

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Geiss, Christel; Steinicke, Alexander

Title: Existence, uniqueness and comparison results for BSDEs with Lévy jumps in an extended monotonic generator setting

Year: 2018

Version: Published version

Copyright: © The Author(s), 2018.

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Geiss, C., & Steinicke, A. (2018). Existence, uniqueness and comparison results for BSDEs with Lévy jumps in an extended monotonic generator setting. *Probability, Uncertainty and Quantitative Risk*, 3(9), 1-33. <https://doi.org/10.1186/s41546-018-0034-y>

Existence, uniqueness and comparison results for BSDEs with Lévy jumps in an extended monotonic generator setting



Christel Geiss · Alexander Steinicke 

Received: 4 January 2018 / Accepted: 29 November 2018 / Published online: 28 December 2018

© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Abstract We show that the comparison results for a backward SDE with jumps established in Royer (Stoch. Process. Appl 116: 1358–1376, 2006) and Yin and Mao (J. Math. Anal. Appl 346: 345–358, 2008) hold under more simplified conditions. Moreover, we prove existence and uniqueness allowing the coefficients in the linear growth- and monotonicity-condition for the generator to be random and time-dependent. In the L^2 -case with linear growth, this also generalizes the results of Kruse and Popier (Stochastics 88: 491–539, 2016). For the proof of the comparison result, we introduce an approximation technique: Given a BSDE driven by Brownian motion and Poisson random measure, we approximate it by BSDEs where the Poisson random measure admits only jumps of size larger than $1/n$.

Keywords Backward stochastic differential equation · Lévy process · comparison theorem · existence and uniqueness

Mathematics Subject Classification: 60H10

C. Geiss

University of Jyväskylä, Department of Mathematics and Statistics, P.O. Box 35, 40014 Jyväskylä, Finland

A. Steinicke (✉)

Department of Mathematics and Information Technology, Montanuniversität Leoben, Leoben, Austria

e-mail: alexander.steinicke@unileoben.ac.at

1 Introduction

In this paper, we study backward stochastic differential equations (BSDEs) of the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, U_s) ds - \int_t^T Z_s dW_s - \int_{[t, T] \times (\mathbb{R} \setminus \{0\})} U_s(x) \tilde{N}(ds, dx), \quad (1)$$

where W denotes a one-dimensional Brownian motion and \tilde{N} a compensated Poisson random measure belonging to a given Lévy process with Lévy measure ν . In particular, our focus lies on comparison results and existence and uniqueness of solutions.

Comparison theorems state that—under certain conditions—if $\xi \leq \xi'$ and $f \leq f'$, then the process Y of the solution satisfies $Y_t \leq Y'_t$ for all $t \in [0, T]$. These types of theorems in the case of one-dimensional, Brownian BSDEs has been treated by Peng (1992), El Karoui et al. (1997, 2009), and Cao and Yan (1999).

In (Barles et al. (1997), Remark 2.7) a counterexample was given, which shows that in the jump case the conditions $\xi \leq \xi'$ and $f \leq f'$ are not sufficient to guarantee $Y \leq Y'$. They propose an additional sufficient condition which has been generalized by Kruse and Popier (2016), Royer (2006), Yin and Mao (2008), Becherer et al. (2018) (allowing more general jump processes), and Cohen et al. (2010) (for BSDEs driven by martingales). The condition of Kruse and Popier (2016) reads (in our L^2 -setting) as follows: for each $s, y, z, u, u' \in [0, T] \times \mathbb{R} \times \mathbb{R} \times L^2(\nu) \times L^2(\nu)$ there is a progressively measurable process $\gamma^{y,z,u,u'} : \Omega \times [0, T] \times \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ such that

$$\begin{aligned} f(s, y, z, u) - f(s, y, z, u') &\leq \int_{\mathbb{R} \setminus \{0\}} (u(x) - u'(x)) \gamma_s^{y,z,u,u'}(x) \nu(dx), \\ -1 &\leq \gamma_s^{y,z,u,u'}(x) \quad \text{and} \quad \sup_{s, \omega, y, z, u, u'} \left| \gamma_s^{y,z,u,u'} \right| \in L^2(\nu). \end{aligned} \quad (2)$$

One of the main results in the present paper is Theorem 3.5 which states that (2) can be replaced by the simpler condition

$$\begin{aligned} f(s, y, z, u) - f(s, y, z, u') &\leq \int_{\mathbb{R} \setminus \{0\}} (u'(x) - u(x)) \nu(dx), \quad \mathbb{P} \otimes \lambda\text{-a.e.} \\ \text{for all } u, u' &\in L^2(\nu) \text{ with } u \leq u'. \end{aligned} \quad (3)$$

Notice that the r.h.s. is infinite for $u'(x) - u(x) \notin L^1(\nu)$. Clearly, (3) is a weaker condition than (2), because one only needs to check the inequality for those $u, u' \in L^2(\nu)$ for which $u \leq u'$ holds. Moreover, we do not need any $L^2(\nu)$ condition for $\gamma_s^{y,z,u,u'}$ but we choose $\gamma_s^{y,z,u,u'}(x) = -1$. Under the constraint $-1 \leq \gamma_s^{y,z,u,u'}(x)$, the choice $\gamma_s^{y,z,u,u'}(x) = -1$ yields for $u' - u \geq 0$ the largest possible expression on the r.h.s. of (2), so that (3) can be seen as the weakest possible condition which (2) could impose on f .

For a finite Lévy measure ν , Theorem 3.5 can be shown using only elementary means.

Another main result is a method of how to approximate a BSDE driven by a Lévy process with an infinite measure ν , by a sequence of BSDEs where the driving processes have a finite Lévy measure. We apply this result to show the comparison theorem for BSDEs driven by a general Lévy process. The proof relies on the Jankov–von Neumann theorem on measurable sections/uniformizations (this theorem is also important for dynamic programming, see El Karoui and Tan (2013)). Under certain conditions on the generator, the approximating solutions can be interpreted as nonlinear conditional expectations (in the sense of Peng (2010)), conditioned on a Lévy process whose jumps are not of arbitrarily small size. (See the comments after Theorem 3.4.)

Studying the existence, uniqueness, and comparison results by Darling and Pardoux (1997), Pardoux and Zhang (1996), Pardoux (1997), Fan and Jiang (2012), Royer (2006), Situ (1997), Yin and Mao (2008), Kruse and Popier (2016, 2017), Yao (2017), and Sow (2014), one notices that one can unify and generalize the assumptions on f .

Indeed, and this is our third main result, in the case of L^2 -solutions, for a progressively measurable generator f with linear growth, it suffices to assume (cf. Theorems 3.1 and 3.5) the following growth- and monotonicity conditions with time-dependent, random coefficients:

- $|f(\omega, s, y, z, u)| \leq F(s, \omega) + K_1(s, \omega)|y| + K_2(s, \omega)(|z| + \|u\|),$
- $(y - y')(f_1(\omega, s, y, z, u) - f_1(\omega, s, y', z', u'))$
 $\leq \alpha(s)\rho(|y - y'|^2) + \beta(s, \omega)|y - y'|(|z - z'| + \|u - u'\|),$

with $\alpha \in L^1([0, T])$ and F being nonnegative and progressively measurable such that $\mathbb{E} \left[\left(\int_0^T F(\omega, t) dt \right)^2 \right] < \infty$. The processes K_1 , K_2 , and β are nonnegative and progressively measurable such that for a constant $c > 0$,

$$\int_0^T \left(K_1(s) + K_2(s)^2 + \beta(s)^2 \right) ds < c, \quad \mathbb{P}\text{-a.s.}$$

The concave function ρ in the monotonicity condition may grow faster than linear at zero and satisfies $\int_{0+} 1/\rho(x) dx = \infty$. This type of function already appeared in context with BSDEs in Mao (1995) in 1997.

These assumptions also extend the monotonicity condition of Kruse and Popier (2016, 2017), for the L^2 -case with linear growth, since the coefficients in our setting take randomness, the function ρ and time-dependence into account. BSDEs with time-dependent coefficients appear, for example, in Gobet and Turkedjiev (2016).

The existence and uniqueness result Theorem 3.1 and the comparison result Theorem 3.5 are basic tools in the forthcoming paper (Geiss and Steinicke 2018) on Malliavin differentiability and boundedness of solutions to BSDEs. To compute the Malliavin derivative for the jump part of the Lévy process, more structure from the generator is required in its dependency on u , usually via an integral w.r.t. $\nu(dx)$, for example,

$$f(s, u) = h \left(s, \int_{\mathbb{R} \setminus \{0\}} u(x) \kappa(s, x) \nu(dx) \right),$$

where $[0, T] \times \mathbb{R} \ni (s, v) \mapsto h(s, v)$. One can find h and κ such that the assumptions of Theorem 3.5 are satisfied while condition (2) does not hold: By the mean value theorem there exists a $\zeta \in]0, 1[$ and

$$v_\zeta := \int_{\mathbb{R} \setminus \{0\}} (\zeta u(x) + (1 - \zeta)u'(x)) \kappa(s, x) v(dx),$$

such that

$$f(s, u) - f(s, u') = \partial_v h(s, v_\zeta) \int_{\mathbb{R} \setminus \{0\}} (u(x) - u'(x)) \kappa(s, x) v(dx).$$

Assumption (3) holds if $\gamma_s^{u, u'}(x) := \partial_v h(s, v_\zeta) \kappa(s, x) \geq -1$ for all (s, u, u', x) . Choosing, for example, a bounded function h such that also $\sup_{s, v} |\partial_v h(s, v)| < \infty$, but $\partial_v h(s, v) \neq 0$ for a.e. s and v , and putting $\kappa(s, x) = s^{-\frac{1}{4}}(|x| \wedge 1)$, then (2) does not hold since

$$\sup_{s, u, u'} |\gamma_s^{u, u'}| \notin L^2(v).$$

However, the Assumptions (A2), (A3) of Section 3 are satisfied for

$$K_2(s) = \beta(s) = \sup_v |\partial_v h(s, v)| \|\kappa(s, \cdot)\|_{L^2(v)} \leq cs^{-\frac{1}{4}}.$$

The paper is structured as follows: Section 2 contains preliminaries and basic definitions. In Section 3, we present the main theorems of this paper about existence and uniqueness of solutions, the approximation using BSDEs based on Lévy processes with finite Lévy measure, and the comparison result. The latter we also prove there. Having stated and proved some auxiliary results in Section 4, including an a-priori estimate for our type of BSDEs, we are able to prove existence and uniqueness and the approximation result from Section 3. In the appendix, we recall the Bihari–LaSalle inequality and the Jankov–von Neumann theorem.

2 Setting

Let $X = (X_t)_{t \in [0, T]}$ be a càdlàg Lévy process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with Lévy measure ν . We will denote the augmented natural filtration of X by $(\mathcal{F}_t)_{t \in [0, T]}$ and assume that $\mathcal{F} = \mathcal{F}_T$. For $0 < p \leq \infty$ we use the notation $(L^p, \|\cdot\|_p) := (L^p(\Omega, \mathcal{F}, \mathbb{P}), \|\cdot\|_{L^p})$. Equations or inequalities for objects of these spaces throughout the paper are considered up to \mathbb{P} -null sets.

The Lévy–Itô decomposition of a Lévy process X can be written as

$$X_t = at + \sigma W_t + \int_{[0, t] \times \{|x| \leq 1\}} x \tilde{N}(ds, dx) + \int_{[0, t] \times \{|x| > 1\}} x N(ds, dx), \quad (4)$$

where $a \in \mathbb{R}$, $\sigma \geq 0$, W is a Brownian motion and $N(\tilde{N})$ is the (compensated) Poisson random measure corresponding to X , see Applebaum (2004) or Sato (1999).

Notation

- Let \mathcal{S}^2 denote the space of all (\mathcal{F}_t) -progressively measurable and càdlàg processes $Y: \Omega \times [0, T] \rightarrow \mathbb{R}$ such that

$$\|Y\|_{\mathcal{S}^2}^2 := \mathbb{E} \sup_{0 \leq t \leq T} |Y_t|^2 < \infty.$$

- We define $L^2(W)$ as the space of all (\mathcal{F}_t) -progressively measurable processes $Z: \Omega \times [0, T] \rightarrow \mathbb{R}$ such that

$$\|Z\|_{L^2(W)}^2 := \mathbb{E} \int_0^T |Z_s|^2 ds < \infty.$$

- Let $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. We define $L^2(\tilde{N})$ as the space of all random fields $U: \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$ which are measurable with respect to $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_0)$ (where \mathcal{P} denotes the predictable σ -algebra on $\Omega \times [0, T]$ generated by the left-continuous (\mathcal{F}_t) -adapted processes) such that

$$\|U\|_{L^2(\tilde{N})}^2 := \mathbb{E} \int_{[0, T] \times \mathbb{R}_0} |U_s(x)|^2 ds \nu(dx) < \infty.$$

- $L^2(\nu) := L^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$, $\|\cdot\| := \|\cdot\|_{L^2(\nu)}$.
- $L^p([0, T]) := L^p([0, T], \mathcal{B}([0, T]), \lambda)$ for $p > 0$, where λ is the Lebesgue measure on $[0, T]$.
- With a slight abuse of the notation, we define

$$L^2(\Omega; L^1([0, T])) \quad (5)$$

$$:= \left\{ F \in L^0(\Omega \times [0, T], \mathcal{F} \otimes \mathcal{B}([0, T]), \mathbb{P} \otimes \lambda) : \mathbb{E} \left[\int_0^T |F(\omega, t)| dt \right]^2 < \infty \right\}$$

For $F \in L^2(\Omega; L^1([0, T]))$, put

$$I_F(\omega) := \int_0^T F(\omega, t) dt \quad \text{and} \quad K_F(\omega, s) := \frac{F(\omega, s)}{I_F(\omega)}. \quad (6)$$

- A **solution to a BSDE** with terminal condition ξ and generator f is a triplet $(Y, Z, U) \in \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{N})$ which satisfies for all $t \in [0, T]$:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, U_s) ds - \int_t^T Z_s dW_s - \int_{[t, T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx). \quad (7)$$

The BSDE (7) itself will be denoted by (ξ, f) .

3 Main results

We start with a result about existence and uniqueness which is proved in Section 5.

Theorem 3.1 *There exists a unique solution to the BSDE (ξ, f) with $\xi \in L^2$ and generator $f: \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times L^2(\nu) \rightarrow \mathbb{R}$ satisfying the properties*

- (A1) For all $(y, z, u) : (\omega, s) \mapsto f(\omega, s, y, z, u)$ is progressively measurable.
 (A2) There are nonnegative, progressively measurable processes K_1, K_2 , and F with

$$C_K := \left\| \int_0^T \left(K_1(\cdot, s) + K_2(\cdot, s)^2 \right) ds \right\|_\infty < \infty \quad (8)$$

and $F \in L^2(\Omega; L^1([0, T]))$ (see (5)) such that for all (y, z, u) ,

$$|f(s, y, z, u)| \leq F(s) + K_1(s)|y| + K_2(s)(|z| + \|u\|), \quad \mathbb{P} \otimes \lambda\text{-a.e.}$$

- (A3) For λ -almost all s , the mapping $(y, z, u) \mapsto f(s, y, z, u)$ is \mathbb{P} -a.s. continuous. Moreover, there is a nonnegative function $\alpha \in L^1([0, T])$, $c > 0$ and a progressively measurable process β with $\int_0^T \beta(\omega, s)^2 ds < c$, \mathbb{P} -a.s. such that for all $(y, z, u), (y', z', u')$,

$$\begin{aligned} & (y - y') (f(s, y, z, u) - f(s, y', z', u')) \\ & \leq \alpha(s) \rho(|y - y'|^2) + \beta(s) |y - y'| (|z - z'| + \|u - u'\|), \quad \mathbb{P} \otimes \lambda\text{-a.e.}, \end{aligned}$$

where ρ is a nondecreasing, continuous and concave function from $[0, \infty[$ to itself, satisfying $\rho(0) = 0$, and $\int_{0+} \frac{1}{\rho(x)} dx = \infty$.

- (A4) The function ρ in (A3) satisfies $\limsup_{x \downarrow 0} \frac{\rho(x^2)}{x} = 0$.

If f satisfies only (A1)–(A3), then there exists at most one solution.

For $\rho(x) = x$, we are in the case of the ordinary monotonicity condition. Another example for a function ρ is given by

$$\rho(x) = 1 - \min\left(x, \frac{1}{e}\right)^{\min\left(x, \frac{1}{e}\right)}, \quad x \geq 0.$$

Remark 3.2

- Condition (A2) implies that $f(s, y, z, u)$ is integrable for a.e. $s \in [0, T]$ since, by Fubini's theorem,

$$\begin{aligned} & \int_0^T \mathbb{E} |f(s, y, z, u)| ds \\ & \leq \mathbb{E} \int_0^T [F(s) + K_1(s)|y| + K_2(s)(|z| + \|u\|)] ds < \infty. \end{aligned} \quad (9)$$

- If $\limsup_{x \downarrow 0} \frac{\rho(x^2)}{x} = 0$ is satisfied one can derive Lipschitz continuity of $f(s, y, z, u)$ in z and u from the monotonicity condition in (A3). We require (A4) since we later want to apply (Yin and Mao (2008), Theorem 2.1), where Lipschitz continuity in u is used to show uniqueness of solutions. If only (A1)–(A3) are satisfied but not (A4), and a Lipschitz condition in z, u holds nevertheless, all of the article's theorems remain valid. One can show that (A4) does not follow from the other conditions imposed on ρ in (A3): Assume a decreasing sequence $(x_n)_{n=0}^\infty$ with $x_0 = 1$ and $\lim_{n \rightarrow \infty} x_n = 0$. Define

$$\rho(x) := \begin{cases} \sqrt{x_n} & \text{if } x = x_n, n = 0, 1, 2, \dots \\ \sqrt{x} & \text{if } x > 1 \text{ or } x = 0. \end{cases}$$

and let ρ be continuous and piecewise linear on $]0, 1]$. The so defined ρ is a concave function with $\limsup_{x \downarrow 0} \frac{\rho(x)}{\sqrt{x}} = 1$. The sequence $(x_n)_{n=0}^\infty$ can be constructed such that $\int_0^1 \frac{1}{\rho(x)} dx = \infty$. For example, choose x_1 such that $\int_{x_1}^1 \frac{1}{\rho(x)} dx \geq 1$, and if x_n has been chosen find x_{n+1} such that

$$\int_{x_{n+1}}^{x_n} \frac{1}{\rho(x)} dx = \frac{1}{2} (\log(x_n) - \log(x_{n+1})) (\sqrt{x_n} + \sqrt{x_{n+1}}) \geq 1.$$

The next result shows how a solution to a BSDE can be approximated by a sequence of solutions of BSDEs which are driven by Lévy processes with a finite Lévy measure. We do this by approximating the underlying Lévy process defined through

$$X_t = at + \sigma W_t + \int_{[0,t] \times \{|x| > 1\}} x N(ds, dx) + \int_{[0,t] \times \{|x| \leq 1\}} x \tilde{N}(ds, dx)$$

for $n \geq 1$ by

$$X_t^n = at + \sigma W_t + \int_{[0,t] \times \{|x| > 1\}} x N(ds, dx) + \int_{[0,t] \times \{1/n \leq |x| \leq 1\}} x \tilde{N}(ds, dx).$$

The process X^n has a finite Lévy measure ν_n . Furthermore, note that the compensated Poisson random measure associated with X^n can be expressed as $\tilde{N}^n = \chi_{\{1/n \leq |x|\}} \tilde{N}$. Let

$$\begin{aligned} \mathcal{J}^0 &:= \{\Omega, \emptyset\} \vee \mathcal{N}, \\ \mathcal{J}^n &:= \sigma(X^n) \vee \mathcal{N}, \quad n \geq 1, \end{aligned} \quad (10)$$

where \mathcal{N} stands for the null sets of \mathcal{F} . Note that $(\mathcal{J}^n)_{n=0}^\infty$ forms a filtration. The notation $(\mathcal{J}^n)_{n=0}^\infty$ was chosen to indicate that this filtration describes the inclusion of smaller and smaller jumps of the Lévy process. We will use

$$\mathbb{E}_n \cdot := \mathbb{E}[\cdot | \mathcal{J}^n]$$

for the conditional expectation.

The intuitive idea now would be to work with a BSDE driven by X^n where one uses the data $(\mathbb{E}_n \xi, \mathbb{E}_n f)$. The problem is that the generator f needs to be progressively, and also jointly measurable w.r.t. (ω, t, y, z, u) , but it is not obvious whether the conditional expectation $\mathbb{E}_n f$ preserves this property from f . For BSDEs driven by a Brownian motion, this problem has been solved in (Ylinen (2017), Proposition 7.3), but this proposition does not apply to our situation. Therefore, we next propose a method for the construction of a unique progressively measurable and jointly measurable w.r.t. (ω, t, y, z, u) version of $\mathbb{E}_n f$.

Definition 3.3 (Definition of f_n) Assume that f satisfies (A1), (A2) and that $\mathbb{J} := (\mathcal{J}^{[s]})_{s \in [0, \infty[}$ is built using (10), where $[\cdot]$ denotes the floor function. Let ${}^o \mathbb{J} f$ be the optional projection of the process

$$[0, \infty[\times \Omega \times [0, T] \times \mathbb{R}^2 \times L^2(v) \rightarrow \mathbb{R},$$

$$(s, \omega, t, y, z, u) \mapsto f(\omega, t, y, z, u)$$

in the variables (s, ω) with respect to \mathbb{J} , and with parameters (t, y, z, u) . For each $n \geq 0$, assume that the filtration $\mathbb{F}^n := (\mathcal{F}_t^n)_{t \in [0, T]}$ is given by $\mathcal{F}_t^n := \mathcal{F}_t \cap \mathcal{J}^n$. Let f_n be the optional projection of

$$(\omega, t, y, z, u) \mapsto {}^{o, \mathbb{J}}f(n, \omega, t, y, z, u)$$

with respect to \mathbb{F}^n with parameters (y, z, u) .

The reason for using the filtration $(\mathcal{J}^{[s]})_{s \in [0, \infty[}$ instead of the $(\mathcal{J}^n)_{n=0}^\infty$ from (10) is that one can apply known measurability results w.r.t. right continuous filtrations instead of proving measurability here directly. Indeed, the optional projection ${}^{o, \mathbb{J}}f$ defined above is jointly measurable in (s, ω, t, y, z, u) . For this we refer to Meyer (1979), where optional and predictable projections of random processes depending on parameters were considered, and their uniqueness up to indistinguishability was shown.

It follows that for all (t, y, z, u) ,

$${}^{o, \mathbb{J}}f(n, t, y, z, u) = \mathbb{E}_n f(t, y, z, u), \quad \mathbb{P}\text{-a.s.}$$

Then, since f is $(\mathcal{F}_t)_{t \in [0, T]}$ -progressively measurable, for all $n \geq 0$, $t \in [0, T]$ and all (y, z, u) , it holds that

$$f_n(t, y, z, u) = \mathbb{E}_n f(t, y, z, u), \quad \mathbb{P}\text{-a.s.} \quad (11)$$

Hence, $f_n(t, y, z, u)$ is a jointly measurable version of $\mathbb{E}_n f(t, y, z, u)$ which is $(\mathcal{F}_t^n)_{t \in [0, T]}$ -optional, so especially it is progressively measurable.

We comment on the compatibility of the solutions (Y^n, Z^n, U^n) from the BSDE corresponding to $(\mathbb{E}_n \xi, f_n)$,

$$Y_t^n = \mathbb{E}_n \xi + \int_t^T f_n(s, Y_s^n, Z_s^n, U_s^n) ds - \int_t^T Z_s^n dW_s$$

$$- \int_{[t, T] \times \mathbb{R}_0} U_s^n(x) \tilde{N}^n(ds, dx)$$

with the space $S^2 \times L^2(W) \times L^2(\tilde{N})$:

The triplet $(Y^n, Z^n, U^n) \in S^2 \times L^2(W) \times L^2(\tilde{N})$ can be canonically embedded in the space $S^2 \times L^2(W) \times L^2(\tilde{N})$, basically by extending $U_s^n(x)$ onto \mathbb{R}_0 by defining $U_s^n(x) := 0$ for $|x| < \frac{1}{n}$. Moreover, recall that $\tilde{N}^n = \chi_{\{1/n \leq |x|\}} \tilde{N}$, so that

$$\int_{[t, T] \times \mathbb{R}_0} U_s^n(x) \tilde{N}^n(ds, dx) = \int_{[t, T] \times \mathbb{R}_0} U_s^n(x) \chi_{\{1/n \leq |x|\}} \tilde{N}(ds, dx).$$

Therefore, $(Y^n, Z^n, U^n \chi_{\mathbb{R} \setminus]-1/n, 1/n[})$ solves $(\mathbb{E}_n \xi, f_n)$ in $S^2 \times L^2(W) \times L^2(\tilde{N})$.

Theorem 3.4 Let $\xi \in L^2$ and let f satisfy (A1)–(A3). Assume that the BSDE driven by X^n with data $(\mathbb{E}_n \xi, f_n)$ (where f_n is given by Definition 3.3) has a unique solution denoted by (Y^n, Z^n, U^n) . If the solution (Y, Z, U) to (ξ, f) exists as well, then,

$$(Y^n, Z^n, U^n) \rightarrow (Y, Z, U)$$

in $L^2(W) \times L^2(W) \times L^2(\tilde{N})$ on $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, if f additionally satisfies (A4), then the mentioned solution triplets exist.

The benefit of this approximation becomes clear in the proof of the comparison theorem which we state next. There, we only need to prove the comparison result assuming a finite Lévy measure, since the general case then follows by approximation.

Another consequence of this approximation result concerns nonlinear expectations. (For a survey article on nonlinear expectations the reader is referred to Peng (2010)). In the case of Lévy processes, provided that $f(s, y, 0, 0) = 0$ for all s and y , the process Y_t has been described by Royer (2006) as a conditional nonlinear expectation, denoted by $\mathbb{E}_t^f \xi := Y_t$. Hence, our theorem implies that

$$\left(\mathbb{E}_t^{f_n} \mathbb{E}_n \xi \right)_{t \in [0, T]} \rightarrow \left(\mathbb{E}_t^f \xi \right)_{t \in [0, T]} \quad \text{in } L^2(W).$$

Theorem 3.5 Let f, f' be two generators satisfying the conditions (A1)–(A3) of Theorem 3.1 (f and f' may have different coefficients). We assume $\xi \leq \xi'$, \mathbb{P} -a.s. and for all (y, z, u) , $f(s, y, z, u) \leq f'(s, y, z, u)$, for $\mathbb{P} \otimes \lambda$ -a.a. $(\omega, s) \in \Omega \times [0, T]$. Moreover, assume that f or f' satisfy the condition (here formulated for f)

$$(A\gamma) \quad f(s, y, z, u) - f(s, y, z, u') \leq \int_{\mathbb{R}_0} (u(x) - u'(x)) \nu(dx), \quad \mathbb{P} \otimes \lambda\text{-a.e.} \\ \text{for all } u, u' \in L^2(\nu) \text{ with } u \leq u'.$$

Let (Y, Z, U) and (Y', Z', U') be the solutions to (ξ, f) and (ξ', f') , respectively. Then, $Y_t \leq Y'_t$, \mathbb{P} -a.s.

Proof The basic idea for this proof was inspired by the one of Theorem 8.3 in El Karoui et al. (2009).

Step 1:

In this step we assume that the Lévy measure ν is finite. We use Tanaka–Meyer’s formula (cf. Protter (2004), Theorem 70) to see that for $\eta(s) := 2\beta(s)^2 + \nu(\mathbb{R}_0)$,

$$\begin{aligned} e^{\int_0^t \eta(s) ds} (Y_t - Y'_t)_+^2 &= e^{\int_0^T \eta(s) ds} (\xi - \xi')_+^2 + M(t) \\ &\quad + \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{Y_s - Y'_s \geq 0\}} \left[2(Y_s - Y'_s)_+ (f(s, Y_s, Z_s, U_s) \right. \\ &\quad \left. - f'(s, Y'_s, Z'_s, U'_s)) \right. \\ &\quad \left. - |Z_s - Z'_s|^2 - \eta(s) |Y_s - Y'_s|^2 \right. \\ &\quad \left. - \int_{\mathbb{R}_0} ((Y_s - Y'_s + U_s(x) - U'_s(x))_+^2 - (Y_s - Y'_s)_+^2 \right. \\ &\quad \left. - 2(U_s(x) - U'_s(x))(Y_s - Y'_s)_+) \nu(dx) \right] ds. \end{aligned}$$

Here, $M(t)$ is a stochastic integral term having zero expectation which follows from $Y, Y' \in \mathcal{S}^2$ (this holds according to Theorem 3.1). Moreover, we used that on the set $\{\Delta Y_s \geq 0\}$ (where $\Delta Y := Y - Y'$) we have $(Y_s - Y'_s)_+ = |Y_s - Y'_s|$. Taking means and denoting the differences by $\Delta \xi := \xi - \xi'$, $\Delta Z := Z - Z'$, $\Delta U := U - U'$ and $\Delta f := f - f'$ leads us to

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 &= \mathbb{E} e^{\int_0^T \eta(s) ds} (\Delta \xi)_+^2 \\ &\quad + \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{\Delta Y_s \geq 0\}} \left[2 (\Delta Y_s)_+ (f(s, Y_s, Z_s, U_s) \right. \right. \\ &\quad \left. \left. - f'(s, Y'_s, Z'_s, U'_s)) - |\Delta Z_s|^2 - \eta(s) |\Delta Y_s|^2 \right. \right. \\ &\quad \left. \left. - \int_{\mathbb{R}_0} \left((\Delta Y_s + \Delta U_s(x))_+^2 - (\Delta Y_s)_+^2 \right. \right. \right. \\ &\quad \left. \left. \left. - 2 (\Delta U_s(x)) (\Delta Y_s)_+ \right) \nu(dx) \right] ds \right\}, \end{aligned} \quad (12)$$

We split up the set \mathbb{R}_0 into

$$B(\omega, s) = B = \{\Delta U_s(x) \geq -\Delta Y_s\} \text{ and } B^c.$$

Taking into account that $\xi \leq \xi'$, we estimate

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 &\leq \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{\Delta Y_s \geq 0\}} \left[2 (\Delta Y_s)_+ (f(s, Y_s, Z_s, U_s) \right. \right. \\ &\quad \left. \left. - f'(s, Y'_s, Z'_s, U'_s)) - |\Delta Z_s|^2 - \eta(s) |\Delta Y_s|^2 - \int_B |\Delta U_s(x)|^2 \nu(dx) \right. \right. \\ &\quad \left. \left. + \int_{B^c} \left((\Delta Y_s)_+^2 + 2(\Delta U_s(x))(\Delta Y_s)_+ \right) \nu(dx) \right] ds \right\}. \end{aligned} \quad (13)$$

We focus on the term $(\Delta Y_s)_+ (f(s, Y_s, Z_s, U_s) - f'(s, Y'_s, Z'_s, U'_s))$, and denoting $((Y, Z), (Y', Z'))$ by (Θ, Θ') , we derive from $f \leq f'$ that

$$\begin{aligned} (\Delta Y_s)_+ (f(s, \Theta_s, U_s) - f'(s, \Theta'_s, U'_s)) &= (\Delta Y_s)_+ (f(s, \Theta_s, U_s) - f(s, \Theta'_s, U'_s) \\ &\quad + f(s, \Theta'_s, U'_s) - f'(s, \Theta'_s, U'_s)) \\ &\leq (\Delta Y_s)_+ (f(s, \Theta_s, U_s) - f(s, \Theta'_s, U'_s)). \end{aligned}$$

We continue with the observation that on $\{\omega : \Delta Y_s > 0\}$ we have

$$B^c = \{\Delta U_s(x) < -\Delta Y_s\} \subseteq \{U'_s(x) > U_s(x)\},$$

so that

$$U'_s \chi_B + U_s \chi_{B^c} \leq U'_s \chi_B + U'_s \chi_{B^c} \quad \text{on} \quad \{\omega : \Delta Y_s > 0\}.$$

Therefore, we split $(\Delta Y_s)_+ (f(s, \Theta_s, U_s) - f(s, \Theta'_s, U'_s))$ into two terms; one we estimate with (A3) and the first inequality of (14), while for the other we use (A7):

$$\begin{aligned} (\Delta Y_s)_+ (f(s, \Theta_s, U_s) - f(s, \Theta'_s, U'_s)) &= (\Delta Y_s)_+ (f(s, \Theta_s, U_s \chi_B + U_s \chi_{B^c}) \\ &\quad - f(s, \Theta'_s, U'_s \chi_B + U'_s \chi_{B^c})) \\ &\quad + (\Delta Y_s)_+ (f(s, \Theta'_s, U'_s \chi_B + U'_s \chi_{B^c}) \\ &\quad - f(s, \Theta'_s, U'_s \chi_B + U'_s \chi_{B^c})) \\ &\leq \alpha(s) \rho \left((\Delta Y_s)_+^2 \right) + \beta(s)^2 (\Delta Y_s)_+^2 \\ &\quad + \frac{|\Delta Z_s|^2}{2} + \frac{\|\Delta U_s \chi_B\|^2}{2} \\ &\quad - \int_{\mathbb{R}_0} (\Delta Y_s)_+ \Delta U_s(x) \chi_{B^c} v(dx). \end{aligned}$$

Thus, by the last two inequalities, (13) evolves to

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 &\leq \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{\Delta Y_s \geq 0\}} \left[2\alpha(s) \rho \left((\Delta Y_s)_+^2 \right) + 2\beta(s)^2 (\Delta Y_s)_+^2 \right. \right. \\ &\quad \left. \left. + |\Delta Z_s|^2 + \|\Delta U_s \chi_B\|^2 - \int_{B^c} 2(\Delta Y_s)_+ (\Delta U_s(x)) v(dx) \right. \right. \\ &\quad \left. \left. - |\Delta Z_s|^2 - \eta(s) |\Delta Y_s|^2 - \int_B |\Delta U_s(x)|^2 v(dx) \right. \right. \\ &\quad \left. \left. + \int_{B^c} \left((\Delta Y_s)_+^2 + 2(\Delta Y_s)_+ (\Delta U_s(x)) \right) v(dx) \right] ds \right\}. \end{aligned}$$

Because of $\|\Delta U_s \chi_B\|^2 = \int_B |\Delta U_s(x)|^2 v(dx)$, we cancel out terms and get

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 &\leq \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{\Delta Y_s \geq 0\}} \left[2\alpha(s) \rho \left((\Delta Y_s)_+^2 \right) + 2\beta(s)^2 (\Delta Y_s)_+^2 \right. \right. \\ &\quad \left. \left. - \eta(s) |\Delta Y_s|^2 + \int_{B^c} (\Delta Y_s)_+^2 v(dx) \right] ds \right\}. \end{aligned}$$

Bounding $\int_{B^c} (\Delta Y_s)_+^2 v(dx)$ by $v(\mathbb{R}_0) (\Delta Y_s)_+^2$, leads us to

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 &\leq \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \chi_{\{\Delta Y_s \geq 0\}} \left[2\alpha(s) \rho \left((\Delta Y_s)_+^2 \right) \right. \right. \\ &\quad \left. \left. + (2\beta(s)^2 + v(\mathbb{R}_0)) (\Delta Y_s)_+^2 - \eta(s) |\Delta Y_s|^2 \right] ds \right\}. \end{aligned}$$

It remains, also using the definition of η ,

$$\mathbb{E} e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 \leq \mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} 2\alpha(s) \rho \left((\Delta Y_s)_+^2 \right) ds.$$

The term $e^{\int_0^T \eta(\tau) d\tau}$ is \mathbb{P} -a.s. bounded by a constant $C > 0$. Thus, by the concavity of ρ , we arrive at

$$\mathbb{E} (\Delta Y_t)_+^2 \leq \mathbb{E} \left[e^{\int_0^t \eta(s) ds} (\Delta Y_t)_+^2 \right] \leq \int_t^T 2C\alpha(s) \rho \left(\mathbb{E} (\Delta Y_s)_+^2 \right) ds.$$

Then, the Bihari–LaSalle inequality (Proposition 5.2)—a generalization of Gronwall’s inequality—shows that $\mathbb{E}(\Delta Y_t)_+^2 = 0$ for all $t \in [0, T]$, which is the desired result for $\nu(\mathbb{R}_0) < \infty$.

Step 2:

The goal of this step is to extend the result of the first step to general Lévy measures. We adapt the notation of Theorem 3.4 for $Y^n, Y^{n'}, f_n$, and f'_n . Now, we claim that for solutions Y^n and $Y^{n'}$ of $(\mathbb{E}_n \xi, f_n)$ and $(\mathbb{E}_n \xi', f'_n)$, Step 1 granted that $Y^n \leq Y^{n'}$: Indeed, $f_n \leq f'_n$ holds by the monotonicity of \mathbb{E}_n , and also (A γ) holds for f_n if it did for f . One notes that the process X^n which is related to $(\mathbb{E}_n \xi, f_n)$ and $(\mathbb{E}_n \xi', f'_n)$ has a finite Lévy measure ν_n satisfying $\nu_n(|x| < \frac{1}{n}) = 0$, while in (A γ) we still have ν . However, the solution processes U^n and $U^{n'}$ are zero for $|x| < \frac{1}{n}$ (see the comment before Theorem 3.4).

Hence, we need (A γ) only for u and u' which are zero for $|x| < \frac{1}{n}$, and for those u and u' we may replace ν by ν_n and then apply Step 1. Finally, the convergence of the sequences to the solutions Y and Y' of (ξ, f) and (ξ', f') , respectively, in $L^2(W)$ shows $Y \leq Y'$, and our theorem is proven. \square

4 Auxiliary results

We will frequently use the following basic algebraic inequalities (special cases of Young’s inequality) which hold for all $R > 0$:

$$ab \leq \frac{a^2}{2R} + \frac{Rb^2}{2} \quad \text{and} \quad ab \leq \frac{Ra}{2} + \frac{ab^2}{2R}. \quad (14)$$

The following proposition states, roughly speaking, that for the BSDEs considered here it is sufficient to find solution processes of a BSDE in the (larger) space $L^2(W) \times L^2(W) \times L^2(\tilde{N})$.

Proposition 4.1 *If $(Y, Z, U) \in L^2(W) \times L^2(W) \times L^2(\tilde{N})$ is a triplet of processes that satisfies the BSDE (ξ, f) with $\xi \in L^2$ and (A1), (A2), then (Y, Z, U) is a solution to (7), i.e., $(Y, Z, U) \in \mathcal{S}^2 \times L^2(W) \times L^2(\tilde{N})$. In particular, there exists a constant $C_1 > 0$ such that*

$$\|Y\|_{\mathcal{S}^2}^2 + \|Z\|_{L^2(W)}^2 + \|U\|_{L^2(\tilde{N})}^2 \leq e^{C_1(1+C_K)^2} (\mathbb{E}|\xi|^2 + \mathbb{E}I_F^2),$$

where C_K was defined in (8) and I_F in (6).

Proof Since (Y, Z, U) satisfies (7), it holds that

$$\begin{aligned} |Y_t|^2 &= Y_t \xi + Y_t \int_t^T f(s, Y_s, Z_s, U_s) ds - Y_t \int_t^T Z_s dW_s \\ &\quad - Y_t \int_{[t, T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx). \end{aligned}$$

We apply the first inequality of (14), where Y_t takes the role of a , to get for an arbitrary $R > 0$:

$$|Y_t|^2 \leq \frac{3|Y_t|^2}{2R} + \frac{R|\xi|^2}{2} + \frac{R}{2} \left(\left| \int_t^T Z_s dW_s \right|^2 + \left| \int_{[t,T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right) \\ + |Y_t| \int_t^T |f(s, Y_s, Z_s, U_s)| ds.$$

Condition (A2) implies

$$|Y_t|^2 \leq \frac{3|Y_t|^2}{2R} + \frac{R|\xi|^2}{2} + \frac{R}{2} \left(\left| \int_t^T Z_s dW_s \right|^2 + \left| \int_{[t,T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right) \\ + |Y_t| \int_t^T (F(s) + K_1(s)|Y_s| + K_2(s)(|Z_s| + \|U_s\|)) ds.$$

We estimate with the help of the inequalities (14),

$$|Y_t|F(s) \leq K_F(s) \left(\frac{|Y_t|^2}{2R} + \frac{RI_F^2}{2} \right), \\ K_1(s)|Y_t||Y_s| \leq K_1(s) \left(\frac{|Y_t|^2}{2R} + \frac{R|Y_s|^2}{2} \right), \\ |Y_t|K_2(s)(|Z_s| + \|U_s\|) \leq \frac{K_2(s)^2|Y_t|^2}{2R} + R(|Z_s|^2 + \|U_s\|^2).$$

Hence,

$$|Y_t|^2 \leq \frac{|Y_t|^2}{2R} \left(4 + \int_t^T (K_F(s) + K_1(s) + K_2(s)^2) ds \right) + \frac{R|\xi|^2}{2} \\ + \frac{R}{2} \left(\left| \int_t^T Z_s dW_s \right|^2 + \left| \int_{[t,T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right) \\ + \frac{R}{2} I_F^2 \int_t^T K_F(s) ds + R \int_t^T (|Z_s|^2 + \|U_s\|^2) ds + \int_t^T \frac{RK_1(s)|Y_s|^2}{2} ds.$$

Note that $\int_0^T K_F(s) ds = 1$ and choose $R = R_0 := 5 + \int_0^T (K_1(s) + K_2(s)^2) ds$ so that

$$|Y_t|^2 \leq R_0 \left[|\xi|^2 + \sup_{t \in [0, T]} \left(\left| \int_t^T Z_s dW_s \right|^2 + \left| \int_{[t,T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right) \right. \\ \left. + I_F^2 + 2 \int_0^T |Z_s|^2 + \|U_s\|^2 ds + \int_t^T K_1(s)|Y_s|^2 ds \right].$$

Since Y is a càdlàg process, we may apply (46) from the appendix which leads to

$$|Y_t|^2 \leq R_0 e^{R_0 \int_0^T K_1(s) ds} \left[|\xi|^2 + I_F^2 + 2 \int_0^T |Z_s|^2 + \|U_s\|^2 ds \right. \\ \left. + \sup_{t \in [0, T]} \left(\left| \int_t^T Z_s dW_s \right|^2 + \left| \int_{[t,T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right) \right].$$

The inequality $(a + b)^2 \leq 2a^2 + 2b^2$ and then Doob's martingale inequality used on

$$\sup_{t \in [0, T]} \left(\left| \int_0^T Z_s dW_s - \int_0^t Z_s dW_s \right|^2 + \left| \int_{[0, T] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) - \int_{[0, t] \times \mathbb{R}_0} U_s(x) \tilde{N}(ds, dx) \right|^2 \right)$$

yield, since a.s. $R_0 \leq 5 + C_K$ and $\int_0^T K_1(s) ds \leq C_K$,

$$\mathbb{E} \sup_{t \in [0, T]} |Y_t|^2 \leq c_1 \left[\mathbb{E} |\xi|^2 + \mathbb{E} I_F^2 + 12 \mathbb{E} \int_0^T (|Z_s|^2 + \|U_s\|^2) ds \right] \quad (15)$$

with

$$c_1 = (5 + C_K) e^{(5 + C_K) C_K}. \quad (16)$$

For a progressively measurable process η , which we will determine later, Itô's formula implies that

$$\begin{aligned} & |Y_0|^2 + \int_0^T e^{\int_0^s \eta(\tau) d\tau} (\eta(s) |Y_s|^2 + |Z_s|^2 + \|U_s\|^2) ds \\ &= M(0) + e^{\int_0^T \eta(s) ds} |\xi|^2 + \int_0^T 2e^{\int_0^s \eta(\tau) d\tau} Y_s f(s, Y_s, Z_s, U_s) ds, \end{aligned} \quad (17)$$

where

$$\begin{aligned} M(t) &= - \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} Y_s Z_s dW_s \\ &\quad - \int_{[t, T] \times \mathbb{R}_0} 2e^{\int_0^s \eta(\tau) d\tau} \left((Y_{s-} + U_s(x))^2 - Y_{s-}^2 \right) \tilde{N}(ds, dx). \end{aligned} \quad (18)$$

Provided that $\left\| \int_0^T \eta(\tau) d\tau \right\|_{L^\infty(\mathbb{P})} < \infty$, one gets $\mathbb{E} M(t) = 0$ as a consequence of (15) and the Burkholder–Davis–Gundy inequality (see, for instance, (He et al. (1992), Theorem 10.36)), where the term $((Y_{s-} + U_s(x))^2 - Y_{s-}^2)^2$ appearing in the integrand can be estimated by

$$(|Y_{s-} + U_s(x)| + |Y_{s-}|)^2 (|Y_{s-} + U_s(x)| - |Y_{s-}|)^2 \leq 4 \sup_{r \in [0, T]} |Y_r|^2 |U_s(x)|^2.$$

By (A2) and (14), we have

$$\begin{aligned} |Y_s| |f(s, Y_s, Z_s, U_s)| &\leq |Y_s| [F(s) + K_1(s) |Y_s| + K_2(s) (|Z_s| + \|U_s\|)] \\ &\leq F(s) |Y_s| + K_1(s) |Y_s|^2 + 2R \frac{K_2(s)^2 |Y_s|^2}{2} \\ &\quad + \frac{|Z_s|^2 + \|U_s\|^2}{2R}. \end{aligned}$$

We use this estimate for $R = 2$, and taking the expectation in (17), we have

$$\begin{aligned} \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} \left(\eta(s) |Y_s|^2 + |Z_s|^2 + \|U_s\|^2 \right) ds &\leq \mathbb{E} e^{\int_0^T \eta(s) ds} |\xi|^2 + \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} \left(\frac{|Z_s|^2 + \|U_s\|^2}{2} \right) ds \\ &\quad + 2 \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} F(s) ds \sup_{t \in [0, T]} |Y_t| \\ &\quad + \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} 2 \left(K_1(s) + 2K_2(s)^2 \right) Y_s^2 ds. \end{aligned} \quad (19)$$

Then, we choose $\eta(s) = 2 \left(K_1(s) + 2K_2(s)^2 \right)$ and subtract the terms containing Y , Z , and U from the left hand side of (19). Moreover, we apply the first inequality of (14) to the term containing the supremum. It follows that

$$\begin{aligned} \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} \left(|Z_s|^2 + \|U_s\|^2 \right) ds &\leq 2 \mathbb{E} \left[e^{\int_0^T \eta(s) ds} |\xi|^2 \right] + 2R \mathbb{E} \left[\int_0^T e^{\int_0^s \eta(\tau) d\tau} F(s) ds \right]^2 \\ &\quad + \frac{2}{R} \mathbb{E} \sup_{t \in [0, T]} |Y_t|^2. \end{aligned} \quad (20)$$

Note that

$$\mathbb{E} \int_0^T \left(|Z_s|^2 + \|U_s\|^2 \right) ds \leq \mathbb{E} \int_0^T e^{\int_0^s \eta(\tau) d\tau} \left(|Z_s|^2 + \|U_s\|^2 \right) ds.$$

Hence, by (20) and $\int_0^T \eta(\tau) d\tau \leq 4C_K$ a.s., we have

$$\mathbb{E} \int_0^T \left(|Z_s|^2 + \|U_s\|^2 \right) ds \leq 2e^{4C_K} \mathbb{E} |\xi|^2 + 2Re^{8C_K} \mathbb{E} I_F^2 + \frac{2}{R} \mathbb{E} \sup_{t \in [0, T]} |Y_t|^2. \quad (21)$$

Now, we can plug in (21) into (15) and vice versa which yields for $R := 48c_1$ that

$$\mathbb{E} \sup_{t \in [0, T]} |Y_t|^2 \leq (2c_1 + 48c_1 e^{4C_K}) \mathbb{E} |\xi|^2 + (2c_1 + (48c_1)^2 e^{8C_K}) \mathbb{E} I_F^2,$$

and

$$\mathbb{E} \int_0^T \left(|Z_s|^2 + \|U_s\|^2 \right) ds \leq \left(\frac{1}{12} + 4e^{4C_K} \right) \mathbb{E} |\xi|^2 + \left(\frac{1}{12} + 192c_1 e^{8C_K} \right) \mathbb{E} I_F^2.$$

Using (16) it is easy to see that there exists a constant $C_1 > 0$ such that each factor in front of the expectations on the right side of the previous two inequalities is less than $e^{C_1(1+C_K)^2}$.

□

Our next proposition will be an L^2 a-priori estimate for BSDEs of our type. For the Brownian case, L^p a-priori estimates are done for $p \in [1, \infty[$ in Briand et al. (2003), and for quadratic BSDEs, for $p \in [2, \infty[$ in Geiss and Ylinen (2018). For BSDEs with jumps, for $p \in]1, \infty[$, see Kruse and Popier (2016, 2017); while Becherer et al. (2018) contains an a-priori estimate w.r.t. L^∞ . The following assertion is similar to (Barles et al. (1997), Proposition 2.2), but fits our extended setting.

Proposition 4.2 Let $\xi, \xi' \in L^2$ and let f, f' be two generator functions satisfying (A1)–(A3), where the bounds in (A2) and the coefficients in (A3) may differ for f and f' . The coefficients of f' in (A3) will be referred to as α' and β' . Moreover, let the triplets (Y, Z, U) and $(Y', Z', U') \in L^2(W) \times L^2(W) \times L^2(\tilde{N})$, satisfy the BSDEs (ξ, f) and (ξ', f') , respectively.

Then,

$$\begin{aligned} & \|Y - Y'\|_{L^2(W)}^2 + \|Z - Z'\|_{L^2(W)}^2 + \|U - U'\|_{L^2(\tilde{N})}^2 \\ & \leq h\left(a, b, \mathbb{E}|\xi - \xi'|^2 + 2\mathbb{E}\int_0^T |Y_t - Y'_t| |f(t, Y_t, Z_t, U_t) - f'(t, Y_t, Z_t, U_t)| dt\right), \end{aligned}$$

where $a = \int_0^T \alpha'(s) ds$, $b = \left\| \int_0^T \beta'(s)^2 ds \right\|_\infty$, and

$$h :]0, \infty[\times]0, \infty[\times [0, \infty[\rightarrow [0, \infty[$$

is a function such that $h(a, b, x) \rightarrow 0 = h(a, b, 0)$ if $x \rightarrow 0$.

Proof We start with the following observation gained by Itô's formula for the difference of the BSDEs (ξ, f) and (ξ', f') . We denote differences of expressions by Δ . If $\eta = 4\beta'(s)^2$, we have analogously to (17)

$$\begin{aligned} & e^{\int_0^t \eta(s) ds} |\Delta Y_t|^2 + \int_t^T e^{\int_0^s \eta(\tau) d\tau} \left(\eta(s) |\Delta Y_s|^2 + |\Delta Z_s|^2 + \|\Delta U_s\|^2 \right) ds \\ & = e^{\int_0^T \eta(s) ds} |\Delta \xi|^2 + M(t) \\ & + \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} \Delta Y_s (f(s, Y_s, Z_s, U_s) - f'(s, Y'_s, Z'_s, U'_s)) ds, \end{aligned} \quad (22)$$

where

$$\begin{aligned} M(t) &= - \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} \Delta Y_s \Delta Z_s dW_s \\ &\quad - \int_{[t, T] \times \mathbb{R}_0} 2e^{\int_0^s \eta(\tau) d\tau} \left((\Delta Y_{s-} + \Delta U_s(x))^2 - \Delta Y_{s-}^2 \right) \tilde{N}(ds, dx). \end{aligned}$$

By the same reasoning as for (18), we have $\mathbb{E}M(t) = 0$. We now proceed with the (standard) arguments similar to those used for (17)–(19). By (A3) and the first inequality from (14),

$$\begin{aligned} \Delta Y_s (f'(s, Y_s, Z_s, U_s) - f'(s, Y'_s, Z'_s, U'_s)) &\leq \alpha'(s) \rho(|\Delta Y_s|^2) + \beta'(s) |\Delta Y_s| (|\Delta Z_s| \\ &\quad + \|\Delta U_s\|) \leq \alpha'(s) \rho(|\Delta Y_s|^2) \\ &\quad + \frac{\beta'(s)^2 |\Delta Y_s|^2}{R} + \frac{R (|\Delta Z_s|^2 + \|\Delta U_s\|^2)}{2}. \end{aligned} \quad (23)$$

Taking the expectation in (22) and then using (23) with $R = 1$ (such that we can cancel out the terms with Z and U on the left side), leads to

$$\begin{aligned} \mathbb{E} e^{\int_0^t \eta(s) ds} |\Delta Y_t|^2 + \mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} \eta(s) |\Delta Y_s|^2 ds &\leq \mathbb{E} e^{\int_0^T \eta(s) ds} |\Delta \xi|^2 \\ &+ \mathbb{E} \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} \Delta Y_s \cdot (\Delta f)(s, Y_s, Z_s, U_s) ds \\ &+ \mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} \left(2\alpha'(s) \rho(|\Delta Y_s|^2) + \beta'(s)^2 |\Delta Y_s|^2 \right) ds. \end{aligned}$$

The choice $\eta(s) = 4\beta'(s)^2$ and the fact that $\int_0^T \beta'(s)^2 ds \leq b$ a.s. leads to

$$\begin{aligned} \mathbb{E} |\Delta Y_t|^2 &\leq e^{4b} \left(\mathbb{E} |\Delta \xi|^2 + \mathbb{E} \int_t^T 2|\Delta Y_s| |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right) \\ &+ e^{4b} \int_t^T 2\alpha'(s) \rho(|\Delta Y_s|^2) ds, \end{aligned}$$

since ρ is a concave function.

By Proposition 5.2, a backward version of the Bihari–LaSalle inequality, shows

$$\begin{aligned} \sup_{t \in [0, T]} \mathbb{E} |\Delta Y_t|^2 &\leq \\ G^{-1} \left\{ G \left[e^{4b} \left(\mathbb{E} |\Delta \xi|^2 + \mathbb{E} \int_0^T 2|\Delta Y_s| |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right) \right] + 2e^{4b} \int_0^T \alpha'(s) ds \right\}, \end{aligned} \quad (24)$$

where $G(x) = \int_1^x \frac{1}{\rho(h)} dh$.

If we take the expectation in (22) but choose this time (23) with $R = \frac{1}{2}$ and omit $\mathbb{E} e^{\int_0^t \eta(s) ds} |\Delta Y_t|^2$, then

$$\begin{aligned} &\mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} \left(\eta(s) |\Delta Y_s|^2 + |\Delta Z_s|^2 + \|\Delta U_s\|^2 \right) ds \\ &\leq \mathbb{E} e^{\int_0^T \eta(s) ds} |\Delta \xi|^2 + \mathbb{E} \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} \Delta Y_s \cdot (\Delta f)(s, Y_s, Z_s, U_s) ds \\ &+ \mathbb{E} \left\{ \int_t^T e^{\int_0^s \eta(\tau) d\tau} \left(2\alpha'(s) \rho(|\Delta Y_s|^2) + 4\beta'(s)^2 |\Delta Y_s|^2 + \frac{|\Delta Z_s|^2 + \|\Delta U_s\|^2}{2} \right) ds \right\}. \end{aligned}$$

We subtract the quadratic terms with ΔY , ΔZ , and ΔU which appear on the right hand side. This results in the inequality

$$\begin{aligned} &\mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} (|\Delta Z_s|^2 + \|\Delta U_s\|^2) ds \\ &\leq 2 \left(\mathbb{E} e^{\int_0^T \eta(s) ds} |\Delta \xi|^2 + \mathbb{E} \int_t^T 2e^{\int_0^s \eta(\tau) d\tau} |\Delta Y_s| \cdot |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right. \\ &\quad \left. + \mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} 2\alpha'(s) \rho(|\Delta Y_s|^2) ds \right). \end{aligned}$$

We continue our estimate by

$$\begin{aligned} & \mathbb{E} \int_t^T e^{\int_0^s \eta(\tau) d\tau} (|\Delta Z_s|^2 + \|\Delta U_s\|^2) ds \\ & \leq 2e^{4b} \left(\mathbb{E} |\Delta \xi|^2 + \mathbb{E} \int_t^T 2|\Delta Y_s| \cdot |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right. \\ & \quad \left. + 2 \int_t^T \alpha'(s) ds \rho \left(\sup_{s \in [0, T]} \mathbb{E} |\Delta Y_s|^2 \right) \right), \end{aligned} \quad (25)$$

since $\eta(s) = 4\beta'(s)^2$. We put

$$\begin{aligned} H := G^{-1} \Big\{ & G \left[e^{4b} \left(\mathbb{E} |\Delta \xi|^2 + \mathbb{E} \int_0^T 2|\Delta Y_s| |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right) \right] \\ & + 2e^{4b} \int_0^T \alpha'(s) ds \Big\} \end{aligned}$$

so that (24) reads now as $\sup_{t \in [0, T]} \mathbb{E} |\Delta Y_t|^2 \leq H$. If we add this inequality to (25) and note that $\rho \left(\sup_{t \in [0, T]} \mathbb{E} |\Delta Y_s|^2 \right) \leq \rho(H)$, we have

$$\begin{aligned} & \sup_{s \in [0, T]} \mathbb{E} |\Delta Y_t|^2 + \mathbb{E} \int_0^T |\Delta Z_s|^2 ds + \mathbb{E} \int_0^T \|\Delta U_s\|^2 ds \\ & \leq 2e^{4b} \left(\mathbb{E} |\Delta \xi|^2 + \mathbb{E} \int_0^T 2|\Delta Y_s| \cdot |(\Delta f)(s, Y_s, Z_s, U_s)| ds \right) \\ & \quad + \left(2e^{4b} \int_0^T \alpha'(s) ds + 1 \right) \cdot (\text{id} + \rho)(H). \end{aligned}$$

Note that the integral condition on ρ implies that, if the argument of G approaches zero, then the right hand side vanishes. \square

The following Lemma will be used to estimate the expectation of integrals which contain $|Y_s|^2$.

Lemma 4.3 *Let $\xi \in L^2$ and assume that (A1) and (A2) hold. If (Y, Z, U) is a solution to (ξ, f) and H is a nonnegative, progressively measurable process with $\left\| \int_0^T H(s) ds \right\|_\infty < \infty$, then*

$$\begin{aligned} \mathbb{E} \int_0^T H(s) |Y_s|^2 ds & \leq e^{2C_K} \mathbb{E} \int_0^T H(s) ds |\xi|^2 \\ & \quad + 2e^{2C_K} \left\| \int_0^T H(s) ds \cdot I_F \right\|_2 \|Y\|_{\mathcal{S}^2}. \end{aligned} \quad (26)$$

Proof From the relations (17), (18) and integration by parts applied to the term $\int_0^T H(s) ds \cdot e^{\int_0^T \eta(s) ds} |Y_T|^2$, we get

$$\begin{aligned} \int_0^T H(s) ds \cdot e^{\int_0^T \eta(r) dr} |Y_T|^2 & = \int_0^T H(s) e^{\int_0^s \eta(\tau) d\tau} |Y_s|^2 ds - \int_0^T \left(\int_0^s H(\tau) d\tau \right) dM(s) \\ & \quad + \int_0^T \left(\int_0^s H(r) dr \right) e^{\int_0^s \eta(\tau) d\tau} \left(\eta(s) |Y_s|^2 + |Z_s|^2 + \|U_s\|^2 \right. \\ & \quad \left. - 2Y_s f(s, Y_s, Z_s, U_s) \right) ds. \end{aligned}$$

We take expectations and rearrange the equation so that

$$\begin{aligned} \mathbb{E} \int_0^T H(s) e^{\int_0^s \eta(\tau) d\tau} |Y_s|^2 ds &\leq \mathbb{E} \left[\int_0^T H(s) ds \cdot e^{\int_0^T \eta(s) ds} |\xi|^2 \right] \\ &\quad + \mathbb{E} \left[\int_0^T \left(\int_0^s H(\tau) d\tau \right) e^{\int_0^s \eta(\tau) d\tau} (2Y_s f(s, Y_s, Z_s, U_s) \right. \\ &\quad \left. - \eta(s) |Y_s|^2 - |Z_s|^2 - \|U_s\|^2) ds \right]. \end{aligned}$$

By Assumption (A2) and (14), we have

$$\begin{aligned} 2Y_s f(s, Y_s, Z_s, U_s) &\leq 2|Y_s|F(s) + 2K_1(s)|Y_s|^2 \\ &\quad + 2K_2(s)|Y_s|(|Z_s| + \|U_s\|) \leq 2|Y_s|F(s) \\ &\quad + 2K_1(s)|Y_s|^2 + 2K_2(s)^2|Y_s|^2 + |Z_s|^2 + \|U_s\|^2, \end{aligned}$$

so that for $\eta(s) = 2K_1(s) + 2K_2(s)^2$ it follows

$$\begin{aligned} \mathbb{E} \int_0^T H(s) |Y_s|^2 ds &\leq \mathbb{E} \left[\int_0^T H(s) ds \cdot e^{\int_0^T \eta(s) ds} |\xi|^2 \right] \\ &\quad + 2\mathbb{E} \left[\int_0^T \left(\int_0^s H(\tau) d\tau \right) e^{\int_0^s \eta(\tau) d\tau} F(s) |Y_s| ds \right] \\ &\leq e^{2C_K} \mathbb{E} \left[\int_0^T H(s) ds \cdot |\xi|^2 \right] \\ &\quad + 2e^{2C_K} \left\| \int_0^T H(s) ds \cdot I_F \right\|_2 \|Y\|_{\mathcal{S}^2}. \end{aligned} \quad (27)$$

□

5 Proofs of Theorems 3.1 and 3.4

5.1 Proof of Theorem 3.1

Step 1: Uniqueness

Uniqueness of the solution is a consequence of Proposition 4.2, since the terms $|\xi - \xi'|$ and $|f(s, Y_s, Z_s, U_s) - f'(s, Y_s, Z_s, U_s)|$ are zero.

The proof of existence will be split up in further steps.

Step 2:

In this step, we construct an approximating sequence of generators $f^{(n)}$ for f and show several estimates for the solution processes (Y^n, Z^n, U^n) to the BSDEs $(\xi, f^{(n)})$.

For $n \geq 1$, define $c_n(z) := \min(\max(-n, z), n)$ and $\tilde{c}_n(u) \in L^2(\nu)$ to be the projection of u onto $\{v \in L^2(\nu) : \|v\| \leq n\}$. Let (Y^n, Z^n, U^n) be the unique solution of the BSDE $(\xi, f^{(n)})$, with the definitions

$$\hat{f}^{(n)}(\omega, s, y, z, u) := f(\omega, s, y, c_n(z), \tilde{c}_n(u)),$$

and

$$f^{(n)}(\omega, s, y, z, u) := \text{sign} \left(\hat{f}^{(n)}(\omega, s, y, z, u) \right) \\ \times \left[F(\omega, s) \wedge n + (K_1(\omega, s) \wedge n)|y| + (K_2(\omega, s) \wedge n)(|c_n(z)| + \|\tilde{c}_n(u)\|) \right] \\ \text{if } |\hat{f}^{(n)}(\omega, s, y, z, u)| > F(\omega, s) \wedge n + (K_1(\omega, s) \wedge n)|y| \\ + (K_2(\omega, s) \wedge n)(|c_n(z)| + \|\tilde{c}_n(u)\|),$$

and

$$f^{(n)}(\omega, s, y, z, u) := \hat{f}^{(n)}(\omega, s, y, z, u) \quad \text{else.}$$

Note that $f^{(n)}$ satisfies (A1)–(A4), with the same coefficients as f . Moreover, by (A4), $f^{(n)}$ satisfies a Lipschitz condition with respect to u (see Remark 3.2). Thus, thanks to (Yin and Mao (2008), Theorem 2.1), $(\xi, f^{(n)})$ has a unique solution (Y^n, Z^n, U^n) . Moreover, by Proposition 4.1, we get that

$$\|Y^n\|_{S^2}^2 + \|Z^n\|_{L^2(W)}^2 + \|U^n\|_{L^2(\tilde{N})}^2 \leq e^{C_1(1+C_K)^2} (\mathbb{E}|\xi|^2 + \mathbb{E}I_F^2) < \infty, \quad (28)$$

uniformly in n . This implies that the families

$$\left(\sup_{t \in [0, T]} |Y_t^n|, n \geq 0 \right), (|Y^n|, n \geq 0) \text{ and } (|Z^n| + \|U^n\|, n \geq 0)$$

are uniformly integrable with respect to \mathbb{P} , $\mathbb{P} \otimes \lambda$ and $\mathbb{P} \otimes \lambda$, respectively.

Step 3 :

The goal of this step is to use Proposition 4.2 to get convergence of $(Y^n, Z^n, U^n)_n$ in $L^2(W) \times L^2(W) \times L^2(\tilde{N})$ for a subsequence $n_k \uparrow \infty$ if $\delta_{n_k, n_l} \rightarrow 0$ for $k > l \rightarrow \infty$, where

$$\delta_{n,m} := \mathbb{E} \int_0^T |Y_s^n - Y_s^m| |f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) - f^{(m)}(s, Y_s^n, Z_s^n, U_s^n)| ds.$$

We observe that the difference of the generators is zero if two conditions are satisfied at the same time: First, if $|Z^n|, \|U_s^n\| < n$, and additionally, by the cut-off procedure for F, K_1, K_2 , if

$$n > \max(F(\omega, s), K_1(\omega, s), K_2(\omega, s)) =: k(\omega, s).$$

Thus, putting

$$\chi_n(s) := \chi_{\{|Z_s^n| > n\} \cup \{\|U_s^n\| > n\} \cup \{k(s) > n\}}, \quad (29)$$

we have

$$\delta_{n,m} = \mathbb{E} \int_0^T |Y_s^n - Y_s^m| |f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) - f^{(m)}(s, Y_s^n, Z_s^n, U_s^n)| \chi_n(s) ds \\ \leq \mathbb{E} \left\{ \int_0^T 2|Y_s^n - Y_s^m| \chi_n(s) \times (F(s) + K_1(s)|Y_s^n| + K_2(s)(|Z_s^n| + \|U_s^n\|)) ds \right\},$$

due to the linear growth condition (A2). We estimate this further by

$$\begin{aligned}\delta_{n,m} &\leq \mathbb{E} \int_0^T \chi_n(s) F(s) ds \left(\sup_{r \in [0,T]} |Y_r^n| + \sup_{r \in [0,T]} |Y_r^m| \right) \\ &\quad + \mathbb{E} \int_0^T \chi_n(s) K_2(s) (|Y_s^n| + |Y_s^m|) (|Z_s^n| + \|U_s^n\|) ds \\ &\quad + \mathbb{E} \int_0^T 2|Y_s^n - Y_s^m| |Y_s^n| \chi_n(s) K_1(s) ds \\ &=: \delta_{n,m}^{(1)} + \delta_{n,m}^{(2)} + \delta_{n,m}^{(3)}.\end{aligned}\quad (30)$$

For $\delta_{n,m}^{(1)}$, we use the Cauchy–Schwarz inequality,

$$\delta_{n,m}^{(1)} \leq 2 \left(\mathbb{E} \left| \int_0^T \chi_n(s) F(s) ds \right|^2 \right)^{\frac{1}{2}} (\|Y^n\|_{\mathcal{S}^2} + \|Y^m\|_{\mathcal{S}^2}).$$

Since $\sup_n \|Y^n\|_{\mathcal{S}^2} < \infty$ according to (28), it remains to show that the integral term converges to 0 for a subsequence.

Since $|Z_s^n|$ and $\|U_s^n\|$ are uniformly integrable w.r.t. $\mathbb{P} \otimes \lambda$, we imply from (29) that $\chi_n \rightarrow 0$ in $L^1(\mathbb{P} \otimes \lambda)$. Hence, there exists a subsequence $(n_k)_{k \geq 1}$ such that

$$\chi_{n_k} \rightarrow 0 \quad k \rightarrow \infty, \quad \mathbb{P} \otimes \lambda\text{-a.e.} \quad (31)$$

By dominated convergence, we have $\mathbb{E} \left| \int_0^T \chi_{n_k}(s) F(s) ds \right|^2 \rightarrow 0$ for $k \rightarrow \infty$ since $F \in L^2(\Omega; L^1([0, T]))$.

For $\delta_{n,m}^{(2)}$, we start with the Cauchy–Schwarz inequality and get

$$\begin{aligned}\delta_{n,m}^{(2)} &\leq 2 \sup_k \left[\|Z^k\|_{L^2(W)} + \|U^k\|_{L^2(\tilde{N})} \right] \\ &\quad \times \left[\mathbb{E} \int_0^T \chi_n(s) K_2(s)^2 (|Y_s^n|^2 + |Y_s^m|^2) ds \right]^{\frac{1}{2}}.\end{aligned}$$

By Lemma 4.3,

$$\begin{aligned}&\mathbb{E} \int_0^T \chi_n(s) K_2(s)^2 (|Y_s^n|^2 + |Y_s^m|^2) ds \\ &\leq 2e^{2C_K} \mathbb{E} \int_0^T \chi_n(s) K_2(s)^2 ds |\xi|^2 \\ &\quad + 2e^{2C_K} \left\| \int_0^T \chi_n(s) K_2(s)^2 ds \cdot I_F \right\|_2 (\|Y^n\|_{\mathcal{S}^2} + \|Y^m\|_{\mathcal{S}^2}).\end{aligned}\quad (32)$$

Hence, (31) implies $\delta_{n_k,m}^{(2)} \rightarrow 0$ for $k \rightarrow \infty$.

Finally,

$$\delta_{n,m}^{(3)} \leq 2 \mathbb{E} \int_0^T (2|Y_s^n|^2 + |Y_s^m|^2) \chi_n(s) K_1(s) ds,$$

so that we can argue like in (32) to get that $\delta_{n_k, m}^{(3)} \rightarrow 0$ for $k \rightarrow \infty$.

Thus $(Y^{n_k}, Z^{n_k}, U^{n_k})_{k \geq 1}$ converges to an object (Y, Z, U) in $L^2(W) \times L^2(W) \times L^2(\tilde{N})$.

Step 4:

In the final step, we want to show that (Y, Z, U) solves (ξ, f) . For the approximating sequence $(Y^{n_k}, Z^{n_k}, U^{n_k})_{k \geq 1}$, the stochastic integrals and the left hand side of the BSDEs $(\xi, f^{(n_k)})$ obviously converge in L^2 to the corresponding terms of (ξ, f) . Therefore, this subsequence of $\left(\int_t^T f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) ds\right)_{n=1}^\infty$ converges to a random variable V_t . We need to show that $V_t = \int_t^T f(s, Y_s, Z_s, U_s) ds$. To achieve this, consider

$$\begin{aligned} \delta_n &:= \mathbb{E} \int_t^T |f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) - f(s, Y_s^n, Z_s^n, U_s^n)| ds \\ &\quad + \mathbb{E} \int_t^T |f(s, Y_s^n, Z_s^n, U_s^n) - f(s, Y_s, Z_s, U_s)| ds. \end{aligned} \quad (33)$$

We start with the first integrand where, by the definition of f_n and (29), and the growth condition (A2),

$$\begin{aligned} &|f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) - f(s, Y_s^n, Z_s^n, U_s^n)| \\ &= |f^{(n)}(s, Y_s^n, Z_s^n, U_s^n) - f(s, Y_s^n, Z_s^n, U_s^n)| \chi_n \\ &\leq 2(F(s)\chi_n(s) + K_1(s)|Y_s^n| \chi_n(s) + K_2(s)\chi_n(s)(|Z_s^n| + \|U_s^n\|)) \\ &=: 2(\kappa_n^{(1)}(s) + \kappa_n^{(2)}(s) + \kappa_n^{(3)}(s)). \end{aligned}$$

The estimates are similar as in the previous step. Thanks to (31), we have $\mathbb{E} \int_t^T \kappa_n^{(1)}(s) ds \rightarrow 0$. For the next term, the Cauchy–Schwarz inequality yields

$$\mathbb{E} \int_t^T \kappa_n^{(2)}(s) ds \leq \left\| \int_0^T \chi_n(s) K_1(s) ds \right\|_2 \sup_l \|Y^l\|_{\mathcal{S}^2},$$

so that by (31) the first factor converges to zero along the subsequence (n_k) . The last term we estimate using the Cauchy–Schwarz inequality w.r.t. $\mathbb{P} \otimes \lambda$,

$$\mathbb{E} \int_t^T \kappa_n^{(3)}(s) ds \leq \left[\mathbb{E} \int_0^T K_2(s)^2 \chi_n(s) ds \right]^{\frac{1}{2}} \sup_l \left[\|Z^l\|_{L^2(W)} + \|U^l\|_{L^2(\tilde{N})} \right],$$

and again by (31), we have convergence to zero along the subsequence (n_k) .

We continue showing the convergence of the second term in (33). We extract a sub-subsequence of $(n_k)_{k \geq 1}$, which we call—slightly abusing the notation—again $(n_k)_{k \geq 1}$ such that $(Y^{n_k}, Z^{n_k}, U^{n_k})$, regarded as a triplet of measurable functions with values in $\mathbb{R} \times \mathbb{R} \times L^2(\nu)$, converges to (Y, Z, U) for $\mathbb{P} \otimes \lambda$ -a.a. (ω, s) . Then, for an arbitrary $K > 0$, we have

$$\begin{aligned} &\mathbb{E} \int_t^T |f(s, Y_s^{n_k}, Z_s^{n_k}, U_s^{n_k}) - f(s, Y_s, Z_s, U_s)| ds \\ &\leq \mathbb{E} \left\{ \int_t^T |f(s, Y_s^{n_k}, Z_s^{n_k}, U_s^{n_k}) - f(s, Y_s, Z_s, U_s)| \right. \\ &\quad \times \left(\chi_{\{|Y_s^{n_k}| \leq K, |Z_s^{n_k}| + \|U_s^{n_k}\| \leq K\}} + \chi_{\{|Y_s^{n_k}| > K\}} + \chi_{\{|Z_s^{n_k}| + \|U_s^{n_k}\| > K\}} \right) ds \Big\}. \end{aligned} \quad (34)$$

By dominated convergence and the continuity of f ,

$$\mathbb{E} \int_t^T |f(s, Y_s^{n_k}, Z_s^{n_k}, U_s^{n_k}) - f(s, Y_s, Z_s, U_s)| \chi_{\{|Y_s^{n_k}| \leq K, |Z_s^{n_k}| + \|U_s^{n_k}\| \leq K\}} ds \rightarrow 0,$$

since by (A2) we can bound the integrand by

$$2F(s) + K_1(s)(K + |Y_s|) + K_2(s)(2K + |Z_s| + \|U_s\|),$$

which is integrable. We let

$$\chi_K(n_k, s) := \chi_{\{|Y_s^{n_k}| > K\}} + \chi_{\{|Z_s^{n_k}| + \|U_s^{n_k}\| > K\}}.$$

Then, the remaining terms of (34) are bounded by

$$\begin{aligned} & \mathbb{E} \int_0^T (2F(s) + K_1(s)|Y_s| + K_2(s)(|Z_s| + \|U_s\|)) \chi_K(n_k, s) ds \\ & \quad + \mathbb{E} \int_0^T K_1(s) |Y_s^{n_k}| \chi_K(n_k, s) ds \\ & \quad + \mathbb{E} \int_0^T K_2(s) (|Z_s^{n_k}| + \|U_s^{n_k}\|) \chi_K(n_k, s) ds \\ & =: \delta_{n_k}^{(1)} + \delta_{n_k}^{(2)} + \delta_{n_k}^{(3)}. \end{aligned}$$

If we choose a K large enough, then $\delta_{n_k}^{(1)}$ can be made arbitrarily small since the families $(|Y_s^n|, n \geq 0)$ and $(|Z_s^n| + \|U_s^n\|, n \geq 0)$ are uniformly integrable with respect to $\mathbb{P} \otimes \lambda$. The same holds for

$$\begin{aligned} (\delta_{n_k}^{(2)})^2 & \leq \mathbb{E} \left| \int_0^T K_1(s) \chi_K(n_k, s) ds \right|^2 \sup_l \|Y^{n_l}\|_{\mathcal{S}^2}^2 \\ & \leq \left\| \int_0^T K_1(s) ds \right\|_\infty \mathbb{E} \left[\int_0^T K_1(s) \chi_K(n_k, s) ds \right] \sup_l \|Y^{n_l}\|_{\mathcal{S}^2}^2, \end{aligned}$$

and

$$(\delta_{n_k}^{(3)})^2 \leq 2\mathbb{E} \left[\int_0^T K_2(s)^2 \chi_K(n_k, s) ds \right] \sup_l \mathbb{E} \int_0^T (|Z_s^{n_l}|^2 + \|U_s^{n_l}\|^2) ds.$$

Hence, for δ_n defined in (33), we have that $\lim_{k \rightarrow \infty} \delta_{n_k} = 0$, which implies

$$\lim_{k \rightarrow \infty} \mathbb{E} \left| \int_t^T f^{(n_k)}(s, Y_s^{n_k}, Z_s^{n_k}, U_s^{n_k}) ds - \int_t^T f(s, Y_s, Z_s, U_s) ds \right| = 0.$$

We infer that for a sub-subsequence $(n_{k_l}, l \geq 0)$ we get the a.s. convergence

$$\int_t^T f^{(n_{k_l})}(s, Y_s^{n_{k_l}}, Z_s^{n_{k_l}}, U_s^{n_{k_l}}) ds \rightarrow \int_t^T f(s, Y_s, Z_s, U_s) ds.$$

Thus, for the original sequence, a.s.

$$\int_t^T f^{(n_k)}(s, Y_s^{n_k}, Z_s^{n_k}, U_s^{n_k}) ds \rightarrow V_t = \int_t^T f(s, Y_s, Z_s, U_s) ds,$$

and therefore the triplet (Y, Z, U) satisfies the BSDE (ξ, f) .

5.2 Proof of Theorem 3.4

We start with a preparatory lemma:

Lemma 5.1 *If f satisfies (A1)–(A4), then for all $n \geq 0$, f_n constructed in Definition 3.3 also satisfies (A1)–(A4) (with different coefficients).*

Proof By definition, $(\omega, t) \mapsto f_n(t, y, z, u)$ is progressively measurable for all (y, z, u) , thus (A1) is satisfied. The inequalities in (A2) and (A3) are a.s. satisfied, with coefficients $\mathbb{E}_n F, \mathbb{E}_n K_1, \mathbb{E}_n K_2, \mathbb{E}_n \beta$. To ensure that these coefficients have a $(\mathcal{F}_t^n)_{t \in [0, T]}$ -progressively measurable version, one applies the procedure from Definition 3.3 to the inequalities in (A2) and (A3) and notes that an equation analogous to (11) holds true.

It remains to show a.s. continuity of f_n in the (y, z, u) -variables required in (A3) for a.e. t . In (Ylinen (2017), Proposition 7.3), this was shown by the fact that the approximation of the generators appearing there can be done using spaces of continuous functions. However, since our situation involves $L^2(\nu)$, a non-locally compact space, we can not easily adapt the proof from Ylinen (2017) and therefore we will use different means.

Let $D[0, T]$ be the space of càdlàg functions endowed by the Skorohod metric (which makes this space a Polish space). The Borel σ -algebra $\mathcal{B}(D[0, T])$ is generated by the coordinate projections $p_t: D[0, T] \rightarrow \mathbb{R}, x \mapsto x(s)$ (see Theorem 12.5 of Billingsley (1968), for instance). On this σ -algebra, let \mathbb{P}_X be the image measure induced by the Lévy process $X: \Omega \rightarrow D[0, T], \omega \mapsto X(\omega)$. We denote by \mathcal{G} the completion with respect to \mathbb{P}_X . For $t \in [0, T]$, the notation

$$x^t(s) := x(t \wedge s), \text{ for all } s \in [0, T]$$

induces the natural identification

$$D[0, t] = \{x \in D[0, T] : x^t = x\}.$$

By this identification, we define a filtration on this space through

$$\mathcal{G}_t = \sigma(\mathcal{B}(D[0, t]) \cup \mathcal{N}_X[0, T]), \quad 0 \leq t \leq T,$$

where $\mathcal{N}_X[0, T]$ denotes the null sets of $\mathcal{B}(D[0, T])$ with respect to the image measure \mathbb{P}_X of the Lévy process X . The same procedure applied to the Lévy process X^n yields a filtration $(\mathcal{G}_t^n)_{t \in [0, T]}$ defined in the same way.

According to (Steinicke (2016), Theorem 3.4), which is a generalization of Doob's factorization lemma to random variables depending on parameters, there is a $\mathcal{G}_t \otimes \mathcal{B}([0, t] \times \mathbb{R}^2 \times L^2(\nu))$ -measurable functional

$$g_f: D[0, t] \times [0, t] \times \mathbb{R}^2 \times L^2(\nu) \rightarrow \mathbb{R}$$

and a $\mathcal{G}_t^n \otimes \mathcal{B}([0, t] \times \mathbb{R}^2 \times L^2(\nu))$ -measurable functional

$$g_{f_n}: D[0, t] \times [0, t] \times \mathbb{R}^2 \times L^2(\nu) \rightarrow \mathbb{R}$$

such that \mathbb{P} -a.s.,

$$g_f(X(\omega), \cdot) = f(\omega, \cdot) \quad \text{and} \quad g_{f_n}(X^n(\omega), \cdot) = f_n(\omega, \cdot). \quad (35)$$

Note also, that if $\mathbb{P}_X(M) = 0$ for $M \in \mathcal{G}$, then also $\mathbb{P}(X^{-1}(M)) = 0$. Thus, without loss of generality, we may assume that $(\Omega, \mathcal{F}, \mathbb{P}) = (D[0, T], \mathcal{G}, \mathbb{P}_X)$ and $(\Omega, \mathcal{F}_t^n, \mathbb{P}) = (D([0, t]), \mathcal{G}_t^n, \mathbb{P}_X)$, which are standard Borel spaces. For more details on $D[0, T]$, see Billingsley (1968) and (Delzeith (2004), Section 4).

Now, fix $N \in \mathbb{N}$ and let $c_0 := \{(a_n)_n \in (\mathbb{R}^2 \times L^2(\nu))^{\mathbb{N}} : a_n \rightarrow 0\}$. For $a \in c_0$, let $\|a\|_{c_0} = \sup_{n \in \mathbb{N}}(|a_n(1)| + |a_n(2)| + \|a_n(3)\|)$, where $a(k)$, $k = 1, 2, 3$ are the components of a in \mathbb{R} , \mathbb{R} and $L^2(\nu)$. The space c_0 is a Polish space. Let B_N be the ball with radius $N \in \mathbb{N}$ in c_0 and let B'_N be the ball of radius N in $\mathbb{R}^2 \times L^2(\nu)$. The balls B_N , B'_N are again Polish spaces.

We consider a Borel set M_T of $t \in [0, T]$ for which f is continuous in (y, z, u) and for which it holds that f has an integrable bound:

$$\mathbb{E}|f(t, y, z, u)| \leq \mathbb{E}F(t) + \mathbb{E}K_1(t)|y| + \mathbb{E}K_2(t)(|z| + \|u\|) < \infty. \quad (36)$$

From (A3) and (9) it follows that one can choose M_T such that $\lambda(M_T) = T$.

For a fixed $t \in M_T$ we define the function

$$H_m : \Omega \times B_N \times B'_N \rightarrow \mathbb{R}, (\omega, a, \varphi) \mapsto f_n(\omega, t, a_m + \varphi),$$

where φ denotes a triplet $(y, z, u) \in \mathbb{R}^2 \times L^2(\nu)$. This function is measurable since $f_n(\cdot, t, \cdot)$ is measurable, $\pi_m : B_N \times B'_N \rightarrow \mathbb{R}^2 \times L^2(\nu)$, $(a, \varphi) \mapsto (a_m + \varphi)$ is continuous and $\text{id} \times \pi_m : \Omega \times B_N \times B'_N \rightarrow \Omega \times \mathbb{R}^2 \times L^2(\nu)$ is measurable.

Next, we consider the map

$$H : \Omega \times B_N \times B'_N \rightarrow \mathbb{R}, (\omega, a, \varphi) \mapsto \begin{cases} \lim_{m \rightarrow \infty} f_n(\omega, t, a_m + \varphi), & \text{if it exists} \\ 0, & \text{else.} \end{cases}$$

The set, where the limit exists is measurable, since it can be written as

$$\bigcap_{k \geq 1} \bigcup_{M \geq 0} \bigcap_{m_1, m_2 \geq M} \left\{ |H_{m_1} - H_{m_2}| < \frac{1}{k} \right\}.$$

Therefore, H can be written as the pointwise limit of measurable functions and is thus measurable.

We now know that, for a fixed pair $(a, \varphi) \in B_N \times B'_N$,

$$f_n(t, a_m + \varphi) = \mathbb{E}_n f(t, a_m + \varphi), \quad \mathbb{P}\text{-a.s.}$$

Thus, by (36)

$$|f_n(t, a_m + \varphi)| \leq \mathbb{E}_n F(t) + 2N\mathbb{E}_n K_1(t) + 4N\mathbb{E}_n K_2(t) < \infty.$$

By the continuity of f and the dominated convergence theorem for conditional expectations, we infer that up to a null set $M(a, \varphi) \in \mathcal{F}_t^n$, we have the relation

$$\begin{aligned} \lim_{m \rightarrow \infty} f_n(t, a_m + \varphi) &= \lim_{m \rightarrow \infty} \mathbb{E}_n f(t, a_m + \varphi) = \mathbb{E}_n \lim_{m \rightarrow \infty} f(t, a_m + \varphi) \\ &= \mathbb{E}_n f(t, \varphi) = f_n(t, \varphi). \end{aligned}$$

In other words, on the complement of $M(a, \varphi)$, we have $H(\omega, a, \varphi) = f_n(\omega, t, \varphi)$. This means that H and $f_n(\cdot, t, \cdot)$ are "versions" of each other. What we need is "indistinguishability" of the processes.

For this purpose, let $(A, \Phi) : \Omega \rightarrow B_N \times B'_N$ be an arbitrary \mathcal{F}_t^n -measurable function. Like above, by the definition of the optional projection, (A2), and the continuity of f , we get the equation

$$\lim_{m \rightarrow \infty} f_n(t, A_m + \Phi) = f_n(t, \Phi),$$

which is also satisfied \mathbb{P} -a.s. This equality means, that

$$H(\omega, A(\omega), \Phi(\omega)) = f_n(\omega, t, \Phi(\omega)), \quad \text{a.s.}$$

All \mathcal{F}_t^n were complete σ -algebras (in fact they contain all null sets of \mathcal{F}) and the spaces B_N, B'_N were Polish. Thus we may use a generalized version of the section theorem, the Jankov-von Neumann theorem (Theorem 5.3), by choosing a uniformizing function $(\hat{A}, \hat{\Phi})$ for the set

$$P = \{(\omega, a, \varphi) : H(\omega, a, \varphi) \neq f_n(\omega, t, \varphi)\}.$$

Note that P is a Borel set and therefore especially analytic, since H and $f_n(\cdot, t, \cdot)$ (interpreted as a constant map w.r.t. a) are measurable functions in (ω, a, φ) . Since for this choice of $(\hat{A}, \hat{\Phi})$ it holds, as seen above, that

$$H(\omega, \hat{A}(\omega), \hat{\Phi}(\omega)) = f_n(\omega, t, \hat{\Phi}(\omega)), \quad \text{a.s.}$$

it follows that the projection of P to Ω is a null set. Therefore, H and f_n are indistinguishable. Hence, we find a null set $M_N \in \mathcal{F}_t^n$, such that for ω outside this set and for all $(a, \varphi) \in B_N \times B'_N$:

$$\lim_{m \rightarrow \infty} f_n(\omega, t, a_m + \varphi) = f_n(\omega, t, \varphi).$$

But this means continuity in all points of B'_N a.s. It remains to unite the sets M_N for all $N \in \mathbb{N}$, to obtain a set such that on its complement the function is continuous in all points of $\mathbb{R}^2 \times L^2(\nu)$. \square

Proof Step 1:

If f satisfies (A1)–(A4), by Lemma 5.1 all f_n do so as well. In this case, for all $n \geq 0$, the equations $(\mathbb{E}_n \xi, f_n)$ have unique solutions by Theorem 3.1. In general, the coefficients in (A2) and β differ dependent on n since F, K_1, K_2, β will be replaced by the coefficients $\mathbb{E}_n F, \mathbb{E}_n K_1, \mathbb{E}_n K_2, \mathbb{E}_n \beta$.

Let us compare the solutions (Y^n, Z^n, U^n) and (Y, Z, U) . We start comparing (Y^n, Z^n, U^n) and $(\mathbb{E}_n Y, \mathbb{E}_n Z, \mathbb{E}_n U)$. Here, for instance, the process $((\mathbb{E}_n Y)_t)_{t \in [0, T]}$ is defined as an optional projection with respect to the filtration $(\mathcal{F}_t^n)_{t \in [0, T]}$, similar to Definition 3.3. The so defined processes are versions of the processes $(\mathbb{E}_n Y_t, \mathbb{E}_n Z_t, \mathbb{E}_n U_t)_{t \in [0, T]}$.

Using the BSDE for (Y, Z, U) , we get \mathbb{P} -a.s.

$$\begin{aligned} \mathbb{E}_n Y_t = & \mathbb{E}_n \xi + \int_t^T \mathbb{E}_n f(s, Y_s, Z_s, U_s) ds - \int_t^T \mathbb{E}_n Z_s dW_s \\ & - \int_{[t, T] \times \{1/n \leq |x|\}} \mathbb{E}_n U_s(x) \tilde{N}(ds, dx), \end{aligned} \quad (37)$$

since

$$\mathbb{E}_n \int_{[t, T] \times \{1/n > |x|\}} U_s(x) \tilde{N}(ds, dx) = 0.$$

Now, to estimate $\|Y^n - \mathbb{E}_n Y\|_{L^2(W)} + \|Z^n - \mathbb{E}_n Z\|_{L^2(W)} + \|U^n - \mathbb{E}_n U\|_{L^2(\tilde{N})}$, we apply Itô's formula to the difference of the BSDE $(\mathbb{E}_n \xi, f_n)$ and (37). Similar to the proof of Proposition 4.2, we get, denoting differences by Δ^n and $\eta := 4\beta(s)^2$,

$$\begin{aligned} & \mathbb{E} \left\{ e^{\int_0^t \eta(s) ds} |\Delta^n Y_t|^2 + \int_t^T e^{\int_t^s \eta(\tau) d\tau} \left(\eta(s) |\Delta^n Y_s|^2 + |\Delta^n Z_s|^2 + \|\Delta^n U_s\|^2 \right) ds \right\} \\ &= \mathbb{E} \int_t^T 2e^{\int_t^s \eta(\tau) d\tau} (\Delta^n Y_s) \left(f_n(s, Y_s^n, Z_s^n, U_s^n) - \mathbb{E}_n f(s, Y_s, Z_s, U_s) \right) ds. \end{aligned}$$

By the measurability of (Y^n, Z^n, U^n) , the equality

$$f_n(s, Y_s^n, Z_s^n, U_s^n) = \mathbb{E}_n f(s, Y_s^n, Z_s^n, U_s^n)$$

holds \mathbb{P} -a.s. for all s . We now estimate

$$\begin{aligned} & \mathbb{E} \left[(\Delta^n Y_s) \left(f_n(s, Y_s^n, Z_s^n, U_s^n) - \mathbb{E}_n f(s, Y_s, Z_s, U_s) \right) \right] \\ &= \mathbb{E} \left[(\Delta^n Y_s) \left(\mathbb{E}_n f(s, Y_s^n, Z_s^n, U_s^n) - \mathbb{E}_n f(s, Y_s, Z_s, U_s) \right) \right] \\ &= \mathbb{E} \left[(\Delta^n Y_s) \left(f(s, Y_s^n, Z_s^n, U_s^n) - f(s, Y_s, Z_s, U_s) \right) \right] \\ &= \mathbb{E} \left[(\Delta^n Y_s) \left(f(s, Y_s^n, Z_s^n, U_s^n) - f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) \right) \right] \\ &\quad + \mathbb{E} \left[(\Delta^n Y_s) \left(f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s) \right) \right] \\ &\leq \mathbb{E} \left[\alpha(s) \rho(|\Delta^n Y_s|^2) + \beta(s) |\Delta^n Y_s| (|\Delta^n Z_s| + \|\Delta^n U_s\|) \right] \\ &\quad + \mathbb{E} \left[|\Delta Y_s^n| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| \right]. \end{aligned}$$

Now, we can conduct exactly the same steps as in the standard procedure used in the proof of Proposition 4.2. This means that $\|\Delta^n Y\|_{L^2(W)} + \|\Delta^n Z\|_{L^2(W)} + \|\Delta^n U\|_{L^2(\tilde{N})}$ converges to zero if

$$\mathbb{E} \int_0^T |\Delta Y_s^n| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| ds \quad (38)$$

does, which we will show in the following steps.

Step 2:

In this step, we show that the solution processes (Y^n, Z^n, U^n) satisfy the estimate

$$\sup_{n \geq 0} \left(\|Y^n\|_{\mathcal{S}^2}^2 + \|Z^n\|_{L^2(W)}^2 + \|U^n\|_{L^2(\tilde{N})}^2 \right) < \infty. \quad (39)$$

This, as in the proof of Theorem 3.1, leads to the uniform integrability of the processes $(|Y^n|, n \geq 0)$ and $(|Z^n| + \|U^n\|, n \geq 0)$ with respect to $\mathbb{P} \otimes \lambda$.

By Proposition 4.1, we get that

$$\|Y^n\|_{\mathcal{S}^2}^2 + \|Z^n\|_{L^2(W)}^2 + \|U^n\|_{L^2(\tilde{N})}^2 \leq e^{C_1(1+C_{K,n})^2} (\mathbb{E} |\mathbb{E}_n \xi|^2 + \mathbb{E} (I_{\mathbb{E}_n F})^2),$$

where $C_{K,n} = \left\| \int_0^T (\mathbb{E}_n K_1(s) + (\mathbb{E}_n K_2(s))^2) ds \right\|_\infty$. By the monotonicity of \mathbb{E}_n and Jensen's inequality, we get that

$$\int_0^T (\mathbb{E}_n K_1(s) + (\mathbb{E}_n K_2(s))^2) ds \leq \mathbb{E}_n \int_0^T (K_1(s) + K_2(s)^2) ds \leq C_K, \quad \mathbb{P}\text{-a.s.}$$

Doob's martingale inequality applied to $n \mapsto \mathbb{E}_n \xi$ and $n \mapsto I_{\mathbb{E}_n F} = \mathbb{E}_n \int_0^T F(s) ds$ yields that

$$\|Y^n\|_{\mathcal{S}^2}^2 + \|Z^n\|_{L^2(W)}^2 + \|U^n\|_{L^2(\tilde{N})}^2 \leq e^{C_1(1+C_K)^2} (\mathbb{E}|\xi|^2 + \mathbb{E}(I_F)^2).$$

Furthermore,

$$\sup_{n \geq 0} \left(\|\mathbb{E}_n Y\|_{\mathcal{S}^2} + \|\mathbb{E}_n Z\|_{L^2(W)}^2 + \|\mathbb{E}_n U\|_{L^2(\tilde{N})}^2 \right) < \infty \quad (40)$$

follows from martingale convergence and Jensen's inequality and implies uniform integrability of the processes $(\|\mathbb{E}_n Y\|, n \geq 0)$ and $(\|\mathbb{E}_n Z\| + \|\mathbb{E}_n U\|, n \geq 0)$ with respect to $\mathbb{P} \otimes \lambda$.

Step 3 :

In this step, we show the convergence (38). From martingale convergence, we get that for all $t \in [0, T]$, $\mathbb{E}_n Y_t \rightarrow Y_t$, $\mathbb{E}_n Z_t \rightarrow Z_t$ and $\mathbb{E}_n U_t \rightarrow U_t$, \mathbb{P} -a.s. This implies that $f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) \rightarrow f(s, Y_s, Z_s, U_s)$ in $\mathbb{P} \otimes \lambda$. Therefore,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \int_0^T |Y_s^n - \mathbb{E}_n Y_s| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| \\ \times \chi_{\{|Y_s^n| + |\mathbb{E}_n Y_s| \leq K\}} ds = 0 \end{aligned}$$

since the integrals form a uniformly integrable sequence with respect to $\mathbb{P} \otimes \lambda$. Indeed, we have, using (A2) for f and the first equation of (14), the estimate

$$\begin{aligned} |Y_s^n - \mathbb{E}_n Y_s| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| \chi_{\{|Y_s^n| + |\mathbb{E}_n Y_s| \leq K\}} \\ \leq 4K(F(s) + K_1(s)) \\ + 2K(K_2(s)^2) + |Z_s|^2 + \|U_s\|^2 + |\mathbb{E}_n Z_s|^2 + \|\mathbb{E}_n U_s\|^2, \end{aligned}$$

where $n \mapsto \mathbb{E}_n Z_s, n \mapsto \mathbb{E}_n U_s$ converge since they are closable martingales.

Next, we will show that

$$\begin{aligned} \delta_n(K) := \mathbb{E} \left\{ \int_0^T |Y_s^n - \mathbb{E}_n Y_s| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| \right. \\ \left. \times \chi_{\{|Y_s^n| + |\mathbb{E}_n Y_s| > K\}} ds \right\} \quad (41) \end{aligned}$$

can be made arbitrarily small by the choice of $K > 0$, uniformly in n . Again by (A2) and using the notation $\chi_K^n(s) := \chi_{\{|Y_s^n| + |\mathbb{E}_n Y_s| > K\}}$, we estimate like in (30)

$$\begin{aligned} |Y_s^n - \mathbb{E}_n Y_s| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| \chi_{\{|Y_s^n| + |\mathbb{E}_n Y_s| > K\}} \\ \leq |Y_s^n - \mathbb{E}_n Y_s| (2F(s) + K_1(s)(|Y_s| + |\mathbb{E}_n Y_s|) \\ + K_2(s)(|Z_s| + |\mathbb{E}_n Z_s| + \|U_s\| + \|\mathbb{E}_n U_s\|)) \chi_K^n(s) \end{aligned}$$

and get

$$\begin{aligned} \delta_n(K) &\leq 2\mathbb{E} \left\{ \int_0^T \chi_K^n(s) F(s) ds \left(\sup_{r \in [0, T]} |Y_r^n| + \sup_{r \in [0, T]} |\mathbb{E}_n Y_r| \right) \right\} \\ &\quad + \mathbb{E} \left\{ \int_0^T \chi_K^n(s) K_2(s) (|Y_s^n| + |\mathbb{E}_n Y_s|) (|Z_s| + |\mathbb{E}_n Z_s| + \|U_s\| + \|\mathbb{E}_n U_s\|) ds \right\} \\ &\quad + \mathbb{E} \int_0^T |Y_s^n - \mathbb{E}_n Y_s| (|Y_s| + |\mathbb{E}_n Y_s|) \chi_K^n(s) K_1(s) ds \\ &=: \delta_{n,K}^{(1)} + \delta_{n,K}^{(2)} + \delta_{n,K}^{(3)}. \end{aligned} \quad (42)$$

For $\delta_{n,K}^{(1)}$, we estimate

$$\delta_{n,K}^{(1)} \leq 2 \left\| \int_0^T \chi_K^n(s) F(s) ds \right\|_2 \sup_{l \geq 0} \left(\|Y^l\|_{\mathcal{S}^2} + \|\mathbb{E}_l Y_s\|_{\mathcal{S}^2} \right)$$

which tends to zero as $K \rightarrow \infty$, since we have $\chi_K^n \rightarrow 0$ in $\mathbb{P} \otimes \lambda$, uniformly in n , as $K \rightarrow \infty$. The latter is implied by the uniform integrability of the families $(|Y^n|)_{n \geq 0}$ and $(|\mathbb{E}_n Y|)_{n \geq 0}$ with respect to $\mathbb{P} \otimes \lambda$. We continue with the next summands,

$$\begin{aligned} \delta_{n,K}^{(2)} &\leq 8 \left(\mathbb{E} \int_0^T (|Y_s^n|^2 + |\mathbb{E}_n Y_s|^2) \chi_K^n(s) K_2(s)^2 ds \right)^{\frac{1}{2}} \\ &\quad \times \left(\|Z\|_{L^2(W)} + \|U\|_{L^2(\tilde{N})} \right) \end{aligned} \quad (43)$$

and

$$\delta_{n,K}^{(3)} \leq \mathbb{E} \int_0^T (|Y_s|^2 + |Y_s^n|^2 + 2|\mathbb{E}_n Y_s|^2) \chi_K^n(s) K_1(s) ds, \quad (44)$$

where, for $\mathbb{E} \int_0^T \chi_K^n(s) (|Y_s|^2 + |Y_s^n|^2) K_1(s) ds$ and $\mathbb{E} \int_0^T \chi_K^n(s) |Y_s^n|^2 K_2(s)^2 ds$, we will apply the estimate (27) from the proof of Lemma 4.3. For example (the other terms can be treated similarly), we get

$$\begin{aligned} \mathbb{E} \int_0^T \chi_K^n(s) |Y_s^n|^2 K_2(s)^2 ds &\leq \mathbb{E} \int_0^T \chi_K^n(s) K_2(s)^2 ds \cdot e^{\int_0^T \eta_n(s) ds} |\mathbb{E}_n \xi|^2 \\ &\quad + 2\mathbb{E} \int_0^T \int_0^s \chi_K^n(s) K_2(\tau)^2 d\tau e^{\int_0^s \eta_n(\tau) d\tau} \mathbb{E}_n F(s) |Y_s^n| ds \\ &\leq e^{2C_K} \mathbb{E} \int_0^T \chi_K^n(s) K_2(s)^2 ds |\mathbb{E}_n \xi|^2 \\ &\quad + 2e^{2C_K} \left\| \int_0^T \chi_K^n(s) K_2(s)^2 ds \cdot I_{\mathbb{E}_n F} \right\|_2 \|Y^n\|_{\mathcal{S}^2} \end{aligned} \quad (45)$$

with $\int_0^T \eta_n(s) ds = \int_0^T \mathbb{E}_n K_1(s) + (\mathbb{E}_n K_2(s))^2 ds \leq C_K$ a.s. Now, one gets that

$$\int_0^T \chi_K^n(s) K_2(s)^2 ds \xrightarrow{\mathbb{P}} 0, \quad K \rightarrow \infty.$$

Furthermore, using $\sup_{n \geq 0} \mathbb{E}_n \int_0^T F(s) ds < \infty$, \mathbb{P} -a.s. (which follows from martingale convergence),

$$\int_0^T \chi_K^n(s) K_2(s)^2 ds \int_0^T \mathbb{E}_n F(s) ds \xrightarrow{\mathbb{P}} 0, \quad K \rightarrow \infty,$$

independently of n . Since, by Doob's maximal inequality,

$$\begin{aligned} & \mathbb{E} \left[\sup_{n \geq 0} \int_0^T K_2(s)^2 ds \int_0^T \mathbb{E}_n F(s) ds \right]^2 \\ & \leq \mathbb{E} \left[\sup_{n \geq 0} C_K \mathbb{E}_n \int_0^T F(s) ds \right]^2 \leq 4C_K^2 \mathbb{E} I_F^2 < \infty, \end{aligned}$$

dominated convergence is applicable to the last expression in (45). The first summand containing ξ can be treated in the same way.

The terms containing $|\mathbb{E}_n Y_s|$ in the inequalities (43) and (44), e.g., the expression $\mathbb{E} \int_0^T \chi_K^n(s) |\mathbb{E}_n Y_s|^2 K_1(s) ds$, can be estimated by

$$\begin{aligned} \mathbb{E} \int_0^T \chi_K^n(s) |\mathbb{E}_n Y_s|^2 K_1(s) ds & \leq \mathbb{E} \int_0^T \chi_K^n(s) K_1(s) \left(\sup_{l \geq 0} |\mathbb{E}_l Y_s| \right)^2 ds \\ & \leq \mathbb{E} \left\{ \int_0^T \chi_K^n(s) K_1(s) ds \left(\sup_{t \in [0, T]} \sup_{l \geq 0} \mathbb{E}_l |Y_t| \right)^2 \right\} \\ & \leq 2C_K \|Y\|_{\mathcal{S}^2}^2, \end{aligned}$$

where we used Doob's maximal inequality again. Since $\int_0^T \chi_K^n(s) K_1(s) ds \rightarrow 0$ in \mathbb{P} as $K \rightarrow \infty$, all the terms in (43) and (44) become small, uniformly in n , if K is large. So the expressions $\delta_{n,K}^{(2)}$ and $\delta_{n,K}^{(3)}$ can be made arbitrarily small by the choice of K , which gives us the desired convergence

$$\mathbb{E} \int_0^T |Y_s^n - \mathbb{E}_n Y_s| |f(s, \mathbb{E}_n Y_s, \mathbb{E}_n Z_s, \mathbb{E}_n U_s) - f(s, Y_s, Z_s, U_s)| ds \rightarrow 0.$$

Step 5 :

Since, by the last step,

$$\|Y^n - \mathbb{E}_n Y\|_{L^2(W)} + \|Z^n - \mathbb{E}_n Z\|_{L^2(W)} + \|U^n - \mathbb{E}_n U\|_{L^2(\tilde{N})} \rightarrow 0,$$

and also, by martingale convergence,

$$\|Y - \mathbb{E}_n Y\|_{L^2(W)} + \|Z - \mathbb{E}_n Z\|_{L^2(W)} + \|U - \mathbb{E}_n U\|_{L^2(\tilde{N})} \rightarrow 0,$$

we get

$$\|Y^n - Y\|_{L^2(W)} + \|Z^n - Z\|_{L^2(W)} + \|U^n - U\|_{L^2(\tilde{N})} \rightarrow 0.$$

□

Appendix

The Bihari–LaSalle inequality. For the Bihari–LaSalle inequality we refer to (Mao (1997), pp. 45–46). Here, we formulate a backward version of it which has been applied in Yin and Mao (2008). The proof is analogous to that in Mao (1997).

Proposition 5.2 *Let $c > 0$. Assume that $\rho : [0, \infty[\rightarrow [0, \infty[$ is a continuous and non-decreasing function such that $\rho(x) > 0$ for all $x > 0$. Let K be a non-negative, integrable Borel function on $[0, T]$, and y a non-negative, bounded Borel function on $[0, T]$, such that*

$$y(t) \leq c + \int_t^T K(s)\rho(y(s))ds.$$

Then, it holds that

$$y(t) \leq G^{-1} \left(G(c) + \int_t^T K(s)ds \right)$$

for all $t \in [0, T]$ such that $G(c) + \int_t^T K(s)ds \in \text{dom}(G^{-1})$. Here

$$G(x) := \int_1^x \frac{dr}{\rho(r)},$$

and G^{-1} is the inverse function of G .

Especially, if $\rho(r) = r$ for $r \in [0, \infty[$, it holds that

$$y(t) \leq ce^{\int_t^T K(s)ds}. \quad (46)$$

The Jankov–von Neumann theorem. If X and Y are sets and $P \subseteq X \times Y$, then $P^* \subseteq P$ is called a *uniformization* of P if and only if P^* is the graph of a function $f : \text{proj}_X(P) \rightarrow Y$, i.e., $P^* = \{(x, f(x)) : x \in \text{proj}_X(P)\}$. Such a function f is called a *uniformizing function* for P . Let $\Sigma_1^1(X)$ denote the class of analytic subsets of X . The following theorem can be found, for example, in (Kechris (1994), Theorem 18.1).

Theorem 5.3 (Jankov–von Neumann theorem) *Assume that X and Y are standard Borel spaces and $P \subseteq X \times Y$ is an analytic set. Then, P has a uniformizing function that is $\sigma(\Sigma_1^1(X))$ -measurable.*

Acknowledgements The authors thank Stefan Geiss and Juha Ylinen, University of Jyväskylä, for fruitful discussions and valuable suggestions.

Moreover, we are sincerely grateful to the anonymous reviewers for their helpful comments and questions. Christel Geiss would like to thank the Erwin Schrödinger Institute, Vienna, for hospitality and support, where a part of this work was written.

Funding

Large parts of this article were written when Alexander Steinicke was member of the Institute of Mathematics and Scientific Computing, University of Graz, Austria, and supported by the Austrian Science Fund (FWF): Project F5508-N26, which is part of the Special Research Program “Quasi-Monte Carlo Methods: Theory and Applications.”

Availability of data and material

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Authors' contributions

Both authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Applebaum, D: Lévy Processes and Stochastic Calculus. Cambridge University Press, Cambridge (2004)
- Barles, G, Buckdahn, R, Pardoux, É: Backward stochastic differential equations and integral-partial differential equations. *Stoch. Stoch. Rep.* **60**(1-2), 57–83 (1997)
- Becherer, D, Büttner, M, Kentia, K: On the monotone stability approach to BSDEs with jumps: Extensions, concrete criteria and examples (2018). <https://arxiv.org/abs/1607.06644>
- Billingsley, P: Convergence of probability measures. Wiley, New York (1968)
- Briand, P, Delyon, B, Hu, Y, Pardoux, E, Stoica, L: L_p solutions of backward stochastic differential equations. *Stoch. Proc. Appl.* **108**, 109–129 (2003)
- Cao, Z, Yan, J: A comparison theorem for solutions of backward stochastic differential equations. *Adv. Math.* **28**(4), 304–308 (1999)
- Cohen, S, Elliott, R, Pearce, C: A General Comparison Theorem for Backward Stochastic Differential Equations. *Adv. Appl. Probab.* **42**(3), 878–898 (2010)
- Darling, R, Pardoux, É: Backwards sde with random terminal time and applications to semilinear elliptic pde. *Ann. Probab.* **25**(3), 1135–1159 (1997)
- Delzeith, O: On Skorohod spaces as universal sample path spaces (2004). <https://arxiv.org/abs/math/0412092v1>
- El Karoui, N, Hamadène, S, Matoussi, A: Backward stochastic differential equations. In: Carmona, R (ed.) *Indifference Hedging: Theory and Applications*, pp. 267–320. Princeton University Press (2009)
- El Karoui, N, Tan, X: Capacities, measurable selection and dynamic programming part I: abstract framework (2013). <https://arxiv.org/abs/1310.3363>
- El Karoui, N, Peng, S, Quenez, M: Backward Stochastic Differential Equations in Finance. *Math. Financ.* **7**(1), 1–71 (1997)
- Fan, S, Jiang, L: A Generalized Comparison Theorem for BSDEs and Its Applications. *J. Theor. Probab.* **25**, 50–61 (2012)
- Geiss, C, Steinicke, A: Existence, Uniqueness and Malliavin Differentiability of Lévy-driven BSDEs with locally Lipschitz Driver (2018). <https://arxiv.org/abs/1805.05851>
- Geiss, S, Ylinen, J: Decoupling on the Wiener Space, Related Besov Spaces, and Applications to BSDEs (2018). is to appear in: *Memoirs AMS*, 2018
- Gobet, E, Turkedjiev, P: Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comp.* **85**, 1359–1391 (2016)

- He, SW, Wang, JG, Yan, JA: Semimartingale Theory and Stochastic Calculus. CRC Press, Boca Raton (1992)
- Kechris, A: Classical Descriptive Set Theory. Springer, New York (1994)
- Kruse, T, Popier, A: BSDEs with monotone generator driven by Brownian and Poisson noises in a general filtration. *Stochastics*. **88**(4), 491–539 (2016)
- Kruse, T, Popier, A: L^p -solution for BSDEs with jumps in the case $p < 2$. *Stochastics* (2017). <http://dx.doi.org/10.1080/17442508.2017.1290095>
- Mao, X: Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients. *Stoch. Process. Appl.* **58**, 281–292 (1995)
- Mao, X: Stochastic Differential Equations and Applications. Woodhead Publishing Limited, Cambridge (1997)
- Meyer, PA: Une remarque sur le calcul stochastique dépendant d'un paramètre. *Séminaire probabilités (Strasbourg)*, tome. **13**, 199–203 (1979)
- Pardoux, É: Generalized discontinuous backward stochastic differential equations. In: El Karoui, N, Mazliak, L (eds.) *Backward Stochastic Differential Equations*, Pitman Res. Notes Math., vol. 364, pp. 207–219. Longman, Harlow (1997)
- Pardoux, É, Zhang, S: Generalized BSDEs and nonlinear Neumann boundary value problems. *Probab. Theory Relat. Fields*. **110**, 535–558 (1996)
- Peng, S: A generalized dynamic programming principle and hamilton-jacobi-bellman equation. *Stoch. Stoch. Rep.* **38**, 119–134 (1992)
- Peng, S: Backward stochastic differential equation, nonlinear expectation and their applications. *Proceedings of the International Congress of Mathematicians, Volume I*, pp. 393–432. Hindustan Book Agency, New Delhi (2010)
- Protter, P: Stochastic Integration and Differential Equations. Springer, Berlin (2004)
- Royer, M: Backward stochastic differential equations with jumps and related non-linear expectations. *Stoch. Process. Appl.* **116**(10), 1358–1376 (2006)
- Sato, K: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (1999)
- Situ, R: On solutions of backward stochastic differential equations with jumps and applications. *Stoch. Process. Appl.* **66**, 209–236 (1997)
- Steinicke, A: Functionals of a Lévy Process on Canonical and Generic Probability Spaces. *J. Theoret. Probab.* **29**, 443–458 (2016)
- Sow, AB: BSDE with jumps and non-Lipschitz coefficients: Application to large deviations. *Braz. J. Probab. Stat.* **28**(1), 96–108 (2014)
- Yao, S: \mathbb{L}^p -solutions of Backward Stochastic Differential Equations with Jumps. *Stoch. Proc. Appl.* **127**(11), 3465–3511 (2017)
- Yin, J, Mao, X: The adapted solution and comparison theorem for backward stochastic differential equations with Poisson jumps and applications. *J. Math. Anal. Appl.* **346**, 345–358 (2008)
- Ylinen, J: Weighted Bounded Mean Oscillation applied to Backward Stochastic Differential Equations (2017). <https://arxiv.org/abs/1501.01183>