $\mathcal{R}_Z^2(\pi)$. In the above classical case, it is obtained by applying martingale-type techniques to a suitable representation of Z.

In the following, we show how these techniques can be adapted to more general situations corresponding to BSDEs with jumps, reflected BSDEs or BSDEs with random time horizon.

3.1 BSDEs with jumps

The analysis of Chapter 2 can be extended in a very natural way to decoupled Forward Backward SDEs with jumps. Such equations appear naturally in optimal control for jump diffusion processes, see [72], hedging problems, see [36], or portfolio management, see [10], in finance. From the PDE point of view, they have two important applications.

Example 3.1.1 (Semilinear PDEs with integral term). *Let us consider the solution* (X, Y, Z, U) *of*

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r + \int_0^t \int_E \beta(X_{r-}, e)\bar{\mu}(de, dr) \;, \\ Y_t &= g(X_1) + \int_t^1 h\left(\Theta_r\right)dr - \int_t^1 Z_r' dW_r - \int_t^1 \int_E U_r(e)\bar{\mu}(de, dr) \end{aligned} \quad (3.1)$$

where $\bar{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt$ *is a compensated compound point measure, independent on* W*, with mark-space* $E \subset \mathbb{R}^\kappa$*, and* $\Theta := (X, Y, Z, \Gamma)$ *with*

$$\Gamma := \int_E \rho(e)U(e)\lambda(de)$$

for some bounded map $\rho : E \mapsto \mathbb{R}^\kappa$*. Then, in the case where all the coefficients are Lipschitz continuous, uniformly in the* e*-component,* (Y, Z, U) *can be associated to the (unique viscosity) solution* v *of*

$$-\mathcal{L}v - h\left(\cdot, v, \sigma' Dv, \mathcal{I}[v]\right) \;=\; 0 \quad , \quad v(1-, \cdot) \;=\; g\,, \quad (3.2)$$

where $\mathcal{L}$ *is the Dynkin operator associated to* X *and*

$$\mathcal{I}[v](t, x) := \int_E \left\{v(t, x + \beta(x, e)) - v(t, x)\right\} \rho(e)\, \lambda(de) \;,$$

though the mapping, $Y_t = v(t, X_t)$*,* $U_t(e) = v(t, X_{t-} + \beta(X_{t-}, e)) - v(t, X_{t-})$ *and* $Z_t = \sigma(X_{t-})' Dv(t, X_{t-})$*, whenever it is well defined, see [5] and [18].*

Example 3.1.2 (Systems of semilinear PDEs). *More interestingly, it was shown by [65] and [71] that BSDEs of the form*

$$\begin{aligned} X_t &= X_0 + \int_0^t b(M_r, X_r)dr + \int_0^t \sigma(M_r, X_r)dW_r \;, \\ M_t &= 1 + \left(\int_0^t \sum_{j=1}^{\kappa-1} j\mu(\{j\}, dt)\right) [\text{modulo } \kappa] \\ Y_t &= g(M_1, X_1) + \int_t^1 h\left(M_r, X_r, Y_r, Z_r, U_r\right)dr - \int_t^1 Z_r dW_r \\ &- \int_t^1 \int_E U_r(e)\bar{\mu}(de, dr) \end{aligned} \quad (3.3)$$

allow to provide a probabilistic representation for systems of semilinear PDEs. Here the mark space is finite $E = \{1, \ldots, \kappa - 1\}$ *and* $\lambda(de) \equiv \lambda \sum_{k=1}^{\kappa-1} \delta_k(e)$*, where* $\lambda > 0$

is a constant. Namely, given a family $(\tilde{b}_m, \tilde{\sigma}_m, \tilde{h}_m, \tilde{g}_m)_{m\leq\kappa}$ *of Lipschitz continuous functions, if we set*

$$(b(m,x), \sigma(m,x), g(m,x)) = (\tilde{b}_m(x), \tilde{\sigma}_m(x), \tilde{g}_m(x))$$

and

$$h(m,x,y,z,\gamma) = \tilde{h}_m\Big(x, \big[y+\gamma^{\kappa-m+1}, \ldots, y+\gamma^{\kappa}, \underbrace{y}_{m}, y+\gamma^1, \ldots\big], z\Big) - \lambda\sum_{j=1}^{\kappa}\gamma^j ,$$

then one can show that $Y_t = v^{M_t}(t, X_t)$ *where* $v = (v^1, \ldots, v^\kappa)$ *is the unique viscosity solution of the system*

$$-\mathcal{L}^m v^m - h_m(\cdot, v, Dv^m\sigma_m) = 0 \quad , \quad v^m(1,\cdot) = g_m(\cdot) ,$$

with $\mathcal{L}^m$ *denoting the Dynkin operator associated to the Forward SDE with drift* $\tilde{b}_m$ *and volatility* $\tilde{\sigma}_m$*.*

Obviously, the formulations (3.1) and (3.3) can be combined so as to provide systems of semilinear PDEs with integral term.

The Backward Euler scheme for BSDEs with jumps of the form (3.1) was introduced by Bouchard and Elie [18] and follows the ideas of Section 2.1. The formulation (3.3) can be treated similarly, see Elie [35] for more details.

First, we define the Euler scheme X^π of X as usual. The Euler scheme $(Y^\pi, Z^\pi, \Gamma^\pi)$ for (Y, Z, Γ) is then constructed as above by the Backward induction

$$\left\{\begin{array}{rcl} \bar{Z}^\pi_{t_i} &=& \frac{1}{\Delta t_i}\,\mathbb{E}\left[Y^\pi_{t_{i+1}}(W_{t_{i+1}} - W_{t_i}) \mid \mathcal{F}_{t_i}\right] \\ \bar{\Gamma}^\pi_{t_i} &=& \frac{1}{\Delta t_i}\,\mathbb{E}\left[Y^\pi_{t_{i+1}}\int_E \rho(e)\bar{\mu}(de, (t_i, t_{i+1}]) \mid \mathcal{F}_{t_i}\right] \\ Y^\pi_{t_i} &=& \mathbb{E}\left[Y^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + h\left(X^\pi_{t_i}, Y^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}\right)\Delta t_i \end{array}\right.$$

with terminal condition $Y^\pi_1 := g(X^\pi_1)$.

Remark 3.1.3. In the case where E is a finite set $\{e_1, \ldots, e_\kappa\}$, one can always choose ρ of the form $\rho^i(e_j) = \mathbf{1}_{i=j}/\lambda(\{e_i\})$ so that $\Gamma = (U(e_i))_{i\leq\kappa}$. The above procedure then allows to approximate U directly. It can obviously be extended to the approximation of any integrated version of $e \mapsto U(e)$ with respect to λ. We only concentrate here on Γ as defined above because this is the term which appears in h and its approximation is therefore necessary to approximate Y.

As in the no-jump case, rather standard BSDE techniques allow to show that the approximation error

$$\begin{aligned} \mathrm{Err}(\pi)^2 \;:=\; & \max_{i\leq n-1}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}|Y_t - Y^\pi_{t_i}|^2\right] + \mathbb{E}\left[\sum_{i\leq n-1}\int_{t_i}^{t_{i+1}}|Z_t - \bar{Z}^\pi_{t_i}|^2 dt\right] \\ & + \mathbb{E}\left[\sum_{i\leq n-1}\int_{t_i}^{t_{i+1}}|\Gamma_t - \bar{\Gamma}^\pi_{t_i}|^2 dt\right] , \end{aligned}$$

can be controlled in terms of $|\pi|$, $\mathcal{R}_Y^2(\pi)$, $\mathcal{R}_Z^2(\pi)$ and an additional term that takes jumps into account

$$\mathcal{R}_\Gamma^2(\pi) := \mathbb{E}\left[\sum_{i\leq n-1}\int_{t_i}^{t_{i+1}} |\Gamma_t - \bar{\Gamma}_{t_i}|^2 dt\right] \text{ with } \bar{\Gamma}_{t_i} := \frac{1}{\Delta t_i}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma_s ds \mid \mathcal{F}_{t_i}\right] .$$

Under our Lipschitz continuity assumptions, the $1/2$-Hölder continuity in time and Lipschitz continuity in space of the deterministic map v is rather standard. In view of the identities $Y_t = v(t, X_t)$ and $U_t(e) = v(t, X_{t-} + \beta(X_{t-}, e)) - v(t, X_{t-})$, see [18] for details, this readily implies that $\mathcal{R}_Y^2(\pi) + \mathcal{R}_\Gamma^2(\pi) \leq C|\pi|$ for some $C > 0$.

It thus remains to study the squared modulus of continuity $\mathcal{R}_Z^2(\pi)$ of Z. Following the approach of Zhang [74], it was shown by Bouchard and Elie [18] that $\mathcal{R}_Z^2(\pi)$ can be controlled by $|\pi|$ whenever one of the following assumptions holds:

$Ha)$ For all $e \in E$, the map $x \in \mathbb{R}^d \mapsto \beta(x, e)$ admits a Jacobian matrix $D\beta(x, e)$ such that

$$(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto a(x, \xi; e) := \xi'(D\beta(x, e) + I_d)\xi$$

satisfies, uniformly in $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d$

$$a(x, \xi; e) \;\geq\; |\xi|^2 K^{-1} \quad \text{or} \quad a(x, \xi; e) \;\leq\; -|\xi|^2 K^{-1} ,$$

for some $K > 0$.

$Hb)$ b, σ, $\beta(\cdot, e)$, h and g are C_b^1 with Lipschitz first derivatives, uniformly in $e \in E$.

Remark 3.1.4. The condition $Ha)$ ensures that the tangent process ∇X is invertible, with inverse satisfying suitable integrability conditions, which is necessary to reproduce the argument presented in Section 2.3, recall the definition of the term η. Under $Hb)$, it can be shown, see [35], that Dv is $1/2$-Hölder-continuous in time and Lipschitz-continuous in space. This allows to show that $\mathcal{R}_Z^2(\pi)$ can be controlled by $|\pi|$ by simply using the $1/2$-Hölder regularity of X in L^2.

This allows to provide a $|\pi|^{\frac{1}{2}}$ convergence rate.

Theorem 3.1.5 ([18]). *Let $Ha)$ or $Hb)$ hold, assume that λ is bounded and that the coefficients are Lipschitz, uniformly in the e-variable. Then,* $\mathrm{Err}(\pi) \leq O(|\pi|^{\frac{1}{2}})$.

Remark 3.1.6. So far, it seems that only the case of jump measures with finite activity has been treated. In [18], the compensatory measure is assumed to be bounded, but it can be generalised without much difficulties under suitable integrability conditions. Moreover, pure jump Lévy processes can be approximated by compound point measures by truncating the small jumps. It would therefore be natural to approximate a BSDE driven by a Lévy process with infinite activity by a sequence of BSDEs driven by compound point measures to which the above analysis would apply. This would introduce an additional error related to the approximation of the original random measure. We refer to [70] and the references therein for the approximation of Lévy processes.

3.2 Reflected BSDEs

Let us now turn to the approximation of the solution (Y, Z, A) of a reflected BSDE of the form

$$\begin{aligned} Y_t &= g(X_1) + \int_t^1 h(X_s, Y_s, Z_s)\, ds - \int_t^1 Z'_s dW_s + A_1 - A_t \\ Y_t &\geq f(X_t)\ ,\ \ 0 \leq t \leq 1\ , \end{aligned} \tag{3.4}$$

where A is a cadlag adapted non-decreasing process satisfying

$$\int_0^1 (Y_t - f(X_t))\, dA_t \;=\; 0\ .$$

In the case where $h = 0$ and $f = g$, Y is the Snell envelope of $g(X)$, which, in finance, corresponds to the super-hedging price of the American option with payoff g. More generally, the solution (Y, Z, A) of (3.4) is related to semilinear PDEs with free boundaries of the form

$$\min\left\{-\mathcal{L}v - h(\cdot, v, \sigma' Dv)\ ,\ v - f\right\} = 0\ ,\ v(1-, \cdot) = g \tag{3.5}$$

see [47].

For such BSDEs, it is natural to modify the Backward Euler scheme (2.4) as follows so as to take into account the reflection:

$$\begin{aligned} \bar{Z}^\pi_{t_i} &= (t_{i+1} - t_i)^{-1}\, \mathbb{E}\left[Y^\pi_{t_{i+1}} \Delta W_{t_i} \mid \mathcal{F}_{t_i}\right] \\ \tilde{Y}^\pi_{t_i} &= \mathbb{E}\left[Y^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + h(X^\pi_{t_i}, \tilde{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i})\Delta t_i \\ Y^\pi_{t_i} &= \mathcal{R}\left(t_i\ ,\ X^\pi_{t_i}\ ,\ \tilde{Y}^\pi_{t_i}\right)\ ,\ i \leq n-1\ , \end{aligned} \tag{3.6}$$

with the terminal condition $Y^\pi_1 = g(X^\pi_1)$ and where

$$\mathcal{R}(t, x, y) \;:=\; y + [f(x) - y]^+ \mathbf{1}_{\{t \in \Re \setminus \{0, T\}\}}\ ,\ \ (t, x, y) \in [0,1] \times \mathbb{R}^{d+1}$$

for a partition $\Re := \{0 = r_0 < r_1 < \ldots < r_\kappa = 1\} \supset \pi$.

Numerical procedures for such Backward Euler schemes have been first studied in Bally and Pagès [3] in the case where $\pi = \Re$, h does not depend on Z and $g = f$. The fact that h is independent on Z allows them to use the formulation of Y in terms of optimal stopping problem to study its discrete time approximation error. They retrieve a $|\pi|^{-\frac{1}{2}}$ control for the convergence rate when the coefficients are Lipschitz continuous and improve it to $|\pi|$ when g is semiconvex and X can be perfectly simulated on the time-grid. The error on Z is not discussed.

A more general analysis has then been performed by Ma and Zhang [60], who obtained a bound for the convergence rate of order $|\pi|^{-\frac{1}{4}}$ when f depends on Z whenever it is C^2_b and σ is uniformly elliptic. Note that the arguments leading to (2.10) cannot be used here, since this would involve taking the Malliavin derivative of A, which has no reason to be well-defined. Instead, they rely on a representation of Z of the form ($d = 1$ to avoid notational complexities):

$$Z_t = \mathbb{E}\left[g(X_1)N_1^t + \int_t^1 h\left(X_r, Y_r, Z_r\right) N_r^t dr + \int_t^1 N_r^t dA_r \mid \mathcal{F}_t\right] \sigma(X_t) \tag{3.7}$$

where

$$N_s^t := \frac{1}{s-t} \int_t^s \sigma(X_r)^{-1} \nabla X_r dW_r \left(\nabla X_t\right)^{-1} .$$

Their argument can be formally explained in three steps. First, they derive a representation of the Z-component associated to a discretely reflected BSDE of the form

$$Y_t^b = g(X_1) + \int_t^1 h(\Theta_r^b) dr - \int_t^1 (Z_r^b)' dW_r + A_1^b - A_t^b \tag{3.8}$$

where $\Theta^b := (X, Y^b, Z^b)$ and $A_t^b := \sum_{j=1}^{\kappa-1} \left[h(X_{r_j}) - Y_{r_j}^b\right]^+ \mathbf{1}_{r_j \leq t}$. The first representation is obtained by considering the Malliavin derivatives of (Y, Z) as in Section 2.3, see Step 1, on each interval $[r_j, r_{j+1})$

$$Z_t^b = D_t Y_t^b = D_t Y_{r_{j+1}-}^b + \int_t^{r_{j+1}} D_t h(\Theta_r^b) dr - \int_t^{r_{j+1}} D_t Z_r^b dW_r \ , \quad t \in [r_j, r_{j+1})$$

where $D_t Y_{r_{j+1}-}^b = D_t Y_{r_{j+1}}^b + D_t \left[h(X_{r_{j+1}}) - Y_{r_{j+1}}^b\right]^+$ since, by construction, $Y_{r_{j+1}}^b = Y_{r_{j+1}}^b + \left[h(X_{r_{j+1}}) - Y_{r_{j+1}}^b\right]^+$. This implies that, for $t < r_{j+1}$,

$$Z_t^b = D_t g(X_1) + \int_t^1 D_t h(\Theta_r^b) dr - \int_t^1 D_t Z_r^b dW_r + \sum_{k=j+1}^{\kappa} D_t \left[h(X_{r_k}) - Y_{r_k}^b\right]^+ \ .$$

The fact that the discretely reflected BSDE is differentiable in the Malliavin sense can be easily proved by using a backward induction argument and the fact that $z \mapsto z^+$ is a Lipschitz function. Second, they perform an integration by parts, in the Malliavin sense, in order to get rid of the Malliavin derivatives, and obtain a formulation of the form

$$Z_t^b \quad = \quad \mathbb{E}\left[g(X_1)N_1^t + \int_t^1 h(\Theta_r^b) N_r^t dr + \sum_{k=0}^{\kappa} N_{r_k}^t \Delta A_{r_k}^b \mathbf{1}_{t < r_k} \mid \mathcal{F}_t\right] \sigma(X_t)$$

for $t \in [r_j, r_{j+1})$, where $\Delta A_{r_k}^b := \left[h(X_{r_k}) - Y_{r_k}^b\right]^+$. Finally, they pass to the limit on the grid of reflection times so as to recover (3.7) for the continuously reflected BSDE.

This formulation is in itself very interesting and can be viewed as a generalisation of the representation obtained in [38] for the linear non reflected case. However, it requires a uniform ellipticity condition on σ. Moreover, the exploding behaviour of N_r^t as $r \to t$ is a major drawback in the regularity analysis of Z.

Another regularity analysis was carried out in Bouchard and Chassagneux [16]. The main difference is that they do not perform the integration by parts but rather provide a representation of Z^b in terms of the next reflection time:

$$Z_t^b = \mathbb{E}\Big[Dg(X_1)\eta_1^b \mathbf{1}_{\tau_j=1} + Df(X_{\tau_j})\eta_{\tau_j}^b \mathbf{1}_{\tau_j<1} + \int_t^{\tau_j} D_x h(\Theta_r^b)\eta_r^b dr \mid \mathcal{F}_t\Big](\eta_t^b)^{-1}\sigma(X_t) \ , \tag{3.9}$$

for $t \in [r_j, r_{j+1})$, where $\tau_j := \inf\{r \in \Re \mid r \geq r_{j+1},\ f(X_r) > Y_r^b\} \wedge 1$ and η^b is defined as in (2.9) above with (X, Y^b, Z^b) in place of (X, Y, Z). This is the analogue of (2.10) up to the stopping time τ_j.

This formulation does not require any uniform ellipticity assumption on σ and turn out to be more flexible.

Let us expose the argument in the simpler case where $d = 1$, $g \equiv f$, $h \equiv 0$, $\nabla X \equiv I_d$ and σ is constant. Set

$$V_t^j := \mathbb{E}\left[Df(X_{\tau_j}) \mid \mathcal{F}_t\right]\sigma \quad \text{so that } \ Z_t^b = V_t^j \ \text{ if } \ t \in [r_j, r_{j+1}) \ ,$$

and note that the martingale property of V^j implies that

$$\mathbb{E}\left[|V_t^j - V_{t_i}^j|^2\right] \quad \leq \quad \mathbb{E}\left[|V_{t_{i+1}}^j|^2 - |V_{t_i}^j|^2\right] \ \text{ for } \ t \in [t_i, t_{i+1}] \ .$$

Letting i_j be defined by $r_j = t_{i_j}$, recall that $\pi \subset \Re$, it follows that

$$|\pi|^{-1}\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Z_t^b - Z_{t_i}^b|^2 dt\right] \leq \Sigma := \sum_{j=0}^{\kappa-1}\sum_{k=i_j}^{i_{j+1}-1}\mathbb{E}\left[|V_{t_{k+1}}^j|^2 - |V_{t_k}^j|^2\right]$$

where, by direct computations,

$$\Sigma \quad = \quad \sum_{j=0}^{\kappa-1}\mathbb{E}\left[|V_{r_{j+1}}^j|^2 - |V_{r_j}^j|^2\right] \leq C + \sum_{j=1}^{\kappa-1}\mathbb{E}\left[|V_{r_j}^{j-1}|^2 - |V_{r_j}^j|^2\right]$$

and $C > 0$ denotes a generic constant independent on π and $\Re$. On the other hand, it follows from the Cauchy–Schwarz inequality that, if Df is bounded,

$$|V_{r_j}^{j-1}|^2 - |V_{r_j}^j|^2 \quad \leq \quad |V_{r_j}^j - V_{r_j}^{j-1}|\ |V_{r_j}^j + V_{r_j}^{j-1}| \leq C\,\mathbb{E}\left[\beta \mid \mathcal{F}_{r_j}\right]\ |V_{r_j}^j - V_{r_j}^{j-1}|$$

where $\beta := 1 + \sup_{t\leq 1}|X_t|$. Moreover, if $Df \in C_b^2$, then

$$V_{r_j}^j - V_{r_j}^{j-1} \quad = \quad \mathbb{E}\left[(Df(X_{\tau_j}) - Df(X_{\tau_{j-1}}) \mid \mathcal{F}_{r_j}\right]\sigma \leq C\,\mathbb{E}\left[(\tau_j - \tau_{j-1}) \mid \mathcal{F}_{r_j}\right]\sigma \ .$$

Combining the above inequalities, this implies

$$\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Z_t^b - Z_{t_i}^b|^2 dt\right] \ \leq \ C|\pi|\left(1 + \sum_{j=1}^{\kappa-1}\mathbb{E}\left[\mathbb{E}\left[\beta \mid \mathcal{F}_{r_j}\right]\mathbb{E}\left[\tau_j - \tau_{j-1} \mid \mathcal{F}_{r_j}\right]\right]\right)$$

which is of order of $|\pi|$ since the expectation on the right-hand side simply equals $\mathbb{E}\left[\beta(\tau_j - \tau_{j-1})\right]$.

The general case is treated along the same main argument: play with the difference of the next reflection times $\tau_j - \tau_{j-1}$. It allows to control the modulus of regularity $\mathcal{R}^2_{Z^b}(\pi)$ of Z^b, defined as $\mathcal{R}^2_Z(\pi)$ for Z^b is place of Z, as follows:

Theorem 3.2.1 ([16]). *Assume that all the coefficients are Lipschitz continuous. Then,*

$$\mathcal{R}^2_{Z^b}(\pi) \le \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z^b_t - Z^b_{t_i}|^2\right] dt \quad \le \quad C\,|\pi|\, \left(1 + \beta|\pi|^{-\frac{1}{2}} + \alpha(\Re)\right)\,,$$

where $\alpha(\Re) = |\Re|^{-\frac{1}{2}}$ and $\beta = 1$ under the assumption
$H_1)$: f is C^1_b with Lipschitz first derivative
and, $\alpha(\Re) = 1$ and $\beta = 0$ under the assumption
$H_2)$: f is C^2_b with Lipschitz second derivative and σ is C^1_b with Lipschitz first derivative.

This regularity property of Z^b then allows to provide a convergence rate for the Euler scheme $(Y^\pi, \bar{Z}^\pi)$ to (Y^b, Z^b).

Theorem 3.2.2 ([16]). *Assume that the coefficients are Lipschitz. Then,*

$$\begin{aligned} \max_{i<n} \mathbb{E}\left[\sup_{t\in(t_i,t_{i+1}]} |Y^\pi_{t_{i+1}} - Y^b_t|^2\right]^{\frac{1}{2}} &\le O\left(\alpha_Y(\Re)\,|\pi|^{\frac{1}{2}} + \beta\,|\pi|^{\frac{1}{4}}\right) \\ \sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\bar{Z}^\pi_{t_i} - Z^b_t|^2 dt\right] &\le O\left(\alpha_Z(\Re)\,|\pi| + \beta\,|\pi|^{\frac{1}{2}}\right) \end{aligned}$$

with $(\alpha_Y(\Re), \alpha_Z(\Re)) = (|\Re|^{-\frac{1}{4}}, |\Re|^{-1})$ and $\beta = 1$ under $H_1)$, and, $(\alpha_Y(\Re), \alpha_Z(\Re)) = (1, |\Re|^{-1})$ and $\beta = 0$ under $H_2)$.

They also show that the above result holds true with $\alpha_Z(\kappa) = |\Re|^{-\frac{1}{2}}$ under $H_1)$, and $\alpha_Z(\kappa) = 1$ under $H_2)$, when $(X^\pi_{t_i})_{i\le n}$ is replaced by $(X_{t_i})_{i\le n}$ in the Backward Euler scheme (3.6).

To obtain a convergence rate of the approximation error of Y, it then essentially suffices to consider the case $\Re = \pi$ and show that Y^b approximates Y at least at a rate $|\Re|^{\frac{1}{2}}$. The approximation of the Z-component is more involved and requires the introduction of another approximation scheme.

Theorem 3.2.3 ([16]). *Assume that the coefficients are Lipschitz. Then, there exists*

$C > 0$, *independent on* π, *such that*

$$\mathbb{E}\left[\max_{i\le n}\ |Y^\pi_{t_i} - Y_{t_i}|^2\right]^{\frac{1}{2}} + \max_{i<n}\ \mathbb{E}\left[\sup_{t\in(t_i,t_{i+1}]} |Y^\pi_{t_{i+1}} - Y_t|^2\right]^{\frac{1}{2}} = O(\alpha(\pi))$$

$$\mathcal{R}^2_Z(\pi) \le \sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\bar{Z}^\pi_{t_i} - Z_t|^2 dt\right] = O(|\pi|^{\frac{1}{2}})\ ,$$

where $\alpha(\pi) = |\pi|^{\frac{1}{4}}$ *under* H_1) *and* $\alpha(\pi) = |\pi|^{\frac{1}{2}}$ *under* H_2).

When H_2) holds and $(X^\pi_{t_i})_{i\le n}$ is replaced by $(X_{t_i})_{i\le n}$ in the Backward Euler scheme (3.6), the (squared) error on Z and the (squared) modulus of regularity $\mathcal{R}^2_Z(\pi)$ are shown to be controlled by $|\pi|$, which corresponds to a convergence rate of order $|\pi|^{\frac{1}{2}}$ as in the non-reflected case.

Remark 3.2.4. The representation (3.9) of Z^b is of own interest as it provides a natural Monte-Carlo estimator for the so-called "delta" of Bermudean options. Such a result was already known in finance for American options, under a uniform ellipticity condition, see [41].

Remark 3.2.5. Extensions to doubly reflected and multivariate reflected BSDEs have been considered in Chassagneux [23] and [24].

3.3 BSDEs with random time horizon

By BSDE with random time horizon, we mean the solution (Y, Z) of

$$Y_t \quad = \quad g(X_\tau) + \int_t^1 \mathbf{1}_{s<\tau} h(X_s, Y_s, Z_s) ds - \int_t^1 Z'_s dW_s\ ,\ t \le 1\ , \qquad (3.10)$$

where τ is the first exit time of $(t, X_t)_{t\le T}$ from a cylindrical domain $D = [0, 1) \times \mathcal{O}$.

The first step toward the definition of a Backward Euler scheme, consists in choosing a suitable approximation of the random time horizon τ. In Bouchard and Menozzi [19], the authors consider the first exit time (on the grid) $\bar\tau$ of the Euler scheme X^π

$$\bar\tau := \inf\{t \in \pi\ :\ (t, \bar{X}_t) \notin D\}\ .$$

They then reproduce the Backward scheme (2.4)

$$\begin{cases} \bar{Z}^\pi_{t_i} &= \frac{1}{\Delta t_i}\,\mathbb{E}\left[Y^\pi_{t_{i+1}}\Delta W_{t_i} \mid \mathcal{F}_{t_i}\right] \\ Y^\pi_{t_i} &= \mathbb{E}\left[Y^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + h\left(X^\pi_{t_i}, Y^\pi_{t_i}, \bar{Z}^\pi_{t_i}\right)\Delta t_i \end{cases}$$

with the terminal condition

$$Y^\pi_1 = g(X^\pi_{\bar\tau})\ .$$

Following the approach of [20] and [74], they show that the discrete-time approximation error can be controlled as:

$$\mathrm{Err}(\pi)^2 \;\le\; C\left(|\pi| + \mathcal{R}_Y^2(\pi) + \mathcal{R}_Z^2(\pi) + \mathbb{E}\left[\xi|\tau-\bar{\tau}| + \mathbf{1}_{\bar{\tau}<\tau}\int_{\bar{\tau}}^{\tau}|Z_s|^2 ds\right]\right) \tag{3.11}$$

where $C>0$ is independent of π and ξ is a positive random variable satisfying $\mathbb{E}\left[\xi^p\right] \le C'_p$ for all $p\ge 2$, for some C'_p independent of π.

Not surprisingly, the error now depends on an additional term which is related to the approximation error of τ by $\bar{\tau}$ which has been widely studied under the assumption that the coefficients are smooth enough and σ is uniformly elliptic. In particular, Gobet [39], and, Gobet and Menozzi [45] have derived an expansion of the form $\mathbb{E}\left[\tau-\bar{\tau}\right] = C|\pi|^{\frac{1}{2}} + o(|\pi|^{\frac{1}{2}})$.

A similar result can be obtained without uniform ellipticity, whenever $\mathcal{O}$ is a C^2 domain with a compact boundary satisfying a uniform non-characteristic condition.

Theorem 3.3.1 ([19]). *Assume that the coefficients of X are Lipschitz continuous and that $\mathcal{O}$ is a C^2 domain with a compact boundary satisfying a uniform non-characteristic condition. Then, for each $\varepsilon\in(0,1)$ and each positive random variable ξ satisfying $\mathbb{E}\left[\xi^p\right]<\infty$ for all $p\ge 1$,*

$$\mathbb{E}\left[\mathbb{E}\left[\xi\;|\tau-\bar{\tau}|\;|\;\mathcal{F}_{\tau_+\wedge\bar{\tau}}\right]^2\right] \quad\le\quad O_\varepsilon^\xi\left(|\pi|^{1-\varepsilon}\right)\;,$$

where $\tau_+ := \inf\{t\in\pi\;:\;\tau\le t\}$.

This shows in particular that, for each $\varepsilon\in(0,1/2)$,

$$\mathbb{E}\left[|\tau-\bar{\tau}|\right] \quad\le\quad O_\varepsilon\left(|\pi|^{\frac{1}{2}-\varepsilon}\right)$$

which, up to the ε term, is consistent with the result of Gobet and Menozzi [45]. Moreover, it is easily extended to the case where

$D1)$ $\mathcal{O}$ is an intersection of C^2 domains with compact boundary, and it satisfies a uniform non-characteristic condition outside the corners and a uniform ellipticity condition on a neighbourhood of the corners.

The next step toward a convergence rate consists in studying the modulus of regularity $\mathcal{R}_Y^2(\pi)$ and $\mathcal{R}_Z^2(\pi)$, and the last term $\mathbb{E}\left[\mathbf{1}_{\bar{\tau}<\tau}\int_{\bar{\tau}}^{\tau}|Z_s|^2 ds\right]$.

In contrast with the case where $\mathcal{O}=\mathbb{R}^d$, the uniform Lipschitz continuity of the (viscosity) solution v of the associated semilinear Cauchy Dirichlet problem

$$-\mathcal{L}v - h(\cdot,v,\sigma' Dv) = 0 \;\text{ on }\; D\;,\; v=g \;\text{ on }\; \partial_p D := ([0,1)\times\partial\mathcal{O})\cup\left(\{1\}\times\bar{\mathcal{O}}\right)\;, \tag{3.12}$$

is by no means obvious, and usually requires additional assumptions than the simple Lipschitz continuity of the coefficients. It is shown in [19], by adapting standard barrier techniques, when g is smooth enough:

$Hg)$ $g \in C_b^{1,2}([0,1] \times \mathbb{R}^d)$,

and the domain satisfies a uniform exterior sphere condition as well as a uniform truncated interior cone condition:

$D2)$ For all $x \in \partial\mathcal{O}$, there is $y(x) \in \mathcal{O}^c$, $r(x) \in [L^{-1}, L]$, $L > 0$, and $\delta(x) \in B(0,1)$ such that $\bar{B}(y(x), r(x)) \cap \bar{\mathcal{O}} = \{x\}$ and $\{x' \in B(x, L^{-1}) \ : \ \langle x' - x, \delta(x)\rangle \geq (1 - L^{-1})|x' - x|\} \subset \bar{\mathcal{O}}$.

Here $B(z,r)$ denotes the open ball of centre z and radius r, $\bar{B}(z,r)$ is its closure.

Theorem 3.3.2 ([19]). *Assume that all the coefficients are Lipschitz continuous and that the conditions $Hg)$, $D1)$ and $D2)$ hold. Then, v is uniformly $1/2$-Hölder continuous in time and Lipschitz continuous in space.*

In particular, this allows to show that $\mathcal{R}_Y^2(\pi)$ is controlled by $|\pi|$ and the last term $\mathbb{E}\left[\mathbf{1}_{\bar{\tau}<\tau}\int_{\bar{\tau}}^{\tau}|Z_s|^2 ds\right]$ is controlled by $\mathbb{E}\left[|\tau - \bar{\tau}|\right] \leq C^{\varepsilon}|\pi|^{\frac{1}{2}-\varepsilon}$, for $\varepsilon > 0$.

It thus remains to study the modulus of continuity of Z. As in the reflected case, it is not possible to apply the techniques leading to (2.10) since it would involve taking the Malliavin derivative of $g(X_\tau)$, which has no reason to be well-defined. Instead, we can remark that, if v is smooth enough, $Dv(\cdot, X)\eta + \int_0^{\cdot} D_x h\left(\Theta_s\right)\eta_s ds$ is a martingale on $[0,\tau]$, recall the definition of η in (2.9). This implies that ($d = 1$ for notational simplicity)

$$Z_t = Dv(t, X_t)\sigma(X_t) = \mathbb{E}\left[Dv(\tau, X_\tau)\eta_\tau + \int_t^{\tau} D_x h\left(\Theta_s\right)\eta_s ds \mid \mathcal{F}_t\right]\eta_t^{-1}\sigma(X_t) \ .$$

This formula is very close to (2.10) except that Dg is replaced by Dv. But, since Dv is bounded, one can actually repeat the martingale-type argument of Step 3 of Section 2.3. The smoothness of v can be obtained for smooth coefficients and domains, under a uniform ellipticity condition of σ, which can then be relaxed by standard approximation arguments.

Theorem 3.3.3 ([19]). *Assume that all the coefficients are Lipschitz continuous and that the conditions $Hg)$, $D1)$ and $D2)$ hold. Then,*

$$\mathcal{R}_Y^2(\pi) + \mathcal{R}_Z^2(\pi) \quad \leq \quad O(|\pi|) \ .$$

This provides the following $|\pi|^{\frac{1}{4}-\varepsilon}$ convergence rate.

Corollary 3.3.4 ([19]). *Assume that all the coefficients are Lipschitz continuous and that the conditions $Hg)$, $D1)$ and $D2)$ hold. Then, for all $\varepsilon \in (0,1)$,*

$$\mathrm{Err}(\pi) \quad \leq \quad O_\varepsilon\left(|\pi|^{\frac{1}{2}-\varepsilon}\right) \ .$$

As a matter of facts, the global error is driven by the approximation error of the exit time which propagates backward thanks to the Lipschitz continuity of v.

Remark 3.3.5. It is shown in [19] that one can actually achieve a convergence rate of order $|\pi|^{\frac{1}{2}-\varepsilon}$ if one computes the approximation error only up to $\bar{\tau} \wedge \tau$.

4 Extension to fully nonlinear PDEs

The discretisation of backward stochastic differential equations can be viewed as a probabilistic approach for the numerical resolution of semilinear partial differential equations in parabolic or elliptic form, and possibly with a nonlocal integral term. This restriction on the nonlinearity of the corresponding PDEs excludes the important general class of Hamilton–Jacobi–Bellman equations which appear naturally as the infinitesimal version of the dynamic programming principle for stochastic control problems.

An extension of the theory of backward SDEs to the second order case was initiated by [25] and is still in progress. An alternative approach for such an extension in the context of a specific form of the nonlinearity was also introduced by Peng in various papers, see e.g. [69], and is based on the new concept of G-Brownian motion.

We shall not elaborate more on this aspect. Instead, our objective is to concentrate on the probabilistic numerical implications as discussed in Section 5 of [25]. Following [37], we provide a natural presentation of the discrete-time approximation without appealing to the theory of backward stochastic differential equations. This presentation shows in particular the important connection between the discretisation schemes of this paper and the finite differences approximations of solutions of PDEs. Finally, we were not able to extend the methods of proofs outlined in the previous sections. The convergence results and the rates of convergence are derived by means of the PDE approach for monotonic schemes.

4.1 Discretisation

Consider the fully nonlinear parabolic PDE:

$$-\mathcal{L}v - F\left(\cdot, v, Dv, D^2 v\right) = 0, \quad \text{on } [0,1)\times\mathbb{R}^d, \tag{4.1}$$

$$v(1,\cdot) = g, \quad \text{on } \in\mathbb{R}^d, \tag{4.2}$$

where, given bounded and continuous maps b and σ from $\mathbb{R}_+\times\mathbb{R}^d$ to $\mathbb{M}^d$ and $\mathbb{R}^d$, the linear second order operator $\mathcal{L}$ is defined by

$$\mathcal{L}v := \frac{\partial v}{\partial t} + \langle b, Dv\rangle + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma' D^2 v\right],$$

and the nonlinearity is isolated in the map $F : (t,x,r,p,\gamma) \in \mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\times\mathcal{S}_d \longmapsto F(t,x,r,p,\gamma)\in\mathbb{R}$.

For a positive integer n, set $h := 1/n$, $t_i := ih$, $i = 1,\ldots,n$, and define:

$$\hat{X}_h^{t,x} := x + b(t,x)h + \sigma(t,x)(W_{t+h} - W_t), \tag{4.3}$$

which is the one-step Euler discretisation of the diffusion corresponding to the operator $\mathcal{L}$.

Assuming that the PDE (4.1) has a classical solution, it follows from Itô's formula that

$$\mathbb{E}_{t_i,x}\left[v\left(t_{i+1}, X_{t_{i+1}}\right)\right] = v\left(t_i, x\right) + \mathbb{E}_{t_i,x}\left[\int_{t_i}^{t_{i+1}} \mathcal{L}v(t, X_t)dt\right]$$

where we ignored the difficulties related to local martingale part, and $\mathbb{E}_{t_i,x} := \mathbb{E}[\cdot|X_{t_i} = x]$ denotes the expectation operator conditional on $\{X_{t_i} = x\}$. Since v solves the PDE (4.1), this provides

$$\mathbb{E}_{t_i,x}\left[v\left(t_{i+1}, X_{t_{i+1}}\right)\right] \quad = \quad v(t_i,x) - \mathbb{E}_{t_i,x}\left[\int_{t_i}^{t_{i+1}} F(\cdot, v, Dv, D^2v)(t, X_t)dt\right].$$

By approximating the Riemann integral, and replacing the process X by its Euler discretisation, this suggests the following approximation of the value function v

$$v^h(1,.) = g \quad \text{and} \quad v^h(t_i,x) = T_h[v^h](t_i,x) \tag{4.4}$$

where, for a function $\psi : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}$ with exponential growth, we denote:

$$\begin{aligned} T_h[\psi](t_i,x) &:= \mathbb{E}\left[\psi(t_{i+1}, \hat{X}_h^{t_i,x})\right] + hF\left(\cdot, \mathcal{D}_h\psi\right)(t_i,x), \\ \mathcal{D}_h\psi &:= \left(\mathcal{D}_h^0\psi, \mathcal{D}_h^1\psi, \mathcal{D}_h^2\psi\right), \end{aligned}$$

and for $k = 0, 1, 2$, it follows from an easy integration by parts argument that:

$$\mathcal{D}_h^k\psi(t_i,x) \quad := \quad \mathbb{E}[D^k\psi(t_{i+1}, \hat{X}_h^{t_i,x})] \ = \ \mathbb{E}[\psi(\hat{X}_h^{t_i,x})H_k^h(t_i,x)], \tag{4.5}$$

where

$$H_0^h = 1, \quad H_1^h = (\sigma')^{-1}\ \frac{W_h}{h}, \quad H_2^h = (\sigma')^{-1}\ \frac{W_hW_h' - h\mathbf{1}_d}{h^2}\ \sigma^{-1}. \tag{4.6}$$

Remark 4.1.1. In the semilinear case, i.e. F is independent of the argument γ, the numerical scheme coincides with the discretisation of BSDEs as studied in the previous sections of this paper.

Remark 4.1.2. Other choices can be made for the above integration by parts. For instance, we also have:

$$\mathcal{D}_2^h\varphi(t,x) \quad = \quad \mathbb{E}\left[\varphi(\hat{X}_h^{t,x})(\sigma')^{-1}\frac{W_{h/2}}{(h/2)}\frac{W'_{h/2}}{(h/2)}\sigma^{-1}\right],$$

which shows that the backward scheme (4.4) is very similar to the probabilistic numerical algorithm suggested in [25].

Remark 4.1.3. So far, the choice of the drift and the diffusion coefficients b and σ in the nonlinear PDE (4.1) is arbitrary, and was only used to fix the underlying diffusion X. Our convergence result will however place some restrictions on this choice.

Remark 4.1.4. The integration by parts in (4.5) is intimately related to finite differences for numerical schemes on deterministic grids. We formally justify this connection for $d = 1$, $b \equiv 0$, and $\sigma = 1$ for simplicity:

- Let $\{w_j, j \geq 1\}$ be independent r.v. distributed as $\frac{1}{2}\left(\delta_{\sqrt{h}} + \delta_{-\sqrt{h}}\right)$. Then, the binomial random walk approximation of the Brownian motion $\hat{W}_{t_k} := \sum_{j=1}^k w_j$, $t_k := kh$, $k \geq 1$, suggests the following approximation:

$$\mathcal{D}_h^1\psi(t,x) := \mathbb{E}\left[\psi(t+h, X_h^{t,x})H_1^h\right] \quad \approx \quad \frac{\psi(t, x+\sqrt{h}) - \psi(t, x-\sqrt{h})}{2\sqrt{h}},$$

which is the centred finite differences approximation of the gradient.

- Let $\{w_j, j \geq 1\}$ be independent r.v. distributed as $\frac{1}{6}\left(\delta_{\{\sqrt{3h}\}} + 4\delta_{\{0\}} + \delta_{\{-\sqrt{3h}\}}\right)$. Then, the trinomial random walk approximation of the Brownian motion $\hat{W}_{t_k} := \sum_{j=1}^k w_j$, $t_k := kh$, $k \geq 1$, suggests the following approximation:

$$\mathcal{D}_h^2\psi(t,x) := \mathbb{E}\left[\psi(t+h, X_h^{t,x})H_2^h\right] \approx \frac{\psi(t, x+\sqrt{3h}) - 2\psi(t,x) + \psi(t, x-\sqrt{3h})}{3h},$$

which is the centred finite differences approximation of the Hessian.

In view of the above interpretation, the numerical scheme studied in this paper can be viewed as a mixed Monte Carlo–Finite Differences algorithm. The Monte Carlo component of the scheme consists in the choice of an underlying diffusion process X. The finite differences component of the scheme consists in approximating the remaining nonlinearity by means of the integration-by-parts (4.5).

4.2 Convergence of the discretisation

Definition 4.2.1. We say that (4.1) has strong comparison of bounded solutions if for any bounded upper semicontinuous viscosity supersolution $\overline{v}$ and any bounded lower semicontinuous subsolution $\underline{v}$ on $[0,1) \times \mathbb{R}^d$, satisfying

- either $\overline{v}(1,\cdot) \geq \underline{v}(1,\cdot)$,
- or the viscosity property of $\overline{v}$ and $\underline{v}$ holds true on $[0,1] \times \mathbb{R}^d$,

we have $\overline{v} \geq \underline{v}$.

The strong comparison principle is an important notion in the theory of viscosity solutions which allows to handle situations where the boundary condition g is not compatible with the equation.

In the sequel, we denote by F_r, F_p and F_γ the partial gradients of F with respect to r, p and γ, respectively. We recall that any Lipschitz function is differentiable a.e.

Assumption F. *The nonlinearity F is Lipschitz-continuous with respect to (r,p,γ) uniformly in (t,x) and $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |F(t,x,0,0,0)| < \infty$. Moreover, F is uniformly elliptic and dominated by the diffusion of the linear operator $\mathcal{L}$, i.e.*

$$\varepsilon I_d \leq F_\gamma \leq a \text{ on } \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \quad \textit{for some} \quad \varepsilon > 0. \tag{4.7}$$

Theorem 4.2.2. *Let Assumption F hold true, and assume that the fully nonlinear PDE* (4.1) *has strong comparison of bounded solutions. Then for every bounded function g, there exists a bounded function v so that*

$$v^h \longrightarrow v \qquad \textit{locally uniformly}.$$

In addition, v is the unique bounded viscosity solution of the relaxed boundary problem (4.1)–(4.2).

Remark 4.2.3. The restriction to bounded terminal data g in the above Theorem 4.2.2 can be relaxed by an immediate change of variable, thanks to the boundedness conditions on the coefficients b and σ, see [37].

Remark 4.2.4. Theorem 4.2.2 states that the right hand-side inequality of (4.7) (i.e. diffusion must dominate the nonlinearity in γ) is sufficient for the convergence of the Monte Carlo–Finite Differences scheme. We do not know whether this condition is necessary. However, in [37], it is shown that this condition is not sharp in the simple linear case, while the numerical experiments reveal that the method may have a poor performance in the absence of this condition.

The rest of this section is dedicated to the proof of Theorem 4.2.2 which is based on the convergence of monotonic schemes of Barles and Souganidis [9]. Since the probabilistic community is not familiar with this result we report a complete proof.

Properties of the discretisation scheme The approach of [9] is based on three main properties of the scheme.

1 The first ingredient is the so-called consistency property, which states that for every smooth function φ with bounded derivatives and $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$\lim_{\substack{(t',x') \to (t,x) \\ (h,c) \to (0,0) \\ t'+h \leq 1}} \frac{\left(c+\varphi - T_h[c+\varphi]\right)(t',x')}{h} = -\left(\mathcal{L}^X \varphi + F(\cdot,\varphi,D\varphi,D^2\varphi)\right)(t,x). \tag{4.8}$$

The verification of this property follows from an immediate application of Itô's formula.

2 The second ingredient is the so-called stability property:

$$\text{the family } (v^h)_h \text{ is } L^\infty - \text{bounded, uniformly in } h. \tag{4.9}$$

This property is implied by the boundedness of g and $F(t,x,0,0,0)$ (Assumption F), together with the stronger property that for every $\mathbb{L}^\infty$-bounded functions $\varphi, \psi : [0,1] \times \mathbb{R}^d \longrightarrow \mathbb{R}$:

$$|T_h[\varphi] - T_h[\psi]|_\infty \leq |\varphi - \psi|_\infty (1+Ch) \quad \text{for some} \quad C>0, \tag{4.10}$$

For ease of presentation, we report the proof of (4.10) in the one-dimensional case; the general multi-dimensional case follows the same line of argument. Set $f := \varphi - \psi$. By the mean value theorem there exists some $\theta = (t, x, \bar{r}, \bar{p}, \bar{\gamma})$ that:

$$\begin{aligned}(T_h[\varphi] - T_h[\psi])\,(t,x) &= \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(1 + hF_r(\theta) + \frac{W_h}{\sigma}F_p(\theta) + \frac{W_h^2 - h}{\sigma^2 h}F_\gamma(\theta)\right)\right] \\ &= \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(1 - \frac{F_\gamma(\theta)}{\sigma^2} + hF_r - h\frac{F_p(\theta)^2}{4F_\gamma(\theta)} + A_h^2\right)\right] \quad (4.11)\end{aligned}$$

where

$$A_h := \frac{W_h}{\sigma\sqrt{h}}\sqrt{F_\gamma(\theta)} + \frac{F_p(\theta)\sqrt{h}}{2\sqrt{F_\gamma(\theta)}}, \quad (4.12)$$

Since $1 - \sigma^{-2}F_\gamma \geq 0$ and $|F_\gamma^{-1}|_\infty < \infty$ by (4.7) of Assumption F, it follows from the Lipschitz property of F that

$$\begin{aligned}|T_h[\varphi] - T_h[\psi]|_\infty &\leq |f|_\infty\left(1 - \sigma^{-2}F_\gamma(\theta) + \mathbb{E}[|A_h|^2] + Ch\right) \\ &= |f|_\infty\left(1 + h(4F_\gamma)^{-1}(\theta)(F_p)^2(\theta) + Ch\right) \leq |f|_\infty(1 + Ch).\end{aligned}$$

3 The final ingredient – and the most important – is the monotonicity. In the context of our scheme (4.4), we have for all functions $\varphi, \psi : [0,1] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ with exponential growth:

$$\varphi \leq \psi \implies T_h[\varphi](t,x) \leq T_h[\psi](t,x) + Ch\,\mathbb{E}[(\psi - \varphi)(t + h, \hat{X}_h^{t,x})] \quad (4.13)$$

for some constant $C > 0$. To prove (4.13), we proceed as in the proof of the stability result to arrive at (4.11). Since $f := \psi - \varphi \geq 0$ in the present context, and $F_\gamma \leq \sigma^2$ by (4.7) of Assumption F, we deduce that:

$$(T_h[\psi] - T_h[\varphi])\,(t,x) \geq \mathbb{E}\left[f(\hat{X}_h^{t,x})\left(hF_r(\theta) - \frac{h}{4}\frac{F_p(\theta)^2}{F_\gamma(\theta)}\right)\right], \quad (4.14)$$

and the required result follows from the Lipschitz property of F and the fact that $|F_\gamma^{-1}|_\infty < \infty$ by (4.7).

The Barles–Souganidis monotone scheme convergence argument Given the above consistency, stability and monotonicity properties, the convergence of the family $(v^h)_h$ towards some function v which is the unique viscosity solution of the PDE (4.1) follows from [9]. We report the full argument for completeness.

From the stability property the semi-relaxed limits

$$V_*(t,x) := \liminf_{(t',x',h)\to(t,x,0)} v^h(t',x') \quad \text{and} \quad V^*(t,x) := \limsup_{(t',x',h)\to(t,x,0)} v^h(t',x')$$

are finite lower-semicontinuous and upper-semicontinuous functions, respectively. Also, they obviously inherit the boundedness of the family $(v^h)_h$. The key-point is to show that V_* and V^* are respectively viscosity supersolution and subsolution of the PDE (4.1)–(4.2), which implies by the strong comparison assumption that $V_* \leq V^*$. Since the converse inequality is trivial, this shows that $V_* = V^*$ is the unique bounded solution of (4.1), thus completing the proof of convergence.

In the rest of this section, we show that V_* is a viscosity supersolution of the PDE (4.1). A symmetric argument applies to prove the corresponding property for V^*. Let $(t_0, x_0) \in [0,1) \times \mathbb{R}^d$ and $\varphi \in C^2\left([0,1] \times \mathbb{R}^d\right)$ be such that

$$0 \;=\; (V_* - \varphi)(t_0, x_0) \quad = \quad \min(V_* - \varphi)$$

Without loss of generality, we may assume that (t_0, x_0) is a strict minimiser of $(V_* - \varphi)$.

Let (h_n, t_n, x_n) be a sequence in $(0,1] \times B_1(t_0, x_0)$ with

$$(h_n, t_n, x_n) \longrightarrow (0, t_0, x_0) \quad \text{and} \quad v^{h_n}(t_n, x_n) \longrightarrow V_*(t_0, x_0),$$

and define (t'_n, x'_n) by

$$c_n \;:=\; (v_*^{h_n} - \varphi)(t'_n, x'_n) \quad = \quad \min_{B_0}(v^{h_n} - \varphi),$$

where $B_0 \subset [0,1] \times \mathbb{R}^d$ is a closed ball containing (t_0, x_0) in its interior. By definition, we have $c_n + \varphi(t'_n, x'_n) = v^{h_n}(t'_n, x'_n) = T_{h_n}[v^{h_n}](t'_n, x'_n)$. Since $v^{h_n} \geq c_n + \varphi$, it follows from the monotonicity property (4.13) that:

$$c_n + \varphi(t'_n, x'_n) \quad \geq \quad T_{h_n}[c_n + v^{h_n}](t'_n, x'_n). \tag{4.15}$$

Since the sequence $(t'_n, x'_n)_n$ is bounded, it converges to some (t_1, x_1) after possibly passing to a subsequence. Observe that

$$(V_* - \varphi)(t_0, x_0) = \lim_{n\to\infty} (v^{h_n} - \varphi)(t_n, x_n) \geq \liminf_{n\to\infty} (v^{h_n} - \varphi)(t'_n, x'_n) \geq (V_* - \varphi)(t_1, x_1).$$

Since (t_0, x_0) is a strict minimiser of the difference $V_* - \varphi$, this shows that $(t_1, x_1) = (t_0, x_0)$, and $c_n = (v_*^{h_n} - \varphi)(t'_n, x'_n) \longrightarrow 0$. We now go back to (4.15), normalise by h_n and send n to infinity. By the consistency property (4.8), this provides the required result:

$$-\mathcal{L}\varphi(t_0, x_0) - F\left(., \varphi, D\varphi, D^2\varphi\right)(t_0, x_0) \quad \geq \quad 0.$$

Remark 4.2.5. In the context of nonlinear PDEs of the HJB type, see Assumption HJB below, Bonnans and Zidani [14] introduced finite differences approximations which obey to the monotonicity condition. The choice of the discretisation turns out to be pretty involved and is specific to each choice of control. Our probabilistic scheme has the nice property to be automatically monotonic.

4.3 Rate of convergence of the discretisation of HJB equations

We next provide an error estimate for our probabilistic numerical scheme. For ease of presentation, we assume that the nonlinearity F satisfies the condition

$$F_r - \frac{1}{4} F'_p F_\gamma^{-1} F_p \quad \geq \quad 0, \tag{4.16}$$

which implies that the monotonicity property (4.13) is strengthened to:

$$\varphi \le \psi \quad \Longrightarrow \quad T_h[\varphi] \le T_h[\psi], \tag{4.17}$$

see (4.14). In [37], it is shown that our main results hold without Condition (4.16).

The subsequent argument for the derivation of the error estimate is crucially based on the following comparison principle satisfied by the scheme.

Lemma 4.3.1. *Let Assumption F holds true, and consider two arbitrary bounded functions φ and ψ satisfying:*

$$h^{-1}\left(\varphi - T_h[\varphi]\right) \;\le\; g_1 \quad \textit{and} \quad h^{-1}\left(\psi - T_h[\psi]\right) \;\ge\; g_2 \tag{4.18}$$

for some bounded functions g_1 and g_2. Then, for every $i = 0, \ldots, n$:

$$(\varphi - \psi)(t_i, x) \le e^{\beta(T-t_i)} |(\varphi - \psi)^+(1, \cdot)|_\infty + (1-h) e^{\beta(1-t_i)} |(g_1 - g_2)^+|_\infty \tag{4.19}$$

for some parameter $\beta > |F_r|_\infty$.

Before stating precise results, let us explain the key idea as introduced by Krylov [52, 53, 54]. We also refer to the lecture notes by Bonnans [13] which provide a clear summary in the context of infinite horizon stochastic control problems.

Key-idea for the lower bound Given the solution v of the nonlinear PDE (4.1), suppose that there is a function $\underline{u}^\varepsilon$ satisfying

$$\underline{u}^\varepsilon \quad \text{is a classical subsolution of (4.1) and} \quad v - \underline{C}(\varepsilon) \le \underline{u}^\varepsilon \le v \tag{4.20}$$

for some function $\underline{C}(\varepsilon)$. Let

$$\underline{R}_h[\underline{u}^\varepsilon] \quad := \quad \frac{\underline{u}^\varepsilon - T_h[\underline{u}^\varepsilon]}{h} + \mathcal{L}\underline{u}^\varepsilon + F\left(., \underline{u}^\varepsilon, D\underline{u}^\varepsilon, D^2\underline{u}^\varepsilon\right),$$

and let $\underline{R}(h, \varepsilon)$ be a bound on $\underline{R}_h[\underline{u}^\varepsilon]$:

$$|\underline{R}_h[\underline{u}^\varepsilon](t,x)| \le \underline{R}(h,\varepsilon) \quad \text{for every} \quad (t,x) \in [0,T] \times \mathbb{R}^n. \tag{4.21}$$

Since $\underline{u}^\varepsilon$ is a subsolution of the nonlinear PDE (4.1), it follows that

$$\underline{u}^\varepsilon(t,x) \quad \le \quad \mathbb{E}\left[\underline{u}^\varepsilon(t+h, \hat{X}_h^{t,x})\right] + h\left[F(., \mathcal{D}\underline{u}^\varepsilon)(t,x) + \underline{R}(h,\varepsilon)\right].$$

We next introduce the function $\underline{U}^\varepsilon$ defined by $\underline{U}^\varepsilon(1,x) = \underline{u}^\varepsilon(1,x)$ and

$$\underline{U}^\varepsilon(t_i,x) \quad = \quad \mathbb{E}\left[\underline{U}^\varepsilon(t_i+h, \hat{X}_h^{t_i,x})\right] + h\left[F(., \mathcal{D}\underline{U}^\varepsilon)(t_i,x) + \underline{R}(h,\varepsilon)\right].$$

Then, it follows from the comparison property of Lemma 4.3.1 that

$$\underline{u}^\varepsilon \quad \le \quad \underline{U}^\varepsilon. \tag{4.22}$$

Moreover, arguing as in the proof of (4.10), we see that

$$\left|\underline{U}^\varepsilon - v^h\right| \quad \le \quad (1+Ch)\mathbb{E}_{t,x}\left[\left|\underline{U}^\varepsilon - v^h\right|(t+h,\hat{X}_h^{t,x})\right] + h\overline{R}(h,\varepsilon),$$

which provides the estimate:

$$\left|\underline{U}^\varepsilon - v^h\right| \quad \le \quad \underline{R}(h,\varepsilon). \tag{4.23}$$

We then deduce from (4.20), (4.22) and (4.23) that

$$v - v^h = v - \underline{u}^\varepsilon + \underline{u}^\varepsilon - v \le v - \underline{u}^\varepsilon + \underline{U}^\varepsilon - v^h \le \underline{C}(\varepsilon) + \underline{R}(h,\varepsilon),$$

and therefore

$$v - v^h \quad \le \quad \inf_{\varepsilon>0}\left(\underline{C}(\varepsilon) + \underline{R}(h,\varepsilon)\right). \tag{4.24}$$

The approximating subsolution The following construction of the function $\underline{u}^\varepsilon$ satisfying (4.20) requires that the nonlinearity $F(t,x,r,p,\gamma)$ be concave in (r,p,γ) with some additional conditions in order to ensure some regularity of the solution. Notice that the case where $F(t,x,r,p,\gamma)$ is convex in (r,p,γ) can also be dealt with by a symmetric argument (inverting the roles of supersolutions and subsolutions). For a bounded function $\psi(t,x)$ Lipschitz in x and $1/2$–Hölder continuous in t, we denote

$$|\psi|_1 \quad := \quad |\psi|_\infty + \sup_{([0,1]\times\mathbb{R}^d)^2} \frac{\psi(t,x)-\psi(t',x')}{|x-x'|+|t-t'|^{1/2}}$$

Assumption HJB *The nonlinearity F is of the Hamilton–Jacobi–Bellman type:*

$$F(t,x,r,p,\gamma) \quad = \quad \inf_{\alpha\in\mathcal{A}}\{\mathcal{L}^\alpha(t,x,r,p,\gamma)\}$$
$$\mathcal{L}^\alpha(t,x,r,p,\gamma) \quad := \quad \frac{1}{2}Tr[\sigma^\alpha\sigma^{\alpha\mathrm{T}}(t,x)\gamma] + b^\alpha(t,x)p + c^\alpha(t,x)r + f^\alpha(t,x)$$

where b, σ, σ^α, b^α, c^α and f^α satisfy:

$$|b|_1 + |\sigma|_1 + \sup_{\alpha\in\mathcal{A}}\left(|\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1\right) \quad < \quad \infty.$$

The above Assumption implies in particular that our nonlinear PDE satisfies a strong comparison result for bounded functions. Let w^ε be the unique viscosity solution of the nonlinear PDE with shaked coefficients (in the terminology of Krylov):

$$\begin{array}{l} -\frac{\partial w^\varepsilon}{\partial t} - \inf_{|e|\le 1}\bar{F}\left(t+\varepsilon e, x+\varepsilon e, w^\varepsilon(t,x), Dw^\varepsilon(t,x), D^2 w^\varepsilon(t,x)\right) = 0, \\ w^\varepsilon(1,.) = g, \end{array} \tag{4.25}$$

where:

$$\bar{F}(t,x,r,p,\gamma) \quad := \quad \langle b(t,x),p\rangle + \frac{1}{2}\mathrm{Tr}[\sigma\sigma'\gamma] + F(t,x,r,p,\gamma).$$

The existence of w^ε follows from a direct identification of this PDE as the HJB equation of a corresponding stochastic control problem. By the definition of w^ε together with a strong comparison argument, we then have that

$$w^\varepsilon \text{ is a subsolution of (4.1) and } |w^\varepsilon - v| \le C\varepsilon, \tag{4.26}$$

for some constant $C > 0$. It also follows from classical estimates that for a Lipschitz–continuous final condition g:

$$w^\varepsilon \text{ is bounded, Lipschitz in } x, \text{ and } 1/2\text{–Hölder continuous in } t. \tag{4.27}$$

Let $\rho(t,x)$ be a C^∞ positive function supported in $\{(t,x) : t \in [0,1], |x| \le 1\}$ with unit mass, and define

$$\underline{u}^\varepsilon(t,x) := w^\varepsilon * \rho^\varepsilon \quad \text{where} \quad \rho^\varepsilon(t,x) := \frac{1}{\varepsilon^{d+2}}\rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) \tag{4.28}$$

so that it follows from (4.27) that

$$\underline{u}^\varepsilon \quad \text{is } C^\infty, \text{ and } \quad \left|\partial_t^{\beta_0} D^\beta \underline{u}^\varepsilon\right| \le C\varepsilon^{1-2\beta_0-|\beta|_1} \tag{4.29}$$

for any $(\beta_0, \beta) \in \mathbb{N} \times \mathbb{N}^d \setminus \{0\}$, where $|\beta|_1 := \sum_{i=1}^d \beta_i$, and $C > 0$ is some constant. Then, it follows from two successive applications of Itô's formula that

$$\underline{R}_h[\underline{u}^\varepsilon] \quad \le \quad R(h,\varepsilon) \ := \ Ch\varepsilon^{-3} \tag{4.30}$$

for some constant C. Finally, by the concavity of F, the function $\underline{u}^\varepsilon$ inherits (4.26):

$$\underline{u}^\varepsilon \quad \text{is a viscosity subsolution of and} \quad |\underline{u}^\varepsilon - v| \le C(\varepsilon) := C\varepsilon. \tag{4.31}$$

Since $\underline{u}^\varepsilon$ satisfies all the requirements of the previous section, we deduce that

$$v - v^h \quad \le \quad \inf_{\varepsilon>0} C\left(\varepsilon + h\varepsilon^{-3}\right) \ \sim \ Ch^{1/4}. \tag{4.32}$$

The upper bound In the context of our scheme, we shall obtain an upper bound on the error by inverting the roles of v and v^h in the above lower bound argument. Let v_ε^h be the solution of the scheme with shaked coefficients:

$$v_\varepsilon^h(T,.) = g, \ v_\varepsilon^h = T_h^\varepsilon[v_\varepsilon^h], \ T_h^\varepsilon[\psi] = \inf_{|e|\le\varepsilon} T_{h,e}[\psi], \ T_{h,e}[\psi] := \mathcal{D}_e^0[\psi] + hF(., \mathcal{D}_e[\psi])$$

where

$$\mathcal{D}_e\psi(t,x) = \mathbb{E}\left[\psi(t+h, x+\mu(t+e,x+e)h + \sigma(t+e,x+e)W_h)H_h(t+e,x+e)\right].$$

Then, $0 \le v^h - v_\varepsilon^h$ by the comparison result for the scheme. When the volatility $\sigma(.)$ is chosen to be constant, we also have $v^h - v_\varepsilon^h \le C\varepsilon$ for some constant C. Moreover, it is proved in [37] that v_ε^h is Lipschitz-continuous in x and $1/2$–Hölder continuous in t, uniformly in h. Then, the mollification $v_\varepsilon^h * \rho_\varepsilon$ of v_ε^h satisfies the estimates (4.29). Collecting all these ingredients, we obtain a lower bound on the error:

$$v - v^h \quad \ge \quad -C\inf_{\varepsilon}\left(\varepsilon + R(h,\varepsilon)\right) \ = \ -Ch^{1/4}.$$

The main result Summing up the above results, we have

Theorem 4.3.2. *Let the nonlinearity F be as in Assumption HJB, and let σ be constant. Then, for any bounded Lipschitz final condition g:*

$$Ch^{1/10} \leq v - v^h \leq Ch^{1/4}$$

Remark 4.3.3. In the PDE Finite Differences literature, the rate of convergence is usually stated in terms of the discretisation in the space variable $|\Delta x|$. In our context of stochastic differential equation, notice that $|\Delta x|$ is or the order of $h^{1/2}$. Therefore, the above upper bound on the rate of convergence corresponds to the classical rate $|\Delta x|^{1/2}$.

Bibliography

[1] F. Antonelli and A. Kohatsu-Higa, *Filtration stability of backward SDE's*, Stochastic Analysis and Its Applications **18** (2000), pp. 11–37.

[2] V. Bally, *An approximation scheme for BSDEs and applications to control and nonlinear PDE's*, Pitman Research Notes in Mathematics Series, Longman **364** (1997).

[3] V. Bally and G. Pagès, *A quantization algorithm for solving discrete time multidimensional optimal stopping problems*, Bernoulli **9** (2002), pp. 1003–1049.

[4] V. Bally, G. Pagès, and J. Printems, *A quantization method for pricing and hedging multi-dimensional American style options*, Mathematical Finance **15** (2005), pp. 119–168.

[5] G. Barles, R. Buckdahn, and E. Pardoux, *Backward stochastic differential equations and integral-partial differential equations*, Stochastics and Stochastics Reports **60** (1997), pp. 57–83.

[6] G. Barles and E. R. Jakobsen, *On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations*, Mathematical Modelling and Numerical Analysis, ESAIM, M2AM **36** (2002), pp. 33–54.

[7] G. Barles and E. R. Jakobsen, *Error bounds For monotone approximation schemes for Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal. **43** (2005), pp. 540–558.

[8] G. Barles and E. R. Jakobsen, *Error bounds For monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations*, Math. Comp. **76** (2007), pp. 1861–1893.

[9] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully non-linear second order equation*, Asymptotic Analysis **4** (1991), pp. 271–283.

[10] D. Becherer, *Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging*, preprint, Imperial College, 2005.

[11] C. Bender and J. Zhang, *Time discretization and markovian iteration for coupled FBSDEs*, Ann. Appl. Probab. **18** (2008), pp. 143–177 .

[12] J.M. Bismut, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl. **44** (1973), pp. 384–404.

[13] F. Bonnans, *Quelques aspects numériques de la commande optimale stochastique*, Lecture Notes, 2009.

[14] F. Bonnans and H. Zidani, *Consistency of generalized finite difference schemes for the stochastic HJB equation*, SIAM J. Numerical Analysis **41** (2003),pp. 1008–1021.

[15] B. Bouchard, *A stochastic target formulation for optimal switching problems in finite horizon*, to appear in Stochastics.

[16] B. Bouchard and J.-F. Chassagneux, *Discrete time approximation for continuously and discretely reflected BSDEs*, Stochastic Processes and their Applications **118** (2008), pp. 2269–2293.

[17] B. Bouchard, I. Ekeland, and N. Touzi, *On the Malliavin approach to Monte Carlo approximation of conditional expectations*, Finance and Stochastics **8** (2004), pp. 45–71.

[18] B. Bouchard and R. Elie, *Discrete time approximation of decoupled forward-backward SDE with jumps*, Stochastic Processes And Their Applications **118** (2008), pp. 53–75

[19] B. Bouchard and S. Menozzi, *Strong approximation of BSDEs in a domain*, to appear in Bernoulli.

[20] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Processes And Their Applications **111** (2004), pp. 175–206.

[21] P. Briand, B. Delyon, and J. Mémin, *Donsker-type theorem for BSDE's*, Electronic Communications In Probability **6** (2001), pp. 1–14.

[22] J. F. Carrière, *Valuation of the early-exercise price for options using simulations and nonparametric regression*, Insurance : Mathematics And Economics **19** (1996), pp. 19–30.

[23] J.-F. Chassagneux, *Discrete-time approximation of doubly reflected BSDEs*, Advances in Applied Probability **41** (2009).

[24] J.-F. Chassagneux, *Processus réfléchis en finance et probabilités numérique*, PhD Thesis, University Paris-Diderot, 2008.

[25] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir, *Second order backward stochastic differential equations and fully non-linear parabolic PDEs*, Communications On Pure And Applied Mathematics **60** (2007), pp. 1081–1110.

[26] D. Chevance, *Numerical methods for backward stochastic differential equations*, in Numerical methods in finance 1997, Edt L.C.G. Rogers and D. Talay, Cambridge University Press, pp. 232–244.

[27] E. Clément, D. Lamberton, and P. Protter, *An analysis of a least squares regression method for American option pricing*, Finance And Stochastics **6** (2002), pp. 449–472.

[28] F. Coquet, J. Mémin, and L. Slominski, *On weak convergence of filtrations*, Séminaire De Probabilités, XXXV, pp. 306–328. Lecture Notes in Math. 1755 (2001) Springer, Berlin.

[29] F. Coquet, V. Mackevičius, and J. Mémin, *Stability in D of martingales and backward equations under discretization of filtration*, Stochastic Processes And Their Applications **75** (1998), pp. 235–248.

[30] R. W. R. Darling and E. Pardoux, *BSDE with random terminal time*, Annals Of Probability **25** (1997), pp. 1135–1159.

[31] F. Delarue and S. Menozzi, *A forward backward algorithm for quasi-linear PDEs*, Annals Of Applied Probability **16** (2006), pp. 140–184.

[32] F. Delarue and S. Menozzi, *An interpolated stochastic algorithm for quasi-linear PDEs*, Mathematics Of Computation **261** (2008), pp. 125–158.

[33] H. Dong, N. V. Krylov, *On the rate of convergence of finite-difference approximations for bellman's equations with constant coefficients*, St. Petersburg Math. J. **17** (2005), pp. 108–132.

[34] J. Jr. Douglas, J. Ma, and P. Protter, *Numerical methods for forward- backward stochastic differential equations*, Annals Of Applied Probability **6** (1996), pp. 940–968.

[35] R. Elie, *Contrôle stochastique et méthodes numériques en finance mathématique*, PHD thesis, University Paris-Dauphine, 2006.

[36] A. Eyraud-Loisel, *Backward stochastic differential equations with enlarged filtration. option hedging of an insider trader in a financial market with jumps*, Stochastic Processes And Their Applications **115** (2005), pp. 1745–1763,

[37] A. Fahim, N. Touzi, and X. Warin, *A probabilistic numerical scheme for fully nonlinear PDEs*, preprint, 2009.

[38] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance And Stochastics **3** (1999), pp. 391–412.

[39] E. Gobet, *Schéma d'Euler pour diffusions tues. Application aux options barrire*, PhD Thesis, University Paris-Diderot, 1998.

[40] E. Gobet, *Weak approximation of killed diffusion using Euler schemes*, Stochastic Processes And Their Applications **87** (2000), pp. 167–197.

[41] E. Gobet, *Revisiting the Greeks for European and American options*, Proceedings of the "International Symposium on Stochastic Processes and Mathematical Finance" at Ritsumeikan University, Kusatsu, Japan, March 2003. Edited by J. Akahori, S. Ogawa, S. Watanabe. World Scientific, pp. 53–71, 2004.

[42] E. Gobet and C. Labart, *Error expansion for the discretization of backward stochastic differential equations*, Stochastic Processes And Applications **117** (2007), pp. 803–829.

[43] E. Gobet, J.P. Lemor, and X. Warin, *Rate of convergence of empirical regression method for solving generalized BSDE*, Bernoulli **12** (2006), pp. 889–916.

[44] E. Gobet and A. Makhlouf, *L^2-time regularity of BSDEs with irregular terminal functions*, preprint 2008.

[45] E. Gobet and S. Menozzi, *Stopped diffusion processes: overshoots and boundary correction*, preprint PMA, University Paris-Diderot, 2007.

[46] Y. Hu, P. Imkeller, and M. Müller, *Utility maximization in incomplete markets*, Ann. Appl. Probab. **15** (2005), pp. 1691–1712.

[47] N. El Karoui, C. Kapoudjan, E. Pardoux, S. Peng, and M.C. Quenez, *Reflected solutions of backward stochastic differential equations and related obstacle problems for PDE's*, Annals of Probability **25** (1997), pp. 702–737.

[48] N. El Karoui, S. Pend and M.-C. Quenez, *Backward stochastic differential equations in finance*, Mathematical Finance **7** (1997), pp. 1–71.

[49] N. El Karoui, S. Peng, and M.-C. Quenez, *A dynamic maximum principle for the optimization of recursive utilities under constraints*, The Annals Of Applied Probability **11** (2001), pp. 664–693

[50] N. El Karoui and R. Rouge, *Pricing via utility maximization and entropy*, Mathematical Finance **10** (2000), pp. 259–276.

[51] I. Kharroubi, H. Pham, J. Ma, and J. Zhang, *Backward SDEs with constrained jumps and quasi-variational inequalities*, preprint 2008.

[52] N. V. Krylov, *On the rate of convergence of finite-difference approximations for bellman's equations*, St. Petersburg Math. J. **9** (1997), pp. 245–256.

[53] N. V. Krylov, *On the rate of convergence of finite-difference approximations for bellman's equations with variable coefficients*, Probability Theory And Related Fields **117** (2000), pp. 1–16.

[54] N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equations with Lipschitz coefficients*, Applied Math. And Optimization **52** (2005), pp. 365–399.

[55] P.-L. Lions and H. Regnier, *Calcul du prix et des sensibilités d'une option américaine par une méthode de Monte Carlo*, preprint 2001.

[56] F. A. Longstaff and R. S. Schwartz, *Valuing american options by simulation : a simple least-square approach*, Review Of Financial Studies **14** (2001), pp. 113–147.

[57] J. Ma, P. Protter, J. San Martin, and S. Torres, *Numerical method for backward stochastic differential equations*, Annals Of Applied Probability **12** (2002), pp. 302–316.

[58] J. Ma, P. Protter, and J. Yong, *Solving forward-backward stochastic differential equations explicitly - a four step scheme*, Probability Theory And Related Fields **98** (1994), pp. 339–359.

[59] J. Ma and J. Zhang, *Path regularity of solutions to backward stochastic differential equations*, Probability Theory And Related Fields **122** (2002), pp. 163–190.

[60] J. Ma and J. Zhang, *Representations and regularities for solutions to BSDEs with reflections*, Stochastic Processes And Their Applications **115** (2005), pp. 539–569.

[61] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, Berlin, 1995.

[62] E. Pardoux, *Backward stochastic differential equations and viscosity solutions of semilinear parabolic and elliptic PDE's of second order*, In Stochastic Analysis and Related Topics: The Geilo Workshop 1996. L. Decreusefond, J. Gjerd, B. Oksendal, and A.S. Ustnel (eds.), Birkhäuser, pp. 79–127, 1998.

[63] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems and control letters **14** (1990), pp. 55–61.

[64] E. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Lecture Notes In Control And Inform. Sci. **176** (1992), pp. 200–217.

[65] E. Pardoux, F. Pradeilles, and Z. Rao, *Probabilistic interpretation for a system of semilinear parabolic partial differential equations*, Ann. Inst. H. Poincare **33** (1997), pp. 467–490.

[66] S. Peng, *Probabilistic interpretation for systems of quasilinear parabolic partial differential equations*, Stochastics And Stochastics Reports **37** (1991), pp. 61–74.

[67] S. Peng, *Backward stochastic differential equations and applications to optimal control*, Appl Math Optim **27** (1993), pp. 125–144.

[68] S. Peng, *Backward stochastic differential equations, nonlinear expectations, nonlinear evaluations and risk measures*, Lecture Notes in Chinese Summer School in Mathematics Weihai, July 19–August 14, 2004.

[69] S. Peng, *G-Brownian motion and dynamic risk measure under volatility uncertainty*, arXiv:0711.2834v1, 2007.

[70] S. Rubenthaler, *Numerical simulation of the solution of a stochastic differential equation driven by a Levy process*, Stochastic Processes And Their Applications **103** (2003), pp. 311–349.

[71] A. B. Sow and E. Pardoux, *Probabilistic interpretation of a system of quasilinear parabolic PDEs*, Stochastics And Stochastics Reports **76** (2004), pp. 429–477.

[72] S. Tang and X. Li, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM J. Control Optim. **32** (1994), pp. 1447–1475.

[73] J. Zhang, *Some fine properties of backward stochastic differential equations*, PhD thesis, Purdue University, 2001.

[74] J. Zhang, *A numerical scheme for BSDEs*, Annals of Applied Probability **14** (2004), pp. 459–488.

Author information

Bruno Bouchard, CEREMADE, Université Paris Dauphine and CREST-ENSAE, France.
Email: `bouchard@ceremade.dauphine.fr`

Romuald Elie, CEREMADE, Université Paris Dauphine and CREST-ENSAE, France.
Email: `elie@ceremade.dauphine.`

Nizar Touzi, CMAP, Ecole Polytechnique Paris, France.
Email: `nizar.touzi@polytechnique.edu`