SFB 823

Discussion Paper

Approximation and error analysis of forward-backward SDEs driven by general Lévy processes using shot noise series representations

Till Massing

Nr. 25/2021

Approximation and Error Analysis of Forward-Backward SDEs driven by General Lévy Processes using Shot Noise Series Representations

Till Massing*

November 22, 2021

Abstract

We consider the simulation of a system of decoupled forward-backward stochastic differential equations (FBSDEs) driven by a pure jump Lévy process L and an independent Brownian motion B. We allow the Lévy process L to have an infinite jump activity. Therefore, it is necessary for the simulation to employ a finite approximation of its Lévy measure. We use the generalized shot noise series representation method by Rosiński (2001) to approximate the driving Lévy process L. We compute the L^p error, $p \geq 2$, between the true and the approximated FBSDEs which arises from the finite truncation of the shot noise series (given sufficient conditions for existence and uniqueness of the FBSDE). We also derive the L^p error between the true solution and the discretization of the approximated FBSDE using an appropriate backward Euler scheme.

Keywords: Decoupled forward-backward SDEs with jumps; Lévy processes; Shot noise series representation; Discrete-time approximation; Euler Scheme
2010 Mathematics Subject Classification: 60H10; 60H35; 65C05

1 Introduction

We consider a system of decoupled forward-backward stochastic differential equations (FBSDE) with jumps of the type

$$X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t a(s, X_s)\mathrm{d}B_s + \int_0^t h(s, X_{s-})\mathrm{d}L_s \tag{1}$$

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s, \Gamma_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}B_s - \int_t^T U_s\mathrm{d}L_s, \tag{2}$$

for $t \in [0, T]$, where B is a Brownian motion and L is an independent pure jump Lévy process and $\Gamma_s = \int_{\mathbb{R}} \rho(e)U_s e\nu(\mathrm{d}e)$. We discuss the case where L has an infinite jump activity, i.e., $\nu(\mathbb{R}) = \infty$, where ν denotes its corresponding Lévy measure. (Assumptions on the functions a, b, f, g, h are discussed below.) For path simulation it is therefore necessary to employ a finite approximation of the Lévy measure first. Afterwards we have to discretize the FBSDE. We are interested in path simulation and the associated error between the true solution and the approximate solution of (1)-(2).

Backward SDEs are a vibrant research topic since the seminal paper of Pardoux & Peng (1990). They proved existence and uniqueness in the L^2 sense of a solution of a BSDE (without jumps) under the assumptions of square integrability of the terminal condition and Lipschitz continuity of the generator f. Since then BSDEs and/or FBSDEs have been analyzed in many directions.

One strand of the literature treats extensions of the existence and uniqueness result of Pardoux & Peng (1990) by relaxing the underlying assumptions or extending the BSDE under consideration. For example, Tang & Li (1994) and Barles et al. (1997) included jumps into the BSDE. Briand et al. (2003) discussed the existence and uniqueness in an L^p sense given a Brownian filtration. Buckdahn & Pardoux (1994) did the same including jumps. Since then several papers have shown L^p existence and uniqueness with a generalized filtration under weak assumptions, e.g., Kruse & Popier (2016), Yao (2017), and Eddahbi et al.

*Faculty of Economics, University of Duisburg-Essen, Universitätsstr. 12, 45117 Essen, Germany.
E-Mail: till.massing@uni-due.de

(2017). FBSDEs are the Markovian special case of BSDEs where the terminal condition is determined by the forward SDE.

Another strand of the literature covers possible areas of applications of FBSDEs. For example, FBSDEs turned out to be useful in mathematical finance, see Karoui et al. (1997) and Delong (2013), in optimal control, see Tang & Li (1994), or for partial differential equations, see Pardoux (1999) or also the book Pardoux & Răşcanu (2014).

A third strand is about the discrete-time approximation of FBSDEs, which this paper aims to contribute to. A popular approach is a backward Euler scheme, see Zhang (2004) and Bouchard & Touzi (2004) who derived the L^2 approximation error of the scheme. Gobet & Labart (2007) generalized this by computing the L^p error. Bouchard & Elie (2008) derived the L^2 error for FBSDEs containing a finite number of jumps. In the case of an infinite jump activity, Aazizi (2013) proposed a two-step approximation by first approximating the small jumps by a Brownian motion to have only finitely many big jumps, and second by discretizing according to Bouchard & Elie (2008). Aazizi (2013) then derived the L^2 approximation-discretization error. The approach follows Kohatsu-Higa & Tankov (2010) who approximated forward SDEs with infinitely many jumps by finitely many jumps.

This paper contributes to literature in the following way. First, we extend the results of Aazizi (2013) for the L^2 approximation-discretization error to a more general L^p, $p \geq 2$ version. Second, instead of partitioning the Lévy measure into jumps larger or smaller than a certain level we allow for various truncation functions using the approach of shot noise series representations by Rosiński (2001), which may be the more efficient way for a certain Lévy process. Third, we enlarge the class of pure jump Lévy processes to these which do not fulfill the Asmussen & Rosiński (2001) assumption for the approximation of small jumps. All in all, we obtain a statement for the L^p error for general Lévy processes. We find that the error depends on $N^{-1/2}$, where N is the number of time steps, and on the pth and second moments of the Lévy measure of the discarded jumps.

The remainder of this paper is organized as follows. In Section 2 we discuss the settings in more detail. In Section 3 we derive an upper bound for the error of the approximation with a finite jump measure. In Section 4 we present the discrete Euler scheme and prove an upper bound for the discretization error.

2 Settings

This section introduces the setting and notation needed throughout this paper. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space such that $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets, $\mathcal{F}_T = \mathcal{F}$, and $(\mathcal{F}_t)$ satisfies the usual assumptions. We assume that $(\mathcal{F}_t)$ is generated by a one-dimensional Brownian motion B and an independent Poisson measure μ on $[0,T] \times \mathbb{R}$ with intensity $\nu(\mathrm{d}e)\mathrm{d}t$, where ν is a Lévy measure on $\mathbb{R}$ and $\mathbb{R}$ is equipped with the Borel set $\mathcal{B} := \mathcal{B}(\mathbb{R})$. We assume that the Lévy measure ν satisfies $\int_{\mathbb{R}}(1 \wedge |e|^2)\nu(\mathrm{d}e) \leq K < \infty$, for a constant $K > 0$ (the Lipschitz constant from below) and that $\nu(\mathbb{R}) = \infty$. Furthermore, we assume that

$$\int_{\mathbb{R}} |e|^p \nu(\mathrm{d}e) < \infty,$$

for $p \geq 2$. This implies that the pth moment of L_t for each $t \in [0,T]$ is finite. We denote by $\widetilde{\mu}(\mathrm{d}e, \mathrm{d}s) = \mu(\mathrm{d}e, \mathrm{d}s) - \nu(\mathrm{d}e)\mathrm{d}s$ the compensated Poisson measure corresponding to μ.

For $p \geq 2$ we define the normed spaces on $[r,t]$, $r \leq t$,

- $\mathcal{S}^p_{[r,t]}$ is the set of real-valued adapted càdlàg processes Y such that

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

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

$$||Z||_{\mathbb{H}^p_{[r,t]}} := \mathbb{E}\left[\left(\int_r^t |Z_s|^2 \mathrm{d}s\right)^{p/2}\right]^{1/p} < \infty.$$

- $\mathbb{L}^p_{\mu,[r,t]}$ is the set of $(\mathcal{P} \otimes \mathcal{B})$-measurable maps $U : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ such that

$$||U||_{\mathbb{L}^p_{\mu,[r,t]}} := \mathbb{E}\left[\left(\int_r^t \int_{\mathbb{R}} |U_s(e)e|^2 \nu(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right]^{1/p} < \infty,$$

where $\mathcal{P}$ is the σ-algebra of $(\mathcal{F}_t)$-predictable subsets of $\Omega \times [0,T]$.

- $\mathbb{L}^p_\nu$ is the set of measurable maps $U : \mathbb{R} \to \mathbb{R}$ such that

$$||U||_{\mathbb{L}^p_\nu} := \left(\int_{\mathbb{R}} |U(e)e|^p \nu(\mathrm{d}e) \right)^{1/p} < \infty.$$

- The space $\mathcal{E}^p_{[r,t]} := \mathcal{S}^p_{[r,t]} \times \mathbb{H}^p_{[r,t]} \times \mathbb{L}^p_{\mu,[r,t]}$ is endowed with the norm

$$||(Y,Z,U)||_{\mathcal{E}^p_{[r,t]}} := \left(||Y||^p_{\mathcal{S}^p_{[r,t]}} + ||Z||^p_{\mathbb{H}^p_{[r,t]}} + ||U||^p_{\mathbb{L}^p_{\mu,[r,t]}} \right)^{1/p}.$$

In the remainder we omit the subscripts if $[r,t] = [0,T]$, e.g., $\mathcal{E}^p := \mathcal{E}^p_{[0,T]}$.

We introduce the set of assumptions needed throughout the proofs. Note that these assumptions are not the minimal ones needed for existence and uniqueness. However they are not overly restrictive and used frequently throughout the literature, e.g., Bouchard & Elie (2008) and Aazizi (2013).

Assumption 1. *(i) Let $a : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $b : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $h : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be Lipschitz continuous functions w.r.t. x and $\frac{1}{2}$-Hölder continuous w.r.t. t, i.e., for a constant $K > 0$*

$$|b(t,x) - b(t',x')| + |a(t,x) - a(t',x')| + |h(t,x) - h(t',x')| \le K \left(|t-t'|^{1/2} + |x-x'| \right) \quad (3)$$

is satisfied for all $(t,x),(t',x') \in [0,T]\times\mathbb{R}$.

(ii) Let $f : [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ such that it is Lipschitz continuous w.r.t. (x,y,z,q) and $\frac{1}{2}$-Hölder continuous w.r.t. t, i.e., for a constant $K > 0$

$$|f(t,x,y,z,q) - f(t,x',y',z',q')| \le K(|t-t'|^{1/2} + |x-x'| + |y-y'| + |z-z'| + |q-q'|) \quad (4)$$

is satisfied for all $(t,x,y,z,q),(t',x',y',z',q') \in [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$.

(iii) Let $\rho : \mathbb{R} \to \mathbb{R}$ be a measurable function such that for a constant $K > 0$

$$\sup_{e\in\mathbb{R}} |\rho(e)| \le K(1 \wedge |e|),$$

for all $e \in \mathbb{R}$.

(iv) For $p \ge 2$ the integrability condition

$$\mathbb{E}\left[|g(X_T)|^p + \int_0^T |f(t,0,0,0,0)|^p \mathrm{d}t \right] < \infty.$$

is satisfied.

To prove Theorem 2 we need the following additional assumption. A discussion about it can be found in Remark 6.

Assumption 2. *For each $e \in \mathbb{R}$, the function $h(x)$ is differentiable with derivative $h'(x)$ such that the function*

$$(x,\xi) \in \mathbb{R}\times\mathbb{R} \mapsto \ell(x,\xi;e) := \xi^2(h'(x)e + 1)$$

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

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

We next mention some important facts on Lévy processes which we need to approximate the infinite Lévy measure. We opt for the approximation using series representations which goes back to Rosiński (2001), see Yuan & Kawai (2021) for a recent overview. Let L be a pure jump Lévy process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ with Lévy measure ν as discussed above. One version of the famous Lévy-Itô decomposition states that L can be written as

$$L_t = \int_0^t \int_{\mathbb{R}} e\widetilde{\mu}(\mathrm{d}e, \mathrm{d}s) + \xi t \quad (5)$$

for $t \in [0,T]$, where $\widetilde{\mu}$ is the associated compensated Poisson measure and $\xi \in \mathbb{R}$. Without loss of generality we assume $\xi = 0$, otherwise we would include it in the functions b and f in (1)-(2).

Rosiński (2001) proved the useful result that it is possible to express jump-type Lévy processes as an infinite series. We now summarize his theory of *generalized shot noise series representations*. We only present the one-dimensional case (it is of course also available in d dimensions). Suppose that the Lévy measure ν can be decomposed as

$$\nu(B) = \int_0^\infty \mathbb{P}[H(r,V) \in B]\mathrm{d}r, \quad B \in \mathcal{B},$$

where V is a random variable in some space $\mathcal{V}$ and $H : (0,\infty) \times \mathcal{V} \to \mathbb{R}$ is a measurable function such that for every $v \in \mathcal{V}$, $r \mapsto |H(r,v)|$ is nonincreasing. Then, it holds that

$$L_t \overset{\mathcal{L}}{=} \sum_{i=1}^\infty H\left(\frac{G_i}{T}, V_i\right) \mathbb{1}_{[0,t]}(T_i) - tc_i, \tag{6}$$

for $t \in [0,T]$, where $\{G_i\}_{i\in\mathbb{N}}$ are the arrival times of a standard Poisson process, $\{V_i\}_{i\in\mathbb{N}}$ are i.i.d. copies of the random variable V independent of $\{G_i\}_{i\in\mathbb{N}}$, $\{T_i\}_{i\in\mathbb{N}}$ are i.i.d. uniforms on $[0,T]$ independent of $\{G_i\}_{i\in\mathbb{N}}$ and $\{V_i\}_{i\in\mathbb{N}}$, and $\{c_k\}_{k\in\mathbb{N}}$ are centering constants such that

$$c_i = \int_{i-1}^i \int_{|x|\le 1} x\mathbb{P}[H(r,V) \in \mathrm{d}x]\mathrm{d}r.$$

Rosiński's theorem offers several choices for different series representations. The most convenient representation is case-dependent given the specific Lévy measure. Well-known special cases include the inverse Lévy measure method, the rejection method or the thinning method, see Yuan & Kawai (2021) for details.

To obtain a feasible numerical algorithm one has to truncate the infinite series in (6). Instead of truncating the series deterministically, i.e., after n summands, we choose a random truncation

$$L_t^n := \sum_{\{i:G_i\le nT\}} H\left(\frac{G_i}{T}, V_i\right) \mathbb{1}_{[0,t]}(T_i) - tc_i,$$

where we cut off all summands if $G_i > nT$, which depends on the random Poisson arrival times $\{G_i\}$. The reason is that with the random truncation L^n itself is a compound Poisson process and hence a proper Lévy process with Lévy measure

$$\nu^n(B) = \int_0^n \mathbb{P}[H(r,U) \in B]\mathrm{d}r, \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Note that the truncated Lévy measure only describes finitely many jumps, i.e., $\nu^n(\mathbb{R}) = n < \infty$. In the following we use the notation $\bar{\nu}^n(\mathrm{d}e) := \nu(\mathrm{d}e) - \nu^n(\mathrm{d}e)$, which is the Lévy measure of the infinitely many small jumps that are discarded. The important quantity which determines the approximation error of FBSDEs will be in terms of the second and pth moments of the Lévy measure $\bar{\nu}^n$, which are defined as

$$\sigma^p(n) := \int_{\mathbb{R}} |e|^p \bar{\nu}^n(\mathrm{d}e),$$

for $p \ge 2$. For example, $\sigma^2(n)$ is the variance of the discarded jumps if n is the level of truncation. By μ^n we denote the Poisson measure with intensity measure $\nu^n(\mathrm{d}e)\mathrm{d}t$ and by $\bar{\mu}^n$ the Poisson measure with intensity measure $\bar{\nu}^n(\mathrm{d}e)\mathrm{d}t$. Let $\widetilde{\mu^n}$ and $\widetilde{\bar{\nu}^n}$ be the corresponding compensated Poisson measures. Clearly, $\nu = \nu^n + \bar{\nu}^n$, $\mu = \mu^n + \bar{\mu}^n$ and $\widetilde{\mu} = \widetilde{\mu^n} + \widetilde{\bar{\mu}^n}$.

Remark 1. The Lévy measures ν and $\bar{\nu}^n$ are assumed to be infinite, i.e., $\nu(\mathbb{R}) = \bar{\nu}^n(\mathbb{R}) = \infty$ and thus we cannot apply Jensen's inequality to integrals of the type $\int \nu(\mathrm{d}e)$. Assumption 1.(iii) provides a necessary bound. Indeed, by Hölder's inequality

$$\left(\int_{\mathbb{R}} \rho(e)U_s(e)e\nu(\mathrm{d}e)\right)^2 \le \int_{\mathbb{R}} U_s(e)^2e^2\nu(\mathrm{d}e) \int_{\mathbb{R}} \rho(e)^2\nu(\mathrm{d}e) \le K^3 \int_{\mathbb{R}} U_s(e)^2e^2\nu(\mathrm{d}e)$$

for $U_s \in \mathbb{L}^2_\mu$, because $\int(1+|e|^2)\nu(\mathrm{d}e) \le K < \infty$ is bounded by a finite constant.

We return to the discussion of FBSDEs. We rephrase (1)-(2) given the Lévy-Itô decomposition (5). We call (X, Y, Z, U) the solution of the original FBSDE

$$X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t a(s, X_s)\mathrm{d}B_s + \int_0^t \int_{\mathbb{R}} h(s, X_{s-})e\widetilde{\mu}(\mathrm{d}e, \mathrm{d}s) \tag{7}$$

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s, \Gamma_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}B_s - \int_t^T \int_{\mathbb{R}} U_s(e)e\widetilde{\mu}(\mathrm{d}e, \mathrm{d}s), \tag{8}$$

where $\Gamma_s = \int_{\mathbb{R}} \rho(e)U_s(e)e\nu(\mathrm{d}e)$. A minor generalization is that the process U now may also depend on $e \in \mathbb{R}$. (We could also let h depend on e. The generalization is straightforward which we omit in this paper.) Note that the last integrand of (8) is written in the product form $U_s(e)e$ because this turns out to be useful in the proofs.

Given the approximation of Lévy processes using truncated series representations we use $\widetilde{\mu^n}$ to approximate the Poisson measure $\widetilde{\mu}$. We call (X^n, Y^n, Z^n, U^n) the solution of the approximate FBSDE

$$X_t^n = X_0 + \int_0^t b(s, X_s^n)\mathrm{d}s + \int_0^t a(s, X_s^n)\mathrm{d}B_s + \int_0^t \int_{\mathbb{R}} h(s, X_{s-}^n)e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s) \tag{9}$$

$$Y_t^n = g(X_T^n) + \int_t^T f(s, X_s^n, Y_s^n, Z_s^n, \Gamma_s^n)\mathrm{d}s - \int_t^T Z_s^n\mathrm{d}B_s - \int_t^T \int_{\mathbb{R}} U_s^n(e)e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s), \tag{10}$$

where $\Gamma_s^n = \int_{\mathbb{R}} \rho(e)U_s^n(e)e\nu^n(\mathrm{d}e)$. The aim of the next section is to compute the approximation error between the original FBSDE (7)-(8) and the approximate FBSDE (9)-(10).

Remark 2. In this paper we restrict ourselves to one-dimensional FBSDEs. The extension to multidimensional FBSDEs (the comparison principle, see Barles et al. (1997), only holds for one-dimensional BSDEs) is straightforward. We can replace Itô's formula by its multidimensional counterpart in the proofs and use the same arguments and bounds to obtain the multidimensional results. We omit these details in the proofs in favor of a simpler notation.

We end this section with a notational remark. Let C_p denote a generic constant depending only on p and further constants including K, T, $a(0)$, $b(0)$, $f(0)$, $g(0)$, $h(0)$ and the starting value X_0, which may vary from step to step.

3 Error of the approximation of the pure jump process

In this section we compute the L^p approximation error between the original backward SDE (8) and the approximate backward SDE (10), defined as

$$Err_n(Y, Z, U) := \mathbb{E}\left[\sup_{0\le t\le T} |Y_t - Y_t^n|^p + \left(\int_0^T |Z_s - Z_s^n|^2\mathrm{d}s\right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} |U_s(e) - U_s^n(e)|^2 e^2\nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} U_s(e)^2 e^2\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right]^{1/p}.$$

Furthermore, we derive an upper bound for the approximation error of the forward SDE defined as

$$\mathbb{E}\left[\sup_{0\le t\le T} |X_t - X_t^n|^p\right].$$

In the following proofs, the standard estimate for the solution of the FBSDEs is useful:

$$||(X, Y, Z, U)||_{\mathcal{S}^p\times\mathcal{E}^p}^p \le C_p(1 + |X_0|^p), \tag{11}$$

for $p \ge 2$, see Bouchard & Elie (2008). In particular, the forward SDE has the estimate

$$\mathbb{E}\left[\sup_{0\le t\le T} |X_t|^p\right] < C_p(1 + |X_0|^p). \tag{12}$$

Because the FBSDEs are decoupled we can analyze the forward and backward components separately. We begin with an error bound for the forward SDE.

Proposition 1. *Let $p \geq 2$. Under Assumption 1 on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$*

- *there exists a unique solution X of (7) on $[0,T]$ with $X_0 = 0$,*
- *for any $n \in \mathbb{N}$, there exists a unique solution X^n of (9) on $[0,T]$ with $X_0^n = 0$,*

Moreover, there exists a constant C_p such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t - X_t^n|^p\right] \leq C_p\left(\sigma^p(n) + \sigma^2(n)^{p/2}\right). \tag{13}$$

Proof. The existence and uniqueness is a standard result, see Applebaum (2009). We thus only prove the bound (13) which is an easy extension of Aazizi (2013).

Let $t \leq T$. We plug in the SDEs and use the Burkholder-Davis-Gundy inequality to obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq r\leq t} |X_r - X_r^n|^p\right] \leq C_p \Bigg(& \mathbb{E}\left[\left(\int_0^t |b(s,X_s) - b(s,X_s^n)|\mathrm{d}s\right)^p\right] \\ & + \mathbb{E}\left[\left(\int_0^t |a(s,X_s) - a(s,X_s^n)|^2\mathrm{d}s\right)^{p/2}\right] \\ & + \mathbb{E}\left[\left(\int_0^t \int_{\mathbb{R}} |h(s,X_s) - h(s,X_s^n)|^2|e|^2\nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right] \\ & + \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} |h(s,X_s) - h(s,X_s^n)|^p|e|^p\nu^n(\mathrm{d}e)\mathrm{d}s\right] \\ & + \mathbb{E}\left[\left(\int_0^t \int_{\mathbb{R}} |h(s,X_s)|^2|e|^2\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right] \\ & + \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} |h(s,X_s)|^p|e|^p\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right]\Bigg). \end{aligned} \tag{14}$$

By Jensen's inequality, the Lipschitz assumption (3) and (11)

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq r\leq t} |X_r - X_r^n|^p\right] &\leq C_p \Bigg(\mathbb{E}\left[\int_0^t |X_s - X_s^n|^p \mathrm{d}s\right] + \mathbb{E}\left[\left(\int_0^t \int_{\mathbb{R}} (1+|X_s|^2)|e|^2\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right] \\ &\qquad + \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} (1+|X_s|^p)|e|^p\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right]\Bigg) \\ &\leq C_p\left(\int_0^t \mathbb{E}\left[\sup_{0\leq r\leq t} |X_s - X_s^n|^p\right]\mathrm{d}s + \sigma^2(n)^{p/2} + \sigma^p(n)\right),\end{aligned}$$

Now the result follows from Gronwall's lemma. □

We now turn to the approximation error of the backward SDE. We start with a remark which gives some insights into the error and next we state and prove our first main result.

Remark 3. Observe that, by the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_t^T \int_{\mathbb{R}} U_s(e)e\widetilde{\mu}(\mathrm{d}e,\mathrm{d}s) - \int_t^T \int_{\mathbb{R}} U_s^n(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|^p\right] \\ &\leq C_p\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_t^T \int_{\mathbb{R}} (U_s(e) - U_s^n(e))e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|^p + \sup_{0\leq t\leq T}\left|\int_t^T \int_{\mathbb{R}} U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)\right|^p\right]\end{aligned}$$

$$\leq C_p\mathbb{E}\left[\left(\left(\int_0^T\int_{\mathbb{R}}|U_s(e)-U_s^n(e)|^2e^2\nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}+\left(\int_0^T\int_{\mathbb{R}}U_s(e)^2e^2\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right]$$

Theorem 1. *Let $p\geq 2$. Under Assumption (1) on $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbb{P})$*

- *there exists a unique solution (Y,Z,U) of (8) in $\mathcal{E}^p$,*
- *for any $n\in\mathbb{N}$, there exists a unique solution (Y^n,Z^n,U^n) of (10) in $\mathcal{E}^p$.*

Moreover, there exists a constant C_p such that

$$Err_n(Y,Z,U)^p\leq C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right).$$

Proof. We omit the proof of existence and uniqueness because it can be found in the vast literature. The standard way is to first show existence and uniqueness in the space $\mathcal{E}^2$ and second show that the solution also belongs to $\mathcal{E}^p$. We refer to Barles et al. (1997), Briand et al. (2003), Buckdahn & Pardoux (1994), Kruse & Popier (2016) and Eddahbi et al. (2017). Some of their techniques also provide to be useful for the derivation of the error.

Define

$$\begin{aligned}\delta Y_t &:= Y_t-Y_t^n\\ &= g(X_T)-g(X_T^n)+\int_t^T f(\Theta_s)\mathrm{d}s-\int_t^T f(\Theta_s^n)\mathrm{d}s-\left(\int_t^T Z_s\mathrm{d}B_s-\int_t^T Z_s^n\mathrm{d}B_s\right)\\ &\quad -\left(\int_t^T\int_{\mathbb{R}}U_s(e)e\widetilde{\mu}(\mathrm{d}e,\mathrm{d}s)-\int_t^T\int_{\mathbb{R}}U_s^n(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right)\\ &= \delta g(X_T)+\int_t^T\delta f(\Theta_s)\mathrm{d}s-\int_t^T\delta Z_s\mathrm{d}B_s\\ &\quad -\left(\int_t^T\int_{\mathbb{R}}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e\mathrm{d}s)+\int_t^T\int_{\mathbb{R}}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right),\end{aligned}$$

where we use the notations $\delta g(X_T):=g(X_T)-g(X_T^n)$, $\delta U_s(e):=U_s(e)-U_s^n(e)$, $\delta f(\Theta_s):=f(\Theta_s)-f(\Theta_s^n)$ and $\delta\Theta_s:=(s,\delta X_s,\delta Y_s,\delta Z_s,\delta\Gamma_s):=(s,X_s-X_s^n,Y_s-Y_s^n,Z_s-Z_s^n,\Gamma_s-\Gamma_s^n)$.

Step 1: We apply the Itô formula with the C^2-function $\eta(y)=|y|^p$ to the process δY_t. We use that

$$\frac{\partial\eta}{\partial y}(y)=py|y|^{p-2},\quad \frac{\partial^2\eta}{\partial y^2}(y)=p|y|^{p-2}+p(p-2)|y|^{p-2}=p(p-1)|y|^{p-2}.$$

Hence

$$\begin{aligned}|\delta Y_t|^p &= |\delta g(X_T)|^p+\int_t^T p\delta Y_s|\delta Y_s|^{p-2}\delta f(\Theta_s)\mathrm{d}s\\ &\quad -p\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s-\frac{1}{2}\int_t^T p(p-1)|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s\\ &\quad -\int_t^T\int_{\mathbb{R}}\Big(|\delta Y_{s-}+U_s(e)e|^p-|\delta Y_{s-}|^p-p\delta Y_{s-}|\delta Y_{s-}|^{p-2}U_s(e)e\Big)\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\\ &\quad -\int_t^T\int_{\mathbb{R}}\Big(|\delta Y_{s-}+\delta U_s(e)e|^p-|\delta Y_{s-}|^p-p\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta U_s(e)e\Big)\mu^n(\mathrm{d}e,\mathrm{d}s)\\ &\quad -p\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)-p\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s) \qquad (15)\\ &= |\delta g(X_T)|^p+\int_t^T p\delta Y_s|\delta Y_s|^{p-2}\delta f(\Theta_s)\mathrm{d}s\\ &\quad -p\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s-\frac{1}{2}\int_t^T p(p-1)|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s\\ &\quad -\int_t^T\int_{\mathbb{R}}\Big(|\delta Y_{s-}+U_s(e)e|^p-|\delta Y_{s-}|^p-p\delta Y_{s-}|\delta Y_{s-}|^{p-2}U_s(e)e\Big)\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\end{aligned}$$

$$
\begin{aligned}
&- \int_t^T \int_{\mathbb{R}} \Big(|\delta Y_{s-} + \delta U_s(e)e|^p - |\delta Y_{s-}|^p - p\delta Y_{s-} |\delta Y_{s-}|^{p-2} \delta U_s(e) e \Big) \nu^n(\mathrm{d}e)\mathrm{d}s \\
&- \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\bar{\mu}^n}(\mathrm{d}e, \mathrm{d}s) - \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + \delta U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s)
\end{aligned}
$$

We use a Taylor expansion of $\eta(x+y)$ around x.

$$
\begin{aligned}
\eta(x+y) - \eta(x) - \frac{\partial \eta}{\partial x}(x) y &= p(p-1) \int_0^1 (1-r)|x+ry|^{p-2}|y|^2 \mathrm{d}r \\
&\geq p(p-1)3^{1-p}|y|^2|x|^{p-2}.
\end{aligned}
$$

The inequality follows by Lemma A.4 of Yao (2010), an earlier version of Yao (2017). This implies

$$
\begin{aligned}
&- \int_t^T \int_{\mathbb{R}} \Big(|\delta Y_{s-} + U_s(e)e|^p - |\delta Y_{s-}|^p - p\delta Y_{s-} |\delta Y_{s-}|^{p-2} U_s(e) e \Big) \bar{\nu}^n(\mathrm{d}e)\mathrm{d}s \\
&- \int_t^T \int_{\mathbb{R}} \Big(|\delta Y_{s-} + \delta U_s(e)e|^p - |\delta Y_{s-}|^p - p\delta Y_{s-} |\delta Y_{s-}|^{p-2} \delta U_s(e) e \Big) \nu^n(\mathrm{d}e)\mathrm{d}s \\
\leq &- p(p-1)3^{1-p} \int_t^T \int_{\mathbb{R}} |\delta Y_{s-}|^{p-2} |U_s(e)|^2 e^2 \bar{\nu}^n(\mathrm{d}e)\mathrm{d}s \\
&- p(p-1)3^{1-p} \int_t^T \int_{\mathbb{R}} |\delta Y_{s-}|^{p-2} |\delta U_s(e)|^2 e^2 \nu^n(\mathrm{d}e)\mathrm{d}s \\
= &- p(p-1)3^{1-p} \int_t^T |\delta Y_s|^{p-2} ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}} \mathrm{d}s - p(p-1)3^{1-p} \int_t^T |\delta Y_s|^{p-2} ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}} \mathrm{d}s.
\end{aligned}
$$

Denote $\kappa_p := p(p-1)3^{1-p}$ and by $\kappa_p \leq \frac{p(p-1)}{2}$ we get that (15) becomes

$$
\begin{aligned}
&|\delta Y_t|^p + \kappa_p \int_t^T |\delta Y_{s-}|^{p-2} \delta Z_s^2 \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}} \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}} \mathrm{d}s \\
\leq &|\delta g(X_T)|^p + \int_t^T p \delta Y_s |\delta Y_s|^{p-2} \delta f(\Theta_s) \mathrm{d}s - p \int_t^T \delta Y_{s-} |\delta Y_{s-}|^{p-2} \delta Z_s \mathrm{d}B_s \\
&- \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\bar{\mu}^n}(\mathrm{d}e, \mathrm{d}s) - \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + \delta U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s).
\end{aligned}
$$

Now we apply the Lipschitz condition (4) of f to obtain

$$
\begin{aligned}
&|\delta Y_t|^p + \kappa_p \int_t^T |\delta Y_s|^{p-2} \delta Z_s^2 \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}} \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}} \mathrm{d}s \\
\leq &|\delta g(X_T)|^p + Kp \int_t^T \delta Y_s |\delta Y_s|^{p-2} |\delta X_s| \mathrm{d}s + Kp \int_t^T \delta Y_s |\delta Y_s|^{p-1} \mathrm{d}s + Kp \int_t^T \delta Y_s |\delta Y_s|^{p-2} |\delta Z_s| \mathrm{d}s \\
&+ Kp \int_t^T \delta Y_s |\delta Y_s|^{p-2} \int_{\mathbb{R}} \rho(e) |U_s(e)| e \bar{\nu}^n(\mathrm{d}e)\mathrm{d}s + Kp \int_t^T \delta Y_s |\delta Y_s|^{p-2} \int_{\mathbb{R}} \rho(e) |\delta U_s(e)| e \nu^n(\mathrm{d}e)\mathrm{d}s \\
&- p \int_t^T \delta Y_{s-} |\delta Y_{s-}|^{p-2} \delta Z_s \mathrm{d}B_s \\
&- \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\bar{\mu}^n}(\mathrm{d}e, \mathrm{d}s) - \int_t^T \int_{\mathbb{R}} \left(|\delta Y_{s-} + \delta U_s(e)e|^p - |\delta Y_{s-}|^p \right) \widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s).
\end{aligned}
$$

Next we use the inequality $xy \leq \alpha x^2 + y^2/\alpha$ for $\alpha > 0$, $x, y \geq 0$, the bound of ρ (recall Remark 1), and that $\delta Y_s \leq |\delta Y_s|$ to derive

$$
\begin{aligned}
&|\delta Y_t|^p + \kappa_p \int_t^T |\delta Y_s|^{p-2} \delta Z_s^2 \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}} \mathrm{d}s + \kappa_p \int_t^T |\delta Y_s|^{p-2} ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}} \mathrm{d}s \\
\leq &|\delta g(X_T)|^p + Kp(1+\alpha+\beta+\gamma+\varepsilon) \int_t^T |\delta Y_s|^p \mathrm{d}s + \frac{Kp}{\alpha} \int_t^T |\delta Y_s|^{p-2} |\delta X_s|^2 \mathrm{d}s + \frac{Kp}{\beta} \int_t^T |\delta Y_s|^{p-2} |\delta Z_s|^2 \mathrm{d}s
\end{aligned}
$$

$$+\frac{K^4p}{\gamma}\int_t^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s+\frac{K^4p}{\varepsilon}\int_t^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s-p\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s$$
$$-\int_t^T\int_{\mathbb{R}}\left(|\delta Y_{s-}+U_s(e)e|^p-|\delta Y_{s-}|^p\right)\widetilde{\tilde{\mu}^n}(\mathrm{d}e,\mathrm{d}s)-\int_t^T\int_{\mathbb{R}}\left(|\delta Y_{s-}+\delta U_s(e)e|^p-|\delta Y_{s-}|^p\right)\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s).$$

We make use of the Lipschitz condition (3) on g and Young's inequality for $|\delta Y_s|^{p-2}|\delta X_s|^2$ to get

$$\begin{aligned}
&|\delta Y_t|^p+\kappa_p\int_t^T|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\\
\le\ & K^p|\delta X_T|^p+Kp(1+\alpha+\beta+\gamma+\varepsilon+\frac{p-2}{\alpha p}+\frac{p-2}{\gamma p})\int_t^T|\delta Y_s|^p\mathrm{d}s\\
&+\frac{2K}{\alpha}\int_t^T|\delta X_s|^p\mathrm{d}s+\frac{Kp}{\beta}\int_t^T|\delta Y_s|^{p-2}|\delta Z_s|^2\mathrm{d}s+\frac{K^4p}{\gamma}\int_t^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s\\
&+\frac{K^4p}{\varepsilon}\int_t^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s-p\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s\\
&-\int_t^T\int_{\mathbb{R}}\left(|\delta Y_{s-}+U_s(e)e|^p-|\delta Y_{s-}|^p\right)\widetilde{\tilde{\mu}^n}(\mathrm{d}e,\mathrm{d}s)\\
&-\int_t^T\int_{\mathbb{R}}\left(|\delta Y_{s-}+\delta U_s(e)e|^p-|\delta Y_{s-}|^p\right)\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s),
\end{aligned}\tag{16}$$

where we choose the constants $\alpha,\beta,\gamma,\varepsilon>0$ arbitrarily such that $\frac{Kp}{\beta}<\kappa_p$, $\frac{K^4p}{\gamma}<\kappa_p$ and $\frac{K^4p}{\varepsilon}<\kappa_p$.

We take expectations of (16) to obtain

$$\begin{aligned}
&\mathbb{E}\left[|\delta Y_t|^p+\kappa_p\int_t^T|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right]\\
\le& C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}+C_p\mathbb{E}\left[\int_t^T|\delta Y_s|^p\mathrm{d}s\right]\right).
\end{aligned}\tag{17}$$

Then Gronwall's lemma implies

$$\mathbb{E}\left[|\delta Y_t|^p\right]\le C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right).\tag{18}$$

We substitute (18) into (17) to get

$$\mathbb{E}\left[\int_0^T|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s+\int_0^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s+\int_0^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right]\le C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right),$$

which implies that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T|\delta Y_s|^p\mathrm{d}s+\int_0^T|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s+\int_0^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\tilde{\nu}^n}}\mathrm{d}s+\int_0^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right]\\
&\le C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right),
\end{aligned}$$

Now we apply the Burkholder-Davis-Gundy inequality and Young's inequality to the martingales in (15). First,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s\right|\right]\le C_p\mathbb{E}\left[\left(\int_0^T|\delta Y_s|^{2p-2}|\delta Z_s|^2\mathrm{d}s\right)^{1/2}\right]\\
&\le\frac{1}{4p}\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^p\right]+pC_p^2\mathbb{E}\left[\int_0^T|\delta Y_s|^{p-2}|\delta Z_s|^2\mathrm{d}s\right].
\end{aligned}\tag{19}$$

Second,

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)\right|\right]\le C_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}|\delta Y_{s-}|^{2p-2}U_s(e)^2e^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\right)^{1/2}\right]$$

$$\le\frac{1}{4p}\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^p\right]+pC_p^2\mathbb{E}\left[\int_0^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s\right].\tag{20}$$

Third,

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|\right]\le C_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}|\delta Y_{s-}|^{2p-2}\delta U_s(e)^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{1/2}\right]$$

$$\le\frac{1}{4p}\mathbb{E}\left[\sup_{0\le t\le T}|\delta Y_t|^p\right]+pC_p^2\mathbb{E}\left[\int_0^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right].\tag{21}$$

We return to (15) and use the convexity of η to get

$$\begin{aligned}|\delta Y_t|^p\le{}&|\delta g(X_T)|^p+\int_t^T p\delta Y_s|\delta Y_s|^{p-2}\delta f(\Theta_s)\mathrm{d}s-p\int_t^T\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta Z_s\mathrm{d}B_s\\&-p\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)-p\int_t^T\int_{\mathbb{R}}\delta Y_{s-}|\delta Y_{s-}|^{p-2}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\\&+\kappa_p\int_t^T|\delta Y_s|^{p-2}\delta Z_s^2\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s+\kappa_p\int_t^T|\delta Y_s|^{p-2}||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s.\end{aligned}$$

When we now follow the previous lines in the proof to bound the $\delta f(\Theta_s)$ integral and use the bounds (19), (20) and (21) by the Burkholder-Davis-Gundy inequality we finally derive

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t|^p\right]\le C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right).$$

Step 2: In the second step we prove that

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^T|\delta Z_s|^2\mathrm{d}s\right)^{p/2}+\left(\int_0^T\int_{\mathbb{R}}|U_s(e)e|^2\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}+\left(\int_0^T\int_{\mathbb{R}}|\delta U_s(e)e|^2\nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right]\\&\le C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right).\end{aligned}$$

Again we apply Itô's formula, this time to $|\delta Y_t|^2$:

$$\begin{aligned}&|\delta Y_0|^2+\int_0^T|\delta Z_s|^2\mathrm{d}s+\int_0^T\int_{\mathbb{R}}|U_s(e)e|^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)+\int_0^T\int_{\mathbb{R}}|\delta U_s(e)e|^2\mu^n(\mathrm{d}e,\mathrm{d}s)\\&=|\delta Y_T|^2+2\int_0^T\delta Y_s\delta f(\Theta_s)\mathrm{d}s-2\int_0^T\delta Y_s\delta Z_s\mathrm{d}B_s\\&\quad-2\int_0^T\int_{\mathbb{R}}\delta Y_{s-}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)-2\int_0^T\int_{\mathbb{R}}\delta Y_{s-}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s).\end{aligned}$$

Next we use the Lipschitz condition (4)

$$\begin{aligned}&\int_0^T|\delta Z_s|^2\mathrm{d}s+\int_0^T\int_{\mathbb{R}}|U_s(e)e|^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)+\int_0^T\int_{\mathbb{R}}|\delta U_s(e)e|^2\mu^n(\mathrm{d}e,\mathrm{d}s)\\&\le|\delta Y_*|^2+2K\int_0^T|\delta Y_s|^2\mathrm{d}s+2K\int_0^T\delta Y_s|\delta X_s|\mathrm{d}s+2K\int_0^T\delta Y_s|\delta Z_s|\mathrm{d}s\end{aligned}$$

$$+ 2K \int_0^T \delta Y_s \int_{\mathbb{R}} \rho(e)|U_s(e)|e\bar{\nu}^n(\mathrm{d}e)\mathrm{d}s + 2K \int_0^T \delta Y_s \int_{\mathbb{R}} \rho(e)|\delta U_s(e)|e\nu^n(\mathrm{d}e)\mathrm{d}s$$
$$- 2\int_0^T \delta Y_s \delta Z_s \mathrm{d}B_s - 2\int_0^T \int_{\mathbb{R}} \delta Y_{s-} U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e, \mathrm{d}s) - 2\int_0^T \int_{\mathbb{R}} \delta Y_{s-}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s),$$

where $\delta Y_* := \sup_{0\le t\le T} |\delta Y_t|$.
We again use the inequality $xy \le \alpha x^2 + y^2/\alpha$ for $\alpha > 0$, $x, y \ge 0$ to get the bound

$$\begin{aligned}
&\int_0^T |\delta Z_s|^2 \mathrm{d}s + \int_0^T \int_{\mathbb{R}} |U_s(e)e|^2 \bar{\mu}^n(\mathrm{d}e, \mathrm{d}s) + \int_0^T \int_{\mathbb{R}} |\delta U_s(e)e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s) \\
&\le |\delta Y_*|^2 + 2K(1+\alpha+\beta+\gamma+\varepsilon)\int_0^T |\delta Y_s|^2\mathrm{d}s + \frac{2K}{\alpha}\int_0^T |\delta X_s|^2 \mathrm{d}s + \frac{2K}{\beta}\int_0^T |\delta Z_s|^2\mathrm{d}s \\
&+ \frac{2K^4}{\gamma}\int_0^T ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s + \frac{2K^4}{\varepsilon}\int_0^T ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s - 2\int_0^T \delta Y_s \delta Z_s \mathrm{d}B_s \\
&- 2\int_0^T \int_{\mathbb{R}} \delta Y_{s-} U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e, \mathrm{d}s) - 2\int_0^T \int_{\mathbb{R}} \delta Y_{s-}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s).
\end{aligned} \tag{22}$$

Next we take powers of (22) (and use Jensen's inequality)

$$\begin{aligned}
&\left(\int_0^T |\delta Z_s|^2 \mathrm{d}s\right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} |U_s(e)e|^2 \bar{\mu}^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} |\delta U_s(e)e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2} \\
&\le C_p|\delta Y_*|^p + C_p\left(2K(1+\alpha+\beta+\gamma+\varepsilon)\right)^{p/2}\int_0^T |\delta Y_s|^p \mathrm{d}s \\
&\quad + C_p\left(\frac{2K}{\alpha}\right)^{p/2}\int_0^T |\delta X_s|^p\mathrm{d}s + C_p\left(\frac{2K}{\beta}\right)^{p/2}\left(\int_0^T |\delta Z_s|^2\mathrm{d}s\right)^{p/2} \\
&\quad + C_p\left(\frac{2K^4}{\gamma}\right)^{p/2}\left(\int_0^T ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s\right)^{p/2} + C_p\left(\frac{2K^4}{\varepsilon}\right)^{p/2}\left(\int_0^T ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right)^{p/2} \\
&\quad + C_p\left(\left|\int_0^T \delta Y_s\delta Z_s\mathrm{d}B_s\right|^{p/2} + \left|\int_0^T\int_{\mathbb{R}} \delta Y_{s-}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)\right|^{p/2} + \left|\int_0^T\int_{\mathbb{R}} \delta Y_{s-}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|^{p/2}\right).
\end{aligned} \tag{23}$$

Because $p/2 \ge 1$, we can apply the Burkholder-Davis-Gundy inequality and Young's inequality to get

$$C_p\mathbb{E}\left[\left|\int_0^T \delta Y_s\delta Z_s\mathrm{d}B_s\right|^{p/2}\right] \le c_p\mathbb{E}\left[\left(\int_0^T |\delta Y_s|^2|\delta Z_s|^2\mathrm{d}s\right)^{p/4}\right] \le \frac{c_p^2}{4}\mathbb{E}\left[|\delta Y_*|^p\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_0^T |\delta Z_s|^2\mathrm{d}s\right)^{p/2}\right],$$

$$\begin{aligned}
&C_p\mathbb{E}\left[\left|\int_0^T\int_{\mathbb{R}} \delta Y_{s-}U_s(e)e\widetilde{\bar{\mu}^n}(\mathrm{d}e,\mathrm{d}s)\right|^{p/2}\right] \le c_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}} |\delta Y_{s-}|^2|U_s(e)|^2e^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/4}\right] \\
&\quad \le \frac{c_p^2}{4}\mathbb{E}\left[|\delta Y_*|^p\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_0^T ||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right], \\
&C_p\mathbb{E}\left[\left|\int_0^T\int_{\mathbb{R}} \delta Y_{s-}\delta U_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|^{p/2}\right] \le c_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}} |\delta Y_{s-}|^2|\delta U_s(e)|^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/4}\right] \\
&\quad \le \frac{c_p^2}{4}\mathbb{E}\left[|\delta Y_*|^p\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_0^T ||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right],
\end{aligned}$$

for some constant c_p. Using this for the expectation of (23), we see

$$\frac{1}{2}\mathbb{E}\left[\left(\int_0^T |\delta Z_s|^2\mathrm{d}s\right)^{p/2}\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}} |U_s(e)e|^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right]$$

$$+\frac{1}{2}\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}|\delta U_s(e)e|^2\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right]$$

$$\leq C_{p,K,T,\alpha,\beta,\gamma,\varepsilon}\mathbb{E}\left[|\delta Y_*|^p\right]+C_{p,K,T,\alpha}\mathbb{E}\left[\sup_{0\leq t\leq T}|\delta X_t|^p\right]+C_p\left(\frac{2K}{\beta}\right)^{p/2}\mathbb{E}\left[\left(\int_0^T|\delta Z_s|^2\mathrm{d}s\right)^{p/2}\right]$$

$$+C_p\left(\frac{2K^4}{\gamma}\right)^{p/2}\mathbb{E}\left[\left(\int_0^T||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s\right)^{p/2}\right]+C_p\left(\frac{2K^4}{\varepsilon}\right)^{p/2}\mathbb{E}\left[\left(\int_0^T||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right)^{p/2}\right].$$

As in Kruse & Popier (2016) (see also Dzhaparidze & Valkeila 1990) we use the bounds

$$\mathbb{E}\left[\left(\int_0^T||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s\right)^{p/2}\right]\leq d_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}|U_s(e)e|^2\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right],$$

$$\mathbb{E}\left[\left(\int_0^T||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right)^{p/2}\right]\leq d_p\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}}|\delta U_s(e)e|^2\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{p/2}\right],$$

for some constant $d_p>0$.
All in all we can choose the constants α, β, γ and ε (only depending on p) such that

$$\mathbb{E}\left[\left(\int_0^T|\delta Z_s|^2\mathrm{d}s\right)^{p/2}\right]+\mathbb{E}\left[\left(\int_0^T||U_s||^2_{\mathbb{L}^2_{\bar{\nu}^n}}\mathrm{d}s\right)^{p/2}\right]+\mathbb{E}\left[\left(\int_0^T||\delta U_s||^2_{\mathbb{L}^2_{\nu^n}}\mathrm{d}s\right)^{p/2}\right]$$

$$\leq C_p\mathbb{E}\left[|\delta Y_*|^p\right]+C_p\mathbb{E}\left[\sup_{0\leq t\leq T}|\delta X_t|^p\right]$$

$$\leq C_p\left(\sigma^p(n)+\sigma^2(n)^{p/2}\right),$$

by Step 1 and Proposition 1. □

Remark 4. So far we have approximated the Lévy process by discarding small jumps. We now discuss an alternative where we approximate the small jumps by a scaled Brownian motion and discuss the approximation error. The idea of approximating small jumps goes back to Asmussen & Rosiński (2001), for a gentle introduction see Yuan & Kawai (2021). Note that this approach is not generally valid for every Lévy process and requires additional assumptions given below. Consider an appropriate method $H(r,V)$ of the series representation of the Lévy process L (e.g., the inverse Lévy measure method). In this case we set the truncation parameter n sufficiently large such that only jumps of magnitude less than ε are discarded. Then, the variance of the discarded jumps is given by

$$\sigma^2(\varepsilon):=\int_{|e|<\varepsilon}|e|^2\nu(\mathrm{d}e).$$

Denote by L_t^ε the large jumps that are simulated and by $L_{\varepsilon,t}$ the small jumps that are discarded. Clearly, $L_t=L_t^\varepsilon+L_{\varepsilon,t}$. We now aim to replace $L_{\varepsilon,t}$ by a normal random variable with the proper variance. In order to do so, let W be a Brownian motion which is independent of L and the Brownian motion B. Asmussen & Rosiński (2001) showed that $\frac{L_{\varepsilon,t}}{\sigma(\varepsilon)}\xrightarrow{d}W_t$ as $\varepsilon\to0$ for each $t\in[0,T]$ if and only if for each $c>0$

$$\sigma(c\sigma(\varepsilon)\wedge\varepsilon)\sim\sigma(\varepsilon),\tag{24}$$

as $\varepsilon\to0$. If (24) holds, we replace the small jumps by the random variable $\sigma(\varepsilon)W_t$ at time t. Asmussen & Rosiński (2001) also showed that $\frac{\sigma(\varepsilon)}{\varepsilon}\to+\infty$ as $\varepsilon\to0$ implies (24). Moreover, if the Lévy measure ν does not have atoms in the neighborhoods around the origin then both conditions are equivalent.

We turn back to the approximation of the FBSDEs. We slightly change the notation of the approximate SDEs now given in dependence of ε instead of n to highlight the dependence on the extra assumptions and to distinguish between the version without Gaussian approximation of small jumps. More precisely, we approximate (7) and (8) by

$$X_t^\varepsilon=X_0+\int_0^t b(s,X_s^\varepsilon)\mathrm{d}s+\int_0^t a(s,X_s^\varepsilon)\mathrm{d}B_s+\int_0^t\int_{|e|>\varepsilon}h(s,X_{s-}^\varepsilon)e\widetilde{\mu}(\mathrm{d}e,\mathrm{d}s)+\sigma(\varepsilon)\int_0^t h(s,X_s^\varepsilon)\mathrm{d}W_s,$$

$$Y_t^\varepsilon = g(X_T^\varepsilon) + \int_t^T f(s, X_s^\varepsilon, Y_s^\varepsilon, Z_s^\varepsilon, \Gamma_s^\varepsilon)\mathrm{d}s - \int_t^T Z_s^\varepsilon \mathrm{d}B_s - \int_t^T \int_{|e|>\varepsilon} U_s^\varepsilon(e) e \widetilde{\mu}(\mathrm{d}e, \mathrm{d}s) - \int_t^T \Psi_s^\varepsilon \mathrm{d}W_s, \quad (25)$$

with $\Lambda_s^\varepsilon := \int_{|e|>\varepsilon} \rho(e) U_s^\varepsilon(e) e \nu(\mathrm{d}e)$. We again want to control the approximation error for the forward and the backward SDEs. For the forward SDE we observe that in (14) we have to add the term

$$\mathbb{E}\left[\left(\int_0^t \sigma^2(\varepsilon) h(s, X_s^\varepsilon)^2 \mathrm{d}s\right)^{p/2}\right], \quad (26)$$

and change the remainder of the terms in (14) correspondingly. Then the same Lipschitz argument as above implies that (26) is bounded by

$$C_p \mathbb{E}\left[\left(\int_0^t \sigma^2(\varepsilon)(1 + |X_s^\varepsilon|^2)\mathrm{d}s\right)^{p/2}\right]. \quad (27)$$

We recall the important bound (12) for X and note that it is possible to prove that the same bound (with a different constant) also holds for X^ε, i.e.,

$$\mathbb{E}\left[\sup_{0\le t\le T} |X_t^\varepsilon|^p\right] < C_p(1 + |X_0|^p),$$

where C_p is independent of ε. This implies that (27) is bounded by $C_p \sigma^2(\varepsilon)^{p/2}$ and hence

$$\mathbb{E}\left[\sup_{0\le t\le T} |X_t - X_t^\varepsilon|^p\right] \le C_p\left(\sigma^p(\varepsilon) + \sigma^2(\varepsilon)^{p/2}\right).$$

For the approximation error of the backward SDE we have to re-define the notion of the error to include the new term in (25). More precisely, we set

$$Err_\varepsilon(Y, Z, U)^p := \mathbb{E}\left[\sup_{0\le t\le T} |Y_t - Y_t^\varepsilon|^p + \left(\int_0^T |Z_s - Z_s^\varepsilon|^2 \mathrm{d}s\right)^{p/2} + \left(\int_0^T \int_{|e|>\varepsilon} |U_s(e) - U_s^\varepsilon(e)|^2 e^2 \nu(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right.$$

$$\left. + \sup_{0\le t\le T}\left|\int_t^T \int_{|e|\le\varepsilon} U_s(\varepsilon)^2 e^2 \nu(\mathrm{d}e)\mathrm{d}s - \int_t^T \Psi_s^\varepsilon \mathrm{d}W_s\right|^p\right].$$

The modification of the proof of Theorem 1 is straightforward: we only have to add the terms

$$p\int_t^T (Y_t - Y_t^\varepsilon)|Y_t - Y_t^\varepsilon|^{p-2}\Psi_s^\varepsilon \mathrm{d}W_s + \frac{p(p-1)}{2}\int_t^T |Y_t - Y_t^\varepsilon|^{p-2}(\Psi_s^\varepsilon)^2 \mathrm{d}s$$

into Itô's formula and proceed analogously to the proof of Theorem 1. It turns out that the approximation error is bounded as above by

$$Err_\varepsilon(Y, Z, U)^p \le C_p\left(\sigma^p(n) + \sigma^2(n)^{p/2}\right).$$

To conclude this remark, the error bounds qualitatively do not change with the additional Gaussian approximation of small jumps as the asymptotic behavior coincides. Of course, the error with Gaussian approximation is smaller than without (but only the constant C_p is smaller).

4 Error of the discretization of the FBSDE

In this section we discretize the approximated FBSDE (X^n, Y^n, Z^n, U^n) and derive error bounds. We use a forward-backward Euler scheme for simulation. First we define the regular grid $\pi :=$

$\left\{t_k := \frac{kT}{N},\ k = 0, \ldots, N\right\}$ on $[0,T]$. We do not discuss the discretization of the original FBSDE because in practice they cannot be simulated and the proofs of this section rely on $\nu^n(\mathbb{R}) < \infty$. Starting with the forward Euler scheme for X^n, we define

$$\begin{cases} X_0^{n,\pi} & := X_0 \\ X_{t_{k+1}}^{n,\pi} & := X_{t_k}^{n,\pi} + \frac{T}{N} b(t_k, X_{t_k}^{n,\pi}) + a(t_k, X_{t_k}^{n,\pi})\Delta B_{k+1} + \int_{\mathbb{R}} h(t_k, X_{t_k}^{n,\pi}) e\widetilde{\mu^n}(\mathrm{d}e, (t_k, t_{k+1}]), \end{cases} \tag{28}$$

where $\Delta B_{k+1} := B_{t_{k+1}} - B_{t_k}$ are normal random variables.

It is well-known that under Assumption 1.(i) the Euler scheme (28) of the forward SDE has the discretization error

$$\max_{k<N} \mathbb{E}\left[\sup_{t\in[t_k,t_{k+1}]} |X_t^n - X_t^{n,\pi}|^p\right] \le C_p n^{-p/2}, \tag{29}$$

for all $p \ge 1$, see, e.g., Aazizi (2013).

Next we introduce the backward implicit scheme to approximate (Y^n, Z^n, Γ^n). We follow Bouchard & Elie (2008) and Elie (2006) and define

$$\begin{cases} \bar{Z}_t^{n,\pi} & := \frac{N}{T}\mathbb{E}\left[\bar{Y}_{t_{k+1}}^{n,\pi}\Delta B_{k+1}|\mathcal{F}_{t_k}\right] \\ \bar{\Gamma}_t^{n,\pi} & := \frac{N}{T}\mathbb{E}\left[\bar{Y}_{t_{k+1}}^{n,\pi}\int_{\mathbb{R}}\rho(e)e\widetilde{\mu^n}(\mathrm{d}e,(t_k,t_{k+1}])|\mathcal{F}_{t_k}\right] \\ \bar{Y}_t^{n,\pi} & := \mathbb{E}\left[\bar{Y}_{t_{k+1}}^{n,\pi}|\mathcal{F}_{t_k}\right] + \frac{T}{N} f\left(t_k, X_{t_k}^{n,\pi}, \bar{Y}_{t_k}^{n,\pi}, \bar{Z}_{t_k}^{n,\pi}, \bar{\Gamma}_{t_k}^{n,\pi}\right), \end{cases} \tag{30}$$

on each interval $[t_k, t_{k+1})$, where $Y_{t_N}^{n,\pi} := g(X_{t_N}^{n,\pi})$. If f depends on Y^n, the last step of (30) requires a fixed point procedure. However, since f is Hölder continuous in t and Lipschitz continuous in the other variables and because f is multiplied by $1/N$ the approximation error can be neglected for large values of N.

Given the backward scheme (30), we will analyze the discretization error

$$Err_\pi(Y^n, Z^n, U^n) := \left(\sup_{0\le t\le T}\mathbb{E}\left[|Y_t^n - \bar{Y}_t^{n,\pi}|^p\right] + ||Z^n - \bar{Z}^{n,\pi}||_{\mathbb{H}^p}^p + ||\Gamma^n - \bar{\Gamma}^{n,\pi}||_{\mathbb{H}^p}^p\right)^{1/p}$$

and we will show that it converges to zero with order $N^{-1/2}$.

In the following we discuss some related processes which will be needed throughout the proofs. By the representation theorem, see Tang & Li (1994), there exist two processes $Z^{n,\pi} \in \mathbb{H}^p$ and $U^{n,\pi} \in \mathbb{L}_{\mu^n}^p$ such that

$$\bar{Y}_{t_{k+1}}^{n,\pi} - \mathbb{E}\left[\bar{Y}_{t_{k+1}}^{n,\pi}|\mathcal{F}_{t_k}\right] = \int_{t_k}^{t_{k+1}} Z_s^{n,\pi}\mathrm{d}B_s + \int_{t_k}^{t_{k+1}}\int_{\mathbb{R}} U_s^{n,\pi}(e)e\bar{\mu}^n(\mathrm{d}e,\mathrm{d}s).$$

Observe that $\bar{Z}_t^{n,\pi}$ and $\bar{\Gamma}_t^{n,\pi}$ in (30) satisfy

$$\bar{Z}_{t_k}^{n,\pi} = \frac{N}{T}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z_s^{n,\pi}\mathrm{d}s\middle|\mathcal{F}_{t_k}\right], \tag{31}$$

$$\bar{\Gamma}_{t_k}^{n,\pi} = \frac{N}{T}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \Gamma_s^{n,\pi}\mathrm{d}s\middle|\mathcal{F}_{t_k}\right],$$

and thus coincide with the best $\mathbb{H}^2_{[t_k,t_{k+1}]}$-approximations of the processes $(Z_t^{n,\pi})$ and $(\Gamma_t^{n,\pi}) := \left(\int_{\mathbb{R}}\rho(e)U_t^{n,\pi}(e)e\nu^n(\mathrm{d}e)\right)$ on $[t_k, t_{k+1})$ by $\mathcal{F}_{t_k}$-measurable random variables (viewed as constant processes on $[t_k, t_{k+1})$), i.e.,

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |Z_s^{n,\pi} - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] = \inf_{Z_k\in L^2(\Omega,\mathcal{F}_{t_k})}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_s^{n,\pi} - Z_k|^2\mathrm{d}s\right],$$

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |\Gamma_s^{n,\pi} - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] = \inf_{\Gamma_k\in L^2(\Omega,\mathcal{F}_{t_k})}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|\Gamma_s^{n,\pi} - \Gamma_k|^2\mathrm{d}s\right].$$

Thus, it holds that

$$\bar{Y}^{n,\pi}_{t_k} = \bar{Y}^{n,\pi}_{t_{k+1}} + \frac{T}{N} f(t_k, X^{n,\pi}_{t_k}, \bar{Y}^{n,\pi}_{t_k}, \bar{Z}^{n,\pi}_{t_k}, \bar{\Gamma}^{n,\pi}_{t_k}) - \int_{t_k}^{t_{k+1}} Z^{n,\pi}_s \mathrm{d}B_s - \int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} U^{n,\pi}_s(e) e \bar{\mu}^n(\mathrm{d}e, \mathrm{d}s).$$

We define the process $Y^{n,\pi}$

$$Y^{n,\pi}_t := \bar{Y}^{n,\pi}_{t_k} - (t - t_k) f(t_k, X^{n,\pi}_{t_k}, \bar{Y}^{n,\pi}_{t_k}, \bar{Z}^{n,\pi}_{t_k}, \bar{\Gamma}^{n,\pi}_{t_k}) + \int_{t_k}^{t} Z^{n,\pi}_s \mathrm{d}B_s + \int_{t_k}^{t} \int_{\mathbb{R}} U^{n,\pi}_s(e) e \bar{\mu}^n(\mathrm{d}e, \mathrm{d}s)$$

on $[t_k, t_{k+1})$ and obtain that

$$\frac{N}{T}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Y^{n,\pi}_s \mathrm{d}s | \mathcal{F}_{t_k}\right] = \mathbb{E}\left[\bar{Y}^{n,\pi}_{t_{k+1}} | \mathcal{F}_{t_k}\right] + \frac{T}{N} f(t_k, X^{n,\pi}_{t_k}, \bar{Y}^{n,\pi}_{t_k}, \bar{Z}^{n,\pi}_{t_k}, \bar{\Gamma}^{n,\pi}_{t_k}) = Y^{n,\pi}_{t_k} = \bar{Y}^{n,\pi}_{t_k}. \tag{32}$$

Thus $\bar{Y}^{n,\pi}_{t_k}$ is the best approximation of $Y^{n,\pi}$ on $[t_k, t_{k+1})$ by $\mathcal{F}_{t_k}$-measurable random variables (viewed as constant processes on $[t_k, t_{k+1})$), which explains the notation $\bar{Y}^{n,\pi}$, consistent with the definition of $\bar{Z}^{n,\pi}$ and $\bar{\Gamma}^{n,\pi}$.

Furthermore, we need to define the processes $(\bar{Z}^n, \bar{\Gamma}^n)$ on each interval $[t_k, t_{k+1})$ by

$$\bar{Z}^n_t := \frac{T}{N}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z^n_s \mathrm{d}s \middle| \mathcal{F}_{t_k}\right], \tag{33}$$

$$\bar{\Gamma}^n_t := \frac{T}{N}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \Gamma^n_s \mathrm{d}s \middle| \mathcal{F}_{t_k}\right].$$

Remark 5. $\bar{Z}^n_{t_k}$ and $\bar{\Gamma}^n_{t_k}$ are the counterparts $\bar{Z}^{n,\pi}_{t_k}$ and $\bar{\Gamma}^{n,\pi}_{t_k}$ for the original backward SDE. They can be interpreted as the best $\mathbb{H}^2_{[t_k,t_{k+1}]}$-approximations of $(Z^n_t)_{t_k \le t < t_{k+1}}$ and $(\Gamma^n_t)_{t_k \le t < t_{k+1}}$ by an $\mathcal{F}_{t_k}$-measurable random variables (viewed as constant processes on $[t_k, t_{k+1})$), i.e.,

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |Z^n_s - \bar{Z}^n_{t_k}|^2 \mathrm{d}s\right] = \inf_{Z_k \in L^2(\Omega, \mathcal{F}_{t_k})} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} |Z^n_s - Z_k|^2 \mathrm{d}s\right],$$

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |\Gamma^n_s - \bar{\Gamma}^n_{t_k}|^2 \mathrm{d}s\right] = \inf_{\Gamma_k \in L^2(\Omega, \mathcal{F}_{t_k})} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} |\Gamma^n_s - \Gamma_k|^2 \mathrm{d}s\right].$$

We now state our second main theorem, which gives a bound for the discretization error.

Theorem 2. *Under Assumptions 1 and 2, the discretization error for the backward SDE is bounded by*

$$Err_\pi(Y^n, Z^n, U^n) \le C_p N^{-1/2}. \tag{34}$$

Proof. The proof is an L^p extension of the proofs of Bouchard & Elie (2008), Elie (2006) and Bouchard & Touzi (2004). For the sake of brevity we set $\delta^n Y_t := Y^n_t - Y^{n,\pi}_t$, $\delta^n Z_t := Z^n_t - Z^{n,\pi}_t$, $\delta^n U_t(e) := U^n_t(e) - U^{n,\pi}_t(e)$, $\delta^n \Gamma_t := \Gamma^n_t - \Gamma^{n,\pi}_t$ and $\delta^n f(\Theta_t) := f(t, X^n_t, Y^n_t, Z^n_t, \Gamma^n_t) - f(t_k, X^{n,\pi}_{t_k}, \bar{Y}^{n,\pi}_{t_k}, \bar{Z}^{n,\pi}_{t_k}, \bar{\Gamma}^{n,\pi}_{t_k})$. Note that $\bar{Y}^{n,\pi}_{t_k} = Y^{n,\pi}_{t_k}$ by (32) which we will use repeatedly.

The proof is divided in four steps. Before turning to the first step, we discuss some bounds which we will need throughout:

$$\mathbb{E}\left[|X^n_s - X^{n,\pi}_{t_k}|\right] \le C_p N^{-p/2}, \tag{35}$$

by (29). Moreover,

$$\mathbb{E}\left[|Y^n_s - \bar{Y}^{n,\pi}_{t_k}|^p\right] \le p\left(\mathbb{E}\left[|Y^n_s - Y^n_{t_k}|^p\right] + \mathbb{E}\left[|\delta^n Y_{t_k}|^p\right]\right)$$

and

$$\mathbb{E}\left[|Z^n_s - \bar{Z}^{n,\pi}_{t_k}|^p\right] \le p\left(\mathbb{E}\left[|Z^n_s - \bar{Z}^n_{t_k}|^p\right] + \mathbb{E}\left[|\bar{Z}^n_{t_k} - \bar{Z}^{n,\pi}_{t_k}|^p\right]\right)$$

$$
\begin{aligned}
&= p\left(\mathbb{E}\left[|Z^n_s - \bar{Z}^n_{t_k}|^p\right] + \mathbb{E}\left[\left|\frac{N}{T}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \delta^n Z_s \mathrm{d}s \middle| \mathcal{F}_{t_k}\right]\right|^p\right]\right)\\
&\le C_p\left(\mathbb{E}\left[|Z^n_s - \bar{Z}^n_{t_k}|^p\right] + \mathbb{E}\left[\mathbb{E}\left[\frac{N}{T}\int_{t_k}^{t_{k+1}} |\delta^n Z_s|^2 \mathrm{d}s \middle| \mathcal{F}_{t_k}\right]^{p/2}\right]\right)\\
&\le C_p\left(\mathbb{E}\left[|Z^n_s - \bar{Z}^n_{t_k}|^p\right] + N^{p/2}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n Z_s|^2 \mathrm{d}s\right)^{p/2}\right]\right)\\
&= C_p\left(\mathbb{E}\left[|Z^n_s - \bar{Z}^n_{t_k}|^p\right] + N^{p/2}||\delta^n Z||^p_{\mathbb{H}^p_{[t_k,t_{k+1}]}}\right). \qquad (36)
\end{aligned}
$$

The second equality follows by (31) and (33) and the third and fourth inequalities by Jensen's inequality. Analogously, using the bound on ρ, we can prove

$$
\mathbb{E}\left[|\Gamma^n_s - \bar{\Gamma}^{n,\pi}_{t_k}|^p\right] \le C_p\left(\mathbb{E}\left[|\Gamma^n_s - \bar{\Gamma}^n_{t_k}|^p\right] + N^{p/2}||\delta^n U||^p_{\mathbb{L}^p_{\mu^n,[t_k,t_{k+1}]}}\right). \qquad (37)
$$

Step 1: We apply Itô's formula to $|\delta^n Y_t|^p$ for $t \in [t_k, t_{k+1})$,

$$
\begin{aligned}
\mathbb{E}\left[|\delta^n Y_t|^p\right] &= \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + p\mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s\ |\delta^n Y_s|^{p-2}\delta^n f(\Theta_s)\mathrm{d}s\right]\\
&- \frac{p(p-1)}{2}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right]\\
&- \mathbb{E}\left[\int_t^{t_{k+1}}\int_{\mathbb{R}}\left(|\delta^n Y_{s-} + \delta^n U_s(e)e|^p - |\delta^n Y_{s-}|^p - p\delta^n Y_{s-}\ |\delta^n Y_{s-}|^{p-2}\delta^n U_s(e)e\right)\nu^n(\mathrm{d}e)\mathrm{d}s\right].
\end{aligned}
$$

As in the proof of Theorem 1 we use

$$
\begin{aligned}
&- \mathbb{E}\left[\int_t^{t_{k+1}}\int_{\mathbb{R}}\left(|\delta^n Y_{s-} + \delta^n U_s(e)e|^p - |\delta^n Y_{s-}|^p - p\delta^n Y_{s-}\ |\delta^n Y_{s-}|^{p-2}\delta^n U_s(e)e\right)\nu^n(\mathrm{d}e)\mathrm{d}s\right]\\
&\le\ - \kappa_p\mathbb{E}\left[\int_t^{t_{k+1}}\int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2\nu^n(\mathrm{d}e)\mathrm{d}s\right],
\end{aligned}
$$

with $\kappa_p = p(p-1)3^{1-p}$, to derive

$$
\begin{aligned}
&\mathbb{E}\left[|\delta^n Y_t|^p\right] + \kappa_p\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right] + \kappa_p\mathbb{E}\left[\int_t^{t_{k+1}}\int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2\nu^n(\mathrm{d}e)\mathrm{d}s\right]\\
&\le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + p\mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s\ |\delta^n Y_s|^{p-2}\delta^n f(\Theta_s)\mathrm{d}s\right].
\end{aligned}
$$

We use the Lipschitz condition (4) to get

$$
\begin{aligned}
&\mathbb{E}\left[|\delta^n Y_t|^p\right] + \kappa_p\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right] + \kappa_p\mathbb{E}\left[\int_t^{t_{k+1}}\int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2\nu^n(\mathrm{d}e)\mathrm{d}s\right]\\
&\le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right]\\
&+ \mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s\ |\delta^n Y_s|^{p-2}\left(N^{-1/2} + |X^n_s - X^{n,\pi}_{t_k}| + |Y^n_s - \bar{Y}^{n,\pi}_{t_k}| + |Z^n_s - \bar{Z}^{n,\pi}_{t_k}| + |\Gamma^n_s - \bar{\Gamma}^{n,\pi}_{t_k}|\right)\mathrm{d}s\right].
\end{aligned}
$$

We rewrite this inequality to have

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} \int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2 \nu^n(\mathrm{d}e)\mathrm{d}s\right]$$

$$\leq \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right]$$

$$+ \mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s \, |\delta^n Y_s|^{p-2} \left(N^{-1/2} + |X_s^n - X_{t_k}^{n,\pi}| + |\delta^n Y_{t_k}| + |Y_s^n - Y_{t_k}^n| + |Z_s^n - \bar{Z}_s^n| + |\Gamma_s^n - \bar{\Gamma}_s^n|\right) \mathrm{d}s\right]$$

$$+ \mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s \, |\delta^n Y_s|^{p-2}|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|\mathrm{d}s\right] + \mathbb{E}\left[\int_t^{t_{k+1}} \delta^n Y_s \, |\delta^n Y_s|^{p-2}|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|\mathrm{d}s\right].$$

We repeatedly use the inequality $ab \leq \alpha a^2 + b^2/\alpha$ to get

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} \int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2 \nu^n(\mathrm{d}e)\mathrm{d}s\right]$$

$$\leq \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + (\alpha+\beta+\gamma)\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^p \mathrm{d}s\right]$$

$$+ \frac{C_p}{\alpha}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2} \left(N^{-1} + |X_s^n - X_{t_k}^{n,\pi}|^2 + |\delta^n Y_{t_k}|^2 + |Y_s^n - Y_{t_k}^n|^2 + |Z_s^n - \bar{Z}_s^n|^2 + |\Gamma_s^n - \bar{\Gamma}_s^n|^2\right) \mathrm{d}s\right]$$

$$+ \frac{1}{\beta}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] + \frac{1}{\gamma}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s\right].$$

Next we apply Young's inequality

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2 \mathrm{d}s\right] + \kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} \int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2 \nu^n(\mathrm{d}e)\mathrm{d}s\right]$$

$$\leq \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(\alpha+\beta+\gamma+\frac{1}{\alpha}\right)\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^p \mathrm{d}s\right]$$

$$+ \frac{C_p}{\alpha}\mathbb{E}\left[\int_t^{t_{k+1}} \left(N^{-p/2} + |X_s^n - X_{t_k}^{n,\pi}|^p + |\delta^n Y_{t_k}|^p + |Y_s^n - Y_{t_k}^n|^p + |Z_s^n - \bar{Z}_s^n|^p + |\Gamma_s^n - \bar{\Gamma}_s^n|^p\right) \mathrm{d}s\right]$$

$$+ \frac{1}{\beta}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] + \frac{1}{\gamma}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s\right].$$

Because we know from above that

$$\mathbb{E}\left[\int_t^{t_{k+1}} |\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] \leq C_2\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |\delta^n Z_s|^2\mathrm{d}s\right]$$

and

$$\mathbb{E}\left[\int_t^{t_{k+1}} |\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s\right] \leq C_2\mathbb{E}\left[\int_{t_k}^{t_{k+1}} |\delta^n U_s|^2\mathrm{d}s\right],$$

for a constant $C_2 > 0$, we can choose $\beta, \gamma > 0$ independent of N such that

$$\kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2\mathrm{d}s\right] \geq \frac{1}{\beta}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s\right]$$

and

$$\kappa_p \mathbb{E}\left[\int_t^{t_{k+1}} \int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2 \nu^n(\mathrm{d}e)\mathrm{d}s\right] \geq \frac{1}{\gamma}\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s\right].$$

This and (35) imply that

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] \le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(\alpha + \frac{1}{\alpha}\right)\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^p \mathrm{d}s\right]$$
$$+ \frac{C_p}{\alpha}\int_t^{t_{k+1}} \mathbb{E}\left[N^{-p/2} + |\delta^n Y_{t_k}|^p + |Y_s^n - Y_{t_k}^n|^p + |Z_s^n - \bar{Z}_s^n|^p + |\Gamma_s^n - \bar{\Gamma}_s^n|^p\right]\mathrm{d}s,$$

for $t \in [t_k, t_{k+1})$ and thus

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] \le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(\alpha + \frac{1}{\alpha}\right)\mathbb{E}\left[\int_t^{t_{k+1}} |\delta^n Y_s|^p \mathrm{d}s\right]$$
$$+ \frac{C_p}{\alpha}\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right),$$

where

$$\bar{B}_k := \int_{t_k}^{t_{k+1}} \left(\mathbb{E}\left[|Y_s^n - Y_{t_k}^n|^p\right] + \mathbb{E}\left[|Z_s^n - \bar{Z}_s^n|^p\right] + \mathbb{E}\left[|\Gamma_s^n - \bar{\Gamma}_s^n|^p\right]\right)\mathrm{d}s.$$

Using Gronwall's Lemma, we can choose α independent of N such that

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] \le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right). \tag{38}$$

If we take $t = t_k$ in (38) we get

$$\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] \le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right). \tag{39}$$

Plugging (39) into (38) iteratively, combined with the Lipschitz condition for the terminal value $g(X_T^n) - g(X_t^{n,\pi})$ and the bound (29) we obtain

$$\mathbb{E}\left[|\delta^n Y_t|^p\right] \le C_p\left(N^{-p/2} + \bar{B}\right),$$

for $t \in [0, T]$, where

$$\bar{B} := \sum_{k=0}^{N-1} \bar{B}_k.$$

We can take the supremum over all t and conclude

$$\sup_{0\le t\le T} \mathbb{E}\left[|\delta^n Y_t|^p\right] \le C_p\left(N^{-p/2} + \bar{B}\right).$$

Step 2: We also can show that (38) holds for taking the supremum over $[t_k, t_{k+1})$, i.e.,

$$\mathbb{E}\left[\sup_{t_k\le t<t_{k+1}} |\delta^n Y_t|^p\right] \le \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right). \tag{40}$$

This follows like in Step 1 by using Itô's formula (without the expectations)

$$|\delta^n Y_t|^p + \kappa_p\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\delta^n Z_s|^2\mathrm{d}s + \kappa_p\int_t^{t_{k+1}}\int_{\mathbb{R}} |\delta^n Y_s|^{p-2}|\delta^n U_s(e)e|^2\nu^n(\mathrm{d}e)\mathrm{d}s$$
$$\le |\delta^n Y_{t_{k+1}}|^p + C_p\left(\alpha + \beta + \gamma + \frac{1}{\alpha}\right)\int_t^{t_{k+1}} |\delta^n Y_s|^p\mathrm{d}s$$
$$+ \frac{C_p}{\alpha}\int_t^{t_{k+1}}\left(N^{-p/2} + |X_s^n - X_{t_k}^{n,\pi}|^p + |\delta^n Y_{t_k}|^p + |Y_s^n - Y_{t_k}^n|^p + |Z_s^n - \bar{Z}_s^n|^p + |\Gamma_s^n - \bar{\Gamma}_s^n|^p\right)\mathrm{d}s$$
$$+ \frac{1}{\beta}\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^2\mathrm{d}s + \frac{1}{\gamma}\int_t^{t_{k+1}} |\delta^n Y_s|^{p-2}|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^2\mathrm{d}s + M_t, \tag{41}$$

where

$$M_t = \int_t^{t_{k+1}} \delta^n Y_{s-}|\delta^n Y_{s-}|^{p-2}\delta^n Z_s\mathrm{d}B_s + \int_t^{t_{k+1}}\int_{\mathbb{R}} \delta^n Y_{s-}|\delta^n Y_{s-}|^{p-2}\delta^n U_s(e)e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s)$$

denotes the martingales which can be handled with the Burkholder-Davis-Gundy inequality:

$$\mathbb{E}\left[\sup_{t_k\le t<t_{k+1}}\left|\int_t^{t_{k+1}}\delta^nY_{s-}|\delta^nY_{s-}|^{p-2}\delta^nZ_s\mathrm{d}B_s\right|\right]\le C_p\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}}|\delta^nY_s|^{2p-2}|\delta^nZ_s|^2\mathrm{d}s\right)^{1/2}\right]$$
$$\le\frac{1}{4p}\mathbb{E}\left[\sup_{t_k\le t<t_{k+1}}|\delta^nY_t|^p\right]+pC_p^2\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|\delta^nY_s|^{p-2}|\delta^nZ_s|^2\mathrm{d}s\right]$$

and

$$\mathbb{E}\left[\sup_{t_k\le t<t_{k+1}}\left|\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nY_{s-}|\delta^nY_{s-}|^{p-2}\delta^nU_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s)\right|\right]$$
$$\le C_p\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}}|\delta^nY_{s-}|^{2p-2}\delta^nU_s(e)^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)\right)^{1/2}\right]$$
$$\le\frac{1}{4p}\mathbb{E}\left[\sup_{t_k\le t<t_{k+1}}|\delta^nY_t|^p\right]+pC_p^2\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}}|\delta^nY_s|^{p-2}|\delta^nU_s(e)|^2e^2\nu^n(\mathrm{d}e)\mathrm{d}s\right].$$

Taking the supremum and expectations of (41), using the above two bounds and proceeding as in Step 1 yields (40).

Step 3: The next step controls

$$\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}}|\delta^nZ_s|^2\mathrm{d}s\right)^{p/2}+\left(\int_{t_k}^{t_{k+1}}\int_{\mathbb{R}}|\delta^nU_s(e)|^2e^2\nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right].$$

We start by applying Itô's formula to $|\delta^nY_t|^2$ on $[t_k,t_{k+1})$

$$|\delta^nY_t|^2+\int_t^{t_{k+1}}\delta^nZ_s^2\mathrm{d}s+\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nU_s(e)^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)$$
$$=|\delta^nY_{t_{k+1}}|^2+2\int_t^{t_{k+1}}\delta^nY_s\ \delta^nf(\Theta_s)\mathrm{d}s$$
$$-2\int_t^{t_{k+1}}\delta^nY_{s-}\ \delta^nZ_s\mathrm{d}B_s-2\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nY_{s-}\ \delta^nU_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s).$$

The Lipschitz and Hölder condition on f then imply

$$|\delta^nY_t|^2+\int_t^{t_{k+1}}\delta^nZ_s^2\mathrm{d}s+\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nU_s(e)^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)$$
$$\le|\delta^nY_{t_{k+1}}|^2+2\int_t^{t_{k+1}}\delta^nY_s\left(N^{-1/2}+|X_s^n-X_s^{n,\pi}|+|Y_s^n-\bar{Y}_{t_k}^{n,\pi}|+|Z_s^n-\bar{Z}_{t_k}^{n,\pi}|+|\Gamma_s^n-\bar{\Gamma}_{t_k}^{n,\pi}|\right)\mathrm{d}s$$
$$-2\int_t^{t_{k+1}}\delta^nY_{s-}\ \delta^nZ_s\mathrm{d}B_s-2\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nY_{s-}\ \delta^nU_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s).$$

Again we use the inequality $ab\le\alpha a^2+b^2/\alpha$ to get

$$|\delta^nY_t|^2+\int_t^{t_{k+1}}\delta^nZ_s^2\mathrm{d}s+\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nU_s(e)^2e^2\mu^n(\mathrm{d}e,\mathrm{d}s)$$
$$\le|\delta^nY_{t_{k+1}}|^2+(\alpha+\beta+\gamma)\int_t^{t_{k+1}}|\delta^nY_s|^2\mathrm{d}s$$
$$+\frac{C_p}{\alpha}\int_t^{t_{k+1}}\left(N^{-1}+|X_s^n-X_s^{n,\pi}|^2+|\delta^nY_{t_k}|^2+|\bar{Z}_{t_k}^n-\bar{Z}_{t_k}^{n,\pi}|^2+|\bar{\Gamma}_{t_k}^n-\bar{\Gamma}_{t_k}^{n,\pi}|^2\right)\mathrm{d}s$$
$$+\frac{C_p}{\alpha}\int_t^{t_{k+1}}|Y_s^n-Y_{t_k}^n|^2\mathrm{d}s+\frac{C_p}{\beta}\int_t^{t_{k+1}}|Z_s^n-\bar{Z}_s^n|^2\mathrm{d}s+\frac{C_p}{\gamma}\int_t^{t_{k+1}}|\Gamma_s^n-\bar{\Gamma}_s^n|^2\mathrm{d}s$$
$$-2\int_t^{t_{k+1}}\delta^nY_{s-}\ \delta^nZ_s\mathrm{d}B_s-2\int_t^{t_{k+1}}\int_{\mathbb{R}}\delta^nY_{s-}\ \delta^nU_s(e)e\widetilde{\mu^n}(\mathrm{d}e,\mathrm{d}s).$$

Next we take powers

$$
\begin{aligned}
&|\delta^n Y_t|^p + \left(\int_t^{t_{k+1}} \delta^n Z_s^2 \mathrm{d}s\right)^{p/2} + \left(\int_t^{t_{k+1}} \int_{\mathbb{R}} \delta^n U_s(e)^2 e^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2} \\
&\leq C_p |\delta^n Y_{t_{k+1}}|^p + C_p(\alpha+\beta+\gamma)^{p/2} \left|\int_t^{t_{k+1}} |\delta^n Y_s|^2 \mathrm{d}s\right|^{p/2} + \frac{C_p}{\alpha^{p/2}} N^{-p} + \frac{C_p}{\alpha^{p/2}} \left|\int_t^{t_{k+1}} |X_s^n - X_s^{n,\pi}|^2 \mathrm{d}s\right|^{p/2} \\
&\quad + \frac{C_p}{\alpha^{p/2}} N^{-p/2} |\delta^n Y_{t_k}|^p + \frac{C_p}{\beta^{p/2}} N^{-p/2} |\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^p + \frac{C_p}{\gamma^{p/2}} N^{-p/2} |\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^p \\
&\quad + \frac{C_p}{\alpha} \left|\int_t^{t_{k+1}} \left(|Y_s^n - Y_{t_k}^n|^2 + |Z_s^n - \bar{Z}_s^n|^2 + |\Gamma_s^n - \bar{\Gamma}_s^n|^2\right) \mathrm{d}s\right|^{p/2} \\
&\quad + C_p \left|\int_t^{t_{k+1}} \delta^n Y_{s-}\, \delta^n Z_s \mathrm{d}B_s\right|^{p/2} + C_p \left|\int_t^{t_{k+1}} \int_{\mathbb{R}} \delta^n Y_{s-}\, \delta^n U_s(e) e \widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s)\right|^{p/2},
\end{aligned}
$$

and expectations to get

$$
\begin{aligned}
&\mathbb{E}\left[|\delta^n Y_t|^p\right] + \mathbb{E}\left[\left(\int_t^{t_{k+1}} \delta^n Z_s^2 \mathrm{d}s\right)^{p/2} + \left(\int_t^{t_{k+1}} \int_{\mathbb{R}} \delta^n U_s(e)^2 e^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right] \\
&\leq C_p \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p(\alpha+\beta+\gamma)^{p/2} \mathbb{E}\left[\left|\int_t^{t_{k+1}} |\delta^n Y_s|^2 \mathrm{d}s\right|^{p/2}\right] + \frac{C_p}{\alpha^{p/2}} N^{-p} \\
&\quad + \frac{C_p}{\alpha^{p/2}} \mathbb{E}\left[\left|\int_t^{t_{k+1}} |X_s^n - X_s^{n,\pi}|^2 \mathrm{d}s\right|^{p/2}\right] + \frac{C_p}{\alpha^{p/2}} N^{-p/2} \mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] \\
&\quad + \frac{C_p}{\beta^{p/2}} N^{-p/2} \mathbb{E}\left[|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^p\right] + \frac{C_p}{\gamma^{p/2}} N^{-p/2} \mathbb{E}\left[|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^p\right] \\
&\quad + \frac{C_p}{\alpha} \mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} \left(|Y_s^n - Y_{t_k}^n|^2 + |Z_s^n - \bar{Z}_s^n|^2 + |\Gamma_s^n - \bar{\Gamma}_s^n|^2\right) \mathrm{d}s\right|^{p/2}\right] \\
&\quad + C_p \mathbb{E}\left[\left|\int_t^{t_{k+1}} \delta^n Y_{s-}\, \delta^n Z_s \mathrm{d}B_s\right|^{p/2}\right] + C_p \mathbb{E}\left[\left|\int_t^{t_{k+1}} \int_{\mathbb{R}} \delta^n Y_{s-}\, \delta^n U_s(e) e \widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s)\right|^{p/2}\right]. \qquad (42)
\end{aligned}
$$

We discuss the terms in (42) separately. First, we recall that by (39) $\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] \leq \mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right)$ and, by (40) and additionally invoking Jensen's inequality,

$$
\begin{aligned}
\mathbb{E}\left[\left|\int_t^{t_{k+1}} |\delta^n Y_s|^2 \mathrm{d}s\right|^{p/2}\right] &\leq C_p N^{-p/2} \mathbb{E}\left[\sup_{t_k \leq s < t_{k+1}} |\delta^n Y_s|^p\right] \\
&\leq C_p\left(N^{-p/2}\mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + N^{-p-1} + N^{-p/2-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right]\right).
\end{aligned}
$$

Second, in a similar manner, the term with the forward SDE X is bounded by $C_p N^{-p}$. Third, recalling (36) and (37) we note that

$$
N^{-p/2}\mathbb{E}\left[|\bar{Z}_{t_k}^n - \bar{Z}_{t_k}^{n,\pi}|^p\right] \leq C_p ||\delta^n Z||^p_{\mathbb{H}^p_{[t_k,t_{k+1}]}}
$$

and

$$
N^{-p/2}\mathbb{E}\left[|\bar{\Gamma}_{t_k}^n - \bar{\Gamma}_{t_k}^{n,\pi}|^p\right] \leq C_p ||\delta^n U||^p_{\mathbb{L}^p_{\mu^n,[t_k,t_{k+1}]}}.
$$

Fourth, by Jensen's inequality

$$
\mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} \left(|Y_s^n - Y_{t_k}^n|^2 + |Z_s^n - \bar{Z}_s^n|^2 + |\Gamma_s^n - \bar{\Gamma}_s^n|^2\right) \mathrm{d}s\right|^{p/2}\right] \leq C_p N^{-p/2+1} \bar{B}_k.
$$

Finally, we can apply the Burkholder-Davis-Gundy inequality and Young's inequality to the martingales in (42):

$$
\begin{aligned}
&C_p\mathbb{E}\left[\left|\int_t^{t_{k+1}} \delta^n Y_{s-}\ \delta^n Z_s \mathrm{d}B_s\right|^{p/2}\right] \\
&\leq C_p\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n Y_{s-}|^2\ |\delta^n Z_s|^2 \mathrm{d}s\right)^{p/4}\right] \\
&\leq \frac{C_p^2}{4}\mathbb{E}\left[\sup_{t_k\leq s<t_{k+1}} |\delta^n Y_s|^p\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n Z_s|^2 \mathrm{d}s\right)^{p/2}\right] \\
&\leq C_p\mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right) + \frac{1}{2}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n Z_s|^2 \mathrm{d}s\right)^{p/2}\right],
\end{aligned}
$$

where the last inequality follows by (40). Analogously,

$$
\begin{aligned}
&C_p\mathbb{E}\left[\left|\int_t^{t_{k+1}} \int_{\mathbb{R}} \delta^n Y_{s-}\ \delta^n U_s(e) e\widetilde{\mu^n}(\mathrm{d}e, \mathrm{d}s)\right|^{p/2}\right] \\
&\leq C_p\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n Y_{s-}|^2\ |\delta^n U_s(e) e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/4}\right] \\
&\leq \frac{C_p^2}{4}\mathbb{E}\left[\sup_{t_k\leq s<t_{k+1}} |\delta^n Y_s|^p\right] + \frac{1}{2}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n U_s(e) e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right] \\
&\leq C_p\mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right) + \frac{1}{2}\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n U_s(e) e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right].
\end{aligned}
$$

We again use

$$
\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n U_s(e) e|^2 \nu^n(\mathrm{d}e)\mathrm{d}s\right)^{p/2}\right] \leq d_p\mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} |\delta^n U_s(e) e|^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right]
$$

and conclude that, for $t = t_k$, we can choose constants α, β and γ independent of N such that (42) can be simplified to

$$
\begin{aligned}
&\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \mathbb{E}\left[\left(\int_{t_k}^{t_{k+1}} \delta^n Z_s^2 \mathrm{d}s\right)^{p/2} + \left(\int_{t_k}^{t_{k+1}} \int_{\mathbb{R}} \delta^n U_s(e)^2 e^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right] \\
&\leq C_p\mathbb{E}\left[|\delta^n Y_{t_{k+1}}|^p\right] + C_p\left(N^{-p/2-1} + N^{-1}\mathbb{E}\left[|\delta^n Y_{t_k}|^p\right] + \bar{B}_k\right). \qquad (43)
\end{aligned}
$$

Now we can sum up equation (43). Together with (39) and the Lipschitz condition for the terminal value we obtain

$$
\mathbb{E}\left[\left(\int_0^T \delta^n Z_s^2 \mathrm{d}s\right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} \delta^n U_s(e)^2 e^2 \mu^n(\mathrm{d}e, \mathrm{d}s)\right)^{p/2}\right] \leq C_p\left(N^{-p/2} + \bar{B}\right).
$$

Joining Step 1 with Step 3 then implies that

$$
Err_\pi(Y^n, Z^n, U^n)^p \leq C_p\left(N^{-p/2} + \bar{B}\right).
$$

Step 4: It remains to show that $\bar{B} \leq C_p N^{-p/2}$. For the first term in $\bar{B}$, we recall that Y^n solves (10) and hence

$$
\mathbb{E}\left[|Y_t^n - Y_{t_k}^n|^p\right] \leq C_p \int_{t_k}^t \mathbb{E}\left[|f(\Omega_s)|^p + |Z_s|^p + \int_{\mathbb{R}} |U_s(e) e|^p \nu(\mathrm{d}e)\right] \mathrm{d}s.
$$

The Lipschitz property of f combined with (11) implies

$$\sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \mathbb{E}\left[|Y_t - Y_{t_k}|^p\right] \mathrm{d}t \leq C_p N^{-p/2}.$$

For the second and third term of $\bar{B}$ we exactly follow the proofs of Bouchard & Elie (2008) given the additional Assumption 2. Bouchard & Elie (2008, Propositions 4.5-4.6 & Theorem 2.1) proved that the regularities of Z^n and Γ^n are bounded by $C_2 N^{-1}$ for $p = 2$. Replacing $p = 2$ with a $p \geq 2$ is a straightforward extension of their proofs. This implies $\bar{B} \leq C_p N^{-p/2}$ and finally the statement follows by joining Steps 1-4. □

Remark 6. Following the argument of Bouchard & Elie (2008), if Assumption 2 does not hold the bound the error bound (34) is not valid anymore because the regularity in Z, i.e., $||Z^n - \bar{Z}^n||^p_{\mathbb{H}^p}$ is not bounded by $C_p N^{-p/2}$ in this case. However, one can show without using Assumption 2 that, for any $\varepsilon > 0$, there exists a constant $C_{p,\varepsilon}$ such that

$$||Z^n - \bar{Z}^n||^p_{\mathbb{H}^p} \leq C_{p,\varepsilon} N^{-p/2+\varepsilon}.$$

Note that the regularities of Y^n and Γ^n remain unaffected whether Assumption 2 is fulfilled or not, i.e., $||Y^n - \bar{Y}^n||^p_{\mathcal{S}^p} \leq C_p N^{-p/2}$ and $||\Gamma^n - \bar{\Gamma}^n||^p_{\mathbb{H}^p} \leq C_p N^{-p/2}$ even without Assumption 2. Furthermore, if either $a \equiv 0$, or the generator f is independent of Z Theorem 2 holds without Assumption 2.

Remark 7. It could be tempting to choose the random times when the jumps of the Lévy process occur given its series representation as the simulation grid for the FBSDE. This is slightly easier to implement because the jumps $H(\frac{G_i}{T}, V_i)$ are simply ordered by the size of the T_i's. However this approach has the disadvantage that we now deal with a non-regular grid $\tilde{\pi} : 0 = \tilde{t}_0 < \tilde{t}_1 < \cdots < \tilde{t}_N = T$ and $|\tilde{\pi}| := \max_{1\leq k\leq N} |\tilde{t}_k - \tilde{t}_{k-1}| > \frac{T}{N} = |\pi|$. Theorem 2 then reads $Err_{\tilde{\pi}}(Y^n, Z^n, U^n) \leq C_p |\tilde{\pi}|^{1/2}$ which is larger than for the regular grid. Thus, employing a regular grid is advisable.

Remark 8. Instead of the implicit scheme (30), one could use an explicit scheme where we replace $\bar{Y}^{n,\pi}_{t_k}$ by $\bar{Y}^{n,\pi}_{t_{k+1}}$ in the argument of h. The advantage is that we do not need a fixed-point procedure in this case. One disadvantage is that the conditional expectations are more difficult to estimate. We refer to Bouchard & Elie (2008) and Elie (2006) for details.

Remark 9. Besides the backward Euler scheme, Aazizi (2013) proposed a second scheme based on Malliavin calculus techniques. The author showed that the L^p error between a BSDE with finitely many jumps and the discrete version using Malliavin derivatives is bounded by $C_p n^{1/2(1/\log n - 1)}$ for $p \geq 2$. Aazizi (2013) derived the L^2 error between the original SDE with infinite jump activity and the discrete scheme. With Theorem 1 at hand, one could easily derive the L^p error adopting Theorem 4.3 of Aazizi (2013). We omit the details here but note that although the Malliavin scheme has a larger error than the Euler scheme it has the advantage that it can be also used when the terminal value is not given by the forward SDE.

Using Theorems 1 and 2, we deduce a bound for the approximation-discretization error between the original backward SDE (8) and the scheme (30) which is defined as

$$Err_{n,\pi}(Y, Z, U) := \left(\sup_{0\leq t\leq T} \mathbb{E}\left[|Y_t - \bar{Y}^{n,\pi}_t|^p\right] + ||Z - \bar{Z}^{n,\pi}||^p_{\mathbb{H}^p} + ||\Gamma - \bar{\Gamma}^{n,\pi}||^p_{\mathbb{H}^p} \right)^{1/p}.$$

The approximation-discretization error for the forward SDE

$$\max_{k<N} \mathbb{E}\left[\sup_{t\in[t_k,t_{k+1}]} |X_t - X^{n,\pi}_t|^p \right] \leq C_p \left(n^{-p/2} + \sigma^p(n) + \sigma^2(n)^{p/2} \right),$$

is straightforward combining (13) with (29).

Corollary 1. *Under Assumptions 1 and 2, the approximation-discretization error is bounded by*

$$Err_{n,\pi}(Y, Z, U) \leq C_p \left(N^{-1/2} + \sigma^p(n)^{1/p} + \sigma^2(n)^{1/2} \right).$$

Proof. This is an easy consequence because

$$Err_{n,\pi}(Y, Z, U)^p \leq C_p \Bigg(\sup_{0\leq t\leq T} \mathbb{E}\left[|Y_t - Y^n_t|^p + |Y^n_t - \bar{Y}^{n,\pi}_t|^p\right] + ||Z - Z^n||^p_{\mathbb{H}^p} + ||Z^n - \bar{Z}^{n,\pi}||^p_{\mathbb{H}^p}$$

$$+ ||\Gamma - \Gamma^n||^p_{\mathbb{H}^p} + ||\Gamma^n - \bar{\Gamma}^{n,\pi}||^p_{\mathbb{H}^p} \Bigg).$$

Using Remark 1 we can show

$$\left(\int_{\mathbb{R}} \rho(e)(U_s(e) - U^n_s(e))e\nu^n(\mathrm{d}e) \right)^2 \le C_p \int_{\mathbb{R}} (U_s(e) - U^n_s(e))^2 e^2 \nu^n(\mathrm{d}e)$$

and

$$\left(\int_{\mathbb{R}} \rho(e)U_s(e)e\bar{\nu}^n(\mathrm{d}e) \right)^2 \le C_p \int_{\mathbb{R}} U_s(e)^2 e^2 \bar{\nu}^n(\mathrm{d}e)$$

which imply

$$||\Gamma - \Gamma^n||^p_{\mathbb{H}^p} \le C_p \mathbb{E} \left[\left(\int_0^T \int_{\mathbb{R}} (U_s(e) - U^n_s(e))^2 e^2 \nu^n(\mathrm{d}e)\mathrm{d}s \right)^{p/2} + \left(\int_0^T \int_{\mathbb{R}} U_s(e)^2 e^2 \bar{\nu}^n(\mathrm{d}e)\mathrm{d}s \right)^{p/2} \right],$$

and thus the result follows. □

We end this paper with some remarks about implementation of the scheme in practice.

Remark 10. If we aim to approximate the FBSDE (1)-(2) by using the Euler scheme together with truncated series representations we can do the following. First, we simulate the forward SDE according to

$$\begin{cases} X^{n,\pi}_0 & := X_0 \\ X^{n,\pi}_{t_{k+1}} & := X^{n,\pi}_{t_k} + \frac{T}{N} b(t_k, X^{n,\pi}_{t_k}) + a(t_k, X^{n,\pi}_{t_k})\Delta B_{k+1} + h(t_k, X^{n,\pi}_{t_k})\Delta L^n_{k+1}, \end{cases}$$

where $\Delta B_{k+1} := B_{t_{k+1}} - B_{t_k}$ and

$$\Delta L^n_{k+1} := L^n_{t_{k+1}} - L^n_{t_k} = \sum_{\{i: G_i \le nT\}} H\left(\frac{G_i}{T}, V_i \right) \mathbb{1}_{[t_k, t_{k+1}]}(T_i) - tc_i.$$

Second, we modify the Euler scheme (30) to

$$\begin{cases} \bar{Z}^{n,\pi}_t & := \frac{N}{T} \mathbb{E}\left[\bar{Y}^{n,\pi}_{t_{k+1}} \Delta B_{k+1} | \mathcal{F}_{t_k} \right] \\ \bar{U}^{n,\pi}_t & := \frac{N}{T} \mathbb{E}\left[\bar{Y}^{n,\pi}_{t_{k+1}} \Delta L^n_{k+1} | \mathcal{F}_{t_k} \right] \\ \bar{\Gamma}^{n,\pi}_t & := \int_{\mathbb{R}} \rho(e) \bar{U}^{n,\pi}_t e\nu^n(\mathrm{d}e) \\ \bar{Y}^{n,\pi}_t & := \mathbb{E}\left[\bar{Y}^{n,\pi}_{t_{k+1}} | \mathcal{F}_{t_k} \right] + \frac{T}{N} f\left(t_k, X^{n,\pi}_{t_k}, \bar{Y}^{n,\pi}_{t_k}, \bar{Z}^{n,\pi}_{t_k}, \bar{\Gamma}^{n,\pi}_{t_k} \right). \end{cases}$$

Remark 11. The proposed scheme is not fully implementable in practice. One key step is the computation of the conditional expectations in (30) which has to be performed numerically. There are several methods to estimate these. Among them there are nonparametric kernel regression (Bouchard & Touzi 2004, Lemor et al. 2006), Malliavin regression (Bouchard & Touzi 2004), quantization (Bally & Pagès 2003) and some other approaches. We discuss the nonparametric regression approach in some more detail which works by simulating $1 \le m \le M$ paths $X^{n,\pi,m}$ of $X^{n,\pi}$ and initialize $\bar{Y}^{n,\pi,m}_T = g(X^{n,\pi,m}_T)$. Then we regress $\bar{Y}^{n,\pi,m}_{t_{k+1}}$ and $\bar{Y}^{n,\pi,m}_{t_{k+1}} \Delta B^m_{k+1}$ and $\bar{Y}^{n,\pi,m}_{t_{k+1}} \int_{\mathbb{R}} \rho(e) e\widetilde{\mu^{n,m}}(\mathrm{d}e, (t_k, t_{k+1}])$ on $X^{n,\pi,m}_{t_k}$. Details are presented in Elie (2006).

To compute the L^p error between the original backward SDE and the numerical backward SDE taking into account approximation of the jump process, discretization and estimation of conditional expectations we have to sum up the error of Corollary 1, the error of a localization procedure and the statistical error by the kernel regression. Elie (2006) derived the L^p error of the localization procedure. Furthermore, Elie (2006) derived the statistical error which is in terms of the Euclidean norm on $\mathbb{R}^M$. Since all norms on $\mathbb{R}^M$ are equivalent it is not much work to deduce a bound for the error in terms of the p-norm. All in all, if we choose some other parameters in the algorithm large enough, we can conclude that the total error is of the order $N^{-1/2} + \sigma^p(n)^{1/p} + \sigma^2(n)^{1/2}$ under Assumptions 1 and 2.

Acknowledgements

Financial support of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) via the Collaborative Research Center "Statistical modelling of nonlinear dynamic processes" (SFB 823, Teilprojekt A4) is gratefully acknowledged.

The author is grateful to Christoph Hanck for valuable comments which helped to substantially improve this paper. Full responsibility is taken for all remaining errors.

References

Aazizi, S. (2013), 'Discrete-Time Approximation of Decoupled Forward–Backward Stochastic Differential Equations Driven by Pure Jump Lévy Processes', *Advances in Applied Probability* **45**(3), 791–821.
URL: *https://doi.org/10.1239/aap/1377868539*

Applebaum, D. (2009), *Lévy processes and stochastic calculus*, Cambridge university press.

Asmussen, S. & Rosiński, J. (2001), 'Approximations of small jumps of Lévy processes with a view towards simulation', *Journal of Applied Probability* **38**(2), 482–493.
URL: *http://dx.doi.org/10.1239/jap/996986757*

Bally, V. & Pagès, G. (2003), 'A quantization algorithm for solving multidimensional discrete-time optimal stopping problems', *Bernoulli : official journal of the Bernoulli Society for Mathematical Statistics and Probability* **9**(6), 1003–1049.
URL: *https://doi.org/10.3150/bj/1072215199*

Barles, G., Buckdahn, R. & Pardoux, E. (1997), 'Backward stochastic differential equations and integral-partial differential equations', *Stochastics and Stochastic Reports* **60**, 57–83.
URL: *https://www.tandfonline.com/doi/full/10.1080/17442509708834099*

Bouchard, B. & Elie, R. (2008), 'Discrete-time approximation of decoupled Forward-Backward SDE with jumps', *Stochastic Processes and their Applications* **118**, 53–75.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S030441490700052X*

Bouchard, B. & Touzi, N. (2004), 'Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations', *Stochastic Processes and their Applications* **111**, 175–206.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0304414904000031*

Briand, P., Delyon, B., Hu, Y., Pardoux, E. & Stoica, L. (2003), 'Lp solutions of backward stochastic differential equations', *Stochastic Processes and their Applications* **108**, 109–129.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0304414903000899*

Buckdahn, R. & Pardoux, E. (1994), 'BSDE's with jumps and associated integro-partial differential equations'.
URL: *https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.4924&rep=rep1&type=pdf*

Delong, Ł. (2013), *Backward stochastic differential equations with jumps and their actuarial and financial applications*, Springer.

Dzhaparidze, K. & Valkeila, E. (1990), 'On the Hellinger type distances for filtered experiments', *Probability theory and related fields* **85**(1), 105–117.
URL: *https://link.springer.com/content/pdf/10.1007/BF01377632.pdf*

Eddahbi, M., Fakhouri, I. & Ouknine, Y. (2017), '$\mathbb{L}^p (p \geq 2)$-solutions of generalized BSDEs with jumps and monotone generator in a general filtration', *Modern Stochastics: Theory and Applications* **4**, 25–63.
URL: *https://www.vmsta.org/doi/10.15559/17-VMSTA73*

Elie, R. (2006), Contrôle stochastique et méthodes numériques en finance mathématique, PhD thesis, University Paris-Dauphine.
URL: *https://pastel.archives-ouvertes.fr/tel-00122883/file/thesis.pdf*

Gobet, E. & Labart, C. (2007), 'Error expansion for the discretization of backward stochastic differential equations', *Stochastic Processes and their Applications* **117**, 803–829.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0304414906001487*

Karoui, N. E., Peng, S. & Quenez, M. C. (1997), 'Backward Stochastic Differential Equations in Finance', *Mathematical Finance* **7**, 1–71.
URL: *http://doi.wiley.com/10.1111/1467-9965.00022*

Kohatsu-Higa, A. & Tankov, P. (2010), 'Jump-adapted discretization schemes for Lévy-driven SDEs', *Stochastic Processes and their Applications* **120**, 2258–2285.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0304414910001717*

Kruse, T. & Popier, A. (2016), 'BSDEs with monotone generator driven by Brownian and Poisson noises in a general filtration', *Stochastics* **88**(4), 491–539.
URL: *https://doi.org/10.1080/17442508.2015.1090990*

Lemor, J.-P., Gobet, E. & Warin, X. (2006), 'Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations', *Bernoulli* **12**(5), 889–916.
URL: *https://doi.org/10.3150/bj/1161614951*

Pardoux, É. (1999), BSDEs, weak convergence and homogenization of semilinear PDEs, *in* 'Nonlinear analysis, differential equations and control', Springer, pp. 503–549.
URL: *https://link.springer.com/chapter/10.1007/978-94-011-4560-2_9*

Pardoux, E. & Peng, S. (1990), 'Adapted solution of a backward stochastic differential equation', *Systems & Control Letters* **14**(1), 55–61.
URL: *https://doi.org/10.1016/0167-6911(90)90082-6*

Pardoux, E. & Răşcanu, A. (2014), *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, Vol. 69, Springer International Publishing.
URL: *http://link.springer.com/10.1007/978-3-319-05714-9*

Rosiński, J. (2001), Series representations of Lévy processes from the perspective of point processes, *in* O. E. Barndorff-Nielsen, S. I. Resnick & T. Mikosch, eds, 'Lévy Processes: Theory and Applications', Birkhäuser, Boston, pp. 401–415.
URL: *https://doi.org/10.1007/978-1-4612-0197-7_18*

Tang, S. & Li, X. (1994), 'Necessary conditions for optimal control of stochastic systems with random jumps', *SIAM Journal on Control and Optimization* **32**(5), 1447–1475.
URL: *https://doi.org/10.1137/S0363012992233858*

Yao, S. (2010), 'Lp solutions of backward stochastic differential equations with jumps', *arXiv preprint arXiv:1007.2226* .
URL: *https://arxiv.org/abs/1007.2226v1*

Yao, S. (2017), 'Lp solutions of backward stochastic differential equations with jumps', *Stochastic Processes and their Applications* **127**(11), 3465–3511.
URL: *https://doi.org/10.1016/j.spa.2017.03.005*

Yuan, S. & Kawai, R. (2021), 'Numerical aspects of shot noise representation of infinitely divisible laws and related processes', *arXiv preprint arXiv:2101.10533* .
URL: *https://arxiv.org/abs/2101.10533*

Zhang, J. (2004), 'A numerical scheme for BSDEs', *The Annals of Applied Probability* **14**, 459–488.
URL: *https://projecteuclid.org/journals/annals-of-applied-probability/volume-14/issue-1/A-numerical-scheme-for-BSDEs/10.1214/aoap/1075828058.full*