Discrete time approximation of decoupled Forward-Backward SDE with jumps

Bruno BOUCHARD
LPMA
CNRS, UMR 7599
Université Paris 6
e-mail: bouchard@ccr.jussieu.fr
and CREST

Romuald ELIE
CEREMADE
CNRS, UMR 7534
Université Paris 9
e-mail: elie@ensae.fr
and CREST-ENSAE

This version : March 2007

Abstract

We study a discrete-time approximation for solutions of systems of decoupled forward-backward stochastic differential equations with jumps. Assuming that the coefficients are Lipschitz-continuous, we prove the convergence of the scheme when the number of time steps n goes to infinity. The rate of convergence is at least $n^{-1/2+\varepsilon}$, for any $\varepsilon > 0$. When the jump coefficient of the first variation process of the forward component satisfies a non-degeneracy condition which ensures its inversibility, we achieve the optimal convergence rate $n^{-1/2}$. The proof is based on a generalization of a remarkable result on the path-regularity of the solution of the backward equation derived by Zhang [25] in the no-jump case.

Key words : Discrete-time approximation, forward-backward SDE's with jumps, Malliavin calculus.

MSC Classification (2000): 65C99, 60H07, 60J75.

1 Introduction

In this paper, we study a discrete time approximation scheme for the solution of a system of decoupled Forward-Backward Stochastic Differential Equations (FBSDE in short) with jumps of the form

$$\left\{ \begin{array}{rcl} X_t &=& X_0 + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r + \int_0^t \int_E \beta(X_{r-}, e)\bar{\mu}(de, dr) \; , \\ Y_t &=& g(X_T) + \int_t^T h\left(\Theta_r\right)dr - \int_t^T Z_r \cdot dW_r - \int_t^T \int_E U_r(e)\bar{\mu}(de, dr) \end{array} \right. \tag{1.1}$$

where $\Theta := (X, Y, Z, \Gamma)$ with $\Gamma := \int_E \rho(e)U(e)\lambda(de)$. Here, W is a d-dimensional Brownian motion and $\bar{\mu}$ an independent compensated Poisson measure $\bar{\mu}(de, dr) = \mu(de, dr) - \lambda(de)dr$. Such equations naturally appear in hedging problems, see e.g. Eyraud-Loisel [13], or in stochastic control, see e.g. Tang and Li [23] and the recent paper Becherer [3] for an application to exponential utility maximization in finance. Under standard Lipschitz assumptions on the coefficients b, σ, β, g and h, existence and uniqueness of the solution have been proved by Tang and Li [23], thus generalizing the seminal paper of Pardoux and Peng [20].

The main motivation for studying discrete time approximations of systems of the above form is that they provide an alternative to classical numerical schemes for a large class of (deterministic) PDE's of the form

$$\mathcal{L}u(t,x) + h\left(t,x,u(t,x),\sigma(t,x)\nabla_x u(t,x),\mathcal{I}[u](t,x)\right) \;=\; 0\;,\; u(T,x) \;=\; g(x)\;, \quad (1.2)$$

where

$$\begin{aligned}\mathcal{L}u(t,x) \;&:=\; \frac{\partial u}{\partial t}(t,x) + \nabla_x u(t,x)b(x) + \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^*(x))^{ij}\frac{\partial^2 u}{\partial x^i \partial x^j}(t,x)\\ &\quad + \int_E \{u(t,x+\beta(x,e)) - u(t,x) - \nabla_x u(t,x)\beta(x,e)\}\,\lambda(de)\;,\\ \mathcal{I}[u](t,x) \;&:=\; \int_E \{u(t,x+\beta(x,e)) - u(t,x)\}\,\rho(e)\,\lambda(de)\;.\end{aligned}$$

Indeed, it is well known that, under mild assumptions on the coefficients, the component Y of the solution can be related to the (viscosity) solution u of (1.2) in the sense that $Y_t = u(t,X_t)$, see e.g. [1] or [9]. Thus solving (1.1) or (1.2) is essentially the same. In the so-called four-steps scheme, this relation allows to approximate the solution of (1.1) by first estimating numerically u, see e.g. [10]. Here, we follow the converse approach. Since classical numerical schemes for PDE's generally do not perform well in high dimension, we want to estimate directly the solution of (1.1) so as to provide an approximation of u.

In the no-jump case, i.e. $\beta = 0$, the numerical approximation of (1.1) has already been studied in the literature, see e.g. Zhang [25], Bally and Pages [2], Bouchard and Touzi [7] or Gobet et al. [16]. In [7], the authors suggest the following implicit scheme. Given a regular grid $\pi = \{t_i = iT/n,\ i = 0,\ldots,n\}$, they approximate X by its Euler scheme X^π and (Y,Z) by the discrete-time process $(\bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i})_{i\le n}$ defined backward by

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

where $\bar{Y}^\pi_{t_n} := g(X^\pi_{t_n})$ and $\Delta W_{i+1} := W_{t_{i+1}} - W_{t_i}$. In the no-jump case, it turns out that the discretization error

$$\mathrm{Err}_n(Y,Z) := \left\{\max_{i<n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\left[|Y_t - \bar{Y}^\pi_{t_i}|^2\right] + \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|Z_t - \bar{Z}^\pi_{t_i}|^2\right]dt\right\}^{\frac{1}{2}}$$

is intimately related to the quantity

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2\right]dt \quad \text{where} \quad \bar{Z}_{t_i} := n\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z_t dt \mid \mathcal{F}_{t_i}\right]\;.$$

Under Lipschitz continuity conditions on the coefficients, Zhang [24] was able to prove that the later is of order of n^{-1}. This remarkable result allows to derive the bound $\mathrm{Err}_n(Y,Z) \le Cn^{-1/2}$. Observe that this rate of convergence can not be improved in general. Consider for example the case where X is equal to the Brownian motion W, g is the identity and $h = 0$. Then, $Y = W$ and $\bar{Y}^\pi_{t_i} = W_{t_i}$.

In this paper, we extend the approach of Bouchard and Touzi [7] and approximate the solution of (1.1) by the backward scheme

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

where $\bar{Y}^\pi_{t_n} := g(X^\pi_{t_n})$. By adapting the arguments of Gobet et al. [16], we first prove that our discretization error $\mathrm{Err}_n(Y,Z,U)$ defined as

$$\left\{ \max_{i<n} \sup_{t\in[t_i,t_{i+1}]} \mathbb{E}\left[|Y_t - \bar{Y}^\pi_{t_i}|^2\right] + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}^\pi_{t_i}|^2 + |\Gamma_t - \bar{\Gamma}^\pi_{t_i}|^2\right] dt \right\}^{\frac{1}{2}}$$

converges to 0 as the discretization step T/n tends to 0. We then provide upper bounds on

$$\max_{i<n} \sup_{t\in[t_i,t_{i+1}]} \mathbb{E}\left[|Y_t - Y_{t_i}|^2\right] + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2 + |\Gamma_t - \bar{\Gamma}_{t_i}|^2\right] dt\ ,$$

where $\bar{\Gamma}_{t_i} := (n/T)\ \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma_t dt \mid \mathcal{F}_{t_i}\right]$. When the coefficients are Lipschitz continuous, we obtain

$$\max_{i<n} \sup_{t\in[t_i,t_{i+1}]} \mathbb{E}\left[|Y_t - Y_{t_i}|^2\right] + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\Gamma_t - \bar{\Gamma}_{t_i}|^2\right] dt \quad \le \quad C\, n^{-1}$$

and

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2\right] dt \quad \le \quad C_\varepsilon\, n^{-1+\varepsilon}\ , \quad \text{for any } \varepsilon > 0\ .$$

Under some additional conditions on the inversibility of $\nabla\beta + I_d$, see **H**, we then prove that the previous inequality holds true for $\varepsilon = 0$. This extends to our framework the remarkable result derived by Zhang [25] in the no-jump case. It allows us to show that our discrete-time scheme achieves, under the standard Lipschitz conditions, a rate of convergence of at least $n^{-1/2+\varepsilon}$, for any $\varepsilon > 0$, and the optimal rate $n^{-1/2}$ under the additional assumption **H**.

Observe that, in opposition to algorithms based on the approximation of the Brownian motion by discrete processes taking a finite number of possible values (see e.g. [17] and the references therein), our scheme does not provide a fully implementable numerical procedure since it involves the computation of a large number of conditional expectations. However, the implementation of the above mentioned schemes in high dimension is questionable and, in our setting, this issue could be solved by approximating the conditional expectation operators numerically in an efficient way, see [2], [7], [16] and [11] for an adaptation to our setting of the technics suggested in [16].

The rest of the paper is organized as follows. In Section 2, we describe the approximation scheme and state our main convergence result. Section 3 contains some results on the Malliavin derivatives of Forward and Backward SDE's. Applying these results in Section 4, we derive some regularity properties for the solution of the backward equation under additional smoothness assumptions on the coefficients. We finally use an approximation argument to conclude the proof of our main theorem.

Notations : Any element $x \in \mathbb{R}^d$ will be identified to a column vector with i-th component x^i and Euclidian norm $|x|$. For $x_i \in \mathbb{R}^{d_i}$, $i \leq n$ and $d_i \in \mathbb{N}$, we define $(x_1, \ldots, x_n)$ as the column vector associated to $(x_1^1, \ldots, x_1^{d_1}, \ldots, x_n^1, \ldots, x_n^{d_n})$. The scalar product on $\mathbb{R}^d$ is denoted by $x \cdot y$. For a $(m \times d)$-dimensional matrix M, we note $|M| := \sup\{|Mx|;\ x \in \mathbb{R}^d\ ,\ |x| = 1\}$, M^* its transpose and we write $M \in \mathbb{M}^d$ if $m = d$. Given $p \in \mathbb{N}$ and a measured space $(A, \mathcal{A}, \mu_A)$, we denote by $L^p(A, \mathcal{A}, \mu_A; \mathbb{R}^d)$, or simply $L^p(A, \mathcal{A})$ or $L^p(A)$ if no confusion is possible, the set of p-integrable $\mathbb{R}^d$-valued measurable maps on $(A, \mathcal{A}, \mu_A)$. For $p = \infty$, $L^\infty(A, \mathcal{A}, \mu_A; \mathbb{R}^d)$ is the set of essentially bounded $\mathbb{R}^d$-valued measurable maps. The set of k-times differentiable maps with bounded derivatives up to order k is denoted by C_b^k and $C_b^\infty := \cap_{k \geq 1} C_b^k$. For a map $b : \mathbb{R}^d \mapsto \mathbb{R}^k$, we denote by ∇b is Jacobian matrix whenever it exists.
In the following, we shall use these notations without specifying the dimension when it is clearly given by the context.

2 Discrete time approximation of decoupled FBSDE with jumps

2.1 Decoupled forward backward SDE's

As in [5], we shall work on a suitable product space $\Omega := \Omega_W \times \Omega_\mu$ where Ω_W is the set of continuous functions w from $[0, T]$ into $\mathbb{R}^d$, and Ω_μ is the set of integer-valued measures on $[0, T] \times E$ with $E := \mathbb{R}^m$ for some $m \geq 1$. For $\omega = (w, \eta) \in \Omega$, we set $W(w, \eta) = w$ and $\mu(w, \eta) = \eta$ and define $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \leq T}$ (resp. $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{t \leq T}$) as the smallest right-continuous filtration on Ω_W (resp. Ω_μ) such that W (resp. μ) is optional. We let $\mathbb{P}_W$ be the Wiener measure on $(\Omega_W, \mathcal{F}_T^W)$ and $\mathbb{P}_\mu$ be the measure on $(\Omega_\mu, \mathcal{F}_T^\mu)$ under which μ is a Poisson measure with intensity $\nu(dt, de) = \lambda(de)dt$, for some finite measure λ on E, endowed with its Borel tribe $\mathcal{E}$. We then define the probability measure $\mathbb{P} := \mathbb{P}_W \otimes \mathbb{P}_\mu$ on $(\Omega, \mathcal{F}_T^W \otimes \mathcal{F}_T^\mu)$. With this construction, W and μ are independent under $\mathbb{P}$. Without loss of generality, we can assume that the natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t \leq T}$ induced by (W, μ) is complete. We denote by

$\bar{\mu} := \mu - \nu$ the compensated measure associated to μ.

Given $K > 0$, two K-Lipschitz continuous functions $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{M}^d$, and a measurable map $\beta : \mathbb{R}^d \times E \to \mathbb{R}^d$ such that

$$\sup_{e \in E} |\beta(0,e)| \;\leq\; K \;\text{ and }\; \sup_{e \in E} |\beta(x,e) - \beta(x',e)| \;\leq\; K|x - x'| \;\; \forall\, x,\; x' \in \mathbb{R}^d \;, \quad (2.1)$$

we define X as the solution on $[0,T]$ of

$$X_t \;=\; X_0 + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r + \int_0^t \int_E \beta(X_{r-},e)\bar{\mu}(de,dr) \;, \quad (2.2)$$

for some initial condition $X_0 \in \mathbb{R}^d$. The existence and uniqueness of such a solution is well known under the above assumptions, see the Appendix for standard estimates for solutions of such SDE.

Before introducing the backward SDE, we need to define some additional notations. Given $s \leq t$ and some real number $p \geq 2$, we denote by $\mathcal{S}^p_{[s,t]}$ the set of real valued adapted *càdlàg* processes Y such that

$$\|Y\|_{\mathcal{S}^p_{[s,t]}} \;:=\; \mathbb{E}\left[\sup_{s \leq r \leq t} |Y_r|^p\right]^{\frac{1}{p}} \;<\; \infty \;,$$

$\mathbf{H}^p_{[s,t]}$ is the set of progressively measurable $\mathbb{R}^d$-valued processes Z such that

$$\|Z\|_{\mathbf{H}^p_{[s,t]}} \;:=\; \mathbb{E}\left[\left(\int_s^t |Z_r|^2 dr\right)^{\frac{p}{2}}\right]^{\frac{1}{p}} \;<\; \infty \;,$$

$\mathbf{L}^p_{\lambda,[s,t]}$ is the set of $\mathcal{P} \otimes \mathcal{E}$ measurable maps $U : \Omega \times [0,T] \times E \to \mathbb{R}$ such that

$$\|U\|_{\mathbf{L}^p_{\lambda,[s,t]}} \;:=\; \mathbb{E}\left[\int_s^t \int_E |U_s(e)|^p \lambda(de)ds\right]^{\frac{1}{p}} \;<\; \infty$$

with $\mathcal{P}$ defined as the σ-algebra of $\mathbb{F}$-predictable subsets of $\Omega \times [0,T]$. The space

$$\mathcal{B}^p_{[s,t]} \;:=\; \mathcal{S}^p_{[s,t]} \times \mathbf{H}^p_{[s,t]} \times \mathbf{L}^p_{\lambda,[s,t]}$$

is endowed with the norm

$$\|(Y,Z,U)\|_{\mathcal{B}^p_{[s,t]}} \;:=\; \left(\|Y\|^p_{\mathcal{S}^p_{[s,t]}} + \|Z\|^p_{\mathbf{H}^p_{[s,t]}} + \|U\|^p_{\mathbf{L}^p_{\lambda,[s,t]}}\right)^{\frac{1}{p}} \;.$$

In the sequel, we shall omit the subscript $[s,t]$ in these notations when $(s,t) = (0,T)$. For ease of notations, we shall sometimes write that an $\mathbb{R}^n$-valued process is in $\mathcal{S}^p_{[s,t]}$ or $\mathbf{L}^p_{\lambda,[s,t]}$ meaning that each component is in the corresponding space. Similarly an element of $\mathbb{M}^m$ is said to belong to $\mathbf{H}^p_{[s,t]}$ if each column belongs to $\mathbf{H}^p_{[s,t]}$. The norms are then naturally extended to such processes.

The aim of this paper is to study a discrete time approximation of the triplet (Y, Z, U) solution on $[0, T]$ of the backward stochastic differential equation

$$Y_t = g(X_T) + \int_t^T h(\Theta_r)\, dr - \int_t^T Z_r \cdot dW_r - \int_t^T \int_E U_r(e)\bar{\mu}(de, dr) \;, \qquad (2.3)$$

where $\Theta := (X, Y, Z, \Gamma)$ and Γ is defined by

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

for some measurable map $\rho : E \to \mathbb{R}^m$ satisfying

$$\sup_{e \in E} |\rho(e)| \;\leq\; K \;. \qquad (2.4)$$

By a solution, we mean a triplet $(Y, Z, U) \in \mathcal{B}^2$ satisfying (2.3).

In order to ensure the existence and uniqueness of a solution to (2.3), we assume that the map $g : \mathbb{R}^d \mapsto \mathbb{R}$ and $h : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ are K-Lipschitz continuous (see Lemma 5.2 in the Appendix).

For ease of notations, we shall denote by C_p a generic constant depending only on p and the constants K, $\lambda(E)$, $b(0)$, $\sigma(0)$, $h(0)$, $g(0)$ and T. We write C_p^0 if it also depends on X_0. In this paper, p will always denote a real number greater than 2.

Remark 2.1 For the convenience of the reader, we have collected in the Appendix standard estimates for the solutions of Forward and Backward SDE's. In particular, they imply

$$\|(X, Y, Z, U)\|^p_{\mathcal{S}^p \times \mathcal{B}^p} \;\leq\; C_p \;(1 + |X_0|^p) \;\;, \;\; p \geq 2 \;. \qquad (2.5)$$

The estimate on X is standard, see (5.2) of Lemma 5.1 in the Appendix. Plugging this in (5.6) of Lemma 5.2 leads to the bound on $\|(Y, Z, U)\|_{\mathcal{B}^p}$. Using (5.3) of Lemma 5.1, we also deduce that

$$\mathbb{E}\left[\sup_{s \leq u \leq t} |X_u - X_s|^p\right] \;\leq\; C_p \;(1 + |X_0|^p)\, |t - s| \;, \qquad (2.6)$$

while the previous estimates on X combined with (5.7) of Lemma 5.2 implies

$$\mathbb{E}\left[\sup_{s \leq u \leq t} |Y_u - Y_s|^p\right] \leq C_p \;\left\{(1 + |X_0|^p)\, |t - s|^p + \|Z\|^p_{\mathbf{H}^p_{[s,t]}} + \|U\|^p_{\mathbf{L}^p_{\lambda,[s,t]}}\right\} \;. \qquad (2.7)$$

2.2 Discrete time approximation

We first fix a regular grid $\pi := \{t_i := iT/n,\; i = 0, \ldots, n\}$ on $[0, T]$ and approximate X by its Euler scheme X^π defined by

$$\begin{cases} X_0^\pi & := \; X_0 \\ X_{t_{i+1}}^\pi & := \; X_{t_i}^\pi + \frac{T}{n} b(X_{t_i}^\pi) + \sigma(X_{t_i}^\pi)\Delta W_{i+1} + \int_E \beta(X_{t_i}^\pi, e)\bar{\mu}(de, (t_i, t_{i+1}]) \end{cases} \qquad (2.8)$$

where $\Delta W_{i+1} := W_{t_{i+1}} - W_{t_i}$. It is well known that

$$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |X_t - X^\pi_{t_i}|^2\right] \;\; \leq \;\; C^0_2 \, n^{-1} \, . \tag{2.9}$$

We then approximate (Y, Z, Γ) by $(\bar{Y}^\pi, \bar{Z}^\pi, \bar{\Gamma}^\pi)$ defined by the backward implicit scheme

$$\begin{aligned} \bar{Z}^\pi_t &:= \frac{n}{T}\,\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}\Delta W_{i+1} \mid \mathcal{F}_{t_i}\right] \;,\; \bar{\Gamma}^\pi_t := \frac{n}{T}\,\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}\int_E \rho(e)\bar{\mu}(de,(t_i,t_{i+1}]) \mid \mathcal{F}_{t_i}\right] \\ \bar{Y}^\pi_t &:= \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{T}{n}\,h\left(X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}\right) \end{aligned} \tag{2.10}$$

on each interval $[t_i, t_{i+1})$, where $\bar{Y}^\pi_{t_n} := g(X^\pi_{t_n})$. Observe that the resolution of the last equation in (2.10) may involve the use of a fixed point procedure. However, h being Lipschitz and multiplied by $1/n$, the approximation error can be neglected for large values of n.

Remark 2.2 The above backward scheme is a natural extension of the one considered in [7] in the case $\beta = 0$.
By the representation theorem, see e.g. Lemma 2.3 in [23], there exist two processes $Z^\pi \in \mathbf{H}^2$ and $U^\pi \in \mathbf{L}^2_\lambda$ satisfying

$$\bar{Y}^\pi_{t_{i+1}} - \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] \;\; = \;\; \int_{t_i}^{t_{i+1}} Z^\pi_s \cdot dW_s + \int_{t_i}^{t_{i+1}}\int_E U^\pi_s(e)\bar{\mu}(ds, de) \, .$$

Observe that $\bar{Z}^\pi$ and $\bar{\Gamma}^\pi$ defined in (2.10) satisfy

$$\bar{Z}^\pi_{t_i} \; = \; \frac{n}{T}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z^\pi_s ds \mid \mathcal{F}_{t_i}\right] \quad \text{and} \quad \bar{\Gamma}^\pi_{t_i} \; = \; \frac{n}{T}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma^\pi_s ds \mid \mathcal{F}_{t_i}\right] \tag{2.11}$$

and therefore coincide with the best $\mathbf{H}^2_{[t_i,t_{i+1}]}$-approximations of $(Z^\pi_t)_{t_i\leq t<t_{i+1}}$ and $(\Gamma^\pi_t)_{t_i\leq t<t_{i+1}} := (\int_E \rho(e)U^\pi_t(e)\lambda(de))_{t_i\leq t<t_{i+1}}$ by $\mathcal{F}_{t_i}$-measurable random variables (viewed as constant processes on $[t_i, t_{i+1})$).
Finally, observe that we can define Y^π on $[t_i, t_{i+1})$ by setting

$$Y^\pi_t \;\; := \;\; \bar{Y}^\pi_{t_i} - (t - t_i)\,h(X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}) + \int_{t_i}^t Z^\pi_s dW_s + \int_{t_i}^t\int_E U^\pi_s(e)\bar{\mu}(ds, de) \, .$$

2.3 Convergence of the approximation scheme

In this subsection, we show that the approximation error

$$\mathrm{Err_n}\,(Y, Z, U) \;\; := \;\; \left\{\sup_{t\leq T}\mathbb{E}\left[|Y_t - \bar{Y}^\pi_t|^2\right] + \|Z - \bar{Z}^\pi\|^2_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}^\pi\|^2_{\mathbf{H}^2}\right\}^{\frac{1}{2}}$$

converges to 0. Let us first introduce the processes $(\bar{Z}, \bar{\Gamma})$ defined on each interval $[t_i, t_{i+1})$ by

$$\bar{Z}_t := \frac{n}{T}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z_s \, ds \mid \mathcal{F}_{t_i}\right] \quad \text{and} \quad \bar{\Gamma}_t := \frac{n}{T}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma_s \, ds \mid \mathcal{F}_{t_i}\right] \, .$$

Remark 2.3 Observe that $\bar{Z}_{t_i}$ and $\bar{\Gamma}_{t_i}$ are the counterparts of $\bar{Z}^{\pi}_{t_i}$ and $\bar{\Gamma}^{\pi}_{t_i}$ for the original backward SDE. They can also be interpreted as the best $\mathbf{H}^2_{[t_i,t_{i+1}]}$-approximations of $(Z_t)_{t_i \le t < t_{i+1}}$ and $(\Gamma_t)_{t_i \le t < t_{i+1}}$ by $\mathcal{F}_{t_i}$-measurable random variables (viewed as constant processes on $[t_i, t_{i+1})$).

Proposition 2.1 *We have*

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Y_t - Y_{t_i}|^2\right] dt \le C_2^0\, n^{-1} \quad \textit{and} \quad \|Z - \bar{Z}\|_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2} \le \epsilon(n) \tag{2.12}$$

where $\epsilon(n) \to 0$ *as* $n \to \infty$.
Moreover,

$$\mathrm{Err}_n\,(Y, Z, U) \;\le\; C_2^0 \left(n^{-1/2} + \|Z - \bar{Z}\|_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2}\right)\,, \tag{2.13}$$

so that $\mathrm{Err}_n\,(Y, Z, U) \underset{n\to\infty}{\longrightarrow} 0$.

Proof. Since Y solves (2.3),

$$\mathbb{E}\left[|Y_t - Y_{t_i}|^2\right] \;\le\; C_2^0 \int_{t_i}^{t} \mathbb{E}\left[|h(X_r, Y_r, Z_r, \Gamma_r)|^2 + |Z_r|^2 + \int_E |U_r(e)|^2 \lambda(de)\right] dr\,.$$

Combining the Lipschitz property of h with (2.5), it follows that

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Y_t - Y_{t_i}|^2\right] dt \;\le\; \frac{C_2^0}{n}\,.$$

This is exactly the first part of (2.12). The proof of (2.13) then follows exactly the same arguments as in [7], see p 99 in [11] for details. It remains to prove the second part of (2.12). Since Z is $\mathbb{F}$-adapted, there is a sequence of adapted processes $(Z^n)_n$ such that $Z^n_t = Z^n_{t_i}$ on each $[t_i, t_{i+1})$ and Z^n converges to Z in $\mathbf{H}^2$. By Remark 2.3, we observe that $\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \le \|Z - Z^n\|^2_{\mathbf{H}^2}$, and applying the same reasoning to Γ concludes the proof. □

2.4 Path-regularity and convergence rate

In view of Proposition 2.1, the discretization error converges to zero. In order to control its speed of convergence, it remains to study $\|Z - \bar{Z}\|^2_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2}$. Before stating our main result, let us introduce the following assumption:

$\mathbf{H}$: For each $e \in E$, the map $x \in \mathbb{R}^d \mapsto \beta(x, e)$ admits a Jacobian matrix $\nabla\beta(x, e)$ such that the function

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

satisfies one of the following condition uniformly in $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d$

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

Remark 2.4 Observe for later use that the condition **H** implies that, for each $(x,e) \in \mathbb{R}^d \times E$, the matrix $\nabla\beta(x,e) + I_d$ is invertible with inverse bounded by K. This ensures the inversibility of the first variation process ∇X of X, see Remark 3.6. Moreover, if q is a smooth density on $\mathbb{R}^d$ with compact support, then the approximating functions β^k, $k \in \mathbb{N}$, defined by $\beta^k(x,e) := \int_{\mathbb{R}^d} k^d \beta(\bar{x},e) q(k[x-\bar{x}]) d\bar{x}$ are smooth and also satisfy **H**. In Section 5 of [9], the authors imposes a similar condition: $|\det(\nabla\beta(x,e) + I_d)| \le 1-\delta \;\;, \; \forall\, x \in \mathbb{R}^d \;\; \lambda(de) - a.e.$ for some $\delta > 0$. Under mild additional assumptions, this allows to prove the existence of a bounded solution, in a suitable weighted Sobolev space, to a PDE of the form (1.2) which can then be related to Y.

Our main theorem is stated for a suitable version of (Z, U, Γ). Observe that it does not change the quantity $\mathrm{Err}_n\,(Y,Z,U)$.

Theorem 2.1 *The following holds.*
(i) *For all $i < n$,*

$$\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |Y_t - Y_{t_i}|^2\right] \le C_2^0\, n^{-1} \quad and \quad \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\Gamma_t - \Gamma_{t_i}|^2\right] \le C_2^0\, n^{-1} \tag{2.14}$$

so that $\|\Gamma - \bar{\Gamma}\|^2_{\mathcal{S}^2} \le C_2^0\, n^{-1}$ and $\|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2} \le C_2^0\, n^{-1}$. Moreover, for any $\varepsilon > 0$,

$$\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \;\le\; C^0_\varepsilon\, n^{-1+\varepsilon}\,. \tag{2.15}$$

(ii) *Assume that **H** holds. Then*

$$\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \;\le\; C^0_2\, n^{-1}\,. \tag{2.16}$$

This regularity property will be proved in the subsequent sections. Combined with Proposition 2.1, it provides an upper bound for the convergence rate of our backward implicit scheme.

Corollary 2.1 *For any $\varepsilon > 0$, $\mathrm{Err}_n\,(Y,Z,U) \le C^0_\varepsilon\, n^{-1/2+\varepsilon}$. If **H** holds, then $\mathrm{Err}_n\,(Y,Z,U) \le C^0_2\, n^{-1/2}$.*

Remark 2.5 One could also use an explicit scheme as in e.g. [2] or [16]. In this case, (2.10) has to be replaced by

$$\begin{aligned}
\tilde{Z}^\pi_{t_i} &:= \frac{n}{T}\,\mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \Delta W_{i+1} \mid \mathcal{F}_{t_i}\right] \;,\; \tilde{\Gamma}^\pi_{t_i} := \frac{n}{T}\,\mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \int_E \rho(e)\bar{\mu}(de,(t_i,t_{i+1}]) \mid \mathcal{F}_{t_i}\right] \\
\tilde{Y}^\pi_{t_i} &:= \mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{T}{n}\,\mathbb{E}\left[h\left(X^\pi_{t_i}, \tilde{Y}^\pi_{t_{i+1}}, \tilde{Z}^\pi_{t_i}, \tilde{\Gamma}^\pi_{t_i}\right) \mid \mathcal{F}_{t_i}\right]
\end{aligned} \tag{2.17}$$

with the terminal condition $\tilde{Y}^\pi_{t_n} = g(X^\pi_{t_n})$. The advantage of this scheme is that it does not require a fixed point procedure. However, from a numerical point of view, adding a term in the conditional expectation defining $\tilde{Y}^\pi_{t_i}$ makes it more difficult to estimate. We therefore think that the implicit scheme may be more tractable in practice. The above convergence results can be easily extended to this scheme, see [11] for details.

Remark 2.6 In the unpublished paper [9], the authors discuss the regularity of (X, Y, Z, U) with respect to the initial condition X_0 in a case where the coefficients b, σ, h and g are C^3, with linear growth and bounded derivatives for the two first ones and derivatives having polynomial growth for the two last ones. Under these regularity assumptions, they show that the map $(t, x) \mapsto u(t, x) = Y_t^{t,x}$ belongs to $C^{0,2}([0, T] \times \mathbb{R}^d)$ with $\frac{1}{2}$-Hölder continuity in time and derivatives having polynomial growth in space, see their Proposition 3.5 and their Corollary 3.6. Similar results are obtained for $(t, x) \mapsto (Z_t^{t,x}, U_t^{t,x})$ which can be identified to $(\nabla u(t, x)\sigma(x), u(t, x + \beta(x, \cdot)) - u(t, x))$.
This readily implies the properties stated in Theorem 2.1 which can be seen as weak versions of the regularity results of [9]. The important point here is that:
1. Our results do not require all the regularity assumptions of [9];
2. This is all we need to provide the convergence rates of Corollary 2.1.

Remark 2.7 It will be clear from the proofs that all the results of this paper hold if we let the maps b, σ, β, and h depend on t whenever these functions are 1/2-Hölder in t and the other assumptions are satisfied uniformly in t. The Euler approximation X^π of X could also be replaced by any other adapted approximation satisfying (2.9).

Remark 2.8 We refer to [11] for extensions to the approximation of systems of semilinear PDEs through their relation with BSDEs with jumps, see [21], and for similar convergence results without **H** but under additionnal regularity assumptions.

3 Malliavin calculus for FBSDE

In this section, we prove that the solution (Y, Z, U) of (2.3) is smooth in the Malliavin sense under the additional assumptions

$$\mathbf{C}^X : \; b, \, \sigma \text{ and } \beta(\cdot, e) \text{ are } C_b^1 \text{ uniformly in } e \in E \quad , \quad \mathbf{C}^Y : \; g \text{ and } h \text{ are } C_b^1.$$

This will allow us to provide representation and regularity results for Y, Z and U in Section 4. Under $\mathbf{C}^X$-$\mathbf{C}^Y$, these results will immediately imply the first assertion of (i) of Theorem 2.1, while the second one (resp. (ii)) will be obtained by adapting the arguments of [6] (resp. [25] under the additional assumption **H**).

3.1 Generalities

The construction of the Malliavin derivatives on the Wiener space is standard, see e.g. [18], and can be easily extended to our setting by observing that there is an isometry between $L^2(\Omega_W \times \Omega_\mu)$ and $L^2(\Omega_W, L^2(\Omega_\mu))$, with obvious notations.
Let **S** denote the set of random variables of the form

$$F = \phi \left(\int_0^T f^1(t) \cdot dW_t, \ldots, \int_0^T f^\kappa(t) \cdot dW_t, \mu \right) ,$$

where $\kappa \geq 1$, $f^i : [0, T] \mapsto \mathbb{R}^d$ is a bounded measurable map for each $i \leq \kappa$, ϕ is a real-valued measurable map on $\mathbb{R}^\kappa \times \Omega_\mu$ and $\phi(\cdot, \eta) \in C_b^\infty$, $\mathbb{P}_\mu(d\eta)$-a.e.

We denote by D the Malliavin derivative operator with respect to the Brownian motion. For $F \in \mathbf{S}$ as above and $s \leq T$, it is defined as

$$D_s F \quad := \quad \sum_{i \leq \kappa} \nabla_i \phi \left(\int_0^T f^1(t) \cdot dW_t, \ldots, \int_0^T f^\kappa(t) \cdot dW_t, \mu \right) f^i(s) \ ,$$

where $\nabla_i \phi$ is the derivative of ϕ with respect to its i-th argument.
We then denote by $\mathbb{D}^{1,2}$ the closure of $\mathbf{S}$ with respect to the norm

$$\|F\|_{\mathbb{D}^{1,2}} \quad := \quad \left\{ \mathbb{E}\left[F^2\right] + \mathbb{E}\left[\int_0^T |D_s F|^2 ds\right] \right\}^{\frac{1}{2}} \ ,$$

and define $\mathbf{H}^2(\mathbb{D}^{1,2})$ as the set of elements $\xi \in \mathbf{H}^2$ such that $\xi_t \in \mathbb{D}^{1,2}$ for almost all $t \leq T$ and such that, after possibly passing to a measurable version,

$$\|\xi\|^2_{\mathbf{H}^2(\mathbb{D}^{1,2})} \quad := \quad \|\xi\|^2_{\mathbf{H}^2} + \int_0^T \|D_s \xi\|^2_{\mathbf{H}^2} ds \ < \ \infty \ .$$

Observe that for ψ in $\mathbf{L}^2_\lambda(\mathbb{F}^\mu)$, the set of elements of $\mathbf{L}^2_\lambda$ which are independent of W, we have $D\psi = 0$. We finally define $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$ as the closure of the set

$$\mathbf{L}'^2_\lambda(\mathbb{D}^{1,2}) \quad := \quad \text{Vect}\left\{ \psi = \xi\vartheta \ : \ \xi \in \mathbf{H}^2_{\mathcal{P}}(\mathbb{D}^{1,2}, \mathbb{F}^W), \ \vartheta \in \mathbf{L}^2_\lambda(\mathbb{F}^\mu), \ \|\psi\|_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} < \infty \right\}$$

for the norm

$$\|\psi\|^2_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} \quad := \quad \|\psi\|^2_{\mathbf{L}^2_\lambda} + \int_0^T \|D_s \psi\|^2_{\mathbf{L}^2_\lambda} ds \ .$$

Here, $\mathbf{H}^2_{\mathcal{P}}(\mathbb{D}^{1,2}, \mathbb{F}^W)$ denotes the set of $\mathbb{F}^W$-predictable elements of $\mathbf{H}^2(\mathbb{D}^{1,2})$ and $D_s(\xi\vartheta) = (D_s\xi)\vartheta$ for $\xi \in \mathbf{H}^2_{\mathcal{P}}(\mathbb{D}^{1,2}, \mathbb{F}^W)$, $\vartheta \in \mathbf{L}^2_\lambda(\mathbb{F}^\mu)$. Here again, we extend the definition of $\| \cdot \|_{\mathbf{H}^2(\mathbb{D}^{1,2})}$ and $\| \cdot \|_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})}$ to processes with values in $\mathbb{M}^d$ and $\mathbb{R}^d$ in a natural way.
From now on, given a matrix A, we shall denote by A^i its i-th column. For $k \leq d$, we denote by D^k the Malliavin derivative with respect to W^k, meaning that $D^k F = (DF)^k$ for $F \in \mathbb{D}^{1,2}$.

Remark 3.1 With this construction, the operator D enjoys the usual properties of the Malliavin derivative operator on Wiener spaces. In particular, if $\xi \in \mathbf{H}^2(\mathbb{D}^{1,2})$ and $f \in C^1_b(\mathbb{R}^d)$, then

$$D_s \left(\int_0^T f(\xi_t) dt + \int_0^T \xi_t \cdot dW_t \right) = \int_s^T \nabla f(\xi_t) D_s \xi_t dt + \xi^*_s + \sum_{j=1}^d \int_s^T D_s \xi^j_t \cdot dW^j_t$$

for all $s \leq T$. Here * denotes transposition. It follows from the same argument as in [18], which we refer to for more details.

Remark 3.2 Fix $\xi \in \mathbf{H}^2_{\mathcal{P}}(\mathbb{D}^{1,2}, \mathbb{F}^W)$. By Lemma 1.3.1 in [18], there exists a family of deterministic measurable kernels $f_m(t_1, \ldots, t_m, t)$ in $L^2([0,T]^{m+1})$, $m \geq 0$, such that $\xi_t = \sum_{m\geq 0} I_m(f_m(\cdot, t))$ and $D_s\xi_t = \sum_{m\geq 1} m I_{m-1}(f_m(\cdot, s, t))$ where I_m denotes the m-iterated Wiener integral, see Proposition 1.2.1 in [18]. Therefore, if τ is a random time bounded by T and independent of W, we have $\xi_\tau = \sum_{m\geq 0} I_m(f_m(\cdot, \tau))$ and, by the same argument as in the proof of Proposition 1.2.1 in [18], $\xi_\tau \in \mathbb{D}^{1,2}$ whenever τ has a bounded density and $D_s(\xi_\tau) = \sum_{m\geq 1} m I_{m-1}(f_m(\cdot, s, \tau)) = (D_s\xi)_\tau$.

Remark 3.3 For $\xi \in \mathbf{H}^2(\mathbb{D}^{1,2})$ the integrals $\int_0^T D_s\xi_t dt$ and $\int_0^T D_s\xi_t dW_t$ have to be understood as integrals written on the process $D_s\xi$, i.e. $\int_0^T (D_s\xi)_t dt$ and $\int_0^T (D_s\xi)_t dW_t$. Similarly, for $\psi \in \mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$, $\int_0^T \int_E D_s\psi_t(e)\bar{\mu}(de, dt)$ has to be understood as
$\int_0^T \int_E (D_s\psi)_t(e)\bar{\mu}(de, dt)$. However, it follows from Remark 3.2 that it coincides with $\int_0^T \int_E D_s(\psi_t(e))\bar{\mu}(de, dt)$, so that there is no ambiguity.

The two following Lemmas are generalizations of Lemma 3.3 and Lemma 3.4 in [21] which correspond to the case where E is finite, see also Lemma 2.3 in [20] for the case of Itô integrals.

Lemma 3.1 *Assume that* $\psi \in \mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$*. Then,* $H := \int_0^T \int_E \psi_t(e)\bar{\mu}(de, dt) \in \mathbb{D}^{1,2}$ *and* $D_sH = \int_0^T \int_E D_s\psi_t(e)\bar{\mu}(de, dt)$ *for all* $s \leq T$.

Proof. First notice that it suffices to prove the required result when $\psi \in \mathbf{L}'^2_\lambda(\mathbb{D}^{1,2})$. Indeed, we can then retrieve the general case by considering a sequence $(\psi^n)_n$ in $\mathbf{L}'^2_\lambda(\mathbb{D}^{1,2})$ which converges to ψ in $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$, so that $H^n := \int_0^T \int_E \psi^n_t(e)\bar{\mu}(de, dt)$ is a Cauchy sequence in $\mathbb{D}^{1,2}$ which converges to H and $(D_sH^n)_{s\leq T}$ converges to $(\int_0^T \int_E D_s\psi_t(e)\bar{\mu}(de, dt)))_{s\leq T}$ in $\mathbf{H}^2$.
We therefore assume that $\psi = \xi\vartheta$ where $\xi \in \mathbf{H}^2_{\mathcal{P}}(\mathbb{D}^{1,2}, \mathbb{F}^W)$, $\vartheta \in \mathbf{L}^2_\lambda(\mathbb{F}^\mu)$ and $\|\psi\|_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} < \infty$. Then,

$$\int_0^T \int_E \psi_t(e)\bar{\mu}(de, dt) = \int_0^T \int_E \xi_t\vartheta_t(e)\mu(de, dt) - \int_0^T \xi_t \int_E \vartheta_t(e)\lambda(de)dt \ ,$$

where, by Remark 3.1 and the fact that $\int_E \vartheta_t(e)\lambda(de)$ is independent of W,

$$\begin{aligned} D_s \int_0^T \xi_t \left(\int_E \vartheta_t(e)\lambda(de) \right) dt &= \int_0^T D_s\xi_t \int_E \vartheta_t(e)\lambda(de)dt \\ &= \int_0^T \int_E (D_s\xi_t)\vartheta_t(e)\lambda(de)dt \ . \end{aligned}$$

It remains to prove that

$$D_s \int_0^T \int_E \xi_t\vartheta_t(e)\mu(de, dt) = \int_0^T \int_E (D_s\xi_t)\vartheta_t(e)\mu(de, dt) \ .$$

To see this, we define N by $N_t := \int_0^t \mu(E, ds)$ for $t \le T$, $(\tau_i)_{i\ge1}$ as the sequence of jump times of N and $(\mathcal{E}_i)_{i\ge1}$ by $\mathcal{E}_i := N_{\tau_i} - N_{\tau_i-}$. With these notations, we have to show that

$$D_s \sum_{i\ge1} \xi_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T} \;=\; \sum_{i\ge1} (D_s\xi)_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}\;, \tag{3.1}$$

see Remark 3.3. Using Remark 3.2, we now oberve that $D_s \sum_{i=1}^n \xi_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}$ $= \sum_{i=1}^n (D_s\xi)_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}$, for each $n \ge 1$. Passing to the limit leads to (3.1) and concludes the proof. Indeed, the previous identity implies that the sequence $(F_n)_{n\ge1}$ defined by $F_n := \sum_{i=1}^n \xi_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}$ belongs to $\mathbf{H}^2(\mathbb{D}^{1,2})$ and satisfies $|F_n| \le \sum_{i\ge1} |\xi_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)|\,\mathbf{1}_{\tau_i\le T}$ as well as
$|D_sF_n| \le \sum_{i\ge1} |(D_s\xi)_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)|\,\mathbf{1}_{\tau_i\le T}$ which implies that $\sup_{n\ge1}\|F_n\|^2_{\mathbf{H}^2(\mathbb{D}^{1,2})}$ is bounded by

$$\begin{aligned}
&2\left(\mathbb{E}\left[\left|\int_0^T\int_E |\xi_t\vartheta_t(e)|\,\bar\mu(de,dt)\right|^2\right] + \mathbb{E}\left[\left|\int_0^T\int_E |\xi_t\vartheta_t(e)|\,\lambda(de)dt\right|^2\right]\right)\\
&+2\int_0^T \mathbb{E}\left[\left|\int_0^T\int_E |D_s\xi_t\vartheta_t(e)|\,\bar\mu(de,dt)\right|^2\right]ds\\
&+2\int_0^T \mathbb{E}\left[\left|\int_0^T\int_E |D_s\xi_t\vartheta_t(e)|\,\lambda(de)dt\right|^2\right]ds\\
&\le C_2\|\psi\|^2_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} < \infty\;,
\end{aligned}$$

where the second and the fourth terms on the right hand-side are bounded by using Jensen's inequality and the assumption $\lambda(E) < \infty$. Moreover, by dominated convergence, $\mathbb{E}[|F_n - F|^2] + \int_0^T \mathbb{E}[\|D_sF_n - D_sF\|^2]ds \to 0$ where $F := \sum_{i\ge1} \xi_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}$ and $D_sF := \sum_{i\ge1}(D_s\xi)_{\tau_i}\vartheta_{\tau_i}(\mathcal{E}_i)\mathbf{1}_{\tau_i\le T}$ satisfy, by the same arguments as above, $\|F\|^2_{\mathbf{H}^2(\mathbb{D}^{1,2})} \le C_2\|\psi\|^2_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} < \infty$. □

Remark 3.4 Similar arguments as in the above proof shows that for $\psi \in \mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$ and $f \in L^\infty(E)$, we have, for almost every $s \le T$, $\int_E \psi_s(e)f(e)\lambda(de) \in \mathbb{D}^{1,2}$ and

$$D_t\left(\int_E \psi_s(e)f(e)\lambda(de)\right) \;:=\; \int_E D_t\psi_s(e)f(e)\lambda(de)\;.$$

Lemma 3.2 *Let $\mathbf{S}(W)$ denote the set of random variables of the form*

$$H^W \;=\; \phi\left(\int_0^T f^1(t)\cdot dW_t, \ldots, \int_0^T f^\kappa(t)\cdot dW_t\right)$$

where $\kappa \ge 1$, $\phi \in C_b^\infty$ and $f^i : [0,T] \mapsto \mathbb{R}^d$ is a bounded measurable map for each $i \le \kappa$. Then, $\mathrm{Vect}\{\mathbf{S}(W) \times L^\infty(\Omega_\mu, \mathcal{F}^\mu_T)\}$ is dense in $\mathbb{D}^{1,2}$ for the norm $\|\cdot\|_{\mathbb{D}^{1,2}}$.

Proof. It suffices to prove that $\text{Vect}\{\mathbf{S}(W)\times L^\infty(\Omega_\mu,\mathcal{F}^\mu_T)\}$ is dense in $\mathbf{S}$. Fix $H\in\mathbf{S}$ of the form

$$H \;=\; \phi\left(\int_0^T f^1(t)\cdot dW_t,\ldots,\int_0^T f^\kappa(t)\cdot dW_t,\mu\right)\;.$$

Observe that Ω_μ can be identified to the space of finite (possibly empty) sequences $(t_i,e_i)_{1\le i\le n}$ of $[0,T]\times E$, $n\ge 0$, such that $(t_i)_{i\ge1}$ is increasing. Let $\mathcal{G}_n$ denote the set of such sequences of length $n\ge0$ and $\mathcal{G}:=\cup_{n\ge0}\mathcal{G}_n$. Given $\eta\in\Omega_u$, we denote by $(t^\eta_i,e^\eta_i)_{i\ge1}$ the associated sequence, and we identify ϕ with a measurable map on $\mathbb{R}^\kappa\times\mathcal{G}$. We denote by ϕ_n its restriction to $\mathbb{R}^\kappa\times\mathcal{G}_n$, $n\ge0$. Let ψ_n denote the gradient of ϕ_n with respect to its first κ components and set $f:=(f^1,\ldots,f^\kappa)$, $G:=\left(\int_0^T f^1(t)\cdot dW_t,\ldots,\int_0^T f^\kappa(t)\cdot dW_t\right)$. Since

$$(H,D_sH) \;=\; \sum_{n\ge0}\left(\phi_n\left(G,(t^\mu_i,e^\mu_i)_{1\le i\le n}\right),\psi_n\left(G,(t^\mu_i,e^\mu_i)_{1\le i\le n}\right)\cdot f(s)\right)\mathbf{1}_{\mu(E,[0,T])=n}\;,$$

it suffices to prove that each $H_n:=\phi_n\left(G,(t^\mu_i,e^\mu_i)_{1\le i\le n}\right)$ can be approximated by linear combinations of elements of $\mathbf{S}(W)\times L^\infty(\Omega_\mu,\mathcal{F}^\mu_T)$. Moreover, we can always assume that ϕ_n is C^∞_b on $\mathbb{R}^\kappa\times\mathcal{G}_n$. Indeed, ϕ is already C^∞_b in its first κ components, a.e., and we can replace ϕ_n by its convolution with a sequence of smooth kernels acting only its last n components. Since both functions are continuous, we can then approximate (ϕ_n,ψ_n) pointwise by linear combinations of functions of the form $(\phi_n,\psi_n)(\cdot,(t_i,e_i)_{1\le i\le n})\mathbf{1}_A$ where A is a Borel set of $\mathcal{G}_n$ and $(t_i,e_i)_{1\le i\le n}\in\mathcal{G}_n$. The required result then follows from the fact that $D_s\phi_n\left(G,(t_i,e_i)_{1\le i\le n}\right)\mathbf{1}_A((t^\mu_i,e^\mu_i)_{1\le i\le n})$ $=\left(\psi_n\left(G,(t_i,e_i)_{1\le i\le n}\right)\cdot f(s)\right)\mathbf{1}_A((t^\mu_i,e^\mu_i)_{1\le i\le n})$. □

Lemma 3.3 *Fix $(\xi,\psi)\in\mathbf{H}^2\times\mathbf{L}^2_\lambda$ and assume that*

$$H \;:=\; \int_0^T \xi_t\cdot dW_t+\int_0^T\int_E \psi_t(e)\bar\mu(de,dt)\;\in\;\mathbb{D}^{1,2}\;.$$

Then, $(\xi,\psi)\in\mathbf{H}^2(\mathbb{D}^{1,2})\times\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$ and

$$D_sH \;:=\; \xi^*_s+\int_0^T\sum_{i=1}^d D_s\xi^i_t\,dW^i_t+\int_0^T\int_E D_s\psi_t(e)\bar\mu(de,dt)\;,$$

where ξ^* denotes the transpose of ξ.

Proof. One easily deduce from Lemma 3.2 that $\mathcal{H}:=\text{Vect}\{H^WH^{\bar\mu}\;:\;H^W\in\mathbf{S}(W)\;,$ $H^{\bar\mu}\in L^\infty(\Omega_\mu,\mathcal{F}^\mu_T)\;,\;\mathbb{E}\left[H^WH^{\bar\mu}\right]=0\}$ is dense in $\mathbb{D}^{1,2}\cap\{H\in L^2(\Omega,\mathcal{F},\mathbb{P})\;:\;\mathbb{E}[H]=0\}$ for $\|\cdot\|_{\mathbb{D}^{1,2}}$. Thus, it suffices to prove the result for H of the form $H^WH^{\bar\mu}$ where $H^W\in\mathbf{S}(W)$, $H^{\bar\mu}\in L^\infty(\Omega_\mu,\mathcal{F}^\mu_T)$ and $\mathbb{E}\left[H^WH^{\bar\mu}\right]=0$. By the representation theorem, see e.g. [8], there exists $\psi\in\mathbf{L}^2_\lambda$ such that $H^{\bar\mu}=\mathbb{E}[H^{\bar\mu}]+\int_0^T\int_E\psi_t(e)\bar\mu(de,dt)$ and by Ocone's formula, see e.g. Proposition 1.3.5 in [18], $H^W=\mathbb{E}\left[H^W\right]+\int_0^T\mathbb{E}\left[D_tH^W\mid\mathcal{F}^W_t\right]dW_t$. Thus it follows from Itô's Lemma that H $=\int_0^T H^{\bar\mu}_t\,\mathbb{E}\left[D_tH^W\mid\mathcal{F}^W_t\right]dW_t+\int_0^T\int_E H^W_t\psi_t(e)\bar\mu(de,dt)$ where $H^{\bar\mu}_t=\mathbb{E}\left[H^{\bar\mu}\mid\mathcal{F}_t\right]$ and $H^W_t=\mathbb{E}\left[H^W\mid\mathcal{F}_t\right]$. Furthermore, easy computations show that the two integrands belong respectively to $\mathbf{H}^2(\mathbb{D}^{1,2})$ and $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$. Thus, Remark 3.1 and Lemma 3.1 conclude the proof. □

3.2 Malliavin calculus on the Forward SDE

In this section, we recall well-known properties concerning the differentiability in the Malliavin sense of the solution of a Forward SDE. In the case where $\beta = 0$ the following result is stated in e.g. [18]. The extension to the case $\beta \neq 0$ is easily obtained by conditioning by μ, see e.g. [14] for explanations in the case where E is finite, or by combining Remark 3.1, Lemma 3.1 with a fixed point procedure as in the proof of Theorem 2.2.1. in [18], see also Proposition 3.2 below.

Proposition 3.1 *Assume that* $\mathbf{C}^X$ *holds, then* $X_t \in \mathbb{D}^{1,2}$ *for all* $t \leq T$. *For all* $s \leq T$ *and* $k \leq d$, $D^k_s X$ *admits a version* $\chi^{s,k}$ *which solves on* $[s,T]$

$$\begin{aligned} \chi^{s,k}_t &= \sigma^k(X_{s-}) + \int_s^t \nabla b(X_r)\chi^{s,k}_r dr + \int_s^t \sum_{j=1}^d \nabla\sigma^j(X_r)\chi^{s,k}_r dW^j_r \\ &+ \int_s^t \int_E \nabla\beta(X_{r-},e)\chi^{s,k}_{r-}\bar{\mu}(dr,de)\ . \end{aligned}$$

Remark 3.5 Fix $p \geq 2$ and $r \leq s \leq t \leq u \leq T$. Under $\mathbf{C}^X$, it follows from Lemma 5.1 applied to X and χ^s that $\|\chi^s\|^p_{\mathcal{S}^p} \leq C_p\ (1+|X_0|^p)$, $\mathbb{E}\left[|\chi^s_u - \chi^s_t|^p\right] \leq C_p\ |u-t|\,(1+|X_0|^p)$ and $\|\chi^s - \chi^r\|^p_{\mathcal{S}^p} \leq C_p\ |s-r|\,(1+|X_0|^p)$.

Remark 3.6 Under $\mathbf{C}^X$, we can define the first variation process ∇X of X which solves on $[0,T]$

$$\begin{aligned} \nabla X_t &= I_d + \int_0^t \nabla b(X_r)\nabla X_r dr + \int_0^t \sum_{j=1}^d \nabla\sigma^j(X_r)\nabla X_r dW^j_r \\ &+ \int_0^t \int_E \nabla\beta(X_{r-},e)\nabla X_{r-}\bar{\mu}(dr,de)\ . \end{aligned} \quad (3.2)$$

Moreover, under $\mathbf{H}$, see Remark 2.4, $(\nabla X)^{-1}$ is well defined and solves on $[0,T]$

$$\begin{aligned} (\nabla X)^{-1}_t &= I_d - \int_0^t (\nabla X)^{-1}_r \left[\nabla b(X_r) - \sum_{j=1}^d \nabla\sigma^j(X_r)\nabla\sigma^j(X_r)\right] dr \\ &+ \int_0^t (\nabla X)^{-1}_r \int_E \nabla\beta(X_r,e)\lambda(de)dr - \int_0^t \sum_{j=1}^d (\nabla X)^{-1}_r \nabla\sigma^j(X_r)dW^j_r \\ &- \int_0^t \int_E (\nabla X)^{-1}_{r-}\left(\nabla\beta(X_{r-},e) + I_d\right)^{-1}\nabla\beta(X_{r-},e)\mu(de,dr)\ . \end{aligned} \quad (3.3)$$

Remark 3.7 Fix $p \geq 2$. Under $\mathbf{H}$-$\mathbf{C}^X$, it follows from Remark 2.4 and Lemma 5.1 applied to ∇X and $(\nabla X)^{-1}$ that $\|\nabla X\|_{\mathcal{S}^p} + \|(\nabla X)^{-1}\|_{\mathcal{S}^p} \leq C_p$.

Remark 3.8 Assume that $\mathbf{H}$-$\mathbf{C}^X$ holds and observe that $\chi^s = (\chi^{s,k})_{k\leq d}$ and ∇X solve the same equation up to the condition at time s. By uniqueness of the solution on $[t,T]$, it follows that $\chi^s_r = \nabla X_r(\nabla X_{s-})^{-1}\sigma(X_{s-})\mathbf{1}_{s\leq r}$ for all $s, r \leq T$.

3.3 Malliavin calculus on the Backward SDE

In this section, we generalize the result of Proposition 3.1 in [21]. Let us denote by $\mathcal{B}^2(\mathbb{D}^{1,2})$ the set of triples $(Y,Z,U) \in \mathcal{B}^2$ such that $Y_t \in \mathbb{D}^{1,2}$ for all $t \le T$ and $(Z,U) \in \mathbf{H}^2(\mathbb{D}^{1,2}) \times \mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$.

Proposition 3.2 *Assume that* $\mathbf{C}^X$-$\mathbf{C}^Y$ *holds. Then, the triples* (Y,Z,U) *belongs to* $\mathcal{B}^2(\mathbb{D}^{1,2})$. *For each* $s \le T$ *and* $k \le d$, *the equation*

$$\Upsilon^{s,k}_t \;=\; \nabla g(X_T)\chi^{s,k}_T + \int_t^T \nabla h(\Theta_r)\Phi^{s,k}_r dr - \int_t^T \zeta^{s,k}_r \cdot dW_r - \int_t^T \int_E V^{s,k}_r(e)\bar{\mu}(de,dr) \tag{3.4}$$

with $\Phi^{s,k} := (\chi^{s,k}, \Upsilon^{s,k}, \zeta^{s,k}, \Gamma^{s,k})$ *and* $\Gamma^{s,k} := \int_E \rho(e)V^{s,k}(e)\lambda(de)$, *admits a unique solution. Moreover,* $(\Upsilon^{s,k}_t, \zeta^{s,k}_t, V^{s,k}_t)_{s,t\le T}$ *is a version of* $(D^k_s Y_t, D^k_s Z_t, D^k_s U_t)_{s,t\le T}$.

Proof. With the help of Lemma 3.3 and the estimates of Lemma 5.2, we can reproduce exactly the proof of Proposition 5.3 in [12] up to minor modifications. See [11] p 113 for details. □

Proposition 3.3 *Assume that* $\mathbf{C}^X$-$\mathbf{C}^Y$ *holds. For each* $k \le d$, *the equation*

$$\nabla Y^k_t \;=\; \nabla g(X_T)\nabla X^k_T + \int_t^T \nabla h(\Theta_r)\nabla\Phi^k_r dr - \int_t^T \nabla Z^k_r \cdot dW_r - \int_t^T \int_E \nabla U^k_r(e)\bar{\mu}(de,dr) \tag{3.5}$$

with $\nabla\Phi^k = (\nabla X^k, \nabla Y^k, \nabla Z^k, \nabla\Gamma^k)$ *and* $\nabla\Gamma^k := \int_E \rho(e)\nabla U^k(e)\lambda(de)$, *admits a unique solution* $(\nabla Y^k, \nabla Z^k, \nabla U^k)$. *Moreover, there is a version of* $(\zeta^{s,k}_t, \Upsilon^{s,k}_t, V^{s,k}_t)_{s,t\le T}$ *given by* $\{(\nabla Y_t, \nabla Z_t, \nabla U_t)(\nabla X_{s-})^{-1}\sigma^k(X_{s-})\mathbf{1}_{s\le t}\}_{s,t\le T}$ *where* ∇Y_t *is the matrix whose* k*-column is given by* ∇Y^k_t *and* ∇Z_t,∇U_t *are defined similarly.*

Proof. In view of Proposition 3.2 and Remark 3.8, this follows immediately from the uniqueness of the solution of (3.4). □

Remark 3.9 Combined with $\mathbf{C}^X$-$\mathbf{C}^Y$ and Remark 3.5, Lemma 5.2 below implies that

$$\sup_{s\le T} \|(\Upsilon^s, \zeta^s, V^s)\|^p_{\mathcal{B}^p} \le C_p \;(1+|X_0|^p) \quad \text{for all} \quad p \ge 2\,. \tag{3.6}$$

It follows from Lemma 5.2 and Remark 3.7 that

$$\|(\nabla Y, \nabla Z, \nabla U)\|_{\mathcal{B}^p} \;\le\; C_p \qquad \text{for all} \quad p \ge 2\,. \tag{3.7}$$

4 Representation results and path regularity for the BSDE

In this section, we use the above results to obtain some regularity for the solution of the BSDE (2.3) under $\mathbf{C}^X$-$\mathbf{C}^Y$, $\mathbf{C}^X$-$\mathbf{C}^Y$-$\mathbf{H}$. Similar results without $\mathbf{C}^X$-$\mathbf{C}^Y$ will then be obtained by using an approximation argument.

Fix $(u, s, t, x) \in [0, T]^3 \times \mathbb{R}^d$ and $k, \ell \leq d$. In the sequel, we shall denote by $X(t, x)$ the solution of (2.2) on $[t, T]$ with initial condition $X(t, x)_t = x$, and by $(Y(t, x), Z(t, x), U(t, x))$ the solution of (2.3) with $X(t, x)$ in place of X. We define similarly $(\Upsilon^{s,k}(t, x),\ \zeta^{s,k}(t, x),\ V^{s,k}(t, x))$ and $(\nabla Y(t, x), \nabla Z(t, x), \nabla U(t, x))$. Observe that, with these notations, we have $(X(0, X_0), Y(0, X_0),\ Z(0, X_0),\ U(0, X_0)) = (X, Y, Z,\ U)$.

4.1 Representation

We start this section by proving useful bounds for the (deterministic) maps defined on $[0, T] \times \mathbb{R}^d$ by $u(t, x) := Y(t, x)_t$, $\nabla u(t, x) := \nabla Y(t, x)_t$, $v^{s,k}(t, x) := \Upsilon^{s,k}(t, x)_t$, where $s \in [0, T]$ and $k \leq d$.

Proposition 4.1 *Assume that* $\mathbf{C}^X$ *and* $\mathbf{C}^Y$ *hold, then,*

$$|u(t, x)| + |v^{s,k}(t, x)| \ \leq\ C_2\,(1 + |x|) \quad \text{and} \quad |\nabla u(t, x)| \ \leq\ C_2 \tag{4.1}$$

for all $s, t \leq T$, $k \leq d$ *and* $x \in \mathbb{R}^d$.

Proof. When $(t, x) = (0, X_0)$, the result follows from (2.5) in Remark 2.1, (3.6) and (3.7). The general case is obtained similarly by changing the initial condition on X. □

Proposition 4.2 *Assume that* $\mathbf{C}^X$ *and* $\mathbf{C}^Y$ *hold. Then, there is a version of* Z *given by* $(\Upsilon^t_t)_{t \leq T}$ *which satisfies* $\|Z\|^p_{\mathcal{S}^p} \leq C_p\,(1 + |X_0|^p)$.

Proof. Here again we only consider the case $d = 1$ and omit the indexes k, ℓ. By Proposition 3.2, (Y, Z, U) belongs to $\mathcal{B}^2(\mathbb{D}^{1,2})$ and it follows from Lemma 3.3 that

$$D_s Y_t = Z_s - \int_s^t \nabla h(\Theta_r) D_s \Theta_r dr + \int_s^t D_s Z_r\, dW_r + \int_s^t D_s U_r(e) \bar{\mu}(de, dr)\ , \tag{4.2}$$

for $0 < s \leq t \leq T$. Taking $s = t$ leads to the representation of Z. Thus, after possibly passing to a suitable version, we have $Z_t = D_t Y_t = \Upsilon^t_t$. By uniqueness of the solution of (2.2)-(2.3)-(3.4) for any initial condition in $L^2(\Omega, \mathcal{F}_t)$ at t, we have $\Upsilon^t_t = v^t(t, X_t)$. The bound on Z then follows from Proposition 4.1 combined with (2.5) of Remark 2.1. □

Proposition 4.3 (i) *Define* $\tilde{U}$ *by* $\tilde{U}_t(e) := u\,(t, X_{t-} + \beta(X_{t-}, e)) - \lim_{r \uparrow t} u\,(r, X_r) = Y_t - Y_{t-}$. *Then* $\tilde{U}$ *is a version of* U *and it satisfies*

$$\|\sup_{e \in E} |\tilde{U}(e)|\ \|^p_{\mathcal{S}^p} \ \leq\ C_p\,(1 + |X_0|^p)\ . \tag{4.3}$$

(ii) *Assume that* $\mathbf{C}^X$ *and* $\mathbf{C}^Y$ *hold. Define* $\nabla\tilde{U}$ *by* $\nabla\tilde{U}_t(e) := \nabla u\,(t, X_{t-} + \beta(X_{t-}, e)) - \lim_{r \uparrow t} \nabla u\,(r, X_r)$. *Then* $\nabla\tilde{U}$ *is a version of* ∇U *and it satisfies*

$$\|\sup_{e \in E} |\nabla\tilde{U}(e)|\ \|^p_{\mathcal{S}^p} \ \leq\ C_p\ . \tag{4.4}$$

Remark 4.1 We will see in Proposition 4.4 below that u is continuous under $\mathbf{C}^X$ and $\mathbf{C}^Y$ so that $\tilde{U}_t(e) = u\left(t, X_{t-} + \beta(X_{t-}, e)\right) - u\left(t, X_{t-}\right)$. A similar representation is derived in [21], in a case where E is finite, and in [9], in the case where E is not finite, but under more regularity assumptions on the coefficients.

Proof of Proposition 4.3. We only provide the proof of (i), the other assertion is proved similarly.

1. By uniqueness of the solution of (2.2)-(2.3) for any initial condition in $L^2(\Omega, \mathcal{F}_t)$ at time t, one has $Y_t = u(t, X_t)$ a.s. for each $t \le T$. We shall prove in step 2. below that u is jointly continuous in x and right-continuous in t. This implies that $(u(t, X_t))_{t \le T}$ is right-continuous so that $Y_t = u(t, X_t)$ and $Y_{t-} = \lim_{r \uparrow t} u(r, X_r)$ for each $t \le T$ a.s., see Theorem I.2 in [22] and recall that X and Y are càdlàg. Thus $\int_E U_t(e)\mu(de, \{t\}) = Y_t - Y_{t-} = u(t, X_t) - \lim_{r \uparrow t} u(r, X_r) = \int_E \tilde{U}_t(e)\mu(de, \{t\})$, for each $t \le T$ a.s. and $\int_0^T \int_E \left|\tilde{U}_t(e) - U_t(e)\right|^2 \mu(de, dt) = 0$, which implies, by taking expectation, $\left\|\tilde{U}_t(e) - U_t(e)\right\|_{\mathbf{L}^2_\lambda} = 0$.

2. We now prove that u is continuous in x and right-continuous on t. Fix $0 \le t_1 \le t_2 \le T$ and $(x_1, x_2) \in \mathbb{R}^{2d}$. For A denoting X,Y,Z or U, we set $A^i := A(t_i, x_i)$ for $i = 1, 2$ and $\delta A := A^1 - A^2$. By (5.4) of Lemma 5.1, we derive

$$\|\delta X\|^2_{\mathcal{S}^2_{[t_2,T]}} \le C_2 \left\{|x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1|\right\} . \tag{4.5}$$

Plugging this estimate in (5.8) of Lemma 5.2 leads to

$$\|(\delta Y, \delta Z, \delta U)\|^2_{\mathcal{B}^2_{[t_2,T]}} \le C_2 \left\{|x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1|\right\} . \tag{4.6}$$

Now, observe that

$$|u(t_1, x_1) - u(t_2, x_2)|^2 = |Y^1_{t_1} - Y^2_{t_2}|^2 \le C_2 \, \mathbb{E}\left[\left|Y^1_{t_2} - Y^1_{t_1}\right|^2 + \left|Y^1_{t_2} - Y^2_{t_2}\right|^2\right] \tag{4.7}$$

Since Y^1 is right-continuous and bounded in $\mathcal{S}^2$, the first term on the right-hand side goes to 0 as $t_2 \to t_1$, while the second is controlled by (4.6). □

4.2 Path regularity

Proposition 4.4 *Assume that* $\mathbf{C}^X$ *and* $\mathbf{C}^Y$ *hold. Then, for* $0 \le t_1 \le t_2 \le T$ *and* $(x_1, x_2) \in \mathbb{R}^{2d}$, $|u(t_1, x_1) - u(t_2, x_2)|^2 \le C_2 \left\{(1 + |x_1|^2)\,|t_2 - t_1| + |x_1 - x_2|^2\right\}$.

Proof. It suffices to plug the estimate of Proposition 4.2 and (4.3) in (2.7), which is possible since the norms in (2.7) do not change after passing to suitable versions, and appeal to (4.6) and (4.7). □

Remark 4.2 A similar result is obtained in [21] when λ has a finite support. The continuity of u is proved in [1] in a case where h is bounded, see also [9].

Corollary 4.1 *Assume that* $\mathbf{C}^X$ *and* $\mathbf{C}^Y$ *hold. Then, there is a version of* U *such that for all* $s \le t \le T$

$$\mathbb{E}\left[\sup_{r\in[s,t]}|Y_r - Y_s|^2\right] + \mathbb{E}\left[\sup_{e\in E}\sup_{r\in[s,t]}|U_r(e) - U_s(e)|^2\right] \;\le\; C_2\;(1+|X_0|^2)\;|t-s|\;.$$

Proof. Recall from the proof of Proposition 4.3 that $Y = u(\cdot, X_\cdot)$ on $[0,T]$. Thus, plugging (2.5) and (2.6) in Proposition 4.4 gives the upper-bound on the quantity $\mathbb{E}\left[\sup_{r\in[s,t]}|Y_r - Y_s|^2\right]$. The upper-bound on $\mathbb{E}\left[\sup_{e\in E}\sup_{r\in[s,t]}|U_r(e) - U_s(e)|^2\right]$ is obtained similarly by passing to the version of U given in Remark 4.1. □

Proposition 4.5 *Assume that* $\mathbf{H}$-$\mathbf{C}^X$-$\mathbf{C}^Y$ *holds. Then there is a version of* Z *such that, for all* $n \ge 1$, $\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] \le C_2^0\, n^{-1}$.

Proof. 1. We denote by $\nabla_x h$ (resp. $\nabla_y h$, $\nabla_z h$, $\nabla_\gamma h$) the gradient of h with respect to its x variable (resp. y, z, γ). We first introduce the processes Λ and M defined by $\Lambda_t := \exp\left(\int_0^t \nabla_y h(\Theta_r)\, dr\right)$, $M_t := 1 + \int_0^t M_r\; \nabla_z h(\Theta_r)\cdot dW_r$. Since h has bounded derivatives, it follows from Itô's Lemma and Proposition 4.2 that

$$\Lambda_t M_t Z_t \;=\; \mathbb{E}\left[M_T\left(\Lambda_T\nabla g(X_T)\chi_T^t + \int_t^T \left(\nabla_x h(\Theta_r)\chi_r^t + \nabla_\gamma h(\Theta_r)\Gamma_r^t\right)\;\Lambda_r\; dr\right)\;\middle|\; \mathcal{F}_t\right]\;.$$

By Remark 3.8 and Proposition 3.3, we deduce that

$$\Lambda_t M_t Z_t \;=\; \mathbb{E}\left[M_T\left(\Lambda_T\nabla g(X_T)\nabla X_T + \int_t^T F_r\;\Lambda_r\; dr\right)\;\middle|\;\mathcal{F}_t\right](\nabla X_{t-})^{-1}\sigma(X_{t-})$$

where the process F is defined by $F_r := \nabla_x h(\Theta_r)\nabla X_r + \nabla_\gamma h(\Theta_r)\nabla\Gamma_r$ for $r \le T$. It follows that

$$\Lambda_t M_t Z_t \;=\; \left\{\mathbb{E}\left[G \mid \mathcal{F}_t\right] - \int_0^t F_r\;\Lambda_r\; dr\right\}(\nabla X_{t-})^{-1}\sigma(X_{t-}) \tag{4.8}$$

where $G := M_T\left(\Lambda_T\nabla g(X_T)\;\nabla X_T + \int_0^T F_r\;\Lambda_r\; dr\right)$. By Remark 3.7 and (4.4), we deduce that

$$\mathbb{E}\left[|G|^p\right] \;\le\; C_p^0 \quad \text{for all}\;\; p \ge 2\;. \tag{4.9}$$

Set $m_s := \mathbb{E}\left[G \mid \mathcal{F}_s\right]$ and let $(\tilde\zeta, \tilde V) \in \mathbf{H}^2\times\mathbf{L}_\lambda^2$ (with values in $\mathbb{M}^d\times\mathbb{R}^d$) be defined such that $m_s = G - \int_s^T \tilde\zeta_r dW_r - \int_s^T\int_E \tilde V_r(e)\bar\mu(de,dr)$. Applying (4.9) and Lemma 5.2 to $(m,\tilde\zeta,\tilde V)$ implies that

$$\|(m,\tilde\zeta,\tilde V)\|_{\mathcal{B}^p} \;\le\; C_p^0 \quad \text{for all}\;\; p \ge 2\;. \tag{4.10}$$

Using $\mathbf{C}^X$, Remark 3.7, (4.4), (4.10), applying Lemma 5.1 to M^{-1} and using Itô's Lemma, we deduce from the last assertion that

$$\tilde Z := (\Lambda M)^{-1}\;\left(m - \int_0^\cdot F_r\;\Lambda_r\; dr\right)(\nabla X)^{-1}$$

can be written as $\tilde Z_t = \tilde Z_0 + \int_0^t \tilde\mu_r dr + \int_0^t \tilde\sigma_r dW_r + \int_0^t \int_E \tilde\beta_r(e)\bar\mu(de,dr)$, where

$$\|\tilde Z\|^p_{\mathcal{S}^p} \quad\leq\quad C^0_p \quad \text{for all} \;\; p \geq 2 \;, \tag{4.11}$$

and $\tilde\mu$, $\tilde\sigma$ and $\tilde\beta$ are adapted processes satisfying

$$A^p_{[0,T]} \quad\leq\quad C^0_p \quad \text{for all} \;\; p \geq 2 \tag{4.12}$$

where $A^p_{[s,t]} := \|\tilde\mu\|^p_{\mathbf{H}^p_{[s,t]}} + \|\tilde\sigma\|^p_{\mathbf{H}^p_{[s,t]}} + \|\tilde\beta\|^p_{\mathbf{L}^p_{\lambda,[s,t]}} \quad , \quad s \leq t \leq T$.

2. Observe that $Z_t = \tilde Z_t\; \sigma(X_t) \quad \mathbb{P}-$a.s. since the probability of having a jump at time t is equal to zero. It follows that, for all $i \leq n$ and $t \in [t_i, t_{i+1}]$,

$$\mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] \quad\leq\quad C_2\left(I^1_{t_i,t} + I^2_{t_i,t}\right) \tag{4.13}$$

where $I^1_{t_i,t} := \mathbb{E}\left[|\tilde Z_t - \tilde Z_{t_i}|^2 |\sigma(X_{t_i})|^2\right]$ and $I^2_{t_i,t} := \mathbb{E}\left[|\sigma(X_t) - \sigma(X_{t_i})|^2 |\tilde Z_t|^2\right]$. Observing that

$$\begin{aligned} I^1_{t_i,t} &= \mathbb{E}\left[\mathbb{E}\left[|\tilde Z_t - \tilde Z_{t_i}|^2 \mid \mathcal{F}_{t_i}\right] |\sigma(X_{t_i})|^2\right] \\ &\leq C_2 \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} \left[|\tilde\mu_r|^2 + |\tilde\sigma_r|^2 + \int_E |\tilde\beta_r(e)|^2 \lambda(de)\right] dr\right) |\sigma(X_{t_i})|^2\right] \;, \end{aligned}$$

we deduce from Hölder inequality, (2.5) and the linear growth assumption on σ that

$$\begin{aligned} &\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} I^1_{t_i,t} dt \\ &\leq C_2\, n^{-1}\, \mathbb{E}\left[\left(\int_0^T \left[|\tilde\mu_r|^2 + |\tilde\sigma_r|^2 + \int_E |\tilde\beta_r(e)|^2 \lambda(de)\right] dr\right) \sup_{t\leq T} |\sigma(X_t)|^2\right] \\ &\leq C^0_2 (A^4_{[0,T]})^{\frac{1}{2}}\, n^{-1} \;. \end{aligned} \tag{4.14}$$

Using the Lipschitz continuity of σ, we obtain

$$I^2_{t_i,t} \quad\leq\quad C_2 \mathbb{E}\left[|X_t - X_{t_i}|^2 |\tilde Z_t|^2\right] \;. \tag{4.15}$$

Now observe that for each $k, l \leq d$

$$\mathbb{E}\left[(X^k_t - X^k_{t_i})^2 (\tilde Z^l_t)^2\right] \leq C_2\left(\mathbb{E}\left[(\tilde Z^l_t - \tilde Z^l_{t_i})^2 (X^k_{t_i})^2\right] + \mathbb{E}\left[(X^k_t \tilde Z^l_t - X^k_{t_i} \tilde Z^l_{t_i})^2\right]\right) \tag{4.16}$$

Arguing as above, we obtain

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[(\tilde Z^l_t - \tilde Z^l_{t_i})^2 (X^k_{t_i})^2\right] \leq C^0_2 \left(1 + (A^4_{[0,T]})^{\frac{1}{2}}\right) n^{-1} \;. \tag{4.17}$$

Moreover, we deduce from the linear growth condition on b, σ, β and (2.5), (4.11) and (4.12) that $X^k \tilde Z^l$ can be written as

$$X^k_t \tilde Z^l_t \quad=\quad X^k_0 \tilde Z^l_0 + \int_0^t \hat\mu^{kl}_r dr + \int_0^t \hat\sigma^{kl}_r dW_r + \int_0^t \int_E \hat\beta^{kl}_r(e)\bar\mu(de,dr)$$

where $\hat{\mu}^{kl}$, $\hat{\sigma}^{kl}$ and $\hat{\beta}^{kl}$ are adapted processes satisfying $\|\hat{\mu}^{kl}\|_{\mathbf{H}^2}+\|\hat{\sigma}^{kl}\|_{\mathbf{H}^2}+\|\hat{\beta}^{kl}\|_{\mathbf{L}^2_\lambda} \le C^0_2$. It follows that

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \mathbb{E}\left[(X^k_t\tilde{Z}^l_t - X^k_{t_i}\tilde{Z}^l_{t_i})^2\right] \quad \le \quad C_2\, n^{-1}\left(\|\hat{\mu}^{kl}\|^2_{\mathbf{H}^2}+\|\hat{\sigma}^{kl}\|^2_{\mathbf{H}^2}+\|\hat{\beta}^{kl}\|^2_{\mathbf{L}^2_\lambda}\right)$$

which combined with (4.15), (4.16) and (4.17) leads to

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} I^2_{t_i,t}dt \le C^0_2(1+(A^4_{[0,T]})^{\frac{1}{2}})\, n^{-1}\,. \tag{4.18}$$

The proof is concluded by plugging (4.14)-(4.18) in (4.13) and recalling (4.12). □

Proposition 4.6 *Assume that* $\mathbf{C}^X$-$\mathbf{C}^Y$ *holds. Then there is a version of* Z *such that, for all* $\varepsilon>0$ *and* $n\ge 1$, $\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|Z_t-Z_{t_i}|^2\right] \le C^0_\varepsilon\, n^{-1+\varepsilon}$.

Proof. We adapt the arguments of [6]. Let Λ and M be defined as in the proof of Proposition 4.5 and recall that, after possibly passing to a suitable version, $Z_t = I^t_t$ where, for $s,t\le T$,

$$I^t_s := \mathbb{E}\left[\frac{M_T}{\Lambda_t M_t}\left(\Lambda_T\nabla g(X_T)\chi^t_T + \int_t^T\left(\nabla_x h(\Theta_r)\chi^t_r + \nabla_\gamma h(\Theta_r)\Gamma^t_r\right)\,\Lambda_r\, dr\right)\mid \mathcal{F}_s\right]\,.$$

For $t\in[t_i,t_{i+1}]$, $i\le n-1$, we therefore have $|Z_t-Z_{t_i}|^2 \le C_2\left(|I^{t_i}_t - I^{t_i}_{t_i}|^2 + |I^t_t - I^{t_i}_t|^2\right)$ where, by Remark 3.5, (5.8) below applied to (3.4), recall that ρ is bounded, and standard estimations on ΛM, $\sup_{i\le n-1,\, t\in[t_i,t_{i+1}]}\mathbb{E}\left[|I^t_t - I^{t_i}_t|^2\right] \le C^0_2 n^{-1}$. Thus it suffices to prove that

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|I^{t_i}_t - I^{t_i}_{t_i}|^2\right] \quad\le\quad C^0_\varepsilon\, n^{-1+\varepsilon}\,,$$

where $\varepsilon>0$ is now fixed. To this purpose, we first observe that I^{t_i} is a martingale on $[t_i,t_{i+1}]$, which implies that

$$\mathbb{E}\left[|I^{t_i}_t - I^{t_i}_{t_i}|^2\right] \quad\le\quad \mathbb{E}\left[|I^{t_i}_{t_{i+1}}|^2 - |I^{t_i}_{t_i}|^2\right]\,. \tag{4.19}$$

Remark that $\sum_{i=0}^{n-1}\mathbb{E}\left[|I^{t_i}_{t_{i+1}}|^2 - |I^{t_i}_{t_i}|^2\right] = \mathbb{E}\left[|Z_T|^2-|Z_0|^2\right] + \sum_{i=1}^{n}\mathbb{E}\left[|I^{t_{i-1}}_{t_i}|^2 - |I^{t_i}_{t_i}|^2\right]$, which, combined with Proposition 4.2 and (4.19), leads to

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|I^{t_i}_t - I^{t_i}_{t_i}|^2\right] \quad=\quad C^0_2 n^{-1}\left(1+\sum_{i=1}^{n}\mathbb{E}\left[|I^{t_{i-1}}_{t_i}|^2 - |I^{t_i}_{t_i}|^2\right]\right)\,.$$

To conclude the proof, it remains to show that

$$\mathbb{E}\left[|I^{t_{i-1}}_{t_i}|^2 - |I^{t_i}_{t_i}|^2\right] \le \mathbb{E}\left[|I^{t_{i-1}}_{t_i} - I^{t_i}_{t_i}|\,|I^{t_{i-1}}_{t_i} + I^{t_i}_{t_i}|\right] \quad\le\quad C^0_\varepsilon n^{-1+\varepsilon}\,.$$

which follows from Hölder inequality, Remark 3.5 and Lemma 5.2 as above. □

We now complete the proof of Theorem 2.1.

Proof of Theorem 2.1. 1. We start with (ii). We first show that (2.16) holds under $\mathbf{H}$ and $\mathbf{C}^Y$. We consider a C_b^∞ density q on $\mathbb{R}^d$ with compact support and set

$$(b^k, \sigma^k, \beta^k(\cdot, e))(x) \;\; = \;\; k^d \int_{\mathbb{R}^d} (b, \sigma, \beta(\cdot, e))(\bar{x})\, q\,(k[x - \bar{x}])\, d\bar{x}\ .$$

For large $k \in \mathbb{N}$, these functions are bounded by $2K$ at 0. Moreover, they are K-Lipschitz and C_b^1. Using the continuity of σ, one also easily checks that σ^k is still invertible. By $\mathbf{H}$ and Remark 2.4, for each $e \in E$ and $x \in \mathbb{R}^d$, $I_d + \nabla\beta^k(x, e)$ is invertible with uniformly bounded inverse. We denote by (X^k, Y^k, Z^k, U^k) the solution of (2.2)-(2.3) with (b, σ, β) replaced by (b^k, σ^k, β^k). Since (b^k, σ^k, β^k) converges pointwise to (b, σ, β), one easily deduces from Lemma 5.1 and Lemma 5.2 that (X^k, Y^k, Z^k, U^k) converges to (X, Y, Z, U) in $\mathcal{S}^2 \times \mathcal{B}^2$. Since the result of Proposition 4.5 holds for (X^k, Y^k, Z^k, U^k) uniformly in k, this shows that (ii) holds under $\mathbf{H}$ and $\mathbf{C}^Y$, recall Remark 2.3.

We now prove that (2.16) holds under $\mathbf{H}$. Let (X, Y^k, Z^k, U^k) be the solution of (2.2)-(2.3) with h^k instead of h, where h^k is constructed by considering a sequence of molifiers as above. For large k, $h^k(0)$ is bounded by $2K$. By Lemma 5.2, (Y^k, Z^k, U^k) converges to (Y, Z, U) in $\mathcal{S}^2 \times \mathcal{B}^2$ which implies (ii) by arguing as above.

2. Since ρ is bounded and $\lambda(E) < \infty$, Corollary 4.1 and Proposition 4.6 imply (2.14) and (2.15) under $\mathbf{C}^X$-$\mathbf{C}^Y$, recall Remark 2.3. Now observe that

$$\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\Gamma_t - \bar{\Gamma}_{t_i}|^2\right] \;\; \le \;\; 2\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\Gamma_t - \Gamma_{t_i}|^2\right] + 2\mathbb{E}\left[|\Gamma_{t_i} - \bar{\Gamma}_{t_i}|^2\right]$$

where, by Jensen's inequality and the fact that Γ_{t_i} is $\mathcal{F}_{t_i}$-measurable,

$$\mathbb{E}\left[|\Gamma_{t_i} - \bar{\Gamma}_{t_i}|^2\right] \;\; \le \;\; \mathbb{E}\left[\left|\frac{n}{T}\int_{t_i}^{t_{i+1}} (\Gamma_{t_i} - \Gamma_s) ds\right|^2\right] \;\; \le \;\; \frac{n}{T}\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\Gamma_{t_i} - \Gamma_s|^2\right] ds\ .$$

Thus, (2.14) implies $\|\Gamma - \bar{\Gamma}\|^2_{\mathcal{S}^2} \le C_2^0\, n^{-1}$ and $\|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2} \le C_2^0\, n^{-1}$. We conclude by the same approximation argument as above. □

5 Appendix: A priori estimates

For sake of completeness, we provide in this section some a priori estimates on solutions of forward and backward SDE's with jumps.

In the following, we consider some measurable maps $\tilde{b}^i : \Omega \times [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$, $\tilde{\sigma}^i : \Omega \times [0, T] \times \mathbb{R}^d \mapsto \mathbb{M}^d$, $\tilde{\beta}^i : \Omega \times [0, T] \times \mathbb{R}^d \times E \mapsto \mathbb{R}^d$ and $\tilde{f}^i : \Omega \times [0, T] \times \mathbb{R} \times$

$\mathbb{R}^d \times L^2(E,\mathcal{E},\lambda;\mathbb{R})$, $i=1,2$. Here $L^2(E,\mathcal{E},\lambda;\mathbb{R})$ is endowed with the natural norm $(\int_E |a(e)|^2\lambda(de))^{\frac{1}{2}}$.

Omitting the dependence of these maps with respect to $\omega \in \Omega$, we assume that for each $t \le T$ $\tilde b^i(t,\cdot)$, $\tilde\sigma^i(t,\cdot)$, $\tilde\beta^i(t,\cdot,e)$ and $\tilde f^i(t,\cdot)$ are a.s. K-Lipschitz continuous uniformly in $e \in E$ for $\tilde\beta^i$. We also assume that $t \mapsto (\tilde f^i(t,\cdot),\tilde b^i(t,\cdot))$ is $\mathbb{F}$-progressively measurable, and $t \mapsto (\tilde\sigma^i(t,\cdot),\tilde\beta^i(t,\cdot))$ is $\mathbb{F}$-predictable, $i=1,2$.

Given some real number $p \ge 2$, we assume that $|\tilde b^i(\cdot,0)|$, $|\tilde\sigma^i(\cdot,0)|$ and $|\tilde f^i(\cdot,0)|$ are in $\mathbf{H}^p$, and that $|\tilde\beta^i(\cdot,0,\cdot)|$ is in $\mathbf{L}^p_\lambda$.

For $t_1 \le t_2 \le T$, $\tilde X^i \in L^2(\Omega,\mathcal{F}_{t_i},\mathbb{P};\mathbb{R}^d)$ for $i=1,2$, we now denote by X^i the solution on $[t_i,T]$ of

$$X^i_t = \tilde X^i + \int_{t_i}^t \tilde b^i(s,X^i_s)ds + \int_{t_i}^t \tilde\sigma^i(s,X^i_s)dW_s + \int_{t_i}^t\int_E \tilde\beta^i(s,e,X^i_{s-})\bar\mu(de,ds)\ . \quad (5.1)$$

Lemma 5.1

$$\|X^1\|^p_{\mathcal{S}^p_{[t_1,T]}} \ \le\ C_p\,\mathbb{E}[|\tilde X^1|^p] + \|\tilde b^1(\cdot,0)\|^p_{\mathbf{H}^p_{[t_1,T]}} + \|\tilde\sigma^1(\cdot,0)\|^p_{\mathbf{H}^p_{[t_1,T]}} + \|\tilde\beta^1(\cdot,0,\cdot)\|^p_{\mathbf{L}^p_{\lambda,[t_1,T]}}\ . \quad (5.2)$$

Moreover, for all $t_1 \le s \le t \le T$,

$$\mathbb{E}\left[\sup_{s\le u\le t}|X^1_u - X^1_s|^p\right] \ \le\ C_p\ A^1_p\,|t-s|\ , \quad (5.3)$$

where A^1_p is defined as

$$\mathbb{E}[|\tilde X^1|^p] + \mathbb{E}\left[\sup_{t_1\le s\le T}|\tilde b^1(s,0)|^p + \sup_{t_1\le s\le T}|\tilde\sigma^1(s,0)|^p + \sup_{t_1\le s\le T}\left\{\int_E|\tilde\beta^1(s,0,e)|^p\lambda(de)\right\}\right]\ ,$$

and, for $t_2 \le t \le T$,

$$\begin{aligned}\|\delta X\|^p_{\mathcal{S}^p_{[t_2,T]}} &\le C_p\left(E|\tilde X^1-\tilde X^2|^p + A^1_p\,|t_2-t_1|\right)\\ &+ C_p\left(E\left(\int_{t_2}^T|\delta\tilde b_t|dt\right)^p + \|\delta\tilde\sigma\|^p_{\mathbf{H}^p_{[t_2,T]}} + \|\delta\tilde\beta\|^p_{\mathbf{L}^p_{\lambda,[t_2,T]}}\right) \quad (5.4)\end{aligned}$$

where $\delta X := X^1 - X^2$, $\delta\tilde b = (\tilde b^1-\tilde b^2)(\cdot,X^1_\cdot)$ and $\delta\tilde\sigma$, $\delta\tilde\beta$ are defined similarly.

Lemma 5.2 (i) *Let $\tilde f$ be equal to $\tilde f^1$ or $\tilde f^2$. Given $\tilde Y \in L^p(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R})$, the backward SDE*

$$Y_t = \tilde Y + \int_t^T \tilde f(s,Y_s,Z_s,U_s)ds + \int_t^T Z_s\cdot dW_s + \int_t^T\int_E U_s(e)\bar\mu(de,ds) \quad (5.5)$$

has a unique solution (Y,Z,U) in $\mathcal{B}^2$. It satisfies

$$\|(Y,Z,U)\|^p_{\mathcal{B}^p} \ \le\ C_p\ \mathbb{E}\left[|\tilde Y|^p + \left(\int_0^T|\tilde f(t,0)|dt\right)^p\right]\ . \quad (5.6)$$

Moreover, if $A_p := \mathbb{E}\left[|\tilde{Y}|^p + \sup_{t\le T}|\tilde{f}(t,0)|^p\right] < \infty$*, then*

$$\mathbb{E}\left[\sup_{s\le u\le t}|Y_u - Y_s|^p\right] \le C_p\ \left\{A_p\,|t-s|^p + \|Z\|^p_{\mathbf{H}^p_{[s,t]}} + \|U\|^p_{\mathbf{L}^p_{\lambda,[s,t]}}\right\}\ . \tag{5.7}$$

(ii) *Fix* $\tilde{Y}^1$ *and* $\tilde{Y}^2$ *in* $L^p(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R})$ *and let* (Y^i,Z^i,U^i) *be the solution of* (5.6) *with* $(\tilde{Y}^i,\tilde{f}^i)$ *in place of* $(\tilde{Y},\tilde{f})$*,* $i=1,2$*. Then, for all* $t\le T$*,*

$$\|(\delta Y,\delta Z,\delta U)\|^p_{\mathcal{B}^p_{[t,T]}} \ \le\ C_p\ \mathbb{E}\left[|\delta\tilde{Y}|^p + \left(\int_t^T|\delta\tilde{f}_r|dr\right)^p\right] \tag{5.8}$$

where $\delta\tilde{Y} := \tilde{Y}^1 - \tilde{Y}^2$*,* $\delta Y := Y^1 - Y^2$*,* $\delta Z := Z^1 - Z^2$*,* $\delta U := U^1 - U^2$ *and* $\delta\tilde{f}_\cdot := (\tilde{f}^1 - \tilde{f}^2)(\cdot,Y^1_\cdot,Z^1_\cdot,U^1_\cdot)$*.*

The proof of the previous results is standard, see e.g. [15] and [12]. We refer to pages 128-129 in [11] for more details. The jump part is handled by using the following proposition which plays the same role as the usual Burkholder-Davis-Gundy's Lemma. It follows from an induction argument already used in [4]. The proof can be found in [11], see page 125.

Proposition 5.1 Given $\psi\in\mathbf{L}^2_\lambda$, let M be defined by $M_t = \int_0^t\int_E\psi_s(e)\bar{\mu}(ds,de)$ on $[0,T]$. Then, for all $p\ge 2$, $k_p\ \|\psi\|^p_{\mathbf{L}^p_{\lambda,[0,T]}} \le \|M\|^p_{\mathcal{S}^p_{[0,T]}} \le K_p\ \|\psi\|^p_{\mathbf{L}^p_{\lambda,[0,T]}}$, where k_p, K_p are positive numbers that depend only on p, $\lambda(E)$ and T.

References

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

[2] Bally V., and G. Pages (2002). A quantization algorithm for solving discrete time multidimensional optimal stopping problems. *Bernoulli*, 9 (6), 1003-1049.

[3] Becherer D. (2005). Bounded solutions to Backward SDE's with Jumps for Utility Optimization and Indifference Hedging. Preprint, Imperial College London.

[4] Bichteler K. and J. Jacod (1983). Calcul de Malliavin pour des diffusions avec saut: existence d'une densité dans le cas unidimensionel. *Séminaire de Probabilité*, 17, 132-157.

[5] Bichteler K., Gravereaux J.-B. and J. Jacod (1987). *Malliavin calculus for processes with jumps*. Gordon and Breach Science Publishers, New York.

[6] Bouchard B. and J.-F. Chassagneux (2006). Discrete time approximation for continuously and discretely reflected BSDE's. Preprint LPMA, University Paris 6.

[7] Bouchard B. and N. Touzi (2004). Discrete-Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stochastic Processes and their Applications*, 111 (2), 175-206.

[8] Brémaud P. (1981). *Point Processes and Queues - Martingale Dynamics*, Springer-Verlag, New-York.

[9] Buckdahn R. and E. Pardoux (1994). BSDE's with jumps and associated integro-partial differential equations. Preprint.

[10] Douglas J. Jr., J. Ma, and P. Protter (1996). Numerical Methods for Forward- Backward Stochastic Differential Equations. *Annals of Applied Probability*, 6, 940-968.

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

[12] El Karoui N., S. Peng and M.-C. Quenez (1997). Backward stochastic differential equations in finance. *Mathematical finance*, 7 (1), 1-71.

[13] Eyraud-Loisel A. (2005). 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 (11), 1745-1763.

[14] Forster B., E. Lütkebohmert and J. Teichmann (2005). Calculation of the greeks for jump-diffusions. Preprint.

[15] Gikhman I. and A. Skorohod (1972). *Stochastic differential equations*. Springer.

[16] Gobet E., J.P. Lemor and X. Warin (2005). Rate of convergence of empirical regression method for solving generalized BSDE. To appear in *Bernoulli* .

[17] Ma J. , P. Protter, J. San Martin, and S. Torres (2002). Numerical Method for Backward Stochastic Differential Equations. *Annals of Applied Probability*, 12 (1), 302-316.

[18] Nualart D. (1995). *The Malliavin Calculus and Related Topics*. Springer Verlag, Berlin.

[19] Nualart D. and E. Pardoux (1988). Stochastic calculus with anticipating integrands. *Prob. Theory and Rel. Fields*, 78, 535-581.

[20] Pardoux E. and S. Peng (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in Control and Inform. Sci*, 176, 200-217.

[21] Pardoux E., F. Pradeilles and Z. Rao (1997). Probabilistic interpretation for a system of semilinear parabolic partial differential equations. *Ann. Inst. H. Poincare*, 33 (4), 467-490.

[22] Protter P. (1990). *Stochastic integration and differential equations*. Springer Verlag, Berlin.

[23] Tang S. and X. Li (1994). Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.*, 32 (5), 1447-1475.

[24] Ma J. and Zhang J. (2002). Path Regularity of Solutions to Backward Stochastic Differential Equations. *Probability Theory and Related Fields*, 122, 163-190.

[25] Zhang J. (2004). A numerical scheme for BSDEs. *Annals of Applied Probability*, 14 (1), 459-488.