Discrete time approximation of decoupled Forward-Backward SDE with jumps

Bruno Bouchard, Romuald Elie

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 : November 2005

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. When the jump coefficient of the first variation process of the forward component satisfies a non-degeneracy condition which ensures its inversibility, we obtain 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 [28, 29] in the no-jump case. A similar result is obtained without the non-degeneracy assumption whenever the coefficients are C_b^1 with Lipschitz derivatives. Several extensions of these results are discussed. In particular, we propose a convergent scheme for the resolution of systems of coupled semilinear parabolic PDE's.

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 [26] and the recent paper Becherer [4] for an application to exponential utility maximisation 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 [26], thus generalizing the seminal paper of Pardoux and Peng [21].

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

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

where

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

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. [2]. 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 [11] and [18]. 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 [29], Bally and Pages [3], Bouchard and Touzi [6] or Gobet et al. [15]. In [6], the authors suggest the following implicit scheme. Given a regular grid $\pi = \{t_i = iT/n,\ i = 0, \dots, n\}$, 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

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

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 [27] was able to prove that the later is of order of n^{-1}. This remarkable result allows to derive the bound $\mathrm{Err}_n(Y, Z) \le C n^{-1/2}$, i.e. the above approximation achieves the optimal convergence rate $n^{-1/2}$.

In this paper, we extend the approach of Bouchard and Touzi [6] and approximate the solution of (1.1) by the backward scheme

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

where $\bar{Y}^\pi_{t_n} := g(X^\pi_{t_n})$. By adapting the arguments of Gobet et al. [15], we first prove that our discretization error

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

converges to 0 as the discretisation 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 $\bar{\Gamma}_{t_i} := n\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma_t dt \mid \mathcal{F}_{t_i}\right]$. We first show that

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

whenever the coefficients are Lipschitz continuous. Under some additional conditions on the inversibility of $\nabla\beta + I_d$, see $\mathbf{H}_1$, or on the regularity of the coefficient, see $\mathbf{H}_2$, we then prove that

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2\right] dt \;\leq\; C\, n^{-1}\,.$$

This extends to our framework the remarkable result derived by Zhang [28, 29] in the no-jump case and allows us to show that our discrete-time scheme achieves the optimal convergence rate $n^{-1/2}$.

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 [1], [8], [9], [10] and [17]), 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 can be solved by approximating the conditional expectation operators numerically in an efficient way. In the no-jump case, Bouchard and Touzi [6] use the Malliavin calculus to rewrite conditional expectations as the ratio of two unconditional expectations which can be estimated by standard Monte-Carlo methods. In the reflected case where h does not depend on Z, Bally and Pages [3] use a quantization approach. Finally, Gobet et al. [15] have suggested an adaptation of the so-called Longstaff and Schwartz algorithm based on non-parametric regressions, see [16], which also works in the case where $\beta \neq 0$ but the driver does not depend on U.
Since this is not the main issue of this paper, we leave the theoretical study and numerical implementation of such methods in our setting for further research.

The rest of the paper is organized as follows. In Section 2, we describe the approximation scheme and state our main convergence result. We also discuss several possible extensions. In particular, we propose a convergent scheme for the resolution of systems of coupled semilinear parabolic PDE's. 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

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \leq T}, \mathbb{P})$ be a stochastic basis such that $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets, $\mathcal{F}_T = \mathcal{F}$ and $\mathbb{F}$ satisfies the usual assumptions. We assume that $\mathbb{F}$ is generated by a d-dimensional standard Brownian motion W and an independent Poisson measure μ on $[0, T] \times E$ where $E = \mathbb{R}^m$ for some $m \geq 1$. We denote by $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \leq T}$ (resp. $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{t \leq T}$) the $\mathbb{P}$-augmentation of the natural filtration of W (resp. μ).

We assume that the compensator ν of μ has the form $\nu(dt, de) = \lambda(de)dt$ for some finite measure λ on E, endowed with its Borel tribe $\mathcal{E}$, and denote by $\bar{\mu} := \mu - \nu$ the compensated measure.

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

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

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

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

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 e.g. [14] and 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

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

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

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

By a solution, we mean an $\mathbb{F}$-adapted 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

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

The estimate on X is standard, see (5.4) of Lemma 5.1 in the Appendix. Plugging this in (5.8) of Lemma 5.2 leads to the bound on $\|(Y, Z, U)\|_{\mathcal{B}^p}$. Using (5.5) of Lemma 5.1, we also deduce that

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

while the previous estimates on X combined with (5.9) of Lemma 5.2 implies

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

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

$$\left\{\begin{array}{rcl} X^\pi_0 & := & X_0 \\ X^\pi_{t_{i+1}} & := & X^\pi_{t_i} + \frac{1}{n} b(X^\pi_{t_i}) + \sigma(X^\pi_{t_i})\Delta W_{i+1} + \int_E \beta(X^\pi_{t_i}, e)\bar{\mu}(de, (t_i, t_{i+1}]) \end{array}\right. \tag{2.8}$$

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

$$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |X_t - X^\pi_{t_i}|^2\right] \;\leq\; C_2^0\ n^{-1} \;. \tag{2.9}$$

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

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

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, which is a natural extension of the one considered in [6] in the case $\beta = 0$, can be understood as follows. On each interval $[t_i, t_{i+1})$, we want to replace the arguments (X, Y, Z, Γ) of h in (2.3) by $\mathcal{F}_{t_i}$-measurable random variables $(\tilde{X}_{t_i}, \tilde{Y}_{t_i}, \tilde{Z}_{t_i}, \tilde{\Gamma}_{t_i})$. It is natural to take $\tilde{X}_{t_i} = X^\pi_{t_i}$. Taking conditional expectation, we obtain the approximation

$$Y_{t_i} \;\cong\; \mathbb{E}\left[Y_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{1}{n} h\left(X^\pi_{t_i}, \tilde{Y}_{t_i}, \tilde{Z}_{t_i}, \tilde{\Gamma}_{t_i}\right) .$$

This leads to a backward implicit scheme for Y of the form

$$\bar{Y}^\pi_{t_i} \;=\; \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{1}{n} h\left(X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \tilde{Z}_{t_i}, \tilde{\Gamma}_{t_i}\right) . \tag{2.11}$$

It remains to choose $\tilde{Z}_{t_i}$ and $\tilde{\Gamma}_{t_i}$ in terms of $\bar{Y}^\pi_{t_{i+1}}$. By the representation theorem, 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 they do not depend on the way $\bar{Y}^\pi_{t_i}$ is defined and that $\bar{Z}^\pi$ and $\bar{\Gamma}^\pi$ defined in (2.10) satisfy

$$\bar{Z}^\pi_{t_i} \;=\; n\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z^\pi_s ds \mid \mathcal{F}_{t_i}\right] \quad \text{and} \quad \bar{\Gamma}^\pi_{t_i} \;=\; n\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma^\pi_s ds \mid \mathcal{F}_{t_i}\right] \tag{2.12}$$

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

$$\begin{array}{rcl} \mathbb{E}\left[\displaystyle\int_{t_i}^{t_{i+1}} |Z^\pi_t - \bar{Z}^\pi_{t_i}|^2 dt\right] & = & \displaystyle\inf_{Z_i \in L^2(\Omega,\mathcal{F}_{t_i})} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Z^\pi_t - Z_i|^2 dt\right] \\ \mathbb{E}\left[\displaystyle\int_{t_i}^{t_{i+1}} |\Gamma^\pi_t - \bar{\Gamma}^\pi_{t_i}|^2 dt\right] & = & \displaystyle\inf_{\Gamma_i \in L^2(\Omega,\mathcal{F}_{t_i})} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\Gamma^\pi_t - \Gamma_i|^2 dt\right] . \end{array}$$

Thus, it is natural to take $(\tilde{Z}_{t_i}, \tilde{\Gamma}_{t_i}) = (\bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i})$ in (2.11), so that

$$\bar{Y}^\pi_{t_i} \;=\; \bar{Y}^\pi_{t_{i+1}} + \frac{1}{n} h\left(X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}\right) - \int_{t_i}^{t_{i+1}} Z^\pi_s \cdot dW_s - \int_{t_i}^{t_{i+1}} \int_E U^\pi_s(e)\bar{\mu}(ds, de)\,.$$

Finally, observe that, if we define Y^π on $[t_i, t_{i+1})$ by setting

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

we obtain

$$n\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Y^\pi_t dt \mid \mathcal{F}_{t_i}\right] \;=\; \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{1}{n} h\left(X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}\right) \;=\; Y^\pi_{t_i} = \bar{Y}^\pi_{t_i}\;.$$

Thus, in this scheme, $\bar{Y}^\pi_{t_i}$ is the best $\mathbf{H}^2_{[t_i,t_{i+1}]}$-approximation of Y^π on $[t_i, t_{i+1})$ by $\mathcal{F}_{t_i}$-measurable random variables (viewed as constant processes on $[t_i, t_{i+1})$). This explains the notation $\bar{Y}^\pi$ which is consistent with the definition of $\bar{Z}^\pi$ and $\bar{\Gamma}^\pi$.

Remark 2.3 One could also use an explicit scheme as in e.g. [3] or [15]. In this case, (2.10) has to be replaced

$$\left\{\begin{array}{rcl} \tilde{Z}^\pi_{t_i} &:=& n\,\mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \Delta W_{i+1} \mid \mathcal{F}_{t_i}\right] \\ \tilde{\Gamma}^\pi_{t_i} &:=& n\,\mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \int_E \rho(e)\bar{\mu}(de, (t_i, t_{i+1}]) \mid \mathcal{F}_{t_i}\right] \\ \tilde{Y}^\pi_{t_i} &:=& \mathbb{E}\left[\tilde{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{1}{n}\,\mathbb{E}\left[h\left(X^\pi_{t_i}, \tilde{Y}^\pi_{t_{i+1}}, \tilde{Z}^\pi_{t_i}, \tilde{\Gamma}^\pi_{t_i}\right) \mid \mathcal{F}_{t_i}\right] \end{array}\right. \tag{2.13}$$

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 convergence of the explicit scheme will be discussed in Remarks 2.6 and 2.8 below.

2.3 Convergence of the approximation scheme

In this subsection, we show that the approximation error

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

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

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

Remark 2.4 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 $[t_i,t_{i+1})$), i.e.

$$\begin{aligned}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}|Z_t-\bar{Z}_{t_i}|^2dt\right] &= \inf_{Z_i\in L^2(\Omega,\mathcal{F}_{t_i})}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}|Z_t-Z_i|^2dt\right]\\ \mathbb{E}\left[\int_{t_i}^{t_{i+1}}|\Gamma_t-\bar{\Gamma}_{t_i}|^2dt\right] &= \inf_{\Gamma_i\in L^2(\Omega,\mathcal{F}_{t_i})}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}|\Gamma_t-\Gamma_i|^2dt\right]\ .\end{aligned}$$

Proposition 2.1 *We have*

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|Y_t-Y_{t_i}|^2\right]dt\le C_2^0\,n^{-1}\quad\textit{and}\quad\|Z-\bar{Z}\|_{\mathbf{H}^2}+\|\Gamma-\bar{\Gamma}\|_{\mathbf{H}^2}\le\epsilon(n)\quad(2.14)$$

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

$$\mathrm{Err}_n\left(Y,Z,U\right)\ \le\ C_2^0\left(n^{-1/2}+\|Z-\bar{Z}\|_{\mathbf{H}^2}+\|\Gamma-\bar{\Gamma}\|_{\mathbf{H}^2}\right)\ ,\qquad(2.15)$$

so that

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

Proof. We adapt the arguments of [6]. Recall from Remark 2.2 that

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

on $[t_i,t_{i+1})$ and that $\bar{Y}^\pi_{t_i}=Y^\pi_{t_i}$. For $L=Y,Z$ or U, we set $\delta L:=L-L^\pi$. It follows from the definition of $\bar{Z}^\pi$ and $\bar{U}^\pi$ in (2.12), Jensen's inequality and the bound on ρ that

$$\mathbb{E}\left[|\bar{Z}_{t_i}-\bar{Z}^\pi_{t_i}|^2\right]+\mathbb{E}\left[|\bar{\Gamma}_{t_i}-\bar{\Gamma}^\pi_{t_i}|^2\right]\ \le\ C_2\,n\left(\|\delta Z\|^2_{\mathbf{H}^2_{[t_i,t_{i+1}]}}+\|\delta U\|^2_{\mathbf{L}^2_{\lambda,[t_i,t_{i+1}]}}\right)\ .(2.16)$$

For $t\in[t_i,t_{i+1})$, we deduce from Itô's Lemma, the Lipschitz property of h, (2.9) and (2.16) that

$$\begin{aligned}\mathbb{E}[|\delta Y_t|^2]+\|\delta Z\|^2_{\mathbf{H}^2_{[t,t_{i+1}]}}+\|\delta U\|^2_{\mathbf{L}^2_{\lambda,[t,t_{i+1}]}} &\le \mathbb{E}[|\delta Y_{t_{i+1}}|^2]+\alpha\int_t^{t_{i+1}}\mathbb{E}[|\delta Y_s|^2]ds\\ &+\ \frac{C_2^0}{\alpha}\left(n^{-2}+\bar{B}_i+B^\pi_i\right)\end{aligned}\qquad(2.17)$$

where α is some positive constant to be chosen later, and $(\bar{B}_i,B^\pi_t)$ is defined as

$$\begin{aligned}\bar{B}_i &:= \int_{t_i}^{t_{i+1}}\left(\mathbb{E}\left[|Y_s-Y_{t_i}|^2\right]+\mathbb{E}\left[|Z_s-\bar{Z}_s|^2\right]+\mathbb{E}\left[|\Gamma_s-\bar{\Gamma}_s|^2\right]\right)ds\\ B^\pi_i &:= n^{-1}\mathbb{E}[|\delta Y_{t_i}|^2]+\|\delta Z\|^2_{\mathbf{H}^2_{[t_i,t_{i+1}]}}+\|\delta U\|^2_{\mathbf{L}^2_{\lambda,[t_i,t_{i+1}]}}\ .\end{aligned}$$

Using Gronwall's Lemma, it follows that

$$\mathbb{E}[|\delta Y_t|^2] \;\leq\; \left(\mathbb{E}[|\delta Y_{t_{i+1}}|^2] + \frac{C_2^0}{\alpha}\left(n^{-2} + \bar{B}_i + B_i^\pi\right)\right) e^{\alpha/n}\,. \tag{2.18}$$

Plugging this inequality in (2.17) and taking α and n large enough leads to

$$\begin{aligned}\mathbb{E}[|\delta Y_{t_i}|^2] + \eta\left(\|\delta Z\|^2_{\mathbf{H}^2_{[t_i,t_{i+1}]}} + \|\delta U\|^2_{\mathbf{L}^2_{\lambda,[t_i,t_{i+1}]}}\right) \;&\leq\; (1+\frac{C_2^0}{n})\mathbb{E}[|\delta Y_{t_{i+1}}|^2] \\ &+\; C_2^0\left(n^{-2} + \bar{B}_i + n^{-1}\mathbb{E}[|\delta Y_{t_i}|^2]\right)\,,\end{aligned} \tag{2.19}$$

with $\eta > 0$. For n large enough, combining the last inequality with the identity $\delta Y_{t_n} = g(X_T) - g(X_T^\pi)$ and the estimate (2.9) leads to

$$\mathbb{E}[|\delta Y_{t_i}|^2] \;\leq\; C_2^0\left(n^{-1} + \bar{B}\right) \quad \text{where} \quad \bar{B} := \sum_{j=0}^{n-1} \bar{B}_j\,, \tag{2.20}$$

which plugged into (2.19) implies

$$\mathbb{E}[|\delta Y_{t_i}|^2] + \eta\left(\|\delta Z\|^2_{\mathbf{H}^2_{[t_i,t_{i+1}]}} + \|\delta U\|^2_{\mathbf{L}^2_{\lambda,[t_i,t_{i+1}]}}\right) \;\leq\; \mathbb{E}[|\delta Y_{t_{i+1}}|^2] + C_2^0\left(n^{-2} + \frac{\bar{B}}{n} + \bar{B}_i\right)\,.$$

Summing up over i and using (2.18) and (2.20) , we finally obtain

$$\mathrm{Err}_n\left(Y,Z,U\right)^2 \leq C_2^0\left(n^{-1} + \bar{B}\right)\,. \tag{2.21}$$

Since Y solves (2.3),

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

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

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

This is exactly the first part of (2.14) which combined with (2.21) leads to (2.15). It remains to prove the second part of (2.14). Since Z is $\mathbb{F}$-adapted, there is a sequence of adapted processes $(Z^n)_n$ such that $Z_t^n = Z_{t_i}^n$ on each $[t_i, t_{i+1})$ and Z^n converges to Z in $\mathbf{H}^2$. By Remark 2.4, we observe that

$$\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \;\leq\; \|Z - Z^n\|^2_{\mathbf{H}^2}\,,$$

and applying the same reasoning to Γ concludes the proof. □

Remark 2.5 If $\sigma = 0$, which implies $Z = 0$, or h does not dependent on Z, the term $\bar{B}_i$ in the above proof reduces to

$$\bar{B}_i \;=\; \int_{t_i}^{t_{i+1}} \left(\mathbb{E}\left[|Y_s - Y_{t_i}|^2 \right] + \mathbb{E}\left[|\Gamma_s - \bar{\Gamma}_s|^2 \right] \right) ds \; .$$

In this case, the assertion (2.15) of Proposition 2.1 can be replaced by

$$\mathrm{Err}_n\left(Y, Z, U\right) \;\leq\; C_2^0 \left(n^{-1/2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2} \right) \; . \tag{2.22}$$

Remark 2.6 In this Remark, we explain how to adapt the proof of Proposition 2.1 to the explicit scheme defined in (2.13). First, we can find some $\hat{Z}^\pi \in \mathbf{H}^2$ and $\hat{U}^\pi \in \mathbf{L}^2_\lambda$ such that

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

We then define $\hat{Y}^\pi$ on $[t_i, t_{i+1}]$ by

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

Observe that $\hat{Y}^\pi_{t_{i+1}} = \tilde{Y}^\pi_{t_{i+1}}$ and

$$\tilde{Z}^\pi_{t_i} \;=\; n\, \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \hat{Z}^\pi_s ds \mid \mathcal{F}_{t_i} \right] \quad , \quad \tilde{\Gamma}^\pi_{t_i} \;=\; n\, \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \hat{\Gamma}^\pi_s ds \mid \mathcal{F}_{t_i} \right] \; ,$$

for all $i < n$. Moreover

$$\begin{aligned} h(X_s, Y_s, Z_s, \Gamma_s) \;&=\; \mathbb{E}\left[h(X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) \mid \mathcal{F}_{t_i} \right] \\ &+\; \mathbb{E}\left[h(X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) - h(X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) \mid \mathcal{F}_{t_i} \right] \\ &+\; \left(h(X_s, Y_s, Z_s, \Gamma_s) - h(X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) \right) \end{aligned}$$

where by the Lipschitz continuity of h and (i) of Theorem 2.1 below

$$\mathbb{E}\left[\left(\mathbb{E}\left[h(X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) - h(X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) \mid \mathcal{F}_{t_i} \right] \right)^2 \right] \;\leq\; C_2^0 / n \; ,$$

and

$$\begin{aligned} &\mathbb{E}\left[\int_t^{t_{i+1}} \left(h(X_s, Y_s, Z_s, \Gamma_s) - h(X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}) \right)^2 ds \right] \\ &\leq C_2^0 \left(n^{-2} + \int_t^{t_{i+1}} \mathbb{E}\left[|Z_s - \bar{Z}_{t_i}|^2 \right] + \mathbb{E}\left[|\Gamma_s - \bar{\Gamma}_{t_i}|^2 \right] ds \right) \end{aligned}$$

by (i) of Theorem 2.1 and (2.6). Using these remarks, the proof of Proposition 2.1 can be adapted in a straightforward way. This implies that the approximation error due to the explicit scheme is also upper-bounded by $C_2^0 \left(n^{-1/2} + \|Z - \bar{Z}\|_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2} \right)$.

2.4 Path-regularity and convergence rate under additional assumptions

In view of Proposition 2.1, the discretization error converges to zero. In order to control its speed of convergence, it remains to study $\|Z - \bar{Z}\|^2_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2}$. In this section, we shall appeal to one of the additional assumptions :

$\mathbf{H}_1$: 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 $(x,\xi) \in \mathbb{R}^d \times \mathbb{R}^d$

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

$\mathbf{H}_2$: σ, b, $\beta(\cdot,e)$, h and g are C^1_b functions with K-Lipschitz continuous derivatives, uniformly in $e \in E$.

Remark 2.7 Observe for later use that the condition $\mathbf{H}_1$ 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 ensure the inversibility of the first variation process ∇X of X, see Remark 3.2. 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 $\mathbf{H}_1$.

We can now state the main result of this paper.

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

$$\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |Y_t - Y_{t_i}|^2\right] \ \leq \ C_2^0 \ n^{-1} \quad \textit{and} \quad \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\Gamma_t - \Gamma_{t_i}|^2\right] \ \leq \ C_2^0 \ n^{-1} \tag{2.23}$$

so that $\|\Gamma - \bar{\Gamma}\|^2_{\mathcal{S}^2} \leq C_2^0 \ n^{-1}$ *and* $\|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2} \leq C_2^0 \ n^{-1}$.

(ii) *Assume that* $\mathbf{H}_1$ *holds. Then*

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] dt \ \leq \ C_2^0 \ n^{-1} \ . \tag{2.24}$$

so that $\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \leq C_2^0\, n^{-1}$.

(iii) *Assume that* $\mathbf{H}_2$ *holds. Then, for all* $i < n$ *and* $t \in [t_i, t_{i+1}]$,

$$\mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] \;\leq\; C_2^0\, n^{-1}\,, \tag{2.25}$$

so that $\|Z - \bar{Z}\|^2_{\mathbf{H}^2} \leq C_2^0\, n^{-1}$.

This regularity property will be proved in the subsequent sections. Combined with Proposition 2.1 and Remark 2.5, it provides an upper bound for the convergence rate of our backward implicit scheme.

Corollary 2.1 *Assume that either* $\mathbf{H}_1$ *holds, or* $\mathbf{H}_2$ *holds, or* $\sigma = 0$*, or* h *is independent of* Z*. Then,*

$$\mathrm{Err}_n\,(Y, Z, U) \;\leq\; C_2^0\, n^{-1/2}\,.$$

Remark 2.8 In view of Remark 2.6, the result of Corollary 2.1 can be extended to the explicit scheme defined in (2.13).

2.5 Possible Extensions

(i) 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. In this case, the backward scheme (2.10) is modified by setting

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

(ii) The Euler approximation X^π of X could be replaced by any other adapted approximation satisfying (2.9).

(iii) Let M be the solution of the SDE

$$M_t \;=\; M_0 + \int_0^t b_M(M_r)dr + \int_0^t \int_E \beta_M(M_{r-}, e)\bar{\mu}(de, dr)$$

where $b_M : \mathbb{R}^k \mapsto \mathbb{R}^k$ and $\beta_M(\cdot, e) : \mathbb{R}^k \mapsto \mathbb{R}^k$, $k \geq 1$, are Lipschitz continuous uniformly in $e \in E$ with $|\beta_M(0, \cdot)|$ bounded, and consider the system

$$\left\{\begin{array}{rcl} X_t &=& X_0 + \int_0^t b(M_r, X_r)dr + \int_0^t \sigma(M_r, X_r)dW_r + \int_0^t \int_E \beta(M_{r-}, X_{r-}, e)\bar{\mu}(de, dr) \\ Y_t &=& g(M_T, X_T) + \int_t^T h\,(M_r, \Theta_r)\,dr - \int_t^T Z_r \cdot dW_r - \int_t^T \int_E U_r(e)\bar{\mu}(de, dr) \end{array}\right. \tag{2.26}$$

where b, σ, $\beta(\cdot, e)$ and h are K-Lispchitz, uniformly in $e \in E$ and $|\beta(0, \cdot)|$ is bounded. Here, the discrete-time approximation of Y is given by

$$\bar{Y}^\pi_{t_n} = g(M^\pi_{t_n}, X^\pi_{t_n}) \ , \ \bar{Y}^\pi_{t_i} \ = \ \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] + \frac{1}{n}\, h\left(M^\pi_{t_i}, X^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{Z}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}\right) ,$$

where (M^π, X^π) is the Euler scheme of (M, X). Considering (M, X) as an $\mathbb{R}^{k+d}$ dimensional forward process, we can clearly apply the results of Proposition 2.1. Moreover, we claim that Theorem 2.1 (i) holds as well as (ii) (resp. (iii)) if $\mathbf{H}_1$ (resp. $\mathbf{H}_2$) holds for $b(m, \cdot)$, $\sigma(m, \cdot)$, $\beta(m, \cdot)$, $g(m, \cdot)$ and $h(m, \cdot)$ as functions of (x, y, z, γ) uniformly in $m \in \mathbb{R}^k$. This comes from the fact that the dynamics of M are independent of X and that the Malliavin derivative of M with respect to the Brownian motion equals zero. This particular feature implies that the proofs of Section 3.3 and Section 4 work without any modification in this context.

(iv) In [22], see also [25], the authors consider a system of the form

$$\left\{ \begin{array}{rcl} X_t & = & X_0 + \int_0^t b(M_r, X_r) dr + \int_0^t \sigma(M_r, X_r) dW_r \\ Y_t & = & g(M_T, X_T) + \int_t^T h\left(M_r, \Theta_r\right) dr - \int_t^T Z_r \cdot dW_r - \int_t^T \int_E U_r(e) \bar{\mu}(de, dr) \end{array} \right. \tag{2.27}$$

where M is an $\mathbb{F}^\mu$-adapted purely discontinuous jump process. In [22], it is shown that a large class of systems of (coupled) semilinear parabolic partial differential equations can be rewritten in terms of systems of BSDE of the form (2.27), where the backward components are decoupled. However, their particular construction implies that b, σ, h and g are not Lipschitz in their first variable m. In this remark, we explain how to consider this particular framework.

Hereafter, we assume that the path of M can be simulated exactly, which is the case in [22]. Then, recalling that $\lambda(E) < \infty$ so that μ has a.s. only a finite number of jumps on $[0, T]$, we can include the jump times of M in the Euler scheme X^π of X. Thus, even if b and σ are not Lipschitz in their first variable m, we can still define an approximating scheme X^π of X such that

$$\mathbb{E}\left[\sup_{t \in [t_i, t_{i+1}]} |X_t - X^\pi_{t_i}|^2\right] \ \leq \ C_2^0 \left|t_{i+1} - t_i\right|$$

whenever $b(m, \cdot)$ and $\sigma(m, \cdot)$ are Lipschitz in x and $|b(m, 0)| + |\sigma(m, 0)|$ is bounded, uniformly in m. We now explain how to construct a convergent scheme for the backward component even when g and h are not Lipschitz in m. We assume that $h(m, \cdot)$ is Lipschitz and $h(m, 0)$ is bounded, uniformly in m. We make the same assumption on $g(m, \cdot)$. The

approximation is defined as follows:

$$\left\{\begin{array}{rcl} \bar{Z}_t^\pi & := & n\,\mathbb{E}\left[\bar{Y}_{t_{i+1}}^\pi \Delta W_{i+1} \mid \mathcal{F}_{t_i}\right] \\ \bar{\Gamma}_t^\pi & := & n\,\mathbb{E}\left[\bar{Y}_{t_{i+1}}^\pi \int_E \rho(e)\bar{\mu}(de,(t_i,t_{i+1}]) \mid \mathcal{F}_{t_i}\right] \\ \bar{Y}_t^\pi & := & \mathbb{E}\left[\bar{Y}_{t_{i+1}}^\pi \mid \mathcal{F}_{t_i}\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}^\pi, \bar{Y}_{t_{i+1}}^\pi, \bar{Z}_{t_i}^\pi, \bar{\Gamma}_{t_i}^\pi\right) ds \mid \mathcal{F}_{t_i}\right] \end{array}\right. \tag{2.28}$$

for $t \in [t_i, t_{i+1})$, with the terminal condition $\bar{Y}_{t_n}^\pi = g(M_{t_n}, X_{t_n}^\pi)$. With this scheme the proof of Proposition 2.1 can be modified as follows. We keep the same definition for Z^π and U^π but we now define Y^π as

$$\begin{array}{rcl} Y_t^\pi & = & \bar{Y}_{t_i}^\pi - (t-t_i)\,\mathbb{E}\left[n \int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}^\pi, \bar{Y}_{t_{i+1}}^\pi, \bar{Z}_{t_i}^\pi, \bar{\Gamma}_{t_i}^\pi\right) ds \mid \mathcal{F}_{t_i}\right] \\ & + & \displaystyle\int_{t_i}^t Z_s^\pi \cdot dW_s + \int_{t_i}^t \int_E U_s^\pi(e)\bar{\mu}(ds,de)\,. \end{array}$$

Let us introduce the processes $(H_t)_{t\le T}$ and $(\bar{H}_t)_{t\le T}$ defined, for $t \in [t_i, t_{i+1}]$, by

$$H_t \;:=\; h(M_t, X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i})\,, \quad \bar{H}_t \;:=\; \mathbb{E}\left[n \int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}, Y_{t_i}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}\right) ds \mid \mathcal{F}_{t_i}\right].$$

Observe that $h(M_t,\Theta_t) - \mathbb{E}\left[n \int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}\right) ds \mid \mathcal{F}_{t_i}\right]$ can be written as

$$h(M_t,\Theta_t) - H_t + H_t - \bar{H}_{t_i} + \bar{H}_{t_i} - \mathbb{E}\left[n \int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}\right) ds \mid \mathcal{F}_{t_i}\right].$$

Recall from (iii) of this section that (i) of Theorem 2.1 holds for (2.27). Following the arguments of Remark 2.6, we get

$$\mathbb{E}\left[\left|\bar{H}_{t_i} - \mathbb{E}\left[n \int_{t_i}^{t_{i+1}} h\left(M_s, X_{t_i}, Y_{t_{i+1}}, \bar{Z}_{t_i}, \bar{\Gamma}_{t_i}\right) ds \mid \mathcal{F}_{t_i}\right]\right|^2\right] \;\le\; \frac{C_2^0}{n}\,.$$

By (i) of Theorem 2.1 and (2.6),

$$\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|h(M_t,\Theta_t) - H_t|^2\right] dt \;\le\; C_2^0\left(n^{-2} + \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - \bar{Z}_{t_i}|^2 + |\Gamma_t - \bar{\Gamma}_{t_i}|^2\right] dt\right).$$

We then deduce from the same arguments as in the proof of Proposition 2.1 that

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

where

$$\|Z-\bar{Z}\|_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2} + \|H - \bar{H}\|_{\mathbf{H}^2} \;\le\; \epsilon(n)$$

for some map ϵ such that $\epsilon(n) \to 0$ when $n \to \infty$. This shows that the approximation scheme is convergent. Recall from (iii) of this section that the results of Theorem 2.1 for this system. Since here $\beta = 0$, it follows that $\|Z-\bar{Z}\|_{\mathbf{H}^2} + \|\Gamma - \bar{\Gamma}\|_{\mathbf{H}^2} \le C_2^0 n^{-\frac{1}{2}}$, see (iii) of this section. We leave the study of $\|H - \bar{H}\|_{\mathbf{H}^2}$ to further research.

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

$$\begin{aligned} \mathbf{C}_1^X &: b, \sigma \text{ and } \beta(\cdot, e) \text{ are } C_b^1 \text{ uniformly in } e \in E \\ \mathbf{C}_1^Y &: g \text{ and } h \text{ are } C_b^1. \end{aligned}$$

We shall also show that their derivatives are smooth under the stronger assumptions

$$\begin{aligned} \mathbf{C}_2^X &: b, \sigma \text{ and } \beta(\cdot, e) \text{ are } C_b^2 \text{ with second derivatives bounded by K, uniformly in } e \in E \\ \mathbf{C}_2^Y &: g \text{ and } h \text{ are } C_b^2 \text{ with second derivatives bounded by K.} \end{aligned}$$

This will allow us to provide representation and regularity results for Y, Z and U in Section 4. Under $\mathbf{C}_1^X$-$\mathbf{C}_1^Y$, these results will immediately imply (i) of Theorem 2.1, while (ii) of Theorem 2.1 will be obtained by adapting the arguments of [29] under the additional assumption $\mathbf{H}_1$. Under $\mathbf{C}_2^X$-$\mathbf{C}_2^Y$, these results will also directly imply (iii) of Theorem 2.1. The proof of Theorem 2.1 will then be completed by appealing to an approximation argument.

This section is organized as follows. First we derive some properties for the Malliavin derivatives of stochastic integrals with respect to $\bar{\mu}$. Next, we recall some well known results on the Malliavin derivatives of the forward process X. Finally, we discuss the Malliavin differentiability of the solution of (2.3).

3.1 Generalities

We start by introducing some additional notations. We denote by D the Malliavin derivative operator with respect to the Brownian motion and by $\mathbb{D}^{1,2}$ the space of random variables $H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; \mathbb{R})$ such that $D_t H$ exists for all $t \leq T$ and satisfy

$$\|H\|^2_{\mathbb{D}^{1,2}} \;:=\; \mathbb{E}\left[H^2\right] + \mathbb{E}\left[\int_0^T |D_s H|^2 ds\right] \;<\; \infty\,.$$

As usual we extend these notations to vector or matrix valued processes by taking the Malliavin derivative componentwise and by considering the suitable norm.

We then 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,

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

We also define $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$ as the completion of the set

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

for the norm

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

Here, $\mathbf{H}^2(\mathbb{D}^{1,2}, \mathbb{F}^W)$ (resp. $\mathbf{L}^2_\lambda(\mathbb{F}^\mu)$) denotes the set of $\mathbb{F}^W$-adapted (resp. $\mathbb{F}^\mu$-adapted) elements of $\mathbf{H}^2(\mathbb{D}^{1,2})$ (resp. $\mathbf{L}^2_\lambda$). Moreover, we extend the definition of $\|\cdot\|_{\mathbf{H}^2}$ and $\|\cdot\|_{\mathbf{L}^2_\lambda}$ to processes with values in $\mathbb{M}^d$ and $\mathbb{R}^d$ in a natural way.

The two following Lemmas are generalizations of Lemma 3.3 and Lemma 3.4 in [22] which correspond to the case where E is finite, see also Lemma 2.3 in [21] 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

$$D_s H \;:=\; \int_0^T \int_E D_s\psi_t(e)\bar{\mu}(de, dt) \quad \textit{for all } s \le T \;.$$

Proof. Assume that $\psi = \xi\vartheta$ where $\xi \in \mathbf{H}^2(\mathbb{D}^{1,2}, \mathbb{F}^W)$, $\vartheta \in \mathbf{L}^2_\lambda(\mathbb{F}^\mu)$ and $\|\psi\|_{\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})} < \infty$. Then,

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

Since $\lambda(E) < \infty$, we obtain by conditioning by μ that

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

while, see [20],

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

This proves the required result when $\psi \in \mathbf{L}'^2_\lambda(\mathbb{D}^{1,2})$. For the general case, we consider a sequence $(\psi^n)_n$ in $\mathbf{L}'^2_\lambda(\mathbb{D}^{1,2})$ which converges in $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$ to ψ. Then $H^n := \int_0^T \int_E \psi^n_t(e)\bar{\mu}(de, dt)$ is a Cauchy sequence in $\mathbb{D}^{1,2}$ which converges to H. Thus, $H \in \mathbb{D}^{1,2}$. Since $(D_s H^n)_{s\le T}$ converges in $\mathbf{H}^2$ to $(\int_0^T \int_E D_s\psi_t(e)\bar{\mu}(de, dt)))_{s\le T}$, this proves the required result. □

Lemma 3.2 *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*

$$D_s H \;:=\; \xi_s^* + \int_0^T \sum_{i=1}^d D_s \xi_t^i \; dW_t^i + \int_0^T \int_E D_s \psi_t(e)\bar{\mu}(de, dt) \; ,$$

where ξ^* *denotes the transpose of* ξ*.*

Proof. Let $S(W)$ denote the set of random variables of the form

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

with $\kappa \geq 1$, $\phi \in C_b^\infty$ and $f^i : [0,T] \mapsto \mathbb{R}^d$ is a bounded measurable map for each $i \leq \kappa$. Then, the set

$$\mathcal{H} \;:=\; \text{Vect}\left\{ H^W H^{\bar{\mu}} \;:\; H^W \in S(W) \; , \; H^{\bar{\mu}} \in L^\infty(\Omega, \mathcal{F}_T^\mu) \; , \; \mathbb{E}\left[H^W H^{\bar{\mu}} \right] = 0 \right\}$$

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 $H^W H^{\bar{\mu}}$ where $H^W \in S(W)$, $H^{\bar{\mu}} \in L^\infty(\Omega, \mathcal{F}_T^\mu)$ and $\mathbb{E}\left[H^W H^{\bar{\mu}} \right] = 0$. By the representation theorem, 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 [19],

$$H^W \;=\; \mathbb{E}\left[H^W \right] + \int_0^T \mathbb{E}\left[D_t H^W \mid \mathcal{F}_t^W \right] dW_t \; .$$

Thus it follows from Itô's Lemma that

$$H \;=\; \int_0^T H_t^{\bar{\mu}} \, \mathbb{E}\left[D_t H^W \mid \mathcal{F}_t^W \right] dW_t + \int_0^T \int_E H_t^W \psi_t(e)\bar{\mu}(de, dt)$$

where $H_t^{\bar{\mu}} = \mathbb{E}[H^{\bar{\mu}} \mid \mathcal{F}_t]$ and $H_t^W = \mathbb{E}\left[H^W \mid \mathcal{F}_t \right]$. Furthermore the two integrands belong respectively to $\mathbf{H}^2(\mathbb{D}^{1,2})$ and $\mathbf{L}^2_\lambda(\mathbb{D}^{1,2})$. Thus, Lemma 3.1 above and (1.46) in [20] 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. [19]. The extension to the case $\beta \neq 0$ is easily obtained by conditioning by μ, see e.g. [24] for explanations in the case where E is finite, or by combining Lemma 3.1 with a fixed point procedure as in the proof of Theorem 2.2.1. in [19], see also Proposition 3.2 below.

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.

Proposition 3.1 *Assume that $\mathbf{C}_1^X$ holds, then $X_t \in \mathbb{D}^{1,2}$ for all $t \leq T$. For all $s \leq T$ and $k \leq d$, $D_s^k X$ admits a version $\chi^{s,k}$ which solves on $[s,T]$*

$$\chi_t^{s,k} \;=\; \sigma^k(X_{s-}) + \int_s^t \nabla b(X_r)\chi_r^{s,k}dr + \int_s^t \sum_{j=1}^d \nabla\sigma^j(X_r)\chi_r^{s,k}dW_r^j + \int_s^t \int_E \nabla\beta(X_{r-},e)\chi_{r-}^{s,k}\bar{\mu}(dr,de)\;.$$

If moreover $\mathbf{C}_2^X$ holds, then $D_s^k X_t \in \mathbb{D}^{1,2}$ for all $s,t \leq T$ and $k \leq d$. For all $u \leq T$ and $\ell \leq d$, $D_u^\ell D_s^k X$ admits a version $\chi^{u,\ell,s,k}$ which solves on $[u \vee s, T]$

$$\begin{aligned}
\chi_t^{u,\ell,s,k} &= \nabla\sigma^k(X_{s-})\chi_{s-}^{u,\ell} + \nabla\sigma^\ell(X_u-)\chi_{u-}^{s,k} \\
&+ \int_s^t \left(\nabla b(X_r)\chi_r^{u,\ell,s,k} + \sum_{i=1}^d \nabla(\nabla b(X_r))^i \chi_r^{u,\ell}(\chi_r^{s,k})^i\right) dr \\
&+ \int_s^t \sum_{j=1}^d \left(\nabla\sigma^j(X_r)\chi_r^{u,\ell,s,k} + \sum_{i=1}^d \nabla(\nabla\sigma^j(X_r))^i \chi_r^{u,\ell}(\chi_r^{s,k})^i\right) dW_r^j \qquad (3.1) \\
&+ \int_s^t \int_E \left(\nabla\beta(X_{r-},e)\chi_{r-}^{u,\ell,s,k} + \sum_{i=1}^d \nabla(\nabla\beta(X_{r-},e))^i \chi_{r-}^{u,\ell}(\chi_{r-}^{s,k})^i\right) \bar{\mu}(dr,de)\;.
\end{aligned}$$

Remark 3.1 Fix $p \geq 2$ and $r \leq s \leq t \leq u \leq T$. Under $\mathbf{C}_1^X$, it follows from Lemma 5.1 applied to X and χ^s that

$$\begin{aligned}
\|\chi^s\|_{\mathcal{S}^p}^p &\leq C_p\;(1+|X_0|^p) \qquad (3.2)\\
\mathbb{E}\left[|\chi_u^s - \chi_t^s|^p\right] &\leq C_p\,|u-t|\,(1+|X_0|^p) \qquad (3.3)\\
\|\chi^s - \chi^r\|_{\mathcal{S}^p}^p &\leq C_p\,|s-r|\,(1+|X_0|^p)\;. \qquad (3.4)
\end{aligned}$$

If moreover $\mathbf{C}_2^X$ holds then similar arguments show that

$$\|\chi^{r,s}\|_{\mathcal{S}^p}^p \leq C_p\,(1+|X_0|^{2p})\;, \qquad (3.5)$$

where $\chi^{r,s} = (\chi^{r,\ell,s,k})_{\ell,k\leq d}$.

Remark 3.2 Under $\mathbf{C}_1^X$, we can define the first variation process ∇X of X which solves on $[0,T]$

$$\begin{aligned}\nabla X_t &= I_d + \int_0^t \nabla b(X_r)\nabla X_r dr + \int_0^t \sum_{j=1}^d \nabla \sigma^j(X_r)\nabla X_r dW_r^j \\ &+ \int_0^t \int_E \nabla\beta(X_{r-},e)\nabla X_{r-}\bar{\mu}(dr,de)\ . \end{aligned} \tag{3.6}$$

Moreover, under $\mathbf{H}_1$, see Remark 2.7, $(\nabla X)^{-1}$ is well defined and solves on $[0,T]$

$$\begin{aligned}(\nabla X)_t^{-1} &= I_d - \int_0^t (\nabla X)_r^{-1}\left[\nabla b(X_r) - \sum_{j=1}^d \nabla\sigma^j(X_r)\nabla\sigma^j(X_r)\right]dr \\ &+ \int_0^t (\nabla X)_r^{-1}\int_E \nabla\beta(X_r,e)\lambda(de)dr - \int_0^t \sum_{j=1}^d (\nabla X)_r^{-1}\nabla\sigma^j(X_r)dW_r^j \\ &- \int_0^t \int_E (\nabla X)_{r-}^{-1}\left(\nabla\beta(X_{r-},e)+I_d\right)^{-1}\nabla\beta(X_{r-},e)\mu(de,dr)\ . \end{aligned} \tag{3.7}$$

This can be checked by simply applying Itô's Lemma to the product $\nabla X(\nabla X)^{-1}$, see [19] p. 109 for the case where $\beta = 0$.

Remark 3.3 Fix $p \geq 2$. Under $\mathbf{H}_1$-$\mathbf{C}_1^X$, it follows from Remark 2.7 and Lemma 5.1 applied to ∇X and $(\nabla X)^{-1}$ that

$$\|\nabla X\|_{\mathcal{S}^p} + \|(\nabla X)^{-1}\|_{\mathcal{S}^p} \leq C_p\ . \tag{3.8}$$

Remark 3.4 Assume that $\mathbf{H}_1$-$\mathbf{C}_1^X$ 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

$$\chi_r^s = \nabla X_r(\nabla X_{s-})^{-1}\sigma(X_{s-})\mathbf{1}_{s\leq r} \quad \text{for all } s,r \leq T\ . \tag{3.9}$$

3.3 Malliavin calculus on the Backward SDE

In this section, we generalize the result of Proposition 3.1 in [22]. 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 $t \leq T$ 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* $\mathbf{C}_1^X$-$\mathbf{C}_1^Y$ *holds.*
(i) *The triples* (Y,Z,U) *belongs to* $\mathcal{B}^2(\mathbb{D}^{1,2})$. *For each* $s \leq T$ *and* $k \leq d$, *the equation*

$$\Upsilon_t^{s,k} = \nabla g(X_T)\chi_T^{s,k} + \int_t^T \nabla h(\Theta_r)\Phi_r^{s,k}dr - \int_t^T \zeta_r^{s,k}\cdot dW_r - \int_t^T\int_E V_r^{s,k}(e)\bar{\mu}(de,dr) \tag{3.10}$$

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,* $(\Upsilon_t^{s,k}, \zeta_t^{s,k}, V_t^{s,k})_{s,t\le T}$ *is a version of* $(D_s^k Y_t, D_s^k Z_t, D_s^k U_t)_{s,t\le T}$.

(ii) *Assume further that* $\mathbf{C}_2^X$-$\mathbf{C}_2^Y$ *holds. Then, for each* $s \le T$ *and* $k \le d$, $(D_s^k Y, D_s^k Z, D_s^k U)$ *belongs to* $\mathcal{B}^2(\mathbb{D}^{1,2})$*. For each* $u \le T$ *and* $\ell \le d$*, the equation*

$$\begin{aligned}\Upsilon_t^{u,\ell,s,k} &= \left(\chi_T^{u,\ell}\right)' [Hg](X_T)\chi_T^{s,k} + \nabla g(X_T)\chi_T^{u,\ell,s,k} \\ &+ \int_t^T \left[\nabla h(\Theta_r)\Phi^{u,\ell,s,k} + \left(D_u^\ell \Theta_r\right)' [Hh](\Theta_r) D_s^k \Theta_r\right] dr \\ &- \int_t^T \zeta^{u,\ell,s,k}\cdot dW_r - \int_t^T V_r^{u,\ell,s,k}(e)\bar{\mu}(de,dr) \qquad (3.11)\end{aligned}$$

where $\Phi^{u,\ell,s,k} := (\chi^{u,\ell,s,k}, \Upsilon^{u,\ell,s,k}, \zeta^{u,\ell,s,k}, \Gamma^{u,\ell,s,k})$ *with* $\Gamma^{u,\ell,s,k} := \int_E \rho(e) V^{u,\ell,s,k}(e)\lambda(de)$*, and* $[Hg]$ *(resp.* $[Hh]$*) denotes the Hessian matrix of* g *(resp.* h*), admits a unique solution. Moreover,* $(\Upsilon_t^{u,\ell,s,k}, \zeta_t^{u,\ell,s,k}, V_t^{u,\ell,s,k})_{u,s,t\le T}$ *is a version of* $(D_u^\ell D_s^k(Y_t,\, Z_t, U_t))_{u,s,t\le T}$.

Proof. For ease of notations, we only consider the case $d = 1$ and omit the indexes k and ℓ in the above notations.

(i) We proceed as in Proposition 5.3 in [12]. Combined with $\mathbf{C}_X^1$-$\mathbf{C}_Y^1$ and (3.2), Lemma 5.2 implies that $(\Upsilon^s, \zeta^s, V^s)$ is well defined for each $s \le T$ and that we have

$$\sup_{s\le T} \|(\Upsilon^s, \zeta^s, V^s)\|_{\mathcal{B}^p}^p \le C_p\ (1 + |X_0|^p) \quad \text{for all} \quad p \ge 2\,. \qquad (3.12)$$

We now define the sequence $\Theta^n := (X, Y^n, Z^n, \Gamma^n)$ as follows. First, we set $(Y^0, Z^0, U^0) := (0,0,0)$. Then, given Θ^{n-1}, we define (Y^n, Z^n, U^n) as the unique solution in $\mathcal{B}^2$ of

$$Y_t^n = g(X_T) + \int_t^T h(\Theta_r^{n-1})dr - \int_t^T Z_r^n dW_r - \int_t^T \int_E U_r^n(e)\bar{\mu}(de,dr)$$

and set $\Gamma^n = \int_E \rho(e)U^n(e)\lambda(de)$. From the proof of Lemma 2.4 in [26], $(Y^n, Z^n, U^n)_n$ is a Cauchy sequence in $\mathcal{B}^2$ which converges to (Y, Z, U).

Moreover, using Lemma 3.2 and an inductive argument, one obtains that $(Y^n, Z^n, U^n) \in \mathcal{B}^2(\mathbb{D}^{1,2})$. For $s \le T$, set

$$\begin{aligned}&(\Upsilon^{s,n}, \zeta^{s,n}, V^{s,n}) := (D_s Y^n, D_s Z^n, D_s U^n)\ \ , \ \ \Phi^{s,n} := (\chi^s, \Upsilon^{s,n}, \zeta^{s,n}, \Gamma^{s,n})\,, \\ &\Xi^{s,n} := (\chi^s, \Upsilon^{s,n}, \zeta^{s,n}, U^{s,n}) \ \text{ and } \ \Xi^s := (\chi^s, \Upsilon^s, \zeta^s, U^s)\,,\end{aligned}$$

where $\Gamma^{s,n} := \int_E \rho(e)V^{s,n}(e)\lambda(de)$. By Lemma 3.2 again, we have

$$\Upsilon_t^{s,n} = \nabla g(X_T)\chi_T^s + \int_t^T \nabla h(\Theta_r^{n-1})\Phi_r^{s,n-1}dr - \int_t^T \zeta_r^{s,n}dW_r - \int_t^T V_r^{s,n}(e)\bar{\mu}(de,dr)\,. \quad (3.13)$$

Fix $I \in \mathbb{N}$ to be chosen later, set $\delta := T/I$ and $\tau_i := i\delta$ for $0 \le i \le I$. By (5.10) of Lemma 5.2, we have

$$G_i^{s,n} := \|\Xi^s - \Xi^{s,n}\|^4_{\mathcal{S}^4 \times \mathcal{B}^4_{[\tau_i,\tau_{i+1}]}} \le C_4 \left(\mathbb{E}\left[|\Upsilon^s_{\tau_{i+1}} - \Upsilon^{s,n}_{\tau_{i+1}}|^4 \right] + A_i^{s,n-1} + B_i^{s,n-1} \right) \quad (3.14)$$

where

$$\begin{aligned} A_i^{s,n-1} &:= \left\| \{\nabla h(\Theta^{n-1}) - \nabla h(\Theta)\} \Phi^s \right\|^4_{\mathbf{H}^4_{[\tau_i,\tau_{i+1}]}} \\ B_i^{s,n-1} &:= \mathbb{E}\left[\left(\int_{\tau_i}^{\tau_{i+1}} \nabla h(\Theta^{n-1}_r) \{\Phi^s_r - \Phi^{s,n-1}_r\} dr \right)^4 \right] . \end{aligned}$$

Recalling that ρ and the derivatives of h are bounded, we deduce from Cauchy-Schwartz and Jensen's inequality that

$$B_i^{s,n-1} \le C_4 \delta^2 \, G_i^{s,n-1} , \quad (3.15)$$

which combined with an inductive argument and (3.12)-(3.14) leads to

$$\sup_{s \le T} G_i^{s,n} < \infty \quad \text{for all } n \ge 0 . \quad (3.16)$$

Since the derivatives of h are also continuous and Θ^{n-1} converges to Θ in $\mathcal{S}^2 \times \mathcal{B}^2$, we deduce from (3.2)-(3.12) that, after possibly passing to a subsequence,

$$\lim_{n \to \infty} \sup_{s \le T} A_i^{s,n-1} = 0 . \quad (3.17)$$

It follows from (3.14)-(3.15)-(3.17) that for I large enough there is some $\alpha < 1$ such that for any $\varepsilon > 0$ we can find $N' \ge 0$, independent of s, such that

$$G_i^{s,n} \le C_4 \mathbb{E}\left[|\Upsilon^s_{\tau_{i+1}} - \Upsilon^{s,n-1}_{\tau_{i+1}}|^4 \right] + \varepsilon + \alpha G_i^{s,n-1} \quad \text{for } n \ge N' . \quad (3.18)$$

Since $\Upsilon^s_T = \Upsilon^{s,n-1}_T$, we deduce that for $i = I-1$ and $n \ge N'$

$$\sup_{s \le T} G^{s,n}_{I-1} \le \varepsilon + \alpha^{n-N'} \sup_{s \le T} G^{s,N'}_{I-1} .$$

By (3.16), it follows that $\sup_{s \le T} G^{s,n}_{I-1} \to 0$ as $n \to \infty$. In view of (3.18), a straightforward induction argument shows that, for all $i \le I-1$, $\sup_{s \le T} G_i^{s,n} \to 0$ as $n \to \infty$ so that, summing up over i, we get

$$\sup_{s \le T} \|(\Xi^s - \Xi^{s,n})\|_{\mathcal{S}^4 \times \mathcal{B}^4} \underset{n \to \infty}{\longrightarrow} 0 . \quad (3.19)$$

Since (Y^n, Z^n, U^n) converges to (Y, Z, U) in $\mathcal{B}^2$, this shows that $(Y, Z, U) \in \mathcal{B}^2(\mathbb{D}^{1,2})$ and that there is a version of (DY, DZ, DU) given by (Υ, ζ, V).

(ii) In view of (3.2)-(3.5)-(3.12) and $\mathbf{C}_2^X$-$\mathbf{C}_2^Y$, it follows from Lemma 5.2 that $(\Upsilon^{u,s}, \zeta^{u,s}, V^{u,s})$ is well defined for $u, s \leq T$ and that we have

$$\sup_{u,s\leq T} \|(\Upsilon^{u,s}, \zeta^{u,s}, V^{u,s})\|_{\mathcal{B}^p}^p \;\leq\; C_p \left(1 + |X_0|^{2p}\right) \quad \text{for all } \; p \geq 2 \,. \tag{3.20}$$

Using Lemma 3.2, (3.13) and an inductive argument, we then deduce that $(DY^n, DZ^n, DU^n) \in \mathcal{B}^2(\mathbb{D}^{1,2})$ and

$$\begin{aligned}\Upsilon_t^{u,s,n} &= \chi_T^u[Hg](X_T)\chi_T^s + \nabla g(X_T)\chi_T^{u,s} + \int_t^T \nabla h(\Theta_r^{n-1})\Phi_r^{u,s,n-1}dr \\ &+ \int_t^T \Phi_r^{u,n-1}[Hh](\Theta_r^{n-1})\Phi_r^{s,n-1}dr - \int_t^T \zeta_r^{u,s,n}dW_r - \int_t^T V_r^{u,s,n}(e)\bar{\mu}(de,dr) \;\;,\end{aligned}$$

where $(\Upsilon^{u,s,n}, \zeta^{u,s,n}, V^{u,s,n}, \Phi^{u,s,n}) := D_u(\Upsilon^{s,n}, \zeta^{s,n}, V^{s,n}, \Phi^{s,n})$. By (i), (Y^n, Z^n, U^n) goes to (Y, Z, U) in $\mathcal{B}^2$ and $(\Upsilon^{s,n}, \zeta^{s,n}, V^{s,n})$ converges to $(\Upsilon^s, \zeta^s, V^s)$ in $\mathcal{B}^4$. Moreover, (3.19) implies

$$\sup_{n\geq 1}\sup_{s\leq T} \|(\Upsilon^{s,n}, \zeta^{s,n}, V^{s,n})\|_{\mathcal{B}^4}^4 \;<\; \infty \,, \tag{3.21}$$

so that, by dominated convergence, $\mathbf{C}_2^Y$ and (3.20),

$$\|\Phi^{u,n}[Hh](\Theta^n)\Phi^{s,n} - \Phi^u[Hh](\Theta)\Phi^s\|_{\mathbf{H}^2} + \|(\nabla h(\Theta^n) - \nabla h(\Theta))\,\Phi^{u,s}\|_{\mathbf{H}^2} \underset{n\to\infty}{\longrightarrow} 0 \,,$$

after possibly passing to a subsequence. The rest of the proof follows step by step the arguments of (i) except that we now work on $\mathcal{S}^2 \times \mathcal{B}^2$ instead of $\mathcal{S}^4 \times \mathcal{B}^4$. □

Proposition 3.3 *Assume that $\mathbf{C}_1^X$-$\mathbf{C}_1^Y$ holds. For each $k \leq d$, the equation*

$$\nabla Y_t^k \;=\; \nabla g(X_T)\nabla X_T^k + \int_t^T \nabla h(\Theta_r)\nabla\Phi_r^k dr - \int_t^T \nabla Z_r^k \cdot dW_r - \int_t^T \int_E \nabla U_r^k(e)\bar{\mu}(de,dr) \tag{3.22}$$

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 $(\zeta_t^{s,k}, \Upsilon_t^{s,k}, V_t^{s,k})_{s,t\leq T}$ given by $\{(\nabla Y_t, \nabla Z_t, \nabla U_t)(\nabla X_{s-})^{-1}\sigma^k(X_{s-})\mathbf{1}_{s\leq t}\}_{s,t\leq T}$ where ∇Y_t is the matrix whose k-column is given by ∇Y_t^k and ∇Z_t,∇U_t are defined similarly.

Proof. In view of Proposition 3.2 and (3.9), this follows immediately from the uniqueness of the solution of (3.10). □

Remark 3.5 It follows from Lemma 5.2 and (3.8) that

$$\|(\nabla Y, \nabla Z, \nabla U)\|_{\mathcal{B}^p} \;\leq\; C_p \quad \text{for all } \; p \geq 2 \,. \tag{3.23}$$

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 $\mathbf{C}_1^X$-$\mathbf{C}_1^Y$-$\mathbf{H}_1$ or $\mathbf{C}_2^X$-$\mathbf{C}_2^Y$. Similar results under $\mathbf{H}_1$ or $\mathbf{H}_2$ 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 $(\Upsilon^{s,k}(t,x),\ \zeta^{s,k}(t,x),\ V^{s,k}(t,x))$, $(\nabla Y(t,x), \nabla Z(t,x), \nabla U(t,x))$ and $(\Upsilon^{u,\ell,s,k}(t,x),\ \zeta^{u,\ell,s,k}(t,x),\ V^{u,\ell,s,k}(t,x))$. 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

$$u(t,x) \ := \ Y(t,x)_t \ , \ \nabla u(t,x) \ := \ \nabla Y(t,x)_t \ , \ v^{s,k}(t,x) \ := \ \Upsilon^{s,k}(t,x)_t$$
$$\text{and} \qquad w^{u,\ell,s,k}(t,x) \ := \ \Upsilon^{u,\ell,s,k}(t,x)_t \ ,$$

where $(u,s) \in [0,T]^2$ and $k, \ell \leq d$.

Proposition 4.1 (i) *Assume that* $\mathbf{C}_1^X$ *and* $\mathbf{C}_1^Y$ *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* $x \in \mathbb{R}^d$.
(ii) *Assume that* $\mathbf{C}_2^X$ *and* $\mathbf{C}_2^Y$ *hold, then,*

$$|w^{u,\ell,s,k}(t,x)| \leq C_2 \ (1 + |x|^2) \ , \tag{4.2}$$

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

Proof. When $(t,x) = (0,X_0)$, the result follows from (2.5) in Remark 2.1, (3.12), (3.20) and (3.23). The general case is obtained similarly by changing the initial condition on X. □

Proposition 4.2 *Assume that* $\mathbf{C}_1^X$ *and* $\mathbf{C}_1^Y$ *hold.*
(i) *There is a version of* Z *given by* $(\Upsilon_t^t)_{t \leq T}$ *which satisfies*

$$\|Z\|_{\mathcal{S}^p}^p \ \leq \ C_p \ (1 + |X_0|^p) \ . \tag{4.3}$$

(ii) *Assume further that* $\mathbf{C}_2^X$ *and* $\mathbf{C}_2^Y$ *hold, then, for each* $k \leq d$*, there is a version of* $(\zeta^{s,k})_t)_{s,t\leq T}$ *given by* $((\Upsilon_t^{t,\ell,s,k})_{\ell\leq d})_{s,t\leq T}$ *which satisfies*

$$\| \sup_{s\leq T} |\zeta^{s,k}| \|_{\mathcal{S}^p}^p \leq C_p \left(1 + |X_0|^{2p}\right) . \tag{4.4}$$

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 3.2 that

$$D_s Y_t = Z_s - \int_s^t \nabla h(\Theta_r) D_s \Theta_r dr + \int_s^t D_s Z_r \, dW_r + \int_s^t D_s U_r(e) \bar{\mu}(de, dr) , \tag{4.5}$$

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)-(3.10) 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. Under $\mathbf{C}_2^X$ and $\mathbf{C}_2^Y$, the same arguments applied to $(\Upsilon^s, \zeta^s, V^s)$ instead of (Y, Z, U) leads to the second claim, see (ii) of Proposition 3.2, (ii) of Proposition 4.1 and recall (2.5). □

Proposition 4.3 (i) *Define* $\tilde{U}$ *by*

$$\tilde{U}_t(e) := u\left(t, X_{t-} + \beta(X_{t-}, e)\right) - \lim_{r\uparrow t} u\left(r, X_r\right) .$$

Then $\tilde{U}$ *is a version of* U *and it satisfies*

$$\| \sup_{e\in E} |U(e)| \|_{\mathcal{S}^p}^p \leq C_p \left(1 + |X_0|^p\right) . \tag{4.6}$$

(ii) *Assume that* $\mathbf{C}_1^X$ *and* $\mathbf{C}_1^Y$ *holds. Define* $\nabla\tilde{U}$ *by*

$$\nabla\tilde{U}_t(e) := \nabla u\left(t, X_{t-} + \beta(X_{t-}, e)\right) - \lim_{r\uparrow t} \nabla u\left(r, X_r\right) .$$

Then $\nabla\tilde{U}$ *is a version of* ∇U *and it satisfies*

$$\| \sup_{e\in E} |\nabla U(e)| \|_{\mathcal{S}^p}^p \leq C_p . \tag{4.7}$$

(iii) *Assume that* $\mathbf{C}_1^X$ *and* $\mathbf{C}_1^Y$ *holds, then, for each* $k \leq d$*, there is a version of* $(V_t^{s,k})_{s,t\leq T}$ *given by* $(\tilde{V}_t^{s,k})_{s,t\leq T}$ *defined as*

$$\tilde{V}_t^{s,k}(e) := v^{s,k}\left(t, X_{t-} + \beta(X_{t-}, e)\right) - \lim_{r\uparrow t} v^{s,k}\left(r, X_r\right) .$$

It satisfies

$$\| \sup_{e\in E} \sup_{s\leq T} |V^{s,k}(e)| \|_{\mathcal{S}^p}^p \leq C_p \left(1 + |X_0|^p\right) . \tag{4.8}$$

Remark 4.1 We will see in Proposition 4.4 below that u is continuous under $\mathbf{C}_1^X$ and $\mathbf{C}_1^Y$ so that

$$U_t(e) \;:=\; u\left(t, X_{t-} + \beta(X_{t-}, e)\right) - u\left(t, X_{t-}\right) \;.$$

One could similarly show that $v^{s,k}$ and ∇u are continuous under $\mathbf{C}_2^X$ and $\mathbf{C}_2^Y$ so that

$$\begin{aligned} V_t^{s,k}(e) \;&:=\; v^{s,k}\left(t, X_{t-} + \beta(X_{t-}, e)\right) - v^{s,k}\left(t, X_{t-}\right) \\ \nabla\tilde{U}_t(e) \;&:=\; \nabla u\left(t, X_{t-} + \beta(X_{t-}, e)\right) - \nabla u\left(t, X_{t-}\right) \;. \end{aligned}$$

However, since this result is not required for our main theorem, we do not provide its proof.

Proof of Proposition 4.3. By uniqueness of the solution of (2.2)-(2.3) for any initial condition in $L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$ at time t,

$$\int_E U_t(e)\mu(de, \{t\}) \;=\; Y_t - Y_{t-} \;=\; \int_E \tilde{U}_t(e)\mu(de, \{t\}) \;.$$

Hence,

$$\int_0^T \int_E \left|\tilde{U}_t(e) - U_t(e)\right|^2 \mu(de, dt) = 0 \;,$$

which, by taking expectation, implies

$$\mathbb{E}\left[\int_0^T \int_E \left|\tilde{U}_t(e) - U_t(e)\right|^2 \lambda(de)dt\right] = 0 \;.$$

The bound on U follows from (4.1) and (2.5). The two other claims are proved similarly by using (4.1). □

4.2 Path regularity

Proposition 4.4 *Assume that* $\mathbf{C}_1^X$ *and* $\mathbf{C}_1^Y$ *hold. Then,*

$$|u(t_1, x_1) - u(t_2, x_2)|^2 \;\leq\; C_2\,\left\{(1 + |x_1|^2)\,|t_2 - t_1| + |x_1 - x_2|^2\right\}$$

for all $0 \leq t_1 \leq t_2 \leq T$ *and* $(x_1, x_2) \in \mathbb{R}^{2d}$.

Proof. 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 (5.6) of Lemma 5.1, we derive

$$\|\delta X\|^2_{\mathcal{S}^2_{[t_2,T]}} \;\leq\; C_2\left\{|x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1|\right\} \;. \tag{4.9}$$

Plugging this estimate in (5.10) of Lemma 5.2 leads to

$$\|(\delta Y, \delta Z, \delta U)\|^2_{\mathcal{B}^2_{[t_2,T]}} \;\leq\; C_2 \left\{ |x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1| \right\} . \tag{4.10}$$

Now, observe that

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

Plugging (4.3) and (4.6) in (2.7), we get

$$\mathbb{E}\left[\left|Y^1_{t_2} - Y^1_{t_1}\right|^2 \right] \;\leq\; C_2 \;\left(1 + |x_1|^2\right) |t_2 - t_1| \,,$$

which, combined with (4.10), leads to the first claim. □

Corollary 4.1 *Assume that* $\mathbf{C}^X_1$ *and* $\mathbf{C}^Y_1$ *hold.*
(i) *There is a version of* (Y, U) *such that*

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

for all $s \leq t \leq T$.

(ii) *If moreover* $\mathbf{C}^X_2$ *and* $\mathbf{C}^Y_2$ *hold, then there is a version of* Z *such that*

$$\mathbb{E}\left[|Z_t - Z_s|^2 \right] \;\leq\; C_2 \left(1 + |X_0|^4\right) |t - s| \;\;,$$

for all $s \leq t \leq T$.

Proof. (i) Observe that $Y_t = u(t, X_t)$ by uniqueness of the solution of (2.2)-(2.3). Thus, plugging (2.5) and (2.6) in Proposition 4.4 gives the upper-bound on $\mathbb{E}\left[\sup_{r\in[s,t]} |Y_r - Y_s|^2\right]$. 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 using the representation of U given in Remark 4.1.
(ii) By Proposition 4.2, a version of (Z_t) is given by (Υ^t_t) so that

$$\mathbb{E}\left[|Z_t - Z_s|^2 \right] \;\leq\; C_2 \;\left(\mathbb{E}\left[|\Upsilon^t_t - \Upsilon^s_t|^2 \right] + \mathbb{E}\left[|\Upsilon^s_t - \Upsilon^s_s|^2 \right] \right) .$$

By (5.9) of Lemma 5.2, (3.2), (4.4) and (4.8), we have

$$\mathbb{E}\left[|\Upsilon^s_t - \Upsilon^s_s|^2 \right] \;\leq\; C_2 \left(1 + |X_0|^4\right) |t - s| \,.$$

By plugging (3.4) in (5.10) of Lemma 5.2, we then deduce that

$$\mathbb{E}\left[|\Upsilon^t_t - \Upsilon^s_t|^2 \right] \;\leq\; C_2 (1 + |X_0|^2) |t - s| \,.$$

□

Proposition 4.5 *Assume that* $\mathbf{H}_1$-$\mathbf{C}_1^X$-$\mathbf{C}_1^Y$ *hold. Then for all* $n \geq 1$

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] \;\leq\; C_2^0\, n^{-1}\,.$$

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

$$\Lambda_t := \exp\left(\int_0^t \nabla_y h(\Theta_r)\, dr\right) \quad , \quad M_t := 1 + \int_0^t M_r\, \nabla_z h(\Theta_r) \cdot dW_r\,.$$

Since h has bounded derivatives, it follows from Itô's Lemma and Proposition 4.2 that

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

By Remark 3.4 and Proposition 3.3, we deduce that

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

where the process F is defined by

$$F_r \;=\; \nabla_x h(\Theta_r)\nabla X_r + \nabla_\gamma h(\Theta_r)\nabla\Gamma_r \quad \text{for } r \leq T\,.$$

It follows that

$$\Lambda_t M_t Z_t \;=\; \left\{\mathbb{E}\left[G \mid \mathcal{F}_t\right] - \int_0^t F_r\, \Lambda_r\, dr\right\} (\nabla X_{t-})^{-1}\sigma(X_{t-}) \tag{4.11}$$

where

$$G \;:=\; M_T\left(\Lambda_T \nabla g(X_T)\, \nabla X_T + \int_0^T F_r\, \Lambda_r\, dr\right)\,.$$

By (3.8) and (4.7), we deduce that

$$\mathbb{E}\left[|G|^p\right] \;\leq\; C_p^0 \quad \text{for all } \; p \geq 2\,. \tag{4.12}$$

Set $m_s := \mathbb{E}\left[G \mid \mathcal{F}_s\right]$ and let $(\tilde\zeta, \tilde V) \in \mathbf{H}^2 \times \mathbf{L}^2_\lambda$ (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 (4.12) and Lemma 5.2 to $(m, \tilde\zeta, \tilde V)$ implies that

$$\|(m, \tilde\zeta, \tilde V)\|_{\mathcal{B}^p} \;\leq\; C_p^0 \quad \text{for all } \; p \geq 2\,. \tag{4.13}$$

Using $\mathbf{C}_1^X$, (3.8), (4.7), (4.13), 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

$$\|\tilde{Z}\|^p_{\mathcal{S}^p} \;\le\; C_p^0 \quad \text{for all} \;\; p \geq 2 \;, \tag{4.14}$$

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

$$A^p_{[0,T]} \;\le\; C_p^0 \quad \text{for all} \;\; p \geq 2 \tag{4.15}$$

where

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

2. Observe that

$$Z_t \;=\; \tilde{Z}_t \, \sigma(X_t) \quad \mathbb{P} - \text{a.s.}$$

since the probability of having a jump at time t is equal to zero. It follows that, for all $i \leq n$ and $t \in [t_i, t_{i+1}]$,

$$\mathbb{E}\left[|Z_t - Z_{t_i}|^2 \right] \;\le\; C_2 \left(I^1_{t_i,t} + I^2_{t_i,t} \right) \tag{4.16}$$

where

$$I^1_{t_i,t} := \mathbb{E}\left[|\tilde{Z}_t - \tilde{Z}_{t_i}|^2 |\sigma(X_{t_i})|^2 \right] \;\text{ and }\; I^2_{t_i,t} := \mathbb{E}\left[|\sigma(X_t) - \sigma(X_{t_i})|^2 |\tilde{Z}_t|^2 \right]$$

Observing that

$$\begin{aligned} I^1_{t_i,t} \;&=\; \mathbb{E}\left[\mathbb{E}\left[|\tilde{Z}_t - \tilde{Z}_{t_i}|^2 \mid \mathcal{F}_{t_i} \right] |\sigma(X_{t_i})|^2 \right] \\ &\le\; C_2 \mathbb{E}\left[\left(\int_{t_i}^{t_{i+1}} \left[|\tilde{\mu}_r|^2 + |\tilde{\sigma}_r|^2 + \int_E |\tilde{\beta}_r(e)|^2 \lambda(de) \right] dr \right) |\sigma(X_{t_i})|^2 \right] \end{aligned}$$

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

$$\begin{aligned} \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} I^1_{t_i,t} dt \;&\le\; C_2 \, n^{-1} \, \mathbb{E}\left[\left(\int_0^T \left[|\tilde{\mu}_r|^2 + |\tilde{\sigma}_r|^2 + \int_E |\tilde{\beta}_r(e)|^2 \lambda(de) \right] dr \right) \sup_{t \leq T} |\sigma(X_t)|^2 \right] \\ &\le\; C_2^0 (A^4_{[0,T]})^{\frac{1}{2}} \, n^{-1} \;. \end{aligned} \tag{4.17}$$

Using the Lipschitz continuity of σ, we obtain

$$I^2_{t_i,t} \;\leq\; C_2\mathbb{E}\left[|X_t - X_{t_i}|^2|\tilde{Z}_t|^2\right] . \tag{4.18}$$

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

$$\mathbb{E}\left[(X^k_t - X^k_{t_i})^2(\tilde{Z}^l_t)^2\right] \;\leq\; C_2\left(\mathbb{E}\left[(\tilde{Z}^l_t - \tilde{Z}^l_{t_i})^2(X^k_{t_i})^2\right] + \mathbb{E}\left[(X^k_t\tilde{Z}^l_t - X^k_{t_i}\tilde{Z}^l_{t_i})^2\right]\right) \tag{4.19}$$

Arguing as above, we obtain

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \mathbb{E}\left[(\tilde{Z}^l_t - \tilde{Z}^l_{t_i})^2(X^k_{t_i})^2\right] \leq C^0_2\left(1 + (A^4_{[0,T]})^{\frac{1}{2}}\right) n^{-1} . \tag{4.20}$$

Moreover, we deduce from the linear growth condition on b, σ, β and (2.5), (4.14) and (4.15) that $X^k\tilde{Z}^l$ can be written as

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

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

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

which combined with (4.18), (4.19) and (4.20) leads to

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} I^2_{t_i,t}dt \leq C^0_2(1 + (A^4_{[0,T]})^{\frac{1}{2}})\, n^{-1} . \tag{4.21}$$

The proof is concluded by plugging (4.17)-(4.21) in (4.16) and recalling (4.15). □

We now complete the proof of Theorem 2.1.

Proof of Theorem 2.1. 1. We first prove (ii). Observe that the second assertion is a direct consequence of (2.24) and Remark 2.4.
We first show that (2.24) holds under $\mathbf{H}_1$ and $\mathbf{C}^Y_1$. We consider a C^∞_b density q on $\mathbb{R}^d$ with compact support and set

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

For large $k \in \mathbb{N}$, these functions are bounded by $2K$ at 0. Moreover, they are K-Lipschitz and C^1_b. Using the continuity of σ, one also easily checks that σ^k is still invertible. By $\mathbf{H}_1$ and Remark 2.7, 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 $\mathbf{H}_1$ and $\mathbf{C}_1^Y$.
We now prove that (2.24) holds under $\mathbf{H}_1$. 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. The same approximation argument shows that (i) of Corollary 4.1 holds true without $\mathbf{C}_1^X$-$\mathbf{C}_1^Y$. Since ρ is bounded and $\lambda(E) < \infty$, this leads to (2.23). Now observe that

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

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

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

Thus, (2.23) implies $\|\Gamma - \bar{\Gamma}\|^2_{\mathcal{S}^2} \leq C_2^0\ n^{-1}$ and $\|\Gamma - \bar{\Gamma}\|^2_{\mathbf{H}^2} \leq C_2^0\ n^{-1}$.

3. Item (iii) is proved similarly by using (ii) of Corollary 4.1. □

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. The proofs being standard, we do not provide all the details.

Proposition 5.1 Given $\psi \in \mathbf{L}^2_\lambda$, let M be defined on $[0, T]$ by $M_t = \int_0^t \int_E \psi_s(e)\bar{\mu}(ds, de)$. Then, for all $p \geq 2$,

$$k_p\ \|\psi\|^p_{\mathbf{L}^p_{\lambda,[0,T]}} \;\leq\; \|M\|^p_{\mathcal{S}^p_{[0,T]}} \;\leq\; K_p\ \|\psi\|^p_{\mathbf{L}^p_{\lambda,[0,T]}}\ . \tag{5.1}$$

where k_p, K_p are positive numbers that depend only on p, $\lambda(E)$ and T.

Proof. 1. We first prove the left hand-side. Observe that for a sequence $(a_i)_{i\in I}$ of non-negative numbers we have

$$\sum_{i\in I} a_i^\alpha \;\leq\; \left(\max_{i\in I} a_i\right)^{\alpha-1} \sum_{i\in I} a_i \leq \left(\sum_{i\in I} a_i\right)^\alpha \quad \text{for all } \alpha \geq 1\ . \tag{5.2}$$

It follows that

$$\|\psi\|^p_{\mathbf{L}^p_{\lambda,[0,T]}} \;=\; \mathbb{E}\left[\int_0^T\int_E |\psi_s(e)|^p\mu(de,ds)\right] \le \mathbb{E}\left[\left|\int_0^T\int_E |\psi_s(e)|^2\mu(de,ds)\right|^{\frac{p}{2}}\right] ,$$

since $p/2 \ge 1$, and the result follows from Burkholder-Davis-Gundy inequality (see e.g. [23] p. 175).

2. We now prove the right hand-side inequality for $p \ge 1$. We follow the inductive argument of [5]. For $p \in [1,2]$, we deduce from Burkholder-Davis-Gundy inequality and (5.2) that

$$\mathbb{E}\left[\sup_{s\le T}|M_s|^p\right] \;\le\; K_p\,\mathbb{E}\left[\left(\int_0^T\int_E |\psi_s(e)|^2\mu(de,ds)\right)^{\frac{p}{2}}\right] \;\le\; K_p\,\mathbb{E}\left[\int_0^T\int_E |\psi_s(e)|^p\mu(de,ds)\right]$$

since $2/p \ge 1$. This implies the required result.

We now assume that the inequality is valid from some $p > 1$ and prove that it is also true for $2p$. Set $\tilde M_t = \int_0^t\int_E \psi_s(e)^2\bar\mu(de,ds)$ for $t \in [0,T]$. Then, $[M,M]_T = \tilde M_T + \int_0^T\int_E \psi_s(e)^2\lambda(de)ds$. Applying Burkholder-Davis-Gundy inequality, we obtain $\mathbb{E}\left[\sup_{s\le T}|M_s|^{2p}\right] \le \mathbb{E}\left[\,[M,M]_T^p\,\right]$ where

$$\mathbb{E}\left[\,[M,M]_T^p\,\right] \;\le\; K_p\,\mathbb{E}\left[|\tilde M_T|^p + \left(\int_0^T\int_E \psi_s(e)^2\lambda(de)ds\right)^p\right]$$

and K_p denotes a generic positive number that depends only on p. Applying (5.1) to $\tilde M$, we obtain

$$\mathbb{E}\left[|\tilde M_T|^p\right] \;\le\; K_p\,\mathbb{E}\left[\int_0^T\int_E |\psi_s(e)|^{2p}\lambda(de)ds\right] .$$

On the other hand, it follows from Hölder inequality that

$$\int_0^T\int_E \psi_s(e)^2\lambda(de)ds \;\le\; \left(\int_0^T\int_E |\psi_s(e)|^{2p}\lambda(de)ds\right)^{\frac{1}{p}} (T\lambda(E))^{\frac{1}{q}}$$

where $q = p/(p-1)$, recall that $p > 1$. Combining the two last inequalities leads to the required result. □

We now consider some measurable maps

$$\begin{aligned}
\tilde b^i &: \Omega\times[0,T]\times\mathbb{R}^d \mapsto \mathbb{R}^d\\
\tilde\sigma^i &: \Omega\times[0,T]\times\mathbb{R}^d \mapsto \mathbb{M}^d\\
\tilde\beta^i &: \Omega\times[0,T]\times\mathbb{R}^d\times E \mapsto \mathbb{R}^d\\
\tilde f^i &: \Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^d\times L^2(E,\mathcal{E},\lambda;\mathbb{R}) ,\ i=1,2 .
\end{aligned}$$

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 $t\le T$

$$\tilde{b}^i(t,\cdot)\ ,\ \tilde{\sigma}^i(t,\cdot)\ ,\ \tilde{\beta}^i(t,\cdot,e) \quad \text{and} \quad \tilde{f}^i(t,\cdot) \text{ are a.s. } K\text{-Lipschitz continuous}$$

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

$$X^i_t \;=\; \tilde{X}^i+\int_{t_i}^t \tilde{b}^i(s,X^i_s)ds+\int_{t_i}^t\tilde{\sigma}^i(s,X^i_s)dW_s+\int_{t_i}^t\int_E\tilde{\beta}^i(s,e,X^i_{s-})\bar{\mu}(de,ds)\ . \tag{5.3}$$

Lemma 5.1

$$\|X^1\|^p_{\mathcal{S}^p_{[t_1,T]}}\ \le\ C_p\ \left\{\mathbb{E}[|\tilde{X}^1|^p]+\|\tilde{b}^1(\cdot,0)\|^p_{\mathbf{H}^p_{[t_1,T]}}+\|\tilde{\sigma}^1(\cdot,0)\|^p_{\mathbf{H}^p_{[t_1,T]}}+\|\tilde{\beta}^1(\cdot,0,\cdot)\|^p_{\mathbf{L}^p_{\lambda,[t_1,T]}}\right\}\ . \tag{5.4}$$

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

$$\mathbb{E}\left[\sup_{s\le u\le t}|X^1_u-X^1_s|^p\right]\ \le\ C_p\ A^1_p\,|t-s|\ , \tag{5.5}$$

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 $t_2\le t\le T$,

$$\begin{aligned}\|\delta X\|^p_{\mathcal{S}^p_{[t_2,T]}}\ &\le\ C_p\ \left(E|\tilde{X}^1-\tilde{X}^2|^p+A^1_p\,|t_2-t_1|\right)\\ &+\ C_p\ \left(E\left(\int_{t_2}^T|\delta\tilde{b}_t|dt\right)^p+\|\delta\tilde{\sigma}\|^p_{\mathbf{H}^p_{[t_2,T]}}+\|\delta\tilde{\beta}\|^p_{\mathbf{L}^p_{\lambda,[t_2,T]}}\right)\end{aligned} \tag{5.6}$$

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*

$$Y_t\ =\ \tilde{Y}+\int_t^T\tilde{f}(s,Y_s,Z_s,U_s)ds+\int_t^T Z_s\cdot dW_s+\int_t^T\int_E U_s(e)\bar{\mu}(de,ds) \tag{5.7}$$

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

$$\|(Y,Z,U)\|^p_{\mathcal{B}^p} \;\leq\; C_p\;\mathbb{E}\left[|\tilde{Y}|^p+\left(\int_0^T|\tilde{f}(t,0)|dt\right)^p\right]\;. \tag{5.8}$$

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

$$\mathbb{E}\left[\sup_{s\leq u\leq t}|Y_u-Y_s|^p\right]\leq C_p\;\left\{A_p\,|t-s|^p+\|Z\|^p_{\mathbf{H}^p_{[s,t]}}+\|U\|^p_{\mathbf{L}^p_{\lambda,[s,t]}}\right\}\;. \tag{5.9}$$

(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* (5.8) *with* $(\tilde{Y}^i,\tilde{f}^i)$ *in place of* $(\tilde{Y},\tilde{f})$*,* $i=1,2$*. Then, for all* $t\leq T$*,*

$$\|(\delta Y,\delta Z,\delta U)\|^p_{\mathcal{B}^p_{[t,T]}} \;\leq\; C_p\;\mathbb{E}\left[|\delta\tilde{Y}|^p+\left(\int_t^T|\delta\tilde{f}_r|dr\right)^p\right] \tag{5.10}$$

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*

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

Proof of Lemma 5.1. Applying Burkholder-Davis-Gundy inequality (see e.g. [23] p 175) and using Proposition 5.1, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[t_1,T]}|X^1_s|^p\right] &\leq C_p\;\mathbb{E}\left[|\tilde{X}^1|^p+\left(\int_{t_1}^T|\tilde{b}^1(s,X^1_s)|ds\right)^p\right]\\ &+ C_p\;\left(\|\tilde{\sigma}^1(\cdot,X^1_\cdot)\|^p_{\mathbf{H}^p_{[t_1,T]}}+\|\tilde{\beta}^1(\cdot,X^1_\cdot,\cdot)\|^p_{\mathbf{L}^p_{\lambda,[t_1,T]}}\right)\;.\end{aligned}$$

The estimate (5.4) is then deduced by using the Lipschitz properties of $\tilde{b}^1$, $\tilde{\sigma}^1$, $\tilde{\beta}^1$ and Gronwall's Lemma. The estimate (5.5) is obtained by applying the same arguments to the process $|X^1_\cdot-X^1_s|^p$ on $[s,t]$. To obtain the last assertion (5.6), we first apply the above argument to $\delta X=X^1-X^2$ on $[t_2,T]$. Then, decomposing $\tilde{b}^1(\cdot,X^1)-\tilde{b}^2(\cdot,X^2)$ as $\delta\tilde{b}+\tilde{b}^2(\cdot,X^1)-\tilde{b}^2(\cdot,X^2)$ and doing the same for $\tilde{\sigma}$ and $\tilde{\beta}^i$, the Lipschitz properties of $\tilde{b}^2$, $\tilde{\sigma}^2$, $\tilde{\beta}^2$ combined with Gronwall's lemma leads to

$$\mathbb{E}\left[\sup_{s\in[t_2,T]}|\delta X_s|^p\right] \;\leq\; C_p\;\left(\mathbb{E}|X^1_{t_2}-\tilde{X}^2|^p+\mathbb{E}\left(\int_{t_2}^T|\delta\tilde{b}_t|dt\right)^p+\|\delta\tilde{\sigma}\|^p_{\mathbf{H}^p_{[t_2,T]}}+\|\delta\tilde{\beta}\|^p_{\mathbf{L}^p_{\lambda,[t_2,T]}}\right)\;.$$

We then conclude by using the (5.5). □

Proof of Lemma 5.2. See [26] and [2] for existence and uniqueness.
(i) We divide $[0,T]$ in N intervals $[\tau_i,\tau_{i+1}]$ of equal length $\delta:=T/N$. For $\tau_i\leq t\leq s\leq\tau_{i+1}$

$$|Y_s|\;\leq\;\mathbb{E}\left[|Y_{\tau_{i+1}}|+\int_t^{\tau_{i+1}}|\tilde{f}(r,Y_r,Z_r,U_r)|dr\,|\,\mathcal{F}_s\right]\;,$$

which, by Doob and Jensen's inequalities, implies

$$\mathbb{E}\left[\sup_{t\le s\le\tau_{i+1}}|Y_s|^p\right] \;\le\; C_p\;\mathbb{E}\left[|Y_{\tau_{i+1}}|^p+\left(\int_t^{\tau_{i+1}}|\tilde f(r,Y_r,Z_r,U_r)|dr\right)^p\right].$$

Moreover, it follows from Burkholder-Davis-Gundy inequality (see e.g. [23] p. 175) and Proposition 5.1 that

$$\|Z\|^p_{\mathbf{H}^p_{[t,\tau_{i+1}]}}+\|U\|^p_{\mathbf{L}^p_{\lambda,[t,\tau_{i+1}]}} \;\le\; C_p\;\mathbb{E}\left[|Y_{\tau_{i+1}}|^p+\left(\int_t^{\tau_{i+1}}|\tilde f(r,Y_r,Z_r,U_r)|dr\right)^p+\sup_{t\le s\le\tau_{i+1}}|Y_s|^p\right].$$

Thus, using Hölder and Jensen's inequalities, we obtain

$$\begin{aligned}\|(Y,Z,U)\|^p_{\mathcal{B}^p_{[t,\tau_{i+1}]}} &\le C_p\;\mathbb{E}\left[|Y_{\tau_{i+1}}|^p+\left(\int_t^{\tau_{i+1}}|\tilde f(r,Y_r,Z_r,U_r)|dr\right)^p\right]\\ &\le C_p\left\{\mathbb{E}\left[|Y_{\tau_{i+1}}|^p+\left(\int_0^T|\tilde f(t,0)|dt\right)^p\right]+\int_t^{\tau_{i+1}}\|Y\|^p_{\mathcal{S}^p_{[u,\tau_{i+1}]}}du\right.\\ &\qquad\left.+\,\delta^{p/2}\left(\|Z\|^p_{\mathbf{H}^p_{[t,\tau_{i+1}]}}+\|U\|^p_{\mathbf{L}^p_{\lambda,[t,\tau_{i+1}]}}\right)\right\}\end{aligned}$$

by the Lipschitz continuity assumption on $\tilde f$. For δ smaller than $(2C_p)^{-2/p}$, we then get

$$\|(Y,Z,U)\|^p_{\mathcal{B}^p_{[t,\tau_{i+1}]}} \;\le\; C_p\;\left\{\mathbb{E}\left[|Y_{\tau_{i+1}}|^p\right]+\left(\int_0^T|\tilde f(t,0)|dt\right)^p+\int_t^{\tau_{i+1}}\|Y\|^p_{\mathcal{S}^p_{[u,\tau_{i+1}]}}du\right\}.$$

Using Gronwall's Lemma, we deduce that

$$\|Y\|^p_{\mathcal{S}^p_{[\tau_i,\tau_{i+1}]}} \;\le\; C_p\;\left\{\mathbb{E}\left[|Y_{\tau_{i+1}}|^p\right]+\left(\int_0^T|\tilde f(t,0)|dt\right)^p\right\}.$$

Plugging this estimate into the previous upper bound, we finally get

$$\|(Y,Z,U)\|^p_{\mathcal{B}^p_{[\tau_i,\tau_{i+1}]}} \;\le\; C_p\,\mathbb{E}\left[|Y_{\tau_{i+1}}|^p+\left(\int_0^T|\tilde f(t,0)|dt\right)^p\right].$$

This leads to (5.8).
By Burkholder-Davis-Gundy inequality and Proposition 5.1, we have

$$\begin{aligned}\mathbb{E}\left[\sup_{s\le u\le t}|Y_u-Y_s|^p\right] &\le C_p\;\mathbb{E}\left[\left(\int_s^t|\tilde f(r,Y_r,Z_r,U_r)|dr\right)^p\right]\\ &+ C_p\;\left\{\|Z\|^p_{\mathbf{H}^p_{[s,t]}}+\|U\|^p_{\mathbf{L}^p_{\lambda,[s,t]}}\right\}.\end{aligned}$$

Using the Lipschitz continuity assumption on $\tilde f$ together with (5.8) leads to (5.9).

(ii) The estimate (5.10) is obtained by applying similar arguments to $(\delta Y,\delta Z,\,\delta U)$. □

References

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

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

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

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

[5] Bitcheler 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.

[6] 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.

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

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

[9] Chevance D. (1997). Numerical Methods for Backward Stochastic Differential Equations. In *Numerical methods in finance*, Edt L.C.G. Rogers and D. Talay, Cambridge University Press, 232-244.

[10] Coquet F. , V. Mackevičius, and J. Mémin (1998). Stability in D of martingales and backward equations under discretization of filtration. *Stochastic Processes and their Applications*, 75, 235-248.

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

[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. To appear in *Stochastic processes and their Applications*.

[14] Fujiwara T. and H. Kunita (1989). Stochastic differential equations of Jump type and Lévy processes in diffeomorphism group. *J. Math. Kyoto Univ.*, 25 (1), 71-106.

[15] Gobet E., J.P. Lemor and X. Warin (2005). Rate of convergence of empirical regression method for solving generalized BSDE. *preprint Ecole Polytechnique*, R.I.574.

[16] Longstaff F. A. and R. S. Schwartz (2001). Valuing American Options By Simulation : A simple Least-Square Approach. *Review of Financial Studies*, 14, 113-147.

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

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

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

[21] 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.

[22] 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.

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

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

[25] Sow A. B. and E. Pardoux (2004). Probabilistic interpretation of a system of quasilinear parabolic PDEs. *Stochastics and Stochastics Reports*, 76 (5), 429-477.

[26] 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.

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

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

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