

DISCRETE-TIME APPROXIMATION OF DECOUPLED FORWARD–BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS DRIVEN BY PURE JUMP LÉVY PROCESSES

SOUFIANE AAZIZI,* *Cadi Ayyad University*

Abstract

We present a new algorithm to discretize a decoupled forward–backward stochastic differential equation driven by a pure jump Lévy process (FBSDEL for short). The method consists of two steps. In the first step we approximate the FBSDEL by a forward–backward stochastic differential equation driven by a Brownian motion and Poisson process (FBSDEBP for short), in which we replace the small jumps by a Brownian motion. Then, we prove the convergence of the approximation when the size of small jumps ε goes to 0. In the second step we obtain the L^p -Hölder continuity of the solution of the FBSDEBP and we construct two numerical schemes for this FBSDEBP. Based on the L^p -Hölder estimate, we prove the convergence of the scheme when the number of time steps n goes to ∞ . Combining these two steps leads to the proof of the convergence of numerical schemes to the solution of FBSDEs driven by pure jump Lévy processes.

Keywords: Discrete-time approximation; Euler scheme; decoupled forward–backward SDE with jumps; small jumps; Malliavin calculus

2010 Mathematics Subject Classification: Primary 60H35

Secondary 60H07; 60J75

1. Introduction and summary

In this paper we are concerned with the discretization of a system of decoupled forward–backward stochastic differential equations (FBSDEs) driven by a pure jump Lévy process:

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_r) dr + \int_0^t \int_{\mathbb{R}} \beta(X_{r-}) \bar{M}(de, dr), \\ Y_t &= g(X_T) + \int_t^T f(\Theta_r) dr - \int_t^T \int_{\mathbb{R}} V_r \bar{M}(de, dr). \end{aligned} \tag{1.1}$$

Here $\Theta := (X, Y, \int_E \rho(e) V e \nu(de))$ and $\bar{M}(E, t) = \int_{E \times [0, t]} e \bar{\mu}(de, dr)$, where $\bar{\mu}(de, dr) := \mu(de, dr) - \nu(de) dr$ is an independent compensated Poisson measure and μ is a Poisson random measure on $\mathbb{R} \times [0, T]$ with intensity ν satisfying $\int 1 \wedge |e|^2 \nu(de) < \infty$.

Numerical discretization schemes for FBSDEs have been studied by many authors. In the no-jump case, Ma *et al.* [19] developed the first step algorithm to solve a class of general FBSDEs. Douglas *et al.* [12] suggested a finite difference approximation of the associated PDE. Other discrete schemes have been considered in [7], [8], and [10], based mainly on the approximation of the Brownian motion by some discrete process; Lemor *et al.* [18] proposed

Received 10 January 2010; revision received 5 September 2012.

* Postal address: Department of Mathematics, Faculty of Sciences Semlalia Cadi Ayyad University, B.P. 2390 Marrakesh, Morocco. Email address: aazizi.soufiane@gmail.com

Supported by the Marie Curie Initial Training Network (ITN) project ‘Deterministic and Stochastic Controlled Systems and Application’, FP7-PEOPLE-2007-1-1-ITN, no. 213841-2.

an adapted Longstaff and Schwartz algorithm based on nonparametric regressions. In the jump case, to the author’s knowledge, there is only the work of Bouchard and Elie [5], who proposed Monte Carlo methods in the case when $\nu(\mathbb{R}) < \infty$.

The main motivation to study the numerical scheme of a system of the above form is to examine the case when $\nu(\mathbb{R}) = \infty$, i.e. the existence of an infinite number of jumps in every interval of nonzero length almost surely. In this sense, we should mention the important work on the approximation of stochastic differential equations studied by Kohatsu-Higa and Tankov [17].

Since we are interested in the case $\nu(\mathbb{R}) = \infty$, we will follow the idea of [17] to approximate (1.2) below without cutoff of jumps smaller than ε , which should improve the approximation scheme. Then by using the approximation result of Asmussen and Rosinski [2] we replace the small jumps of the Lévy-driven process with $\sigma(\varepsilon)W$, where W is a standard Brownian motion and $\sigma^2(\varepsilon) := \int_{E^\varepsilon} e^2 \nu(de)$.

To approximate (1.1), we cut the jumps at ε as

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_r) dr + \int_0^t \beta(X_{r-}) dR_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-}) \bar{M}(de, dr), \\ Y_t &= g(X_T) + \int_t^T f(\Theta_r) dr - \int_t^T V_r dR_r - \int_t^T \int_{E_\varepsilon} V_r \bar{M}(de, dr), \end{aligned} \tag{1.2}$$

where $R_t = \int_0^t \int_{|e| \leq \varepsilon} e \bar{M}(de, dr)$, $E^\varepsilon := \{e \in \mathbb{R} \text{ such that } |e| \leq \varepsilon\}$, $E_\varepsilon := \{e \in \mathbb{R} \text{ such that } |e| > \varepsilon\}$, and $E := \mathbb{R} = E^\varepsilon \cup E_\varepsilon$.

The idea we propose is to discretize the solution of (1.1) in two steps. In the first step we approximate (1.2) by the following FBSDE:

$$\begin{aligned} X_t^\varepsilon &= X_0^\varepsilon + \int_0^t b(X_r^\varepsilon) dr + \sigma(\varepsilon) \int_0^t \beta(X_r^\varepsilon) dW_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-}^\varepsilon) \bar{M}(de, dr), \\ Y_t^\varepsilon &= g(X_T^\varepsilon) + \int_t^T f(\Theta_r^\varepsilon) dr - \int_t^T Z_r^\varepsilon dW_s - \int_t^T \int_{E_\varepsilon} U_r^\varepsilon(e) \bar{M}(de, dr). \end{aligned} \tag{1.3}$$

Here $\Theta^\varepsilon := (X^\varepsilon, Y^\varepsilon, \Gamma^\varepsilon)$ and $\Gamma^\varepsilon := \int_{E_\varepsilon} \rho(e) U^\varepsilon(e) e \nu(de)$. Furthermore, we show that, for a finite measure m defined by $m(E) := \int_E e^2 \nu(de)$, our error

$$\begin{aligned} \text{Err}_\varepsilon(Y, V)^2 &:= \mathbb{E} \left[\sup_{t \leq T} |Y_t - Y_t^\varepsilon|^2 \right] + \mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t V_r dR_r - \int_0^t Z_r^\varepsilon dW_r \right|^2 \right] \\ &\quad + \mathbb{E} \left[\int_0^T \int_{E_\varepsilon} |V_r - U_r^\varepsilon(e)|^2 m(de) dr \right], \end{aligned}$$

is controlled by $\sigma(\varepsilon)^2$, which means that the solution of (1.3) converges to the solution of (1.1), as the size of the small jumps ε goes to 0 (see Remark 2.1 below). We also derive the upper bound

$$\mathbb{E} \left[\sup_{t \leq T} |X_t - X_t^\varepsilon|^2 \right] \leq C \sigma(\varepsilon)^2.$$

In the second step we discretize the approximated FBSDE (1.3) and study its convergence to (1.2). For this purpose, we consider two numerical schemes. The first scheme is based on the discrete-time approximation of decoupled FBSDEs derived by Bouchard and Elie [5]. More precisely, for a fixed ε , given a regular grid $\pi = \{t_i = iT/n, i = 0, 1, \dots, n\}$, the authors

approximated X^ε by its Euler scheme $\bar{X}^{n,\varepsilon}$ and $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ by the discrete-time process $(\bar{Y}_t^{n,\varepsilon}, \bar{Z}_t^{n,\varepsilon}, \bar{\Gamma}_t^{n,\varepsilon})$, i.e.

$$\begin{aligned} \bar{X}_{t_{i+1}}^{n,\varepsilon} &= \bar{X}_{t_i}^{n,\varepsilon} + \frac{1}{n}b(\bar{X}_{t_i}^{n,\varepsilon}) + \sigma(\varepsilon)\beta(\bar{X}_{t_i}^{n,\varepsilon})\Delta W_{i+1} + \int_{E_\varepsilon} \beta(\bar{X}_{t_i}^{n,\varepsilon})\bar{M}(de, (t_i, t_{i+1})), \\ \bar{Z}_t^{n,\varepsilon} &= n\mathbb{E}[\bar{Y}_{t_{i+1}}^{n,\varepsilon}\Delta W_{i+1} \mid \mathcal{F}_{t_i}], \\ \bar{\Gamma}_t^{n,\varepsilon} &= n\mathbb{E}\left[\bar{Y}_{t_{i+1}}^{n,\varepsilon} \int_{E_\varepsilon} \rho(e)\bar{M}(de, (t_i, t_{i+1})) \mid \mathcal{F}_{t_i}\right], \\ \bar{Y}_t^{n,\varepsilon} &= \mathbb{E}[\bar{Y}_{t_{i+1}}^{n,\varepsilon} \mid \mathcal{F}_{t_i}] + \frac{1}{n}f(\bar{X}_{t_i}^{n,\varepsilon}, \bar{Y}_{t_i}^{n,\varepsilon}, \bar{\Gamma}_{t_i}^{n,\varepsilon}), \end{aligned}$$

on each interval $[t_i, t_{i+1})$, where the terminal value $\bar{Y}_n^{n,\varepsilon} := g(\bar{X}_n^{n,\varepsilon})$. By adapting the proof of [5], we prove, under Lipschitz continuity of the solution, that the discretization error

$$\overline{\text{Err}}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)^2 := \sup_{t \leq T} \mathbb{E}[|Y_t^\varepsilon - \bar{Y}_t^{n,\varepsilon}|^2] + \int_0^T \mathbb{E}[|Z_t^\varepsilon - \bar{Z}_t^{n,\varepsilon}|^2 + |\Gamma_t^\varepsilon - \bar{\Gamma}_t^{n,\varepsilon}|^2] dt$$

achieves the optimal convergence rate $n^{-1/2}$. Finally, we derive the first main result of this paper (Proposition 3.1), showing that the approximation-discretization error

$$\begin{aligned} \overline{\text{Err}}_{(n,\varepsilon)}(Y, V)^2 &:= \sup_{t \leq T} \mathbb{E}\left[|Y_t - \bar{Y}_t^{n,\varepsilon}|^2 + \left|\int_0^t V_r dR_r - \int_0^t \bar{Z}_r^{n,\varepsilon} dW_r\right|^2\right] \\ &\quad + \|\Gamma - \bar{\Gamma}^{n,\varepsilon}\|_{H^2}^2 \end{aligned}$$

is bounded by $C(n^{-1} + \sigma(\varepsilon)^2)$ and converges to 0 as (n, ε) tends to $(\infty, 0)$, where $\Gamma := \int_{E_\varepsilon} \rho(e)Vev(de)$.

The second numerical scheme is inspired by the work of Hu *et al.* [16], who studied a backward stochastic differential equation driven by a Brownian motion with general terminal variable ξ . They proposed a new scheme using the representation of Z^ε as the trace of the Malliavin derivative of Y^ε . Their discretization scheme is based on the L^p -Hölder continuity of the solution Z^ε , and leads to an estimate of the form

$$\mathbb{E}|Z_t^\varepsilon - Z_s^\varepsilon|^p \leq K|t - s|^{p/2},$$

which implies the existence of a γ -Hölder continuous version of the process Z^ε for any $\gamma < \frac{1}{2} - 1/p$. In this sense, our paper extends the work done in [16] to an FBSDE with jumps and terminal value $g(X_T^\varepsilon)$. Similarly to [16], we obtain the following regularity of Γ^ε :

$$\mathbb{E}|\Gamma_t^\varepsilon - \Gamma_s^\varepsilon|^p \leq C|t - s|^{p/2}.$$

This allows us to deduce the existence of a γ -Hölder continuous version of the process Γ^ε for any $\gamma < \frac{1}{2} - 1/p$. Finally, on the one hand, we use the representations of Z^ε and Γ^ε as the trace of the Malliavin derivative of Y^ε to derive our new extended discretization scheme for the solution $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ of (1.3). On the other hand, we approximate X^ε by $X^{n,\varepsilon}$, the

continuous-time version of the Euler scheme, that is, for a fixed $\varepsilon > 0$,

$$\begin{aligned}
 X_t^{n,\varepsilon} &= X_{\phi_t^n}^{n,\varepsilon} + b(X_{\phi_t^n}^{n,\varepsilon})(t - \phi_t^n) + \sigma(\varepsilon)\beta(X_{\phi_t^n}^{n,\varepsilon})(W_t - W_{\phi_t^n}) \\
 &\quad + \int_{E_\varepsilon} \beta(X_{\phi_t^n}^{n,\varepsilon})\bar{M}(de, (t, \phi_t^n]), \\
 Y_{t_i}^{n,\varepsilon} &= \mathbb{E}[Y_{t_{i+1}}^{n,\varepsilon} + f(\Theta_{t_{i+1}}^{n,\varepsilon})\Delta t_{i+1} \mid \mathcal{F}_{t_i}], \\
 Z_{t_i}^{n,\varepsilon} &= \mathbb{E}\left[\mathfrak{E}_{t_{i+1},t_n}^{n,\varepsilon} \partial_x g(X_T^{n,\varepsilon})D_{t_i} X_T^{n,\varepsilon} \right. \\
 &\quad \left. + \sum_{k=i}^{n-1} \mathfrak{E}_{t_{i+1},t_{k+1}}^{n,\varepsilon} \partial_x f(\Theta_{t_{k+1}}^{n,\varepsilon})D_{t_i} X_{t_k}^{n,\varepsilon} \Delta t_k \mid \mathcal{F}_{t_i}\right], \\
 \Gamma_{t_i}^{n,\varepsilon} &= \mathbb{E}\left[\int_{E_\varepsilon} \rho(e) \left[\mathfrak{E}_{t_{i+1},t_n}^{e,n} D_{t_i,e} g(X_T^{n,\varepsilon}) \right. \right. \\
 &\quad \left. \left. + \sum_{k=i}^{n-1} \mathfrak{E}_{t_{i+1},t_{k+1}}^{e,n} \alpha_{t_i,t_{k+1}}^{n,\varepsilon} D_{t_i,e} X_{t_{k+1}}^{n,\varepsilon} \Delta t_k\right] \nu(de) \mid \mathcal{F}_{t_i}\right],
 \end{aligned}$$

with terminal values $Y_n^{n,\varepsilon} = g(X_T^{n,\varepsilon})$, $Z_n^{n,\varepsilon} = \sigma(\varepsilon)\partial_x g(X_T^{n,\varepsilon})\beta(X_T^{n,\varepsilon})$, and $U_n^{n,\varepsilon} = g(X_T^{n,\varepsilon} + \beta(X_T^{n,\varepsilon})) - g(X_T^{n,\varepsilon})$, where ϕ_t^n , $\mathfrak{E}_{t_i,t_j}^{n,\varepsilon}$, and $\mathfrak{E}_{t_i,t_j}^{e,n}$ are detailed in Section 4.

The key ingredient for the computation of the discretization error is the L^p -Hölder continuity of the solution $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$. This allows us to prove that

$$\text{Err}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)^2 := \mathbb{E} \max_{0 \leq i \leq n} [|Y_{t_i}^\varepsilon - Y_{t_i}^{n,\varepsilon}|^2 + |Z_{t_i}^\varepsilon - Z_{t_i}^{n,\varepsilon}|^2 + |\Gamma_{t_i}^\varepsilon - \Gamma_{t_i}^{n,\varepsilon}|^2]$$

is controlled by $n^{1/\log n-1}$. Then we obtain the second main result of this paper (Theorem 4.3), which states that

$$\begin{aligned}
 \text{Err}_{n,\varepsilon}(Y, V)^2 &:= \max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t - Y_{t_i}^{n,\varepsilon}|^2] + \mathbb{E}\left| \int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z_{t_i}^{n,\varepsilon} \Delta W_{t_i} \right|^2 \\
 &\quad + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}|\Gamma_t - \Gamma_{t_i}^{n,\varepsilon}|^2 dt
 \end{aligned}$$

is of the order $\sigma(\varepsilon)^2 + n^{1/\log n-1}$ and converges to 0 as the discretization step (ε, n) tends to $(0, \infty)$.

The importance of the above scheme is that it can be adapted to the case where the terminal value is not given by the forward diffusion equation X^ε , as is the case in [16]. However, this scheme remains to be further investigated.

The two numerical schemes above are not directly implemented in practice and require an important procedure to simulate the conditional expectation. However, there exist different techniques which can be adapted to our setting to compute this conditional expectation; see, for example, [3], [6], [9], and [18].

The paper is organized as follows. In Section 2 we prove the convergence of the approximated scheme. In Section 3 we describe the discrete-time scheme introduced in [5] and state our first main convergence result. In Section 4 we extend the new discrete scheme of [16] and state our second main result. We also discuss a general case of backward stochastic differential equations (BSDEs). Section 5 is devoted to Malliavin calculus for a class of FBSDEs with jumps. We obtain the L^p -Hölder continuity of Z^ε and Γ^ε via the trace of the Malliavin derivatives of Y^ε .

2. Approximation of decoupled FBSDEs driven by pure jump Lévy processes

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t \leq T}, \mathbb{P})$ 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 one-dimensional Brownian motion W and an independent Poisson measure μ on $[0, T] \times E$. We denote by $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \leq T}$ and $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{t \leq T}$ the \mathbb{P} -augmentation of the natural filtration of W and μ , respectively. As usual, we denote by $\mathcal{B}(X)$ the Borel set of the topological set X . We introduce the subsets $E^\varepsilon := \{e \in \mathbb{R} \text{ such that } |e| \leq \varepsilon\}$ and $E_\varepsilon := \{e \in \mathbb{R} \text{ such that } |e| > \varepsilon\}$, $E := \mathbb{R} = E^\varepsilon \cup E_\varepsilon$.

The martingale measure $\bar{\mu}$ is the compensated measure corresponding to the Poisson random measure μ such that $\bar{\mu}(de, dr) = \mu(de, dr) - \nu(de) dr$, where ν is a Lévy measure on E endowed with its Borel tribe \mathcal{E} . The Lévy measure ν will be assumed to satisfy $\nu(\mathbb{R}) = \infty$ and $\int_{\mathbb{R}} |e|^2 \nu(de) < \infty$. Throughout this paper we deal with the measure \bar{M} defined by

$$\bar{M}(t, B) = \int_{[0,t] \times B} e \bar{\mu}(dr, de), \quad B \in \mathcal{B}(E),$$

which can be considered as a compensated Poisson random measure on $[0, T] \times E$ and $\int_{[0,t] \times E} e \mu(dr, de)$ is a compound Poisson random variable. We associate to \bar{M} the σ -finite measure

$$m(B) := \int_B e^2 \nu(de), \quad B \in \mathcal{B}(E). \tag{2.1}$$

In particular, we have $\sigma(\varepsilon)^2 = m(E^\varepsilon)$.

The measure \bar{M} is taken to drive the jump noise instead of $\bar{\mu}$, in order to adopt the concept of Malliavin calculus on the canonical Lévy space from [11].

For some constant $K > 0$, we consider the four K -Lipschitz functions $\beta: \mathbb{R} \rightarrow \mathbb{R}$, $b: \mathbb{R} \rightarrow \mathbb{R}$, $g: \mathbb{R} \rightarrow \mathbb{R}$, and $f: \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(E, \mathcal{E}, \nu, \mathbb{R}) \rightarrow \mathbb{R}$ such that

$$\|b(x) - b(y)\| + \|\beta(x) - \beta(y)\| + \|g(x) - g(y)\| \leq K \|x - y\| \tag{2.2}$$

and

$$\|f(x, x', x'') - f(y, y', y'')\| \leq K (\|x - y\| + \|x' - y'\| + \|x'' - y''\|). \tag{2.3}$$

Define ρ to be a measurable function $\rho: E \rightarrow \mathbb{R}$ such that

$$\sup_{e \in E} |\rho(e)| < K.$$

For any $p \geq 2$, we consider the following class of processes.

- S^p is the set of real-valued adapted RCLL processes Y such that

$$\|Y\|_{S^p} := \mathbb{E} \left(\sup_{0 \leq t \leq T} |Y_t|^p \right)^{1/p} < \infty.$$

- H^p is the set of progressively measurable \mathbb{R} -valued processes Z such that

$$\|Z\|_{H^p} := \left(\mathbb{E} \left(\int_0^T |Z_r|^2 dr \right)^{p/2} \right)^{1/p} < \infty.$$

- L^p is the set of $(\mathcal{P} \otimes \mathcal{E})$ -measurable maps $U : \Omega \times [0, T] \times E \rightarrow \mathbb{R}$ such that

$$\|U\|_{L^p} := \left(\mathbb{E} \int_0^T \int_E |U_r(e)|^p \nu(\mathrm{d}e) \mathrm{d}r \right)^{1/p} < \infty.$$

- The space $\mathcal{B}^p := \mathcal{S}^p \times H^p \times L^p$ is endowed with the norm

$$\|(Y, Z, U)\|_{\mathcal{B}^p} := (\|Y\|_{S^p}^p + \|Z\|_{H^p}^p + \|U\|_{L^p}^p)^{1/p}.$$

- $M^{2,p}$ is the class of square-integrable random variables F of the form

$$F = \mathbb{E}F + \int_0^T U_r \mathrm{d}W_r + \int_0^T \int_E \psi(r, e) \bar{\mu}(\mathrm{d}e, \mathrm{d}r),$$

where u and ψ are respectively progressively measurable and measurable processes satisfying $\sup_{t \leq T} \mathbb{E}|u_t|^p < \infty$ and $\sup_{t \leq T} \mathbb{E} \int_E |\psi(t, e)|^p \nu(\mathrm{d}e) < \infty$, respectively.

To simplify the notation, we shall denote by C_p a generic constant depending only on p and the constants $K, \nu(E_\varepsilon), b(0), \beta(0), g(0), f(0, 0, 0)$, and T . In the rest of the paper, the real number p will be always greater than 2.

2.1. Approximation scheme

In this subsection we show that the approximation error

$$\begin{aligned} \text{Err}_\varepsilon(Y, V)^2 &:= \mathbb{E} \left[\sup_{t \leq T} |Y_t - Y_t^\varepsilon|^2 \right] + \mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t V_r \mathrm{d}R_r - \int_0^t Z_r^\varepsilon \mathrm{d}W_r \right|^2 \right] \\ &+ \mathbb{E} \left[\int_0^T \int_{E_\varepsilon} |V_r - U_r^\varepsilon(e)|^2 m(\mathrm{d}e) \mathrm{d}r \right] \end{aligned}$$

converges to 0 as ε goes to 0.

Theorem 2.1. *Under the space $(\Omega, \mathcal{F}, \mathbb{P})$,*

1. *there exists a solution X on $[0, T]$ of*

$$X_t = X_0 + \int_0^t b(X_r) \mathrm{d}r + \int_0^t \beta(X_{r-}) \bar{M}(\mathrm{d}r, \mathrm{d}e),$$

where $X_0 \in \mathbb{R}$,

2. *there exists a solution X^ε on $[0, T]$ of*

$$X_t^\varepsilon = X_0^\varepsilon + \int_0^t b(X_r^\varepsilon) \mathrm{d}r + \int_0^t \beta(X_r^\varepsilon) \sigma(\varepsilon) \mathrm{d}W_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-}^\varepsilon) \bar{M}(\mathrm{d}r, \mathrm{d}e),$$

where $X_0^\varepsilon \in \mathbb{R}$.

Moreover,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X_t^\varepsilon|^2 \right] \leq C \sigma(\varepsilon)^2. \tag{2.4}$$

To prove Theorem 2.1, we require the following lemma.

Lemma 2.1. *On the space $(\Omega, \mathcal{F}, \mathbb{P})$, fixing $\varepsilon > 0$, we have, for $p \geq 2$,*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t|^p \right] < \infty, \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t^\varepsilon|^p \right] < \infty.$$

Proof. Using Jensen’s inequality, the Burkholder–Davis–Gundy inequality, and Lipschitz’s property of b and β , we have

$$\begin{aligned} \mathbb{E} \sup_{0 \leq s \leq t} |X_s|^p &\leq C_p \mathbb{E} \sup_{0 \leq s \leq t} \left(|X_0|^p + \left(\int_0^s b(X_r) \, dr \right)^p + \left(\int_0^s \int_E \beta(X_r) \bar{M}(dr, de) \right)^p \right) \\ &\leq C_p \left(|X_0|^p + \int_0^t \mathbb{E}[|b(X_0)| + |X_r|]^p \, dr \right. \\ &\quad \left. + \int_0^t \int_E \mathbb{E}[|\beta(0)| + |X_r|]^p |e|^p \nu(de) \, dr \right) \\ &\leq C \left(|X_0|^p + |b(0)|^p + |\beta(0)|^p + \int_0^t \mathbb{E} \left[\sup_{0 \leq u \leq r} |X_u|^p \, dr \right] \right). \end{aligned}$$

We conclude the first assertion by Gronwall’s lemma. Following the same arguments, we obtain the second assertion.

Proof of Theorem 2.1. The existence and uniqueness of such stochastic differential equations has been studied in the literature; see, e.g. [15] and [1]. Thus, it remains to prove estimate (2.4).

Using Jensen’s inequality, we obtain

$$\begin{aligned} \mathbb{E} \sup_{0 \leq u \leq t} |X_u - X_u^\varepsilon|^2 &\leq C_2 \left(\mathbb{E} \left[\int_0^t |b(X_r) - b(X_r^\varepsilon)| \, dr \right]^2 \right. \\ &\quad \left. + \mathbb{E} \left[\int_0^t \int_{E_\varepsilon} |\beta(X_r) - \beta(X_r^\varepsilon)| \bar{M}(dr, de) \right]^2 \right. \\ &\quad \left. + \mathbb{E} \left[\int_0^t \int_{E_\varepsilon} |\beta(X_r)| \bar{M}(de, dr) \right]^2 \right. \\ &\quad \left. + \mathbb{E} \left[\sigma(\varepsilon) \int_0^t |\beta(X_r^\varepsilon)| \, dW_r \right]^2 \right). \end{aligned}$$

By the Burkholder–Davis–Gundy inequality we obtain

$$\begin{aligned} \mathbb{E} \sup_{0 \leq u \leq t} |X_u - X_u^\varepsilon|^2 &\leq C \left(t \mathbb{E} \left[\int_0^t (b(X_r) - b(X_r^\varepsilon))^2 \, dr \right] \right. \\ &\quad \left. + \mathbb{E} \left[\int_0^t \int_{E_\varepsilon} (\beta(X_r) - \beta(X_r^\varepsilon))^2 m(de) \, dr \right] \right. \\ &\quad \left. + \mathbb{E} \left[\int_0^t \int_{E_\varepsilon} \beta(X_r)^2 m(de) \, dr \right] + C \mathbb{E} \left[\int_0^t \beta(X_r^\varepsilon)^2 \sigma(\varepsilon)^2 \, dr \right] \right). \end{aligned}$$

By the Lipschitz property of b and β ,

$$\begin{aligned} \mathbb{E} \sup_{0 \leq u \leq t} |X_u - X_u^\varepsilon|^2 &\leq C \left(\left[\int_0^t \mathbb{E} |X_r - X_r^\varepsilon|^2 \, dr \right] \right. \\ &\quad \left. + \sigma(\varepsilon)^2 \mathbb{E} \int_0^t [\beta^2(0) + \beta^2(0) + |X_r^\varepsilon|^2 + |X_r|^2] \, dr \right) \\ &\leq C \left[\int_0^t \mathbb{E} \left[\sup_{u \leq r} |X_u - X_u^\varepsilon|^2 \right] \, dr + \sigma(\varepsilon)^2 \right]. \end{aligned}$$

The result follows from Gronwall’s lemma.

Finally, we can now state the main result of this section.

Theorem 2.2. *Under the space $(\Omega, \mathcal{F}, \mathbb{P})$,*

1. *there exists a unique pair $(Y, V) \in \mathcal{B}^2 \times H^2$ which solves the BSDE*

$$Y_t = g(X_T) + \int_t^T f(\Theta_r) \, dr - \int_t^T V_r \bar{M}(dr, de), \tag{2.5}$$

where $\Theta := (X, Y, \int_E \rho(e) V \, ev(de))$,

2. *for a fixed $\varepsilon > 0$, there exists a unique triplet $(Y^\varepsilon, Z^\varepsilon, U^\varepsilon) \in \mathcal{B}^2$ solution to the BSDE*

$$Y_t^\varepsilon = g(X_T^\varepsilon) + \int_t^T f(\Theta_r^\varepsilon) \, dr - \int_t^T Z_r^\varepsilon \, dW_s - \int_t^T \int_{E_\varepsilon} U_r^\varepsilon(e) \bar{M}(dr, de), \tag{2.6}$$

with $\Theta^\varepsilon := (X^\varepsilon, Y^\varepsilon, \Gamma^\varepsilon)$ and $\Gamma^\varepsilon := \int_{E_\varepsilon} \rho(e) U^\varepsilon(e) \, ev(de)$.

Moreover, if $\sup_{t \leq T} \mathbb{E}|V_t|^2 < \infty$ then there exists a constant C_p such that

$$\text{Err}_\varepsilon(Y, V)^2 \leq C_p \sigma(\varepsilon)^2. \tag{2.7}$$

Remark 2.1. Observe that

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \leq T} |Y_t - Y_t^\varepsilon|^2 \right] + \mathbb{E} \left[\sup_{t \leq T} \left| \int_t^T \int_{\mathbb{R}} V_r \bar{M}(de, dr) \right. \right. \\ & \qquad \qquad \qquad \left. \left. - \int_0^t Z_r^\varepsilon \, dW_r - \int_0^T \int_{E_\varepsilon} U_r^\varepsilon(e) \bar{M}(de, dr) \right|^2 \right] \\ & \leq \text{Err}_\varepsilon^2(Y, V) \\ & \leq \sigma(\varepsilon)^2. \end{aligned}$$

This clearly shows the convergence of the approximated scheme (1.3) to BSDE (1.1).

Proof of Theorem 2.2. Existence and uniqueness of the solutions of BSDEs (2.5) and (2.6) has already been proved; see, e.g. [4].

We are now going to prove inequality (2.7). Define $\delta Y := Y - Y^\varepsilon$ and Apply Itô's formula to $|\delta Y|^2$ to obtain

$$\begin{aligned} & |\delta Y_t|^2 + \int_t^T Z_r^{\varepsilon 2} \, dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) \, dr \\ & = |g(X_T) - g(X_T^\varepsilon)|^2 + \sigma(\varepsilon)^2 \int_t^T V_r^2 \, dr - \int_t^T \delta Y_r Z_r^\varepsilon \, dW_r \\ & \quad + 2 \int_t^T \delta Y_r (f(\Theta_r) - f(\Theta_r^\varepsilon)) \, dr + 2 \int_t^T \int_{E^\varepsilon} [(\delta Y_{s-} + V_r)^2 - \delta Y_{s-}^2] \bar{M}(de, dr) \\ & \quad - 2 \int_t^T \int_{E_\varepsilon} [(\delta Y_{s-} + U_r^\varepsilon(e) - V_r)^2 - \delta Y_{s-}^2] \bar{M}(de, dr). \end{aligned}$$

Taking the expectation on both sides of the above equality yields

$$\begin{aligned} & \mathbb{E} \left[\delta Y_t^2 + \int_t^T Z_r^{\varepsilon^2} dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \right] \\ &= \mathbb{E} \left[|g(X_T) - g(X_T^\varepsilon)|^2 + \sigma(\varepsilon)^2 \int_t^T V_r^2 dr + 2 \int_t^T \delta Y_r (f(\Theta_r) - f(\Theta_r^\varepsilon)) dr \right]. \end{aligned}$$

From Lemma 2.1, the Lipschitz property of g , and Jensen’s inequality we obtain

$$\begin{aligned} & \mathbb{E} \left[|\delta Y_t|^2 + \int_t^T Z_r^{\varepsilon^2} dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \right] \\ & \leq C \left(\sigma(\varepsilon)^2 + K \mathbb{E} \int_t^T (\delta Y_r) (|X_r - X_r^\varepsilon| + |\delta Y_r|) dr \right. \\ & \quad + K \mathbb{E} \int_t^T \left[(\delta Y_r) \int_{E_\varepsilon} \rho(e) e |U_r^\varepsilon(e) - V_r| \nu(de) \right] dr \\ & \quad \left. + K \mathbb{E} \int_t^T \left[(\delta Y_r) \int_{E^\varepsilon} \rho(e) V_r e \nu(de) \right] dr \right). \end{aligned}$$

Using the fact that $ab \leq \alpha a^2 + b^2/\alpha$ for some $\alpha > 0$, we obtain

$$\begin{aligned} & \mathbb{E} \left[|\delta Y_t|^2 + \int_t^T |Z_r^\varepsilon|^2 dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \right] \\ & \leq C \left(\sigma(\varepsilon)^2 + K(1 + \alpha^2 + \gamma^2 + \eta^2) \mathbb{E} \int_t^T |\delta Y_r|^2 dr + \frac{K}{\alpha^2} \mathbb{E} \int_t^T |X_r - X_r^\varepsilon|^2 dr \right. \\ & \quad + \frac{K}{\gamma^2} \mathbb{E} \int_t^T \int_{E_\varepsilon} \rho(e)^2 |U_r^\varepsilon(e) - V_r|^2 m(de) dr \\ & \quad \left. + \frac{K}{\eta^2} \mathbb{E} \int_t^T \int_{0 \leq |e| \leq \varepsilon} \rho(e)^2 V_r^2 m(de) dr \right), \end{aligned}$$

where α and γ are two constants taken such that $K/\alpha^2 = K^3/\gamma^2 = \frac{1}{2}$. We then obtain

$$\begin{aligned} & \mathbb{E} \left[|\delta Y_t|^2 + \int_t^T |Z_r^\varepsilon|^2 dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \right] \\ & \leq C \left(\sigma(\varepsilon)^2 + (K + 2K^2 + 2KK^2) \mathbb{E} \int_t^T |\delta Y_r|^2 dr \right). \end{aligned}$$

Using Gronwall’s lemma, we deduce that

$$\mathbb{E} |Y_t - Y_t^\varepsilon|^2 \leq C_p \sigma(\varepsilon)^2.$$

Substituting this estimate into the previous upper bound, we obtain

$$\mathbb{E} \int_0^T |Z_r^\varepsilon|^2 dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \leq C \sigma(\varepsilon)^2.$$

It follows that

$$\mathbb{E} |\delta Y_t|^2 + \mathbb{E} \int_0^T |Z_r^\varepsilon|^2 dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \leq C \sigma(\varepsilon)^2.$$

Now using the Burkholder–Davis–Gundy inequality, we have

$$\mathbb{E} \sup_{t \leq T} |\delta Y_t|^2 + \mathbb{E} \int_0^T |Z_r^\varepsilon|^2 dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 m(de) dr \leq C\sigma(\varepsilon)^2. \tag{2.8}$$

It follows by the Burkholder–Davis–Gundy inequality and (2.8) that

$$\begin{aligned} \mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t Z_r^\varepsilon dW_r - \int_0^t V_r dR_r \right|^2 \right] &\leq C_2 \mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t Z_r^\varepsilon dW_r \right|^2 + \sup_{t \leq T} \left| \int_0^t V_r dR_r \right|^2 \right] \\ &\leq C_2 \mathbb{E} \left[\int_0^T |Z_r^\varepsilon|^2 dr + \sigma(\varepsilon^2) \int_0^T V_r^2 dr \right] \\ &\leq \mathbb{E} \int_0^T |Z_r^\varepsilon|^2 dr + C_2 \sigma(\varepsilon^2). \end{aligned} \tag{2.9}$$

The result now follows by combining (2.8) and (2.9).

3. Forward–backward Euler scheme

In this section we first present the discrete-time solution $(X^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ of (1.3) by $(X^{n,\varepsilon}, Y^{n,\varepsilon}, Z^{n,\varepsilon}, \Gamma^{n,\varepsilon})$, defined by induction in (3.1) below, and then show the convergence of $(X^{n,\varepsilon}, Y^{n,\varepsilon}, Z^{n,\varepsilon}, \Gamma^{n,\varepsilon})$ to the solution of (1.2).

Let us recall some definitions and notation. For each $t \in [t_i, t_{i+1})$, we define

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

and

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

The processes \bar{Z}_t and $\bar{\Gamma}_t$ (respectively $\bar{Z}_t^{n,\varepsilon}$ and $\bar{\Gamma}_t^{n,\varepsilon}$) can be interpreted as the best approximations of Z_t and Γ_t (respectively $Z_t^{n,\varepsilon}$ and $\Gamma_t^{n,\varepsilon}$). We know from Bouchard and Elie [5] that FBSDE (2.5) has a backward Euler scheme that takes the form

$$\bar{X}_{t_{i+1}}^{n,\varepsilon} = \bar{X}_{t_i}^{n,\varepsilon} + \frac{1}{n} b(\bar{X}_{t_i}^{n,\varepsilon}) + \sigma(\varepsilon) \Delta W_{i+1} + \int_{E_\varepsilon} \beta(\bar{X}_{t_i}^{n,\varepsilon}) \bar{M}(de, (t_i, t_{i+1}]), \tag{3.1a}$$

$$\bar{Z}_t^{n,\varepsilon} = n \mathbb{E} [\bar{Y}_{t_{i+1}}^{n,\varepsilon} \Delta W_{i+1} \mid \mathcal{F}_{t_i}], \tag{3.1b}$$

$$\bar{\Gamma}_t^{n,\varepsilon} = n \mathbb{E} \left[\bar{Y}_{t_{i+1}}^{n,\varepsilon} \int_{E_\varepsilon} \rho(e) \bar{M}(de, (t_i, t_{i+1}]) \mid \mathcal{F}_{t_i} \right], \tag{3.1c}$$

$$\bar{Y}_t^{n,\varepsilon} = \mathbb{E} [\bar{Y}_{t_{i+1}}^{n,\varepsilon} \mid \mathcal{F}_t] + \frac{1}{n} f(\bar{X}_{t_i}^{n,\varepsilon}, \bar{Y}_{t_i}^{n,\varepsilon}, \bar{\Gamma}_t^{n,\varepsilon}). \tag{3.1d}$$

3.1. Forward Euler scheme

In order to explicitly compute the constant which appears before the rate of convergence of the Euler scheme, we introduce a continuous-time version of the Euler scheme. We first define the function ϕ for each $t \in [0, T]$ as

$$\phi_t^n := \max\{t_i, i = 0, \dots, n \mid t_i \leq t\},$$

for which we associate

$$\begin{aligned} \bar{X}_t^{n,\varepsilon} &:= \bar{X}_{\phi_t^n}^{n,\varepsilon} + b(\bar{X}_{\phi_t^n}^{n,\varepsilon})(t - \phi_t^n) + \sigma(\varepsilon)\beta(\bar{X}_{\phi_t^n}^{n,\varepsilon})(W_t - W_{\phi_t^n}) \\ &\quad + \int_{E_\varepsilon} \beta(\bar{X}_{\phi_t^n}^{n,\varepsilon})\bar{M}(de, (t, \phi_t^n)). \end{aligned}$$

It could be written as

$$\bar{X}_t^{n,\varepsilon} := \bar{X}_0 + \int_0^t b(\bar{X}_{\phi_r^n}^{n,\varepsilon}) dr + \sigma(\varepsilon) \int_0^t \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon}) dW_r + \int_0^t \int_{E_\varepsilon} \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon})\bar{M}(de, dr). \tag{3.2}$$

Lemma 3.1. *Under hypothesis (2.2), we have*

$$\sup_{n \geq 1} \mathbb{E} \left[\sup_{t \in [0, T]} |\bar{X}_t^{n,\varepsilon}|^p \right]^{1/p} \leq \infty.$$

Proof. The proof straightforwardly follows from Jensen’s inequality, the Burkholder–Davis–Gundy inequality, and Gronwall’s lemma.

Lemma 3.2. *Under hypothesis (2.2), we have*

$$\max_{0 \leq i \leq n-1} \mathbb{E} \left[\sup_{t \in [t_i, t_{i+1}]} |\bar{X}_t^{n,\varepsilon} - \bar{X}_{\phi_t}^{n,\varepsilon}|^p \right]^{1/p} \leq C_p(1 + \sigma(\varepsilon))n^{-1/2}.$$

Proof. By Jensen’s inequality and the convexity of the application $x \mapsto \|x\|^p$, we have

$$\begin{aligned} &\sup_{t \in [t_i, t_{i+1}]} |\bar{X}_t^{n,\varepsilon} - \bar{X}_{\phi_t}^{n,\varepsilon}|^p \\ &\leq C \left(n^{-p/2+1} \int_{t_i}^{t_{i+1}} \left[|b(\bar{X}_{\phi_r^n}^{n,\varepsilon})|^p + \sigma(\varepsilon)^p \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon})^p + \int_{E_\varepsilon} |\beta(\bar{X}_{\phi_r^n}^{n,\varepsilon})e|^p \nu(de) \right] dr \right). \end{aligned}$$

From hypotheses (2.2) and Lemma 3.1, we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [t_i, t_{i+1}]} |\bar{X}_t^{n,\varepsilon} - \bar{X}_{\phi_t}^{n,\varepsilon}|^p \right] &\leq C \left(n^{-p/2+1} \int_{t_i}^{t_{i+1}} C(1 + (1 + \sigma(\varepsilon))^p) \mathbb{E} |\bar{X}_{\phi_r^n}^{n,\varepsilon}|^p dr \right) \\ &\leq C(1 + \sigma(\varepsilon))^p n^{-p/2}. \end{aligned}$$

Theorem 3.1. *Under hypothesis (2.2), we have*

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t^\varepsilon - \bar{X}_t^{n,\varepsilon}|^2 \right] \leq C_p(1 + \sigma(\varepsilon))n^{-1/2}.$$

Proof. Using the same argument as Lemma 3.2, we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{s \leq t} |\bar{X}_s^{n,\varepsilon} - X_s|^p \right] &\leq C \mathbb{E} \left[\int_0^t |\bar{X}_{\phi_s}^{n,\varepsilon} - X_s|^p ds \right] \\ &\leq C_p \mathbb{E} \left[\int_0^t |\bar{X}_{\phi_s}^{n,\varepsilon} - \bar{X}_s^{n,\varepsilon}|^p + |\bar{X}_s^{n,\varepsilon} - X_s|^p ds \right] \\ &\leq C_p \left((1 + \sigma(\varepsilon))n^{-1/2} + \int_0^t \mathbb{E} \left[\sup_{r \leq s} |\bar{X}_r^{n,\varepsilon} - X_r|^p \right] ds \right). \end{aligned}$$

Using Gronwall’s lemma concludes the proof.

Remark 3.1. Since the estimation of Lemma 3.1 is uniform with respect to ε , we can write

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t^\varepsilon - \bar{X}_t^{n, \varepsilon}|^2 \right] \leq C_p n^{-1/2}.$$

3.2. Backward Euler scheme

In this subsection we prove the main result of the section. However, we first need to compute the approximation error

$$\overline{\text{Err}}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) := \left\{ \sup_{t \leq T} \mathbb{E}[|Y_t^\varepsilon - \bar{Y}_t^{n, \varepsilon}|^2] + \|Z^\varepsilon - \bar{Z}^{n, \varepsilon}\|_{H^2}^2 + \|\Gamma^\varepsilon - \bar{\Gamma}^{n, \varepsilon}\|_{H^2}^2 \right\}^{1/2}.$$

We should note that the rate of convergence of the approximated BSDE depends wholly on the rate of convergence of the approximated scheme of the forward stochastic differential equation (FSDE). Thus, by adapting the argument of [5] and using Remark 3.1, we easily prove that

$$\overline{\text{Err}}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \leq C_2 n^{-1/2} \tag{3.3}$$

converges to 0 as the discretization step T/n tends to 0. This means that the discretization scheme (3.1) achieves the optimal convergence rate $n^{-1/2}$. The regularity of Z^ε and Γ^ε has been studied in the L^2 sense in [5] when the terminal value is a functional of the forward diffusion.

Our aim in this subsection is to compute the approximation-discretization error between BSDEs (1.2) and (3.1), and to show that

$$\begin{aligned} \overline{\text{Err}}_{(n, \varepsilon)}^2(Y, V) &:= \sup_{t \leq T} \mathbb{E}[|Y_t - \bar{Y}_t^{n, \varepsilon}|^2] + \sup_{t \leq T} \mathbb{E} \left| \int_0^t V_r dR_r - \int_0^t \bar{Z}_r^{n, \varepsilon} dW_r \right|^2 \\ &\quad + \|\Gamma - \bar{\Gamma}^{n, \varepsilon}\|_{H^2}^2 \end{aligned} \tag{3.4}$$

converges to 0 as $(\varepsilon, n) \rightarrow (0, \infty)$.

The first main result of this paper is the following.

Proposition 3.1. *Under hypotheses (2.2) and (2.3), the approximation-discretization error defined in (3.4) is bounded by*

$$\overline{\text{Err}}_{(n, \varepsilon)}(Y, V) \leq C_2(n^{-1/2} + \sigma(\varepsilon)). \tag{3.5}$$

This means that

$$\overline{\text{Err}}_{(n, \varepsilon)}(Y, V) \rightarrow 0 \text{ as } (n, \varepsilon) \rightarrow (\infty, 0).$$

Proof. From (3.4), Jensen’s inequality, and the Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned} \overline{\text{Err}}_{(n, \varepsilon)}(Y, V)^2 &\leq C_2 \left(\sup_{t \leq T} \mathbb{E}[|Y_t - Y_t^\varepsilon|^2 + |Y_t^\varepsilon - \bar{Y}_t^{n, \varepsilon}|^2] + \|\Gamma - \Gamma^\varepsilon\|_{H^2}^2 + \|\Gamma^\varepsilon - \bar{\Gamma}^{n, \varepsilon}\|_{H^2}^2 \right. \\ &\quad \left. + \sup_{t \leq T} \mathbb{E} \left| \int_0^t V_r dR_r - \int_0^t Z_r^\varepsilon dW_r \right|^2 + \|Z^\varepsilon - \bar{Z}^{n, \varepsilon}\|_{H^2}^2 \right). \end{aligned}$$

It follows from Hölder’s inequality that

$$\int_{E_\varepsilon} \rho(e) e(V_r - U_r^\varepsilon(e)) \nu(de) \leq K \left(\int_{E_\varepsilon} (V_r - U_r^\varepsilon(e))^2 m(de) \right)^{1/2} \nu(E_\varepsilon)^{1/2}.$$

Recall that $\nu(E_\varepsilon) < \infty$ so that ν almost surely has only a finite number of big jumps on $[0, T]$. Combining the two last inequalities with (3.3) leads to

$$\begin{aligned} \overline{\text{Err}}_{(n,\varepsilon)}(Y, V)^2 \leq & C \left(n^{-1} + \sup_{t \leq T} \mathbb{E}[|Y_t - Y_t^\varepsilon|^2] + \sup_{t \leq T} \mathbb{E} \left| \int_0^t V_r dR_r - \int_0^t Z_r dW_r \right|^2 \right. \\ & \left. + \int_0^T \int_{E^\varepsilon} \mathbb{E}|V_r|^2 m(de) dr + \mathbb{E} \left[\int_0^T \int_{E^\varepsilon} |U_r^\varepsilon(e) - V_r|^2 m(de) dr \right] \right). \end{aligned}$$

By Theorem 2.2 we obtain

$$\overline{\text{Err}}_{(n,\varepsilon)}(Y, V) \leq C(n^{-1/2} + \sigma(\varepsilon)).$$

Remark 3.2. In the general case, as we neglect the small jumps, the Brownian part in (1.3) disappears. In this case assertion (3.5) can be replaced by

$$\overline{\text{Err}}_{(n,\varepsilon)}(Y, V) \leq Cn^{-1/2}.$$

4. A discrete scheme via Malliavin derivatives

In this section we generalize the new discrete scheme recently introduced by Hu *et al.* [16] from a general BSDE to our framework of decoupled FBSDEs with jumps. To this end, we use the Malliavin derivatives of Y to derive the discrete scheme. We first fix a regular grid $\pi := \{t_i := iT/n, i = 0, \dots, n\}$ on $[0, T]$ and approximate the FSDE X^ε in (1.3) by its Euler scheme $X^{n,\varepsilon}$ already defined in (3.1).

It is hard to prove the existence and convergence of Malliavin derivatives of $\bar{X}^{n,\varepsilon}$. To avoid this problem, we can instead consider the continuous-time version of the Euler scheme defined in (3.2).

The Malliavin derivatives of the continuous-time version of Euler scheme for $\theta \leq s$ almost everywhere are

$$\begin{aligned} D_\theta \bar{X}_t^{n,\varepsilon} &= \int_\theta^t \partial_x b(\bar{X}_r^{n,\varepsilon}) D_\theta \bar{X}_r^{n,\varepsilon} dr + \int_\theta^t \int_{E_\varepsilon} \partial_x \beta(\bar{X}_r^{n,\varepsilon}) D_\theta \bar{X}_r^{n,\varepsilon} \bar{M}(de, dr) \\ &\quad + \sigma(\varepsilon) \beta(\bar{X}_\theta^{n,\varepsilon}) + \sigma(\varepsilon) \int_\theta^t \partial_x \beta(\bar{X}_r^{n,\varepsilon}) D_\theta \bar{X}_r^{n,\varepsilon} dW_r, \\ D_{\theta,e} \bar{X}_t^{n,\varepsilon} &= \int_\theta^t D_{\theta,e} b(\bar{X}_r^{n,\varepsilon}) dr + \int_\theta^t \int_{E_\varepsilon} D_{\theta,e} \beta(\bar{X}_r^{n,\varepsilon}) \bar{M}(de, dr) \\ &\quad + \beta(\bar{X}_\theta^{n,\varepsilon}) + \sigma(\varepsilon) \int_\theta^t D_{\theta,e} \beta(\bar{X}_r^{n,\varepsilon}) dW_r. \end{aligned}$$

We introduce some additional assumptions.

(A1) $f(t, y, \gamma)$ does not depend on x .

(A2) The first derivatives of b, β , and g form a K -Lipschitz function:

$$|\partial b(x) - \partial b(y)| + |\partial \beta(x) - \partial \beta(y)| + |\partial g(x) - \partial g(y)| \leq K|x - y|.$$

(A3) $f(t, y, \gamma)$ is linear with respect to t, y , and γ . Moreover, there exist three bounded functions f_1, f_2 , and f_3 such that

$$f(t, y, u) = f_1(t) + f_2(t)y + f_3(t)\gamma.$$

Assumption 4.1. For $2 \leq p \leq q/2$, the following assertions hold.

1. The generator f has continuous and uniformly bounded first- and second-order partial derivatives with respect to x, y , and γ .
2. For each $(x, y, \gamma) \in \mathbb{R}^3$, $\partial_x f(\Theta)$, $\partial_y f(\Theta)$, and $\partial_\gamma f(\Theta)$ belong to $\mathbb{L}^{1,2}$ and satisfy

$$\sup_{0 \leq \theta \leq T} \mathbb{E} \left(\int_{\theta}^T |D_{\theta} \partial_i f(\Theta_r)|^2 dr \right)^{q/2} < \infty,$$

$$\sup_{0 \leq \theta \leq T} \sup_{0 \leq u \leq T} \mathbb{E} \left(\int_{\theta \vee u}^T |D_u D_{\theta} \partial_i f(\Theta_r)|^2 dr \right)^{q/2} < \infty,$$

where $i := x, y, \gamma$. There exists a constant $K > 0$ such that, for any $e \in (\mathbb{R} - \{0\})$, $t \in [0, T]$, and $0 \leq \theta, u \leq t \leq T$,

$$\mathbb{E} |D_{\theta} g(X_T) - D_u g(X_T)|^p \leq K |\theta - u|^{p/2}, \tag{4.1}$$

$$\mathbb{E} |D_{\theta, e} g(X_T) - D_{u, e} g(X_T)|^p \leq K |\theta - u|^{p/2}, \tag{4.2}$$

$$\mathbb{E} \left(\int_t^T |D_{\theta} f(\Theta_r) - D_u f(\Theta_r)|^2 dr \right)^{p/2} \leq K |\theta - u|^{p/2}, \tag{4.3}$$

$$\mathbb{E} \left(\int_t^T |D_{\theta, e} f(\Theta_r) - D_{u, e} f(\Theta_r)|^2 dr \right)^{p/2} \leq K |\theta - u|^{p/2}. \tag{4.4}$$

Lemma 4.1. Under the Lipschitz continuity of b and β , we have, for any $q \geq 1$,

$$\sup_{0 \leq \theta \leq T} \sup_{n \geq 1} \mathbb{E} \left[\sup_{-\theta \leq t \leq T} \|D_{\theta} \bar{X}_t^{n, \varepsilon}\|^{2q} \right] < \infty, \quad \sup_{0 \leq \theta \leq T} \sup_{n \geq 1} \mathbb{E} \left[\sup_{-\theta \leq t \leq T} \|D_{\theta, e} \bar{X}_t^{n, \varepsilon}\|^{2q} \right] < \infty.$$

Proof. See Appendix A.

We now derive the following theorem.

Theorem 4.1. Under assumption (A2), the Lipschitz continuity of b and β , and, for any $p \geq 2$, we have

$$\mathbb{E} \left[\sup_{t \in [0, T]} |D_{\theta} \bar{X}_t^{\varepsilon} - D_{\theta} \bar{X}_t^{n, \varepsilon}|^p \right]^{1/p} \leq C_p n^{-1/2}, \tag{4.5}$$

$$\mathbb{E} \left[\sup_{t \in [0, T]} |D_{\theta, e} \bar{X}_t^{\varepsilon} - D_{\theta, e} \bar{X}_t^{n, \varepsilon}|^p \right]^{1/p} \leq C_p n^{-1/2}.$$

Proof. Using the Burkholder–Davis–Gundy inequality, Jensen’s inequality, Remark 3.1, and Lemma 4.1, we obtain

$$\begin{aligned} & \mathbb{E} \left[\sup_{s \in [0, t]} |D_{\theta} X_t^{\varepsilon} - D_{\theta} \bar{X}_t^{n, \varepsilon}|^p \right] \\ & \leq C_p \mathbb{E} \left[\int_0^t \left| \partial_x b(\bar{X}_{\phi_r^n}^{n, \varepsilon}) D_{\theta} \bar{X}_{\phi_r^n}^{n, \varepsilon} - \partial_x b(X_r) D_{\theta} X_r^{\varepsilon} \right|^p dr \right. \\ & \quad + \int_0^t \int_{E_{\varepsilon}} \left| \partial_x \beta(\bar{X}_{\phi_r^n}^{n, \varepsilon}) D_{\theta} \bar{X}_{\phi_r^n}^{n, \varepsilon} - \partial_x \beta(X_r^{\varepsilon}) D_{\theta} X_r^{\varepsilon} \right|^p \nu(de) dr \\ & \quad + \sigma(\varepsilon)^p \left(\int_0^t \left| \partial_x \beta(\bar{X}_{\phi_r^n}^{n, \varepsilon}) D_{\theta} \bar{X}_{\phi_r^n}^{n, \varepsilon} - \partial_x \beta(X_r^{\varepsilon}) D_{\theta} X_r^{\varepsilon} \right|^2 dr \right)^{p/2} \\ & \quad \left. + \sigma(\varepsilon)^p |\beta(\bar{X}_{\theta}^{n, \varepsilon}) - \beta(X_{\theta}^{\varepsilon})|^p \right] \end{aligned}$$

$$\begin{aligned} &\leq C_p \mathbb{E} \left[\int_0^t |D_\theta X_r^\varepsilon|^p [|\partial_x b(X_r^\varepsilon) - \partial_x b(\bar{X}_{\phi_r^n}^{n,\varepsilon})|^p + |\partial_x \beta(X_r^\varepsilon) - \partial_x \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon})|^p] dr \right. \\ &\quad + \int_0^t |D_\theta X_r^\varepsilon - D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon}|^p [|\partial_x b(\bar{X}_{\phi_r^n}^{n,\varepsilon})|^p + |\partial_x \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon})|^p] dr \\ &\quad \left. + \sigma(\varepsilon)^p |\bar{X}_\theta^{n,\varepsilon} - X_\theta^\varepsilon|^p \right] \\ &\leq C_p \mathbb{E} \left[n^{-p/2} + \int_0^t \sup_{u \in [0,r]} |D_\theta X_u^\varepsilon - D_\theta \bar{X}_u^{n,\varepsilon}|^p dr \right]. \end{aligned}$$

We conclude by using Gronwall’s lemma. Following the same arguments, we prove the second assertion.

Now we derive the discrete scheme using the expressions of Z^ε and U^ε as the trace of the Malliavin derivatives of Y . From (5.2) and (5.4), the two Malliavin derivatives $D_\theta Y_t$ and $D_{\theta,e} Y_t$ could be expressed as

$$D_\theta Y_t^\varepsilon = \mathbb{E} \left[\mathfrak{E}_{t,T} \partial_x g(X_T^\varepsilon) D_\theta X_T^\varepsilon + \int_t^T \mathfrak{E}_{t,r} \partial_x f(\Theta_r^\varepsilon) D_\theta X_r^\varepsilon dr \mid \mathcal{F}_t \right], \tag{4.6}$$

$$D_{\theta,e} Y_t^\varepsilon = \mathbb{E} \left[\mathfrak{E}_{t,T}^e D_{\theta,e} g(X_T^\varepsilon) + \int_t^T \mathfrak{E}_{t,r}^e \alpha_{\theta,r} D_{\theta,e} X_r^\varepsilon dr \mid \mathcal{F}_t \right], \tag{4.7}$$

where

$$\begin{aligned} \mathfrak{E}_{t,r} &= \exp \left\{ \int_t^r \left(\partial_y f(\Theta_u^\varepsilon) - \frac{1}{2} \int_{E_\varepsilon} \partial_\gamma f^2(\Theta_u^\varepsilon) \rho^2(e) m(de) \right) du \right. \\ &\quad \left. + \int_t^r \int_{E_\varepsilon} \partial_\gamma f(\Theta_u^\varepsilon) \rho(e) \bar{M}(de, du) \right\}, \\ \mathfrak{E}_{t,r}^e &:= \exp \left\{ \int_t^r \left(\alpha_{\theta,u} - \frac{1}{2} \alpha_{\theta,u}^2 \int_{E_\varepsilon} \rho^2(e) m(de) \right) du + \int_t^r \int_{E_\varepsilon} \alpha_{\theta,u} \rho(e) \bar{M}(de, du) \right\}, \end{aligned}$$

and

$$\alpha_{\theta,r} := \frac{f(\Theta_r^\varepsilon + D_{\theta,e} \Theta_r^\varepsilon) - f(\Theta_r^\varepsilon)}{D_{\theta,e} X_r^\varepsilon + D_{\theta,e} Y_r + D_{\theta,e} \Gamma_r} \mathbf{1}_{\{D_{\theta,e} X_r^\varepsilon + D_{\theta,e} Y_r + D_{\theta,e} \Gamma_r \neq 0\}}.$$

Thus, we define our discrete scheme for $i = n - 1, \dots, 1, 0$ and $t \in [t_i, t_{i+1})$ by induction, i.e.

$$Y_{t_i}^{n,\varepsilon} = \mathbb{E}[Y_{t_{i+1}}^{n,\varepsilon} + f(\Theta_{t_{i+1}}^{n,\varepsilon}) \Delta t_{i+1} \mid \mathcal{F}_{t_i}], \tag{4.8a}$$

$$\begin{aligned} Z_{t_i}^{n,\varepsilon} &= \mathbb{E} \left[\mathfrak{E}_{t_{i+1},t_n}^{n,\varepsilon} \partial_x g(\bar{X}_T^{n,\varepsilon}) D_{t_i} \bar{X}_T^{n,\varepsilon} \right. \\ &\quad \left. + \sum_{k=i}^{n-1} \mathfrak{E}_{t_{i+1},t_{k+1}}^{n,\varepsilon} \partial_x f(\Theta_{t_{k+1}}^{n,\varepsilon}) D_{t_i} X_{t_{k+1}}^{n,\varepsilon} \Delta t_k \mid \mathcal{F}_{t_i} \right], \end{aligned} \tag{4.8b}$$

$$\begin{aligned} \Gamma_{t_i}^{n,\varepsilon} &= \mathbb{E} \left[\int_{E_\varepsilon} \rho(e) \left[\mathfrak{E}_{t_{i+1},t_n}^{e,n} D_{t_i,e} g(\bar{X}_T^{n,\varepsilon}) \right. \right. \\ &\quad \left. \left. + \sum_{k=i}^{n-1} \mathfrak{E}_{t_{i+1},t_{k+1}}^{e,n} \alpha_{t_i,t_{k+1}}^{n,\varepsilon} D_{t_i,e} \bar{X}_{t_{k+1}}^{n,\varepsilon} \Delta t_k \right] v(de) \mid \mathcal{F}_{t_i} \right], \end{aligned} \tag{4.8c}$$

with terminal conditions

$$\begin{aligned}
 Y_{t_n}^{n,\varepsilon} &= g(\bar{X}_T^{n,\varepsilon}), & Z_{t_n}^{n,\varepsilon} &= \sigma(\varepsilon)\partial_x g(\bar{X}_T^{n,\varepsilon})\beta(\bar{X}_T^{n,\varepsilon}), \\
 \text{and } U_{t_n,e}^{n,\varepsilon} &= g(\bar{X}_T^{n,\varepsilon} + \beta(\bar{X}_T^{n,\varepsilon})) - g(\bar{X}_T^{n,\varepsilon}),
 \end{aligned}$$

where, for any $0 \leq i < j \leq n$,

$$\begin{aligned}
 \mathfrak{E}_{t_i,t_j}^{n,\varepsilon} &= \exp\left\{\sum_{k=i}^{j-1} \int_{t_k}^{t_{k+1}} \left(\partial_y f(\Theta_{t_k}^{n,\varepsilon}) - \frac{1}{2} \int_{E_\varepsilon} \partial_\gamma f^2(\Theta_{t_k}^{n,\varepsilon}) \rho^2(e) m(\mathrm{d}e)\right) \mathrm{d}r \right. \\
 &\quad \left. + \sum_{k=i}^{j-1} \int_{t_k}^{t_{k+1}} \int_{E_\varepsilon} \partial_\gamma f(\Theta_{t_k}^{n,\varepsilon}) \rho(e) \bar{M}(\mathrm{d}e, \mathrm{d}r)\right\}, \\
 \mathfrak{E}_{t_i,t_j}^{e,n} &= \exp\left\{\sum_{k=i}^{j-1} \int_{t_k}^{t_{k+1}} \left(\alpha_{\theta,r,t_k}^{n,\varepsilon} - \frac{1}{2} \alpha_{\theta,r,t_k}^{n,\varepsilon 2} \int_{E_\varepsilon} \rho^2(e) m(\mathrm{d}e)\right) \mathrm{d}r \right. \\
 &\quad \left. + \sum_{k=i}^{j-1} \int_{t_k}^{t_{k+1}} \int_{E_\varepsilon} \alpha_{\theta,r,t_k}^{n,\varepsilon} \rho(e) \bar{M}(\mathrm{d}e, \mathrm{d}r)\right\},
 \end{aligned}$$

and

$$\begin{aligned}
 \alpha_{\theta,r,t_k}^{n,\varepsilon} &:= \frac{f(\Theta_{t_k}^{n,\varepsilon} + D_{\theta,e}\Theta_{t_k}^{n,\varepsilon}) - f(\Theta_{t_k}^{n,\varepsilon})}{D_{\theta,e}X_{t_k}^{n,\varepsilon} D_{\theta,e}Y_{t_k}^{n,\varepsilon} + D_{\theta,e}\Gamma_{t_k}^{n,\varepsilon}} \mathbf{1}_{\{D_{\theta,e}\bar{X}_{t_k}^{n,\varepsilon} + D_{\theta,e}Y_{t_k}^{n,\varepsilon} + D_{\theta,e}\Gamma_{t_k}^{n,\varepsilon} \neq 0\}}, \\
 \Gamma_{t_k}^{n,\varepsilon} &:= \int_{E_\varepsilon} U_{t_k,e}^{n,\varepsilon} \rho(e) \nu(\mathrm{d}e),
 \end{aligned}$$

with $\Theta_{t_k}^{n,\varepsilon} = (r, \bar{X}_{t_k}^{n,\varepsilon}, Y_{t_k}^{n,\varepsilon}, \Gamma_{t_k}^{n,\varepsilon})$. We are going to compute the discretization error of our discrete scheme and prove the convergence. We recall the expression of the error between the solutions of (1.3) and (4.8):

$$\text{Err}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)^p := \mathbb{E} \max_{0 \leq i \leq n} [|Y_{t_i}^\varepsilon - Y_{t_i}^{n,\varepsilon}|^p + |Z_{t_i}^\varepsilon - Z_{t_i}^{n,\varepsilon}|^p + |\Gamma_{t_i}^\varepsilon - \Gamma_{t_i}^{n,\varepsilon}|^p].$$

Here $\Gamma_{t_i}^{n,\varepsilon} = \int_{E_\varepsilon} \rho(e) U_{t_i,e}^{n,\varepsilon} \nu(\mathrm{d}e)$. We also recall the expression for the discretization-approximation error between (1.2) and (4.8):

$$\begin{aligned}
 \text{Err}_{n,\varepsilon}^2(Y, V) &:= \max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t - Y_{t_i}^{n,\varepsilon}|^2] + \mathbb{E} \left| \int_0^T V_r \mathrm{d}R_r - \sum_{i=0}^{n-1} Z_{t_i}^{n,\varepsilon} \Delta W_{t_i} \right|^2 \\
 &\quad + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E} |\Gamma_t - \Gamma_{t_i}^{n,\varepsilon}|^2 \mathrm{d}t.
 \end{aligned}$$

We conclude this section with the following theorems.

Theorem 4.2. *Under Assumption 4.1, we assume the existence of a constant $L_3 > 0$ such that*

$$|f(t_2, y, u) - f(t_1, y, u)| \leq L_3 |t_2 - t_1|^{1/2}.$$

Then there exists a positive constant C independent of n such that

$$\text{Err}_n^p(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \leq C p n^{p/2 \log n - p/2}. \tag{4.9}$$

Proof. We adapt the proof of Theorem 5.2 of [16]. Let $i = n - 1, \dots, 1, 0$.

Step 1: show that $\mathbb{E}[\sup_{0 \leq i \leq n} |\delta Z_i^{n,\varepsilon}|^p] \leq C_p n^{1-p}$. Define

$$\delta Z_i^{n,\varepsilon} = Z_{t_i} - Z_{t_i}^{n,\varepsilon}.$$

Combining (4.6) and (4.8),

$$\begin{aligned} |\delta Z_i^{n,\varepsilon}| &\leq |\mathbb{E}[\mathcal{E}_{t_i,T} \partial_x g(X_T^\varepsilon) D_{t_i} X_T^\varepsilon - \mathcal{E}_{t_{i+1},t_n}^{n,\varepsilon} \partial_x g(\bar{X}_T^{n,\varepsilon}) D_{t_i} \bar{X}_T^{n,\varepsilon} \mid \mathcal{F}_{t_i}]| \\ &= |\mathbb{E}[\mathcal{E}_{t_i,T} ([\partial_x g(X_T^\varepsilon) - \partial_x g(\bar{X}_T^{n,\varepsilon})] D_{t_i} \bar{X}_T^{n,\varepsilon} + \partial_x g(X_T^\varepsilon) [D_{t_i} X_T^\varepsilon - D_{t_i} \bar{X}_T^{n,\varepsilon}]) \\ &\quad + \partial_x g(\bar{X}_T^{n,\varepsilon}) D_{t_i} \bar{X}_T^{n,\varepsilon} [\mathcal{E}_{t_i,T} - \mathcal{E}_{t_i,T}^{n,\varepsilon}] \mid \mathcal{F}_{t_i}]|. \end{aligned}$$

From Lemma 4.1, inequality (4.5), and assumption (A2),

$$\begin{aligned} \mathbb{E} \sup_{0 \leq i \leq n} |\delta Z_i^{n,\varepsilon}|^p &\leq \left(\mathbb{E} \left(\sup_{0 \leq i \leq n} D_{t_i} \bar{X}_T^{n,\varepsilon} \right)^{p/(p-1)} \right)^{p-1} [\mathbb{E}(\partial_x g(X_T^\varepsilon) - \partial_x g(\bar{X}_T^{n,\varepsilon}))^p] \\ &\quad + (\mathbb{E}(\partial_x g(X_T^\varepsilon))^{p/(p-1)})^{p-1} \left(\mathbb{E} \left[\sup_{0 \leq i \leq n} |D_{t_i} X_T^\varepsilon - D_{t_i} \bar{X}_T^{n,\varepsilon}|^p \right] \right) \\ &\quad + \mathbb{E} \left[\sup_{0 \leq i \leq n} \mathbb{E}[\partial_x g(\bar{X}_T^{n,\varepsilon}) D_{t_i} \bar{X}_T^{n,\varepsilon} (\mathcal{E}_{t_i,T} - \mathcal{E}_{t_i,T}^{n,\varepsilon}) \mid \mathcal{F}_{t_i}]^p \right] \\ &\leq C_p \left(n^{-p/2} + \mathbb{E} \sup_{0 \leq i \leq n} |I_i|^p \right). \end{aligned}$$

Using the fact that $|e^x - e^y| \leq (e^x + e^y)|x - y|$, combined with assumption (A3), we obtain

$$\begin{aligned} I_i &\leq C \mathbb{E} \left[(D_{t_i} \bar{X}_T^{n,\varepsilon} \mathcal{E}_{t_i,T} + \mathcal{E}_{t_{i+1},t_n}^{n,\varepsilon}) \right. \\ &\quad \times \left| \int_{t_i}^T \left(f_2(r) - \frac{1}{2} \int_{E_\varepsilon} [f_3(r)]^2 \rho^2(e) m(de) \right) dr + \int_{t_i}^T \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right. \\ &\quad - \sum_{k=i+1}^{n-1} \int_{t_k}^{t_{k+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \\ &\quad \left. \left. - \sum_{k=i+1}^{n-1} \int_{t_k}^{t_{k+1}} \left(f_2(r) + \frac{1}{2} \int_{E_\varepsilon} [f_3(r)]^2 \rho^2(e) m(de) \right) dr \right| \mathcal{F}_{t_i} \right]. \end{aligned}$$

It follows that

$$\begin{aligned} I_i &\leq C \mathbb{E} \left[D_{t_i} \bar{X}_T^{n,\varepsilon} (\mathcal{E}_{t_i,T} + \mathcal{E}_{t_{i+1},t_n}^{n,\varepsilon}) \left| \int_{t_i}^{t_{i+1}} \left[f_2(r) - \frac{1}{2} \int_{E_\varepsilon} [f_3(r)]^2 \rho^2(e) m(de) \right] dr \right. \right. \\ &\quad \left. \left. + \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right| \mathcal{F}_{t_i} \right]. \end{aligned}$$

Again, using assumption (A.3), we have

$$\begin{aligned} |D_{t_i} \bar{X}_T^{n,\varepsilon} \mathcal{E}_{t_i,T}| &\leq |D_{t_i} \bar{X}_T^{n,\varepsilon}|^r \exp \left\{ \int_{t_i}^T f_2(u) du - \frac{1}{2} \int_{t_i}^T \int_{E_\varepsilon} f_3(u) \rho^2(e) m(de) du \right. \\ &\quad \left. + \int_{t_i}^T \int_{E_\varepsilon} f_3(u) \rho(e) \bar{M}(de, du) \right\} \\ &\leq C \left(\sup_{0 \leq \theta \leq T} |D_\theta \bar{X}_T^{n,\varepsilon}| \right) \left(\sup_{0 \leq t \leq T} \exp \left\{ \int_t^T \int_{E_\varepsilon} f_3(u) \rho(e) \bar{M}(de, du) \right\} \right). \end{aligned}$$

Similarly, we have

$$|D_{t_i} \bar{X}_T^{n,\varepsilon} | \mathcal{E}_{t_i, t_n}^{n,\varepsilon} \leq C \left(\sup_{0 \leq \theta \leq T} |D_\theta \bar{X}_T^{n,\varepsilon}| \right) \left(\sup_{0 \leq t \leq T} \exp \left\{ \int_t^T \int_{E_\varepsilon} f_3(u) \rho(e) \bar{M}(de, du) \right\} \right).$$

For any $r \geq 0$, we have, by Hölder’s inequality and Proposition 5.1,

$$\begin{aligned} & \mathbb{E} \left(\sup_{0 \leq t \leq T} \exp \left\{ \int_t^T \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right\} \right)^r \\ & \leq \mathbb{E} \left(\exp \left\{ 2r \int_0^T \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right\} \right)^{1/2} \\ & \quad \times \mathbb{E} \left(\sup_{0 \leq t \leq T} \exp \left\{ -2r \int_0^t \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right\} \right)^{1/2} \\ & \leq \mathbb{E} \left(\exp \left\{ r^2 \int_0^T \int_{E_\varepsilon} f_3^3(r) \rho^2(e) m(de) dr \right\} \right) \\ & \quad \times \mathbb{E} \left(\sup_{0 \leq t \leq T} \exp \left\{ -2r \int_0^t \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right\} \right)^{1/2} \\ & < \infty. \end{aligned}$$

Thus, for $p' \in (p, q/2)$,

$$\begin{aligned} \mathbb{E} \sup_{0 \leq i \leq n-1} I_i^p & \leq C \mathbb{E} \left(\sup_{0 \leq \theta \leq T} |D_{t_i} \bar{X}_T^{n,\varepsilon}| \right)^p \left(\sup_{0 \leq t \leq T} \exp \left\{ \int_t^T \int_{E_\varepsilon} f_3(u) \rho(e) \bar{M}(de, du) \right\} \right)^p \\ & \quad \times \left[\sup_{0 \leq i \leq n-1} \int_{t_i}^{t_{i+1}} |f_2(r)| dr + \frac{1}{2} \sup_{0 \leq i \leq n-1} \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} [f_3(r)]^2 \rho^2(e) m(de) dr \right. \\ & \quad \left. + \sup_{0 \leq i \leq n-1} \left| \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right| \right]^p \\ & \leq C \left[\mathbb{E} \left(\sup_{0 \leq \theta \leq T} |D_{t_i} \bar{X}_T^{n,\varepsilon}| \right)^{2pp'/(p'-p)} \right]^{p'/2(p'-p)} \\ & \quad \times \left[\mathbb{E} \left(\sup_{0 \leq t \leq T} \exp \left\{ \int_t^T \int_{E_\varepsilon} f_3(t) \rho(e) \bar{M}(de, du) \right\} \right)^{2pp'/(p'-p)} \right]^{p'/2(p'-p)} \\ & \quad \times \left[\mathbb{E} \sup_{0 \leq i \leq n-1} \left(\int_{t_i}^{t_{i+1}} |f_2(r)| dr \right)^{p'} \right. \\ & \quad \left. + \mathbb{E} \sup_{0 \leq i \leq n-1} \left(\int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} [f_3(r)]^2 \rho^2(e) m(de) dr \right)^{p'} \right. \\ & \quad \left. + \mathbb{E} \sup_{0 \leq i \leq n-1} \left| \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right|^{p'} \right]^{p/p'} \\ & \leq C [I_1 + I_2 + I_3]^{p/p'}. \end{aligned}$$

We first estimate I_3 .

By Holder’s inequality for $r > 1$, Jensen’s inequality, and the Burkholder–Davis–Gundy inequality,

$$\begin{aligned}
 I_3^{p/p'} &= \left[\mathbb{E} \sup_{0 \leq i \leq n-1} \left| \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right|^{rp'} \right]^{p/rp'} \\
 &\leq \mathbb{E} \left(\sum_{0 \leq i \leq n-1} \left| \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} f_3(r) \rho(e) \bar{M}(de, dr) \right|^{rp'} \right)^{p/rp'} \\
 &\leq C_p \left(\sum_{0 \leq i \leq n-1} \mathbb{E} \left| \int_{t_i}^{t_{i+1}} \int_{E_\varepsilon} |f_3(r)|^2 \rho^2(e) m(de) dr \right|^{rp'/2} \right)^{p/rp'} \\
 &\leq C_p n^{p/rp' - p/2}.
 \end{aligned}$$

For small enough π , we take $r = 2 \log n / p'$. Then

$$I_3^{p/p'} \leq Cn^{p/2 \log n - p/2}$$

and

$$\mathbb{E}[I_1 + I_2]^{p/p'} \leq C_p n^{-p}.$$

Consequently,

$$\mathbb{E} \sup_{0 \leq i \leq n} |\delta Z_{t_i}^{n,\varepsilon}|^p \leq C_p n^{p/2 \log n - p/2}. \tag{4.10}$$

Step 2: show that $\mathbb{E} \sup_{0 \leq i \leq n} |\delta \Gamma_{t_i}^{n,\varepsilon}|^p \leq C_p n^{1-p}$, where $\delta \Gamma_{t_i}^{n,\varepsilon} = \Gamma_{t_i} - \Gamma_{t_i}^{n,\varepsilon}$. In fact,

$$\begin{aligned}
 \mathbb{E} \sup_{0 \leq i \leq n} |\delta \Gamma_{t_i}^{n,\varepsilon}|^p &= \mathbb{E} \sup_{0 \leq i \leq n} \left| \int_{E_\varepsilon} \rho(e) \delta U_{t_i, e}^{n,\varepsilon} \nu(de) \right|^p \\
 &\leq C_p \mathbb{E} \sup_{0 \leq i \leq n} \int_{E_\varepsilon} \rho^p(e) |\delta U_{t_i, e}^{n,\varepsilon}|^p \nu(de) \\
 &\leq C_p \mathbb{E} \sup_{0 \leq i \leq n} |\delta U_{t_i, e}^{n,\varepsilon}|^p.
 \end{aligned}$$

However, following exactly the same arguments as in step 1, we can prove that

$$\mathbb{E} \sup_{0 \leq i \leq n} |\delta U_{t_i, e}^{n,\varepsilon}|^p \leq Cn^{p/2 \log n - p/2}$$

and that

$$\mathbb{E} \sup_{0 \leq i \leq n} |\delta \Gamma_{t_i}|^p \leq Cn^{p/2 \log n - p/2}. \tag{4.11}$$

Step 3: show that $\mathbb{E} \sup_{0 \leq i \leq n} |\delta Y_{t_i}^{n,\varepsilon}|^p \leq Cn^{1-p}$. We have

$$\begin{aligned}
 Y_{t_i}^{n,\varepsilon} &= \mathbb{E} \left[g(\bar{X}_T^{n,\varepsilon}) + \sum_{k=i+1}^{n-1} f(\Theta_{t_{k+1}}^{n,\varepsilon}) \Delta t_k \mid \mathcal{F}_{t_i} \right], \\
 Y_{t_i}^\varepsilon &= \mathbb{E} \left[g(X_T^\varepsilon) + \sum_{k=i+1}^{n-1} f(\Theta_{t_{k+1}}) \Delta t_k \mid \mathcal{F}_{t_i} \right].
 \end{aligned}$$

Hence, by again adapting the argument of [16] to our setting, we have, for $i = n - 1, n - 2, \dots, 0$,

$$|\delta Y_i^{n,\varepsilon}| \leq \mathbb{E} \left[\sum_{k=i+1}^{n-1} |f(\Theta_{t_{k+1}}^{n,\varepsilon}) - f(\Theta_{t_{k+1}}^\varepsilon)| \Delta t_k + |R_i^{n,\varepsilon}| + |\delta g^{n,\varepsilon}(X_T^\varepsilon)| \mid \mathcal{F}_i \right],$$

where $\delta g^{n,\varepsilon}(X_t^\varepsilon) = g(X_T^\varepsilon) - g(\bar{X}_T^{n,\varepsilon})$ and

$$|R_i^{n,\varepsilon}| = \left| \int_t^T f(\Theta_r^\varepsilon) dr - \sum_{k=i+1}^{n-1} f(\Theta_{t_{k+1}}^{n,\varepsilon}) \Delta t_k \right|.$$

For $j = n - 1, n - 2, \dots, i$,

$$|\delta Y_{t_j}^{n,\varepsilon}| \leq \mathbb{E} \left[\sum_{k=i+1}^{n-1} |f(\Theta_{t_{k+1}}^{n,\varepsilon}) - f(\Theta_{t_{k+1}})| \Delta t_k + \sup_{0 \leq t \leq T} |R_t^{n,\varepsilon}| + |\delta g^{n,\varepsilon}(X_T^\varepsilon)| \mid \mathcal{F}_{t_j} \right].$$

Since we know from [16] that $\mathbb{E} \sup_{0 \leq t \leq T} |R_t^{n,\varepsilon}|^p \leq C_p n^{-p/2}$, combining this with the standard estimate of $\delta \bar{X}^{n,\varepsilon}$, and the Lipschitz property of the generator f , we have

$$\begin{aligned} & \mathbb{E} \sup_{0 \leq j \leq n} |\delta Y_{t_j}^{n,\varepsilon}|^p \\ & \leq C \mathbb{E} \left[\left(\sum_{k=i+1}^{n-1} |f(\Theta_{t_{k+1}}^{n,\varepsilon}) - f(\Theta_{t_{k+1}})| \Delta t_k \right)^p + \sup_{0 \leq t \leq T} |R_t^{n,\varepsilon}|^p + |\delta g^{n,\varepsilon}(X_T)|^p \right] \\ & \leq C \mathbb{E} \left[\left(\sum_{k=i+1}^n |\delta \bar{X}_{t_k}^{n,\varepsilon}| \Delta t_k \right)^p + \left(\sum_{k=i+1}^n |\delta Y_{t_k}^{n,\varepsilon}| \Delta t_k \right)^p + \left(\sum_{k=i+1}^n |\delta \Gamma_{t_k}^{n,\varepsilon}| \Delta t_k \right)^p \right. \\ & \quad \left. + \sup_{0 \leq t \leq T} |R_t^{n,\varepsilon}|^p + |\delta g^{n,\varepsilon}(X_T)|^p \right] \\ & \leq C \left\{ (T - t_i)^p \mathbb{E} \sup_{i+1 \leq k \leq T} |\delta Y_{t_k}^{n,\varepsilon}|^p + (n^{p/2 \log n - p/2} + n^{-p/2}) \right\}. \end{aligned}$$

Using similar recursive methods as in Theorem 4.2 of [16], we obtain the estimate

$$\mathbb{E} \sup_{0 \leq j \leq n} |\delta Y_{t_j}^{n,\varepsilon}|^p \leq C_p n^{p/2 \log n - p/2}. \tag{4.12}$$

Finally, combining (4.10)–(4.11) and (4.12) completes the proof.

The second main result of this paper is summarized in the following theorem.

Theorem 4.3. *Under the same assumptions as in Theorem 4.2, we have*

$$\text{Err}_{n,\varepsilon}(Y, V)^2 \leq C(\sigma(\varepsilon)^2 + n^{1/\log n - 1}).$$

Proof. For any $i = n - 1, n - 2, \dots, 0$, observe that

$$\begin{aligned} & \max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t^\varepsilon - Y_{t_i}^{n,\varepsilon}|^2] \\ & \leq C_2 \max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t^\varepsilon - Y_t|^2 + |Y_t - Y_{t_i}|^2 + |Y_{t_i} - Y_{t_i}^{n,\varepsilon}|^2]. \end{aligned}$$

Combining (2.7) and (4.9) with Remark 2.1 applied to our setting, we have

$$\max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t^\varepsilon - Y_t^{n,\varepsilon}|^2] \leq C(\sigma(\varepsilon)^2 + n^{1/\log n-1}). \tag{4.13}$$

Combining (2.7) and (4.10) we obtain

$$\begin{aligned} \mathbb{E} \left| \int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z_{t_i}^{n,\varepsilon} \Delta W_{t_i} \right|^2 &\leq C \left(\mathbb{E} \left| \int_0^T V_r dR_r - \int_0^T Z_r^\varepsilon dW_r \right|^2 \right. \\ &\quad \left. + \mathbb{E} \left| \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} [Z_r^\varepsilon - Z_{t_i}^{n,\varepsilon}] dW_r \right|^2 \right) \\ &\leq C(\sigma(\varepsilon)^2 + n^{1/\log n-1}). \end{aligned} \tag{4.14}$$

Arguing as above, we obtain

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E} |\Gamma_t - \Gamma_{t_i}^{n,\varepsilon}|^2 dt \leq \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E} [|\Gamma_t - \Gamma_t^\varepsilon|^2 + |\Gamma_t^\varepsilon - \Gamma_{t_i}^\varepsilon|^2 + |\Gamma_{t_i}^\varepsilon - \Gamma_{t_i}^{n,\varepsilon}|^2] dt.$$

From (2.7), (5.6) below, and (4.11), we have

$$\begin{aligned} \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E} |\Gamma_t - \Gamma_{t_i}^{n,\varepsilon}|^2 dt &\leq \int_0^T \mathbb{E} [|\Gamma_r - \Gamma_r^\varepsilon|^2] dr + Cn^{1/\log n-1} \\ &\leq C \left(\int_0^T \int_{E_\varepsilon} \mathbb{E} |V_r - U_r^\varepsilon(e)|^2 m(de) dr \right. \\ &\quad \left. + \int_0^T \int_{E^\varepsilon} \mathbb{E} |V_r|^2 m(de) dr + n^{1/\log n-1} \right) \\ &\leq C(\sigma(\varepsilon)^2 + n^{1/\log n-1}). \end{aligned} \tag{4.15}$$

Combining (4.13), (4.14), and (4.15) yields the required results.

Remark 4.1. The importance of the above scheme lies in its adeptness to a BSDE when the generator does not depend on the terminal value of a forward equation. Consider the following BSDE driven by a pure jump Lévy process:

$$\hat{Y}_t = \hat{\xi} + \int_t^T f\left(r, \hat{Y}_r, \int_E \rho(e) \hat{V}_r e v(de)\right) dr - \int_t^T \hat{V}_r \bar{M}(dr, de). \tag{4.16}$$

We approximate this process by

$$\hat{Y}_t^\varepsilon = \hat{\xi} + \int_t^T f\left(r, \hat{Y}_r^\varepsilon, \int_{E_\varepsilon} \rho(e) \hat{V}_r e v(de)\right) dr - \int_t^T \hat{Z}_r^\varepsilon dW_s - \int_t^T \int_{E_\varepsilon} \hat{U}_r^\varepsilon(e) \bar{M}(dr, de).$$

We now propose the following discrete-time scheme, defined by terminal values $\hat{Y}_{t_n}^{n,\varepsilon} = \hat{\xi}$, $\hat{Z}_{t_n}^{n,\varepsilon} = D_T \hat{\xi}$ and $\hat{U}_{t_i}^{n,\varepsilon} = D_{T,\varepsilon} \hat{\xi}$:

$$\hat{Y}_{t_i}^{n,\varepsilon} = \mathbb{E}[\hat{Y}_{t_{i+1}}^{n,\varepsilon} + f(\hat{\Theta}_{t_{i+1}}^{n,\varepsilon}) \Delta t_{i+1} \mid \mathcal{F}_{t_i}], \tag{4.17a}$$

$$\hat{Z}_{t_i}^{n,\varepsilon} = \mathbb{E} \left[\mathcal{E}_{t_{i+1}, t_n}^{n,\varepsilon} D_{t_i} \hat{\xi} + \sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1}, t_{k+1}}^{n,\varepsilon} \partial_x f(\hat{\Theta}_{t_{k+1}}^{n,\varepsilon}) D_{t_i} \bar{X}_{t_{k+1}}^{n,\varepsilon} \Delta t_k \mid \mathcal{F}_{t_i} \right], \tag{4.17b}$$

$$\hat{\Gamma}_{t_i}^{n,\varepsilon} = \mathbb{E} \left[\int_{E_\varepsilon} \rho(e) [\mathcal{E}_{t_{i+1},t_n}^{\varepsilon,n} D_{t_i,e} \hat{\xi} + \sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1},t_{k+1}}^{\varepsilon,n} \alpha_{t_i,t_{k+1}}^{n,\varepsilon} D_{t_i,e} \bar{X}_{t_{k+1}}^{n,\varepsilon} \Delta t_k] v(de) \mid \mathcal{F}_{t_i} \right], \tag{4.17c}$$

with $\hat{\Theta}_{t_k}^{n,\varepsilon} = (r, \hat{Y}_{t_k}^{n,\varepsilon}, \hat{\Gamma}_{t_k}^{n,\varepsilon})$. Under the same assumptions as in Theorem 4.2, we prove the convergence of scheme (4.17) to BSDE (4.16). Moreover, we obtain the upper bound

$$\begin{aligned} & \max_{0 \leq i \leq n} \sup_{t \in [t_i, t_{i+1}]} \mathbb{E}[|Y_t - Y_t^{n,\varepsilon}|^2] + \mathbb{E} \left| \int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z_{t_i}^{n,\varepsilon} \Delta W_{t_i} \right|^2 \\ & + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E} |\Gamma_t - \Gamma_{t_i}^{n,\varepsilon}|^2 dt \\ & \leq C(\sigma(\varepsilon)^2 + n^{1/\log n - 1}). \end{aligned}$$

5. Malliavin calculus for FBSDEs

For ease of notation, we shall denote throughout this section the process $(X^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ by (X, Y, Z, Γ) .

In this section we study some regularity properties of the solution (X, Y, Z, Γ) . We recall system (1.3) using the new notation:

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_r) dr + \int_0^t \beta(X_r) \sigma(\varepsilon) dW_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-}) \bar{M}(dr, de), \\ Y_t &= g(X_T) + \int_t^T f(\Theta_r) dr - \int_t^T Z_r dW_s - \int_t^T \int_{E_\varepsilon} U_r(e) \bar{M}(dr, de). \end{aligned} \tag{5.1}$$

In fact, there are many methods to develop Malliavin calculus for Lévy processes. In this paper we opt for the approach of Solé *et al.* [22], based on a chaos decomposition in terms of multiple stochastic integrals with respect to the random measure \bar{M} . Adopting the notation of [11], we will recall the suitable canonical space we adapt to our setting.

We start by introducing some additional notation and definitions. We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the product of two canonical spaces: $(\Omega_W \times \Omega_\mu, \mathcal{F}_W \times \mathcal{F}_\mu, \mathbb{P}_W \times \mathbb{P}_\mu)$ and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$, the canonical filtration completed for \mathbb{P} (for details concerning this construction, see Section 2 of [11]).

We consider the finite measure q defined on $[0, T] \times \mathbb{R}$ by

$$q(B) = \int_{B(0)} dt + \int_{B'} e^2 v(de) dt, \quad B \in \mathcal{B}([0, T] \times \mathbb{R}),$$

where $B(0) = \{t \in [0, T]; (t, 0) \in B\}$, $B' = B - B(0)$, and the random measure $Q \in [0, T] \times \mathbb{R}$ is defined by

$$Q(B) = \int_{B(0)} dW_t + \int_{B'} e \bar{\mu}(dt, de), \quad B \in \mathcal{B}([0, T] \times \mathbb{R}).$$

For $n \in \mathbb{N}$ and a simple function $h_n = \mathbf{1}_{E_1 \times \dots \times E_n}$ with pairwise disjoint sets $E_1, \dots, E_n \in \mathcal{B}([0, T] \times \mathbb{R})$, we define

$$I_n(h_n) = \int_{([0, T] \times \mathbb{R})^n} h((t_1, e_1), \dots, (t_n, e_n)) Q(dt_1, de_1) \dots Q(dt_n, de_n).$$

We also define the following spaces.

- $\mathbb{L}_{T,q,n}^2(\mathbb{R})$, the space of product measurable deterministic functions $h : ([0, T] \times \mathbb{R})^n \rightarrow \mathbb{R}$ satisfying $\|h\|_{\mathbb{L}_{T,q,n}^2}^2 < \infty$, where

$$\|h\|_{\mathbb{L}_{T,q,n}^2}^2 =: \int_{([0,T] \times \mathbb{R})^n} |h((t_1, e_1), \dots, (t_n, e_n))|^2 q(dt_1, de_1) \cdots q(dt_n, de_n).$$

- $\mathbb{D}^{1,2}(\mathbb{R})$ denotes the space of \mathbb{F} -measurable random variables $H \in \mathbb{L}^2(\mathbb{R})$ with the representation $H = \sum_{n=0}^\infty I_n(h_n)$ and satisfying

$$\sum_{n=0}^\infty nn! \|h_n\|_{\mathbb{L}_{T,q,n}^2}^2 < \infty.$$

- $\mathbb{L}^{1,2}(\mathbb{R})$ denotes the space of product measurable and \mathbb{F} -adapted processes $G : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ satisfying

$$\mathbb{E} \left[\int_{[0,T] \times \mathbb{R}} |G(s, y)|^2 q(ds, dy) \right] < \infty,$$

$$G(s, y) \in \mathbb{D}^{1,2}(\mathbb{R}), \text{ for } q \text{ almost every } (s, y) \in [0, T] \times \mathbb{R},$$

$$\mathbb{E} \left[\int_{([0,T] \times \mathbb{R})^2} |D_{t,z} G(s, y)|^2 q(ds, dy) q(dt, dz) \right] < \infty.$$

This space is endowed with the norm

$$\|G\|_{\mathbb{L}^{1,q}}^2 = \mathbb{E} \left[\int_{[0,T] \times \mathbb{R}} |G(s, y)|^2 q(ds, dy) \right] + \mathbb{E} \left[\int_{([0,T] \times \mathbb{R})^2} |D_{t,z} G(s, y)|^2 q(ds, dy) q(dt, dz) \right].$$

We should mention that the derivative $D_{t,0}$ coincides with D_t , the classical Malliavin derivative with respect to Brownian motion.

To study the regularity of Z and U , we shall also introduce the following assumption.

Assumption 5.1. For any $\lambda > 0$ and $q \geq 1$, we consider three progressive measurable processes $\{\alpha_t\}_{0 \leq t \leq T}$, $\{\beta_t\}_{0 \leq t \leq T}$ and $\{\gamma_t\}_{0 \leq t \leq T}$ such that

$$\mathbb{E} \exp \left\{ \lambda \int_0^T (|\beta_r| + \gamma_r^2) dr \right\} < \infty, \quad \sup_{0 \leq t \leq T} \mathbb{E}[|\alpha_t|^q + |\gamma_t|^q] < \infty.$$

Proposition 5.1. Under Assumption 5.1, the discontinuous semimartingale \mathcal{E}_t ,

$$d\mathcal{E}_t = \mathcal{E}_t \beta_t dt + \mathcal{E}_t \gamma_t \int_{E_e} \rho(e) \bar{M}(de, dt),$$

has the following properties.

1. $\mathbb{E} \sup_{0 \leq t \leq T} \mathcal{E}_t^n < \infty$ for any $n \in \mathbb{R}$.
2. The process $\mathcal{Z}_t := \mathcal{E}_t^{-1}$ satisfies the linear stochastic differential equation

$$\frac{d\mathcal{Z}_t}{\mathcal{Z}_t} = \left(-\beta_t + \gamma_t^2 \int_{E_e} \rho^2(e) m(de) \right) dt - \gamma_t \int_{E_e} \rho(e) \bar{M}(de, dt).$$

Moreover, we have, for any $p \geq 2$, $\mathbb{E}|\mathcal{Z}_t - \mathcal{Z}_s|^p \leq C|t - s|^p$.

Proof. We could write \mathcal{E} as

$$\mathcal{E}_t = \exp \left\{ \int_0^t \left(\beta_r - \frac{1}{2} \int_{E_\varepsilon} \gamma_r^2 \rho^2(e) m(de) \right) dr + \int_0^t \int_{E_\varepsilon} \gamma_r \rho(e) \bar{M}(de, dr) \right\}.$$

Under Assumption 5.1, we get the first assertion. The second assertion is deduced from the first assertion, Hölder’s inequality, and the Burkholder–Davis–Gundy inequality.

The following theorem constitutes the main tool used to prove Theorem 5.3 below.

Theorem 5.1. *Suppose that $\mathcal{E}_T X_T$ and $\int_0^T \alpha_r D_\theta X_r dr$ are in $M^{2,q}$. The linear BSDE*

$$Y_t = g(X_T)X_T + \int_t^T [\alpha_r X_r + \beta_r Y_r + \gamma_r \Gamma_r] dr - \int_t^T Z_r dW_r - \int_t^T \int_{E_\varepsilon} U_r(e) \bar{M}(dr, de), \quad 0 \leq t \leq T,$$

has a unique solution (Y, Z, U) and there exists a constant $C > 0$ such that

$$\mathbb{E}|Y_t - Y_s|^p \leq C|t - s|^{p/2} \quad \text{for all } s, t \in [0, T].$$

Proof. Applying Itô’s formula to $\mathcal{E}_t Y_t$, we obtain

$$d(\mathcal{E}_t Y_t) = -\mathcal{E}_t \alpha_t X_t dt + \mathcal{E}_t Z_t dW_t + \mathcal{E}_t \int_{E_\varepsilon} (Y_t \gamma_t \rho(e) + U_t(e)) \bar{M}(de, dt).$$

Then

$$Y_t = \mathbb{E} \left[\mathcal{E}_{t,T} g(X_T) X_T + \int_t^T \mathcal{E}_{t,r} \alpha_r X_r dr \mid \mathcal{F}_t \right],$$

where $\mathcal{E}_{t,r} = Z_t \mathcal{E}_r$.

For $0 \leq s \leq t \leq T$, we have

$$\begin{aligned} \mathbb{E}|Y_t - Y_s|^p &\leq 3^{p-1} \mathbb{E} \left| \mathbb{E}[\mathcal{E}_{t,T} g(X_T) X_T \mid \mathcal{F}_t] - \mathbb{E}[\mathcal{E}_{s,T} g(X_T) X_T \mid \mathcal{F}_s] \right|^p \\ &\quad + 3^{p-1} \mathbb{E} \left| \mathbb{E} \left[\int_t^T \mathcal{E}_{t,r} \alpha_r X_r dr \mid \mathcal{F}_t \right] - \mathbb{E} \left[\int_s^T \mathcal{E}_{s,r} \alpha_r X_r dr \mid \mathcal{F}_s \right] \right|^p \\ &= 3^{p-1} (I_1 + I_2). \end{aligned}$$

By adapting the argument of Theorem 2.3 of [16] and recalling Remark 5.1, we can immediately show that $I_1 \leq C|t - s|^{p/2}$ and $I_2 \leq C|t - s|^{p/2}$.

5.1. Malliavin calculus on the FSDE

In this section we recall some well-known properties of FSDEs concerning the Malliavin derivatives of the solution of an FSDE with jumps, stated in [20] for the case of stochastic differential equations without jumps and in [21] for the case of a Lévy process. The following theorem can be found in [21].

Theorem 5.2. *Let X be the solution of FSDE (5.1). Then, for all $t \in [0, T]$ and $(\theta, e) \in [0, T] \times (\mathbb{R} \setminus \{0\})$, the Malliavin derivatives of X satisfy*

$$\begin{aligned} D_\theta X_t &= \int_\theta^t \partial_x b(X_r) D_\theta X_r dr + \int_\theta^t \int_{E_\varepsilon} \partial_x \beta(X_r) D_\theta X_r \bar{M}(dr, de) \\ &\quad + \sigma(\varepsilon) \beta(X_\theta) + \sigma(\varepsilon) \int_\theta^t \partial_x \beta(X_r) D_\theta X_r dW_r, \quad 0 \leq \theta \leq t \leq T, \end{aligned}$$

and

$$D_{\theta,e}X_t = \int_{\theta}^t D_{\theta,e}b(X_r) dr + \int_{\theta}^t \int_{E_{\varepsilon}} D_{\theta,e}\beta(X_r)\bar{M}(dr, de) + \beta(X_{\theta}) + \sigma(\varepsilon) \int_{\theta}^t D_{\theta,e}\beta(X_r) dW_r, \quad 0 \leq \theta \leq t \leq T.$$

For all $\theta > t$, we have $D_{\theta,e}X_t = D_{\theta}X_t = 0$ almost surely.

Remark 5.1. Using standard arguments as in the proof of Lemma 4.1, we can prove the following a priori estimates:

$$\begin{aligned} \sup_{0 \leq \theta \leq T} \mathbb{E} \left[\sup_{0 \leq t \leq T} D_{\theta}X_t \right] &< \infty, \\ \sup_{0 \leq \theta \leq T, e \in \mathbb{R}^*} \mathbb{E} \left[\sup_{0 \leq t \leq T} D_{\theta,e}X_t \right] &< \infty, \\ \sup_{0 \leq u \leq T} \sup_{0 \leq \theta \leq T} \mathbb{E} \left[\sup_{0 \leq t \leq T} D_u D_{\theta}X_t \right] &< \infty, \\ \sup_{0 \leq u \leq T} \sup_{0 \leq \theta \leq T, e \in \mathbb{R}^*} \mathbb{E} \left[\sup_{0 \leq t \leq T} D_u D_{\theta,e}X_t \right] &< \infty. \end{aligned}$$

5.2. Malliavin calculus on the BSDE

In this section we recall some results of Malliavin derivatives applied to BSDEs, established in [13] and [11], in order to generalize the result of Theorem 2.6 of [16].

Theorem 5.3. *Suppose that Assumption 4.1 holds. There exists a unique solution $\{(Y_t, Z_t, U_t(e))\}_{0 \leq t \leq T, e \in (\mathbb{R} - \{0\})}$ to BSDE (5.1) such that the following assertions hold.*

1. *The first version of the Malliavin derivative $(D_{\theta}Y_t, D_{\theta}Z_t, D_{\theta}U_t(e))_{0 \leq \theta, t \leq T, e \in (\mathbb{R} - \{0\})}$ of the solution $(Y_t, Z_t, U_t(e))_{0 \leq t \leq T, e \in (\mathbb{R} - \{0\})}$ satisfies the linear BSDE*

$$D_{\theta}Y_t = \partial_x g(X_T) D_{\theta}X_T + \int_t^T f^{\theta_r}(\Theta_r) dr - \int_t^T D_{\theta}Z_r dW_s - \int_t^T \int_{E_{\varepsilon}} D_{\theta}U_r(e)\bar{M}(dr, de), \quad 0 \leq \theta \leq t \leq T, \tag{5.2}$$

where $f^{\theta}(\Theta) := \partial_x f(\Theta) D_{\theta}X_r + \partial_y f(\Theta) D_{\theta}Y_r + \partial_{\gamma} f(\Theta) D_{\theta}\Gamma_r$. Moreover, $(D_t Y_t)_{0 \leq t \leq T}$ is a version of $(Z_t)_{0 \leq t \leq T}$:

$$Z_t = D_t Y_t \quad \text{almost surely.} \tag{5.3}$$

2. *The second version of the Malliavin derivative*

$$(D_{\theta,z}Y_t, D_{\theta,z}Z_t, D_{\theta,z}U_t(e))_{0 \leq \theta, t \leq T, (e,z) \in (\mathbb{R} - \{0\})^2}$$

of the solution $(Y_t, Z_t, U_t(z))_{0 \leq t \leq T, z \in (\mathbb{R} - \{0\})}$ satisfies the linear BSDE

$$D_{\theta,z}Y_t = g(X_T + D_{\theta,z}X_T) - g(X_T) + \int_t^T [f(\Theta_r + D_{\theta,z}\Theta_r) - f(\Theta_r)] dr - \int_t^T D_{\theta,z}Z_r dW_r - \int_t^T \int_{E_{\varepsilon}} D_{\theta,z}U_r(e)\bar{M}(dr, de), \quad 0 \leq \theta \leq t \leq T. \tag{5.4}$$

Moreover, $(D_{t,e}Y_t)_{0 \leq t \leq T, e \in (\mathbb{R} - \{0\})}$ is a version of $(U_t(z))_{0 \leq t \leq T, z \in (\mathbb{R} - \{0\})}$:

$$U_t(z) = D_{t,e}Y_t \text{ almost surely.}$$

For $(\theta, e) \in [0, T] \times \mathbb{R}$,

$$D_{\theta,e}Y_t = D_{\theta,e}Z_t = D_{\theta,e}U_t(z) = 0, \quad 0 \leq t < \theta, (e, z) \in \mathbb{R} \times (\mathbb{R} - \{0\}).$$

3. There exists a constant $C > 0$ such that, for all $s, t \in [0, T]$,

$$\mathbb{E}|Z_t - Z_s|^p \leq C|t - s|^{p/2}, \tag{5.5}$$

$$\mathbb{E}|\Gamma_t - \Gamma_s|^p \leq C|t - s|^{p/2}. \tag{5.6}$$

Proof. The proof of the existence and uniqueness of the solution is similar to that of [13, Proposition 5.3] and [11, Theorem 4.1]. We therefore focus our attention on proving inequalities (5.5) and (5.6).

We first prove that $\mathbb{E}|Z_t - Z_s|^p \leq C|t - s|^{p/2}$. From (5.3) we have

$$Z_t - Z_s = D_tY_t - D_sY_s.$$

Then

$$\mathbb{E}|Z_t - Z_s|^p \leq \mathbb{E}|D_tY_t - D_sY_t|^p + \mathbb{E}|D_sY_t - D_sY_s|^p.$$

Step 1: estimate $\mathbb{E}|D_tY_t - D_sY_t|^p$. From Lemma B.1, (5.2), and assumptions (4.1)–(4.3), we obtain

$$\begin{aligned} &\mathbb{E}|D_tY_t - D_sY_t|^p + \mathbb{E}\left(\int_t^T |D_tZ_r - D_sZ_r|^2 dr\right)^{p/2} \\ &+ \mathbb{E}\left(\int_t^T \int_{E_\varepsilon} |D_tU_r(e) - D_sU_r(e)|^2 m(de) dr\right)^{p/2} \\ &\leq C\mathbb{E}(|D_tg(X_T) - D_sg(X_T)|^p) \\ &\quad + C\mathbb{E}\left(\int_t^T |D_t f(r, X_r, Y_r, \Gamma_r) - D_s f(r, X_r, Y_r, \Gamma_r)|^2 dr\right)^{p/2} \\ &\leq C|t - s|^{p/2}. \end{aligned} \tag{5.7}$$

Step 2: estimate $\mathbb{E}|D_sY_t - D_sY_s|^p$. We recall the expression for \mathcal{E}_t :

$$\mathcal{E}_t = \exp\left\{\int_0^t \left(\beta_r - \frac{1}{2} \int_{E_\varepsilon} \gamma_r^2 \rho^2(e) m(de)\right) dr + \int_0^t \int_{E_\varepsilon} \gamma_r \rho(e) \bar{M}(de, dr)\right\}.$$

Define $\beta_r = \partial_y f(\Theta_r)$ and $\gamma_r = \partial_\gamma f(\Theta_r)$. For any $0 \leq \theta \leq t \leq T$, we have

$$\begin{aligned} D_\theta \mathcal{E}_t &= \mathcal{E}_t \left\{ \int_\theta^t \int_{E_\varepsilon} \rho(e) [\partial_{\gamma x} f(\Theta_r) D_\theta X_r + \partial_{\gamma y} f(\Theta_r) D_\theta Y_r + \partial_{\gamma \gamma} f(\Theta_r) D_\theta \Gamma_r] \bar{M}(de, dr) \right. \\ &\quad + \int_\theta^t \int_{E_\varepsilon} [\partial_{xy} f(\Theta_r) - \rho^2(e) \gamma_r \partial_{xy} f(\Theta_r)] D_\theta X_r m(de) dr \\ &\quad + \int_\theta^t \int_{E_\varepsilon} [\partial_{yy} f(\Theta_r) - \rho^2(e) \gamma_r \partial_{yy} f(\Theta_r)] D_\theta Y_r m(de) dr \\ &\quad \left. + \int_\theta^t \int_{E_\varepsilon} [\partial_{\gamma y} f(\Theta_r) - \rho^2(e) \gamma_r \partial_{\gamma \gamma} f(\Theta_r)] D_\theta \Gamma_r m(de) dr \right\}. \end{aligned}$$

By induction on the chain rule,

$$D_{\theta,e}f^2(\Theta) = f^2(\Theta + D_{\theta,e}\Theta) - f^2(\Theta).$$

Using this in the above equality, we obtain

$$D_{\theta,e}\mathcal{E}_t = \mathcal{E}_t \left(\exp \left\{ \int_{\theta}^t \left[\partial_y f(\Theta_r + D_{\theta,e}\Theta_r) - \partial_y f(\Theta_r) - \frac{1}{2} \int_{E_\varepsilon} \rho^2(e) [(\partial_y f(\Theta_r + D_{\theta,e}\Theta_r))^2 - \gamma_r^2] m(de) \right] dr + \gamma_\theta \rho(e) + \int_{\theta}^t \int_{E_\varepsilon} \rho(e) [\partial_y f(\Theta_r + D_{\theta,e}\Theta_r) - \gamma_r] \bar{M}(de, dr) \right\} - 1 \right).$$

From Proposition 5.1, Assumption 4.1, Hölder’s inequality, and the Burkholder–Davis–Gundy inequality, we can show that, for any $p < q$,

$$\sup_{\theta \in [0, T]} \mathbb{E} \sup_{e \in \mathbb{R}} |D_{\theta,e}\mathcal{E}_t|^p < \infty. \tag{5.8}$$

Now, by the Clark–Ocone formula (see [14]) applied to $\mathcal{E}_T D_\theta X_T$,

$$\begin{aligned} \mathcal{E}_T D_\theta X_T &= \mathbb{E}(\mathcal{E}_T D_\theta X_T) + \int_0^T \mathbb{E}[D_r(\mathcal{E}_T D_\theta X_T) \mid \mathcal{F}_r] dW_r \\ &\quad + \int_0^T \int_{E_\varepsilon} \mathbb{E}[D_{r,e}(\mathcal{E}_T D_\theta X_T) \mid \mathcal{F}_r] \bar{M}(dr, de) \\ &= \mathbb{E}(\mathcal{E}_T D_\theta X_T) + \int_0^T u_r^\theta dW_r + \int_0^T \int_{E_\varepsilon} v_{r,e}^\theta \bar{M}(dr, de), \end{aligned}$$

where

$$\begin{aligned} u_r^\theta &:= \mathbb{E}[D_r \mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_r D_\theta X_T \mid \mathcal{F}_r], \\ v_{r,e}^\theta &:= \mathbb{E}[D_{r,e} \mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_{r,e} D_\theta X_T + D_{r,e} \mathcal{E}_T D_{r,e} D_\theta X_T \mid \mathcal{F}_r]. \end{aligned}$$

Thus, it remains to prove that

$$\sup_{\theta \in [0, T]} \sup_{r \in [0, T]} |u_r^\theta|^p < \infty \quad \text{and} \quad \sup_{\theta \in [0, T]} \sup_{r \in [0, T], e \in \mathbb{R}} |v_{r,e}^\theta|^p < \infty.$$

By Hölder’s inequality,

$$\begin{aligned} \mathbb{E}|v_{r,e}^\theta|^p &= \mathbb{E}|\mathbb{E}[D_{r,e} \mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_{r,e} D_\theta X_T + D_{r,e} \mathcal{E}_T D_{r,e} D_\theta X_T \mid \mathcal{F}_r]|^p \\ &\leq 3^{p-1} (\mathbb{E}|D_{r,e} \mathcal{E}_T D_\theta X_T|^p + \mathbb{E}|\mathcal{E}_T D_{r,e} D_\theta X_T|^p + \mathbb{E}|D_{r,e} \mathcal{E}_T D_{r,e} D_\theta X_T|^p) \\ &\leq 3^{p-1} (\mathbb{E}|D_{r,e} \mathcal{E}_T|^{pq/(q-p)})^{(q-p)/q} (\mathbb{E}|D_\theta X_T|^q)^{p/q} \\ &\quad + (\mathbb{E}|\mathcal{E}_T|^{pq/(q-p)})^{(q-p)/q} (\mathbb{E}|D_{r,e} D_\theta X_T|^q)^{p/q} \\ &\quad + (\mathbb{E}|D_{r,e} \mathcal{E}_T|^{pq/(q-p)})^{(q-p)/q} (\mathbb{E}|D_{r,e} D_\theta X_T|^q)^{p/q}. \end{aligned}$$

Combining (5.8) and Remark 5.1, we deduce that $\sup_{\theta \in [0, T]} \sup_{r \in [0, T], e \in \mathbb{R}} |v_{r,e}^\theta|^p < \infty$. Following the same arguments we conclude that $\sup_{\theta \in [0, T]} \sup_{r \in [0, T]} |u_r^\theta|^p < \infty$. As a consequence, $\mathcal{E}_T D_\theta X_T$ belongs to $M^{2,p}$. Therefore, by Theorem 5.1 we conclude that

$$\mathbb{E}|D_s Y_t - D_s Y_s|^p \leq C|t - s|^{p/2}. \tag{5.9}$$

Now, combining (5.7) and (5.9), we finally obtain, for some constant $C > 0$,

$$\mathbb{E}|Z_t - Z_s|^p = \mathbb{E}|D_s Y_t - D_s Y_s|^p \leq C|t - s|^{p/2}.$$

We now prove that $\mathbb{E}|\Gamma_t - \Gamma_s|^p \leq C|t - s|^{p/2}$. By Hölder’s inequality,

$$\begin{aligned} \mathbb{E}|\Gamma_t - \Gamma_s|^p &= \mathbb{E} \left| \int_{E_\varepsilon} \rho(e)(U_t(e) - U_s(e))\nu(\mathrm{d}e) \right|^p \\ &\leq \mathbb{E} \left[\left(\int_{E_\varepsilon} |U_t(e) - U_s(e)|^p \nu(\mathrm{d}e) \right) \left(\int_{E_\varepsilon} |\rho(e)|^{p/(p-1)} \nu(\mathrm{d}e) \right)^{p-1} \right] \\ &\leq C \int_{E_\varepsilon} \mathbb{E}|U_t(e) - U_s(e)|^p \nu(\mathrm{d}e) \\ &= C \int_{E_\varepsilon} \mathbb{E}|D_{t,e}Y_t - D_{s,e}Y_s|^p \nu(\mathrm{d}e) \\ &\leq C \int_{E_\varepsilon} \mathbb{E}[|D_{t,e}Y_t - D_{s,e}Y_t|^p + |D_{s,e}Y_t - D_{s,e}Y_s|^p] \nu(\mathrm{d}e). \end{aligned}$$

Step 3: we prove that $\mathbb{E}|D_{t,e}Y_t - D_{s,e}Y_t|^p \leq C|t - s|^{p/2}$. Under assumptions (4.2) and (4.4), and using Lemma B.1, we obtain

$$\begin{aligned} &\mathbb{E}|D_{t,e}Y_t - D_{s,e}Y_t|^p + \mathbb{E} \left(\int_t^T |D_{t,e}Z_r - D_{s,e}Z_r|^2 \mathrm{d}r \right)^{p/2} \\ &+ \mathbb{E} \left(\int_t^T \int_{E_\varepsilon} |D_{t,e}U_r(e) - D_{s,e}U_r(e)|^2 \nu(\mathrm{d}e) \mathrm{d}r \right)^{p/2} \\ &\leq C \mathbb{E}[|D_{t,e}g(X_T) - D_{s,e}g(X_T)|^p] \\ &\quad + C \mathbb{E} \left(\int_t^T |D_{t,e}f(r, Y_r, U_r) - D_{s,e}f(r, Y_r, U_r)|^2 \mathrm{d}r \right)^{p/2} \\ &\leq C|t - s|^{p/2}. \end{aligned}$$

Step 4: we prove that $\mathbb{E}|D_{s,e}Y_t - D_{s,e}Y_s|^p \leq C|t - s|^{p/2}$. We can write BSDE (5.4) as

$$\begin{aligned} D_{\theta,e}Y_t &= G(X_T)D_{\theta,e}X_T + \int_t^T \alpha_{\theta,r}[D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r] \mathrm{d}r - \int_t^T D_{\theta,e}Z_r \mathrm{d}W_s \\ &\quad - \int_t^T \int_{E_\varepsilon} D_{\theta,e}U_r(e)\bar{M}(\mathrm{d}r, \mathrm{d}e), \end{aligned} \tag{5.10}$$

where

$$\begin{aligned} G(X_T) &:= \frac{g(X_T + D_{\theta,e}X_T) - g(X_T)}{D_{\theta,e}X_T} \mathbf{1}_{\{D_{\theta,e}X_T \neq 0\}}, \\ \alpha_{\theta,r} &:= \frac{f(\Theta_r + D_{\theta,e}\Theta) - f(\Theta_r)}{D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r} \mathbf{1}_{\{D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r \neq 0\}}. \end{aligned}$$

Then, from the Lipschitz continuity of f , we have $\sup_{0 \leq t \leq T} \mathbb{E}|\alpha_{\theta,t}|^p < \infty$. It remains to show that $\mathcal{E}_T D_{\theta,e}X_T$ belongs to $M^{2,p}$. Again, by the Clark–Ocone formula applied to $\mathcal{E}_T D_{\theta,e}X_T$,

$$\begin{aligned} \mathcal{E}_T D_{\theta,e}X_T &= \mathbb{E}(\mathcal{E}_T D_{\theta,e}) + \int_0^T \mathbb{E}[D_r(\mathcal{E}_T D_{\theta,e}X_T) \mid \mathcal{F}_r] \mathrm{d}W_r \\ &\quad + \int_0^T \int_{E_\varepsilon} \mathbb{E}[D_{r,e}(\mathcal{E}_T D_{\theta,e}X_T) \mid \mathcal{F}_r] \bar{M}(\mathrm{d}r, \mathrm{d}e) \\ &= \mathbb{E}(\mathcal{E}_T D_{\theta,e}X_T) + \int_0^T \tilde{u}_r^\theta \mathrm{d}W_r + \int_0^T \int_{E_\varepsilon} \tilde{v}_{r,e}^\theta \bar{M}(\mathrm{d}r, \mathrm{d}e), \end{aligned}$$

with

$$\begin{aligned} \tilde{u}_r^\theta &:= \mathbb{E}[D_r \mathcal{E}_T D_{\theta,e} X_T + \mathcal{E}_T D_r D_{\theta,e} X_T \mid \mathcal{F}_r], \\ \tilde{v}_{r,e}^\theta &:= \mathbb{E}[D_{r,e} \mathcal{E}_T D_{\theta,e} X_T + \mathcal{E}_T D_{r,e} D_{\theta,e} X_T + D_{r,e} \mathcal{E}_T D_{r,e} D_{\theta,e} X_T \mid \mathcal{F}_r]. \end{aligned}$$

Following the same argument as in step 2 and using Remark 5.1, we prove that

$$\sup_{\theta \in [0, T], e \in \mathbb{R}} \mathbb{E}[|\tilde{u}_r^\theta|^p + |\tilde{v}_{r,e}^\theta|^p] < \infty.$$

Therefore, $\mathcal{E}_T D_{\theta,e} X_T$ belongs to $M^{2,p}$. Finally, we once again apply the result of Theorem 5.1 to BSDE (5.10) to obtain

$$\mathbb{E}|D_{s,e} Y_t - D_{s,e} Y_s| \leq C|t - s|^{p/2}.$$

The result then follows.

Appendix A. Proof of Lemma 4.1

By Jensen’s inequality,

$$\begin{aligned} \sup_{\theta \leq s \leq t} \|D_\theta \bar{X}_t^{n,\varepsilon}\|^2 &\leq 3^{2p-1} \sup_{\theta \leq s \leq t} \left\| \int_\theta^s \partial_x b(X_{\phi_r^n}^{n,\varepsilon}) D_\theta X_{\phi_r^n}^{n,\varepsilon} dr \right\|^{2p} \\ &\quad + 3^{2p-1} \sup_{\theta \leq s \leq t} \left\| \sigma(\varepsilon) \int_\theta^s \partial_x \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon}) D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon} dW_r \right\|^{2p} \\ &\quad + 3^{2p-1} \sup_{\theta \leq s \leq t} \left\| \int_\theta^s \int_{E_\varepsilon} \partial_x \beta(\bar{X}_{\phi_r^n}^{n,\varepsilon}) D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon} \bar{M}(de, dr) \right\|^{2p} \\ &\quad + 3^{2p-1} |\sigma(\varepsilon) \beta(\bar{X}_\theta^{n,\varepsilon})|^{2p}. \end{aligned}$$

Taking the expectation of both sides and using Hölder’s inequality, we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{\theta \leq s \leq t} \|D_\theta \bar{X}_t^{n,\varepsilon}\|^2 \right] &\leq C_p \left(\mathbb{E} \left[\int_\theta^t \|D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon}\|^2 dr \right] + \mathbb{E} \left[\int_\theta^t \|D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon}\|^2 dr \right]^p \right. \\ &\quad \left. + \mathbb{E} \left[\int_0^T \|D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon}\|^2 dr \right] + \mathbb{E}[|\sigma(\varepsilon) \beta(\bar{X}_\theta^{n,\varepsilon})|^{2p}] \right) \\ &\leq C \left(\mathbb{E} \left[\int_\theta^t \|D_\theta \bar{X}_{\phi_r^n}^{n,\varepsilon}\|^2 dr \right] + B \right) \\ &\leq C \left(\mathbb{E} \left[\int_\theta^t \sup_{0 \leq u \leq r} \|D_\theta \bar{X}_u^{n,\varepsilon}\|^2 dr \right] + B \right), \end{aligned}$$

where $B := \mathbb{E}[\sup_{0 \leq s \leq T} \|\beta(\bar{X}_s^{n,\varepsilon})\|^2]$. Since the constant C does not depend on θ and n , we conclude by Gronwall’s lemma that $\mathbb{E}[\sup_{\theta \leq t \leq T} \|D_\theta \bar{X}_t^{n,\varepsilon}\|^2]$ is bounded and, therefore, $\sup_{0 \leq \theta \leq T} \sup_{n \geq 1} \mathbb{E}[\sup_{\theta \leq t \leq T} \|D_\theta \bar{X}_t^{n,\varepsilon}\|^2]$ is finite.

By the same arguments, we prove that

$$\sup_{0 \leq \theta \leq T} \sup_{n \geq 1} \mathbb{E} \left[\sup_{\theta \leq t \leq T} \|D_{\theta,e} \bar{X}_t^{n,\varepsilon}\|^2 \right] < \infty.$$

Appendix B. A priori estimates

Lemma B.1. *Let $\xi \in L^q(\Omega)$, and let $f : \Omega \times [0, T] \times \mathbb{R} \times L^2(E, \mathcal{E}, \nu; \mathbb{R}) \rightarrow \mathbb{R}$ be $\mathcal{P} \times \mathcal{B} \times \mathcal{B}(L^2(E, \mathcal{E}, \nu; \mathbb{R}))$ measurable, satisfy $\mathbb{E} \int_0^T |f(t, 0, 0)|^2 dt < \infty$, and be uniformly Lipschitz with respect to (y, z) such that, for some constant $K > 0$, we have*

$$|f(t, y_1, u_1) - f(t, y_2, u_2)| \leq K(|y_1 - y_2| + \|u_1 - u_2\|)$$

for all $y_1, y_2, \in \mathbb{R}$ and $u_1, u_2 \in L^2(E, \mathcal{E}, \nu; \mathbb{R})$. Then, there exists a unique triple $(Y, Z, U) \in \mathcal{B}^2$ solution to BSDEs (1.3). Moreover, for $q \geq 2$, we have the following a priori estimate:

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t|^q \right] + \mathbb{E} \left(\int_0^T |Z_t|^2 \right)^{q/2} + \mathbb{E} \left[\int_0^T \int_E |U_t(e)|^q \nu(de) dt \right] \\ & \leq C \left(\mathbb{E} |\xi|^q + \mathbb{E} \left(\int_0^T |f(t, 0, 0)|^2 dt \right)^{q/2} \right). \end{aligned} \quad (\text{B.1})$$

Proof. The existence and uniqueness of solutions to BSDEs with jumps was proved in [4], and estimate (B.1) is a direct consequence of Proposition 2.2 in the same reference, with $(f', Q') = (0, 0)$.

Acknowledgements

The author thanks the anonymous referee for his/her very constructive comments to improve the presentation of this paper. He is deeply grateful to Professor Bruno Bouchard for fruitful discussions concerning Section 2 and Section 3. Finally, the author thanks Dr Rachid Belfadli for helpful suggestions to improve the readability of the paper.

References

- [1] APPLEBAUM, D. (2004). *Lévy Processes and Stochastic Calculus* (Camb. Stud. Adv. Math. **93**). Cambridge University Press.
- [2] ASMUSSEN, S. AND ROSIŃSKI, J. (2001). Approximations of small jumps of Levy processes with a view towards simulation. *J. Appl. Prob.* **38**, 482–493.
- [3] BALLY, V. AND PAGÈS, G. (2003). Error analysis of the optimal quantization algorithm for obstacle problems. *Stoch. Process. Appl.* **106**, 1–40.
- [4] BARLES, G., BUCKDAHN, R. AND PARDOUX, E. (1997). Backward stochastic differential equations and integral-partial differential equations. *Stoch. Stoch. Reports* **60**, 57–83.
- [5] BOUCHARD, B. AND ELIE, R. (2008). Discrete-time approximation of decoupled forward-backward SDE with jumps. *Stoch. Process. Appl.* **118**, 53–75.
- [6] BOUCHARD, B. AND TOUZI, N. (2004). Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stoch. Process. Appl.* **111**, 175–206.
- [7] BRIAND, P., DELYON, B. AND MÉMIN, J. (2001). Donsker-type theorem for BSDEs. *Electron. Commun. Probab.* **6**, 1–14.
- [8] CHEVANCE, D. (1997). Numerical methods for backward stochastic differential equations. In *Numerical Methods in Finance*, eds L. C. G. Rogers and D. Talay, Cambridge University Press, pp. 232–244.
- [9] CLÉMENT, E., LAMBERTON, D. AND PROTTER, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance Stoch.* **6**, 449–472.
- [10] COQUET, F., MACKEVICIUS, V. AND MÉMIN, J. (1998). Stability in D of martingales and backward equations under discretization of filtration. *Stoch. Process. Appl.* **75**, 235–248.
- [11] DELONG, L. AND IMKELLER, P. (2010). On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures. *Stoch. Process. Appl.* **120**, 1748–1775.
- [12] DOUGLAS, J., JR., MA, J. AND PROTTER, P. (1996). Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.* **6**, 940–968.
- [13] EL KAROUL, N., PENG, S. AND QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7**, 1–71.

- [14] ES-SEBAIY, K. AND TUDOR, C. A. (2008). Lévy processes and Itô Skorokhod integrals. *Theory Stoch. Process.* **14**, 10–18.
- [15] FUJIWARA, T. AND KUNITA, H. (1985). Stochastic differential equations of jump type and Lévy processes in diffeomorphism group. *J. Math. Kyoto Univ.* **25**, 71–106.
- [16] HU, Y., NUALART, D. AND SONG, X. (2011). Malliavin calculus for backward stochastic differential equations and application to numerical solutions. *Ann. Appl. Prob.* **21**, 2379–2424.
- [17] KOHATSU-HIGA, A. AND TANKOV, P. (2010). Jump-adapted discretization schemes for Levy-driven SDEs. *Stoch. Process. Appl.* **120**, 2258–2285.
- [18] LEMOR, J. P., GOBET, E. AND WARIN, X. (2006). Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli* **12**, 889–916.
- [19] MA, J., PROTTER, P. AND YONG, J. M. (1994). Solving forward-backward stochastic differential equations—a four step scheme. *Prob. Theory Relat. Fields* **98**, 339–359.
- [20] NUALART, D. (1995). *The Malliavin Calculus and Related Topics*. Springer, New York.
- [21] PETROU, E. (2008). Malliavin calculus in Lévy spaces and applications to finance. *Electron. J. Prob.* **13**, 852–879.
- [22] SOLÉ, J. L., UTZET, F. AND VIVES, J. (2007). Canonical Lévy process and Malliavin calculus. *Stoch. Process. Appl.* **117**, 165–187.