Electronic Journal of Probability

Electron. J. Probab. **21** (2016), no. 10, 1–28.
ISSN: 1083-6489 DOI: 10.1214/16-EJP4140

Malliavin derivative of random functions and applications to Lévy driven BSDEs

Christel Geiss* Alexander Steinicke†

Abstract

We consider measurable $F : \Omega \times \mathbb{R}^d \to \mathbb{R}$ where for any x the random variable $F(\cdot, x)$ belongs to the Malliavin Sobolev space $\mathbb{D}_{1,2}$ (with respect to a Lévy process) and provide sufficient conditions on F and $G_1, \ldots, G_d \in \mathbb{D}_{1,2}$ such that $F(\cdot, G_1, \ldots, G_d) \in \mathbb{D}_{1,2}$.

The above result is applied to show Malliavin differentiability of solutions to BSDEs (backward stochastic differential equations) driven by Lévy noise where the generator is given by a progressively measurable function $f(\omega, t, y, z)$.

Keywords: Malliavin calculus for Lévy processes; Lévy driven BSDEs.
AMS MSC 2010: 60H07; 60G51; 60H10.
Submitted to EJP on February 24, 2015, final version accepted on January 18, 2016.
Supersedes arXiv:1404.4477v3.

1 Introduction

Backward stochastic differential equations (BSDEs) have been studied with growing interest and from various perspectives. They appear in stochastic control theory, as Feynman-Kac representation of second order semilinear PDEs, and have many applications in Finance and Insurance (see, for instance, El Karoui et al. [18], the survey paper from Bouchard et al. [11] or Delong [13], and the references therein).

Pardoux and Peng have considered in [28] and [29] Forward Backward SDEs (FBSDEs) of the form

$$\begin{aligned} X_s &= x + \int_t^s a(X_r)dr + \int_t^s b(X_r)dW_r \\ Y_s &= g(X_T) + \int_s^T f(X_r, Y_r, Z_r)dr + \int_s^T Z_r dW_r, \quad t \le s \le T, \end{aligned}$$

where W denotes the Brownian motion. Under suitable smoothness and boundedness conditions on the coefficients they have shown that the two-parameter process $\mathcal{D}_\theta Y_s$ is

*Department of Mathematics and Statistics, University of Jyväskylä, Finland. E-mail: christel.geiss@jyu.fi

†Department of Mathematics, University of Innsbruck, Austria.
E-mail: alexander.steinicke@uibk.ac.at

a.s. continuous in $s \in [\theta, T]$ and, moreover, $\{\mathcal{D}_\theta Y_\theta := \lim_{s\downarrow\theta} \mathcal{D}_\theta Y_s : \theta \in [t,T]\}$ is a version of the process $\{Z_s : s \in [t,T]\}$. In this way, using the relation

$$Y_s = \mathbb{E}\left[g(X_T) + \int_s^T f(X_r, Y_r, Z_r)dr \middle| \mathcal{F}_s^W \right]$$

it is possible to represent Z (with the right interpretation) as

$$\left(\mathcal{D}_s \mathbb{E}\left[g(X_T) + \int_s^T f(X_r, Y_r, Z_r)dr \middle| \mathcal{F}_s^W \right] \right)_{s\in[t,T]}.$$

These representations turned out to be useful in regularity estimates for Y and Z which play an important role for estimates of convergence rates of time-discretizations (see, for example, [10], [12], [11], [14]).

El Karoui et al. generalized in [18] this result to a class of progressively measurable generators $(\omega, t) \mapsto f(\omega, t, y, z)$. Also in the Brownian setting, Ankirchner et al. [3] and Mastrolia et al. [24] extended the result to generators of BSDEs with quadratic growth.

On the canonical Lévy space, Malliavin differentiability of BSDEs with jumps has been considered by Delong in [13] and by Delong and Imkeller for delayed BSDEs in [14].

In this paper, we first consider a measurable function $F : \Omega \times \mathbb{R}^d \to \mathbb{R}$ where $F(\cdot, x)$ belongs to the Malliavin Sobolev space $\mathbb{D}_{1,2}$ for any $x \in \mathbb{R}^d$. We ask for sufficient conditions on F and $G_1, \dots, G_d \in \mathbb{D}_{1,2}$ such that $F(\cdot, G_1, \dots, G_d) \in \mathbb{D}_{1,2}$. Our aim was to find very general conditions such that the result is also applicable for BSDEs with non-Lipschitz generators. As we work in the Lévy setting, the results hold of course especially for the Brownian case. In this respect, we could generalize the conditions given in [18, Theorem 5.3] by not imposing the finiteness of fourth moments on the generator and the terminal condition (see Theorem 4.4 below). Moreover, we provide a rigorous proof of the extended chain rule for the Malliavin derivative of $F(\cdot, G_1, \dots, G_d)$ in the Brownian case (see Theorem 3.12). Such a chain rule was already used in [18]. Compared with [13] or [14], we do not require a canonical Lévy space to state Malliavin differentiability of BSDEs (Theorem 4.4).

The paper is organized as follows: Section 2 contains the setting and a collection of used notation.

Section 3 starts with the definition of the Malliavin derivative in the Lévy setting. The Malliavin calculus based on chaos expansions in the Lévy case has been treated in various papers, e.g. by Løkka [23], Lee and Shih [22], Di Nunno et al. [17].In our paper, we recall a method used in [32] which is related to Picard's difference operator approach [31]. It allows to compute the Malliavin derivative $\mathcal{D}_{t,x}$ for $x \neq 0$ without knowing the chaos expansion and without imposing the condition that the underlying probability space is specified, e.g. as the canonical Lévy space from [35] or the probability space of Section 4 in [23]. Based on the fact that $\mathcal{D}_{t,x}$ for $x \neq 0$ and $\mathcal{D}_{t,0}$ are of different nature we solve the question about the Malliavin differentiability of $F(\cdot, G_1, \dots, G_d) \in \mathbb{D}_{1,2}$ in two steps: In Subsection 3.3.1 we treat the question concerning $\mathcal{D}_{t,x}, x \neq 0$, while Subsection 3.3.2 contains the case $\mathcal{D}_{t,0}$. In the latter, we use the result from [36] that for the Brownian motion the Malliavin Sobolev spaces $\mathbb{D}_{1,p}^W(E)$ with $p > 1$ (E denotes a separable Hilbert space) coincide with the Kusuoka-Stroock Sobolev spaces which are defined using the concept of ray absolute continuity and stochastic Gateaux differentiability.

In Section 4 we formulate the conditions on the BSDE such that it is Malliavin differentiable, present the proof and give an example.

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 $\mathcal{F} = \mathcal{F}_T$.

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

$$X_t = \gamma t + \sigma W_t + \int_{]0,t]\times\{|x|\le 1\}} x\tilde{N}(ds, dx) + \int_{]0,t]\times\{|x|>1\}} xN(ds, dx), \tag{2.1}$$

where $\sigma \ge 0$, W is a Brownian motion and N ($\tilde{N}$) is the (compensated) Poisson random measure corresponding to X.

The process

$$\left(\int_{]0,t]\times\{|x|\le 1\}} x\tilde{N}(ds, dx) + \int_{]0,t]\times\{|x|>1\}} xN(ds, dx)\right)$$

is the jump part of X and will be denoted by J. Note that the $\mathbb{P}$-augmented filtrations $(\mathcal{F}_t^W)_{t\in[0,T]}$ resp. $(\mathcal{F}_t^J)_{t\in[0,T]}$ generated by the processes W resp. J satisfy

$$\mathcal{F}_t^W \vee \mathcal{F}_t^J = \mathcal{F}_t,$$

(see [35, Lemma 3.1]) thus spanning the original filtration generated by X again. Throughout the paper we will use the notation $X(\omega) = (X_t(\omega))_{t\in[0,T]}$ for sample trajectories. Let ΔX given by $\Delta X_t := X_t - \lim_{s\nearrow t} X_s$ denote the process of the jumps of X.

Let

$$\mu(dx) := \sigma^2\delta_0(dx) + \nu(dx)$$

and

$$\mathbb{m}(dt, dx) := (\lambda \otimes \mu)(dt, dx)$$

where λ denotes the Lebesgue measure. We define the independent random measure (in the sense of [19, p. 256]) M by

$$M(dt, dx) := \sigma dW_t\delta_0(dx) + \tilde{N}(dt, dx) \tag{2.2}$$

on sets $B \in \mathcal{B}([0, T] \times \mathbb{R})$ with $\mathbb{m}(B) < \infty$. It holds $\mathbb{E}M(B)^2 = \mathbb{m}(B)$.

In [35], Solé et al. consider the independent random measure $\sigma dW_t\delta_0(dx)+x\tilde{N}(dt, dx)$. Here, in order to match the notation used for BSDEs, we work with the *equivalent* approach where the Poisson random measure is not multiplied with x.

We close this section with notation for càdlàg processes on the path space and for BSDEs.

Notation: Skorohod space

- With $D[0, T]$ we denote the Skorohod space of càdlàg functions on the interval $[0, T]$ equipped with the Skorohod topology. The σ-algebra $\mathcal{B}(D[0, T])$ is the Borel σ-algebra i.e. it is generated by the open sets of $D[0, T]$. It coincides with the σ-algebra generated by the family of coordinate projections $(p_t : D[0, T] \to \mathbb{R},\ \mathrm{x} \mapsto \mathrm{x}(t),\ t \ge 0)$ (see Theorem 12.5 of [8] for instance).
- For a measurable mapping $\mathrm{Y} : \Omega \to D[0, T], \omega \mapsto \mathrm{Y}(\omega)$, the probability measure $\mathbb{P}_\mathrm{Y}$ on $(D[0, T], \mathcal{B}(D[0, T]))$ denotes the image measure of $\mathbb{P}$ under Y.

- For a fixed $t \in [0,T]$ the notation

$$\mathtt{x}^t(s) := \mathtt{x}(t \wedge s), \text{ for all } s \in [0,T] \tag{2.3}$$

induces the natural identification

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

By this identification we define a filtration on this space by

$$\mathcal{G}_t = \sigma\left(\mathcal{B}\left(D[0,t]\right) \cup \mathcal{N}_X[0,T]\right), \quad 0 \le t \le T, \tag{2.4}$$

where $\mathcal{N}_X[0,T]$ denotes the null sets of $\mathcal{B}\left(D[0,T]\right)$ with respect to the image measure $\mathbb{P}_X$ of the Lévy process X. For more details on $D[0,T]$, see [8] and [15, Section 4].

Notation for BSDEs

- For $1 \le p \le \infty$ let $\mathcal{S}_p$ denote the space of all $(\mathcal{F}_t)$-progressively measurable and càdlàg processes $Y : \Omega \times [0,T] \to \mathbb{R}$ such that

$$\|Y\|_{\mathcal{S}_p} := \| \sup_{0 \le t \le T} |Y_t| \|_{\mathrm{L}_p} < \infty.$$

- We define $\mathrm{L}_2(W)$ as the space of all $(\mathcal{F}_t)$-progressively measurable processes $Z : \Omega \times [0,T] \to \mathbb{R}$ such that

$$\|Z\|^2_{\mathrm{L}_2(W)} := \mathbb{E} \int_0^T |Z_s|^2 \, ds < \infty.$$

- Let $\mathbb{R}_0 := \mathbb{R} \backslash \{0\}$. We define $\mathrm{L}_2(\tilde{N})$ as the space of all random fields $U : \Omega \times [0,T] \times \mathbb{R}_0 \to \mathbb{R}$ 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

$$\|U\|^2_{\mathrm{L}_2(\tilde{N})} := \mathbb{E} \int_{[0,T] \times \mathbb{R}_0} |U_s(x)|^2 \, ds \nu(dx) < \infty.$$

- We define $\mathrm{L}_2(M)$ by $\mathrm{L}_2(M) := \mathrm{L}_2(W) \oplus \mathrm{L}_2(\tilde{N})$ which is the space of all random fields $\underline{Z} : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ which are measurable with respect to $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$ such that

$$\|\underline{Z}\|^2_{\mathrm{L}_2(M)} := \mathbb{E} \int_{[0,T] \times \mathbb{R}} \left| \underline{Z}_{s,x} \right|^2 \mathbb{m}(ds, dx) < \infty.$$

- $\mathrm{L}_2(\nu) := \mathrm{L}_2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$.
- $|\cdot|$ denotes a norm in $\mathbb{R}^n$.
- For later use we recall the notion of the predictable projection of a stochastic process depending on parameters.

According to [33, Proposition 3] (see also [25, Proposition 3] or [2, Lemma 2.2]) for any $z \in \mathrm{L}_2(\mathbb{P} \otimes \mathbb{m}) := \mathrm{L}_2(\Omega \times [0,T] \times \mathbb{R}, \mathcal{F}_T \otimes \mathcal{B}([0,T] \times \mathbb{R}), \mathbb{P} \otimes \mathbb{m})$ there exists a process

$$^p z \in \mathrm{L}_2\left(\Omega \times [0,T] \times \mathbb{R}, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}), \mathbb{P} \otimes \mathbb{m}\right)$$

such that for any fixed $x \in \mathbb{R}$ the function $({}^p z)_{\cdot,x}$ is a version of the predictable projection (in the classical sense, see e.g. [2, Definition 2.1]) of $z_{\cdot,x}$. In the following we will always use this result to get predictable projections which are measurable w.r.t. a parameter. Again, we call ${}^p z$ the predictable projection of z.

3 Malliavin calculus

3.1 Definition of $\mathbb{D}_{1,2}$ using chaos expansions

The random measure M defined in (2.2) allows to introduce the Malliavin derivative defined via chaos expansions (see, for example, [34]) as follows: Any $\xi \in \mathrm{L}_2 := \mathrm{L}_2(\Omega, \mathcal{F}, \mathbb{P})$ has a unique chaos expansion (see [19, Theorem 2])

$$\xi = \sum_{n=0}^{\infty} I_n(\tilde{f}_n)$$

and it holds

$$\mathbb{E}\xi^2 := \|\xi\|_{\mathrm{L}_2}^2 = \sum_{n=0}^{\infty} n! \left\|\tilde{f}_n\right\|_{\mathrm{L}_2^n}^2$$

where the $\tilde{f}_n \in \widetilde{\mathrm{L}}_2^n := \widetilde{\mathrm{L}}_2\left(([0,T]\times\mathbb{R})^n, \mathbb{m}^{\otimes n}\right)$, the subspace of symmetric functions from $\mathrm{L}_2^n := \mathrm{L}_2\left(([0,T]\times\mathbb{R})^n, \mathbb{m}^{\otimes n}\right)$, and I_n denotes the n-th multiple integral with respect to M from (2.2). The multiple integrals with respect to M can be defined as follows: If $n = 0$ set $\mathrm{L}_2^0 := \mathbb{R}$ and $I_0(f_0) := f_0$ for $f_0 \in \mathbb{R}$. For $n \geq 1$ we start with a simple function $f_n \in \mathrm{L}_2^n$ given by

$$f_n\left((t_1,x_1),\dots,(t_n,x_n)\right) = \sum_{k=1}^{m} a_k \prod_{i=1}^{n} \mathbb{1}_{B_i^k}(t_i,x_i),$$

where the sets $B_i^k \in \mathcal{B}([0,T]\times\mathbb{R})$ for $k = 1,\dots,m, i = 1,\dots,n$ are disjoint for fixed k, and $\mathbb{m}(B_i^k) < \infty$ for all i and k. Then

$$I_n(f_n) := \sum_{k=1}^{m} a_k \prod_{i=1}^{n} M(B_i^k).$$

By denseness of these simple functions in L_2^n and by linearity and continuity of I_n, one extends the domain of the n-fold multiple stochastic integral I_n to become a mapping $I_n \colon \mathrm{L}_2^n \to \mathrm{L}_2$. It holds $I_n(f_n) = I_n(\tilde{f}_n)$ where $\tilde{f}_n$ denotes the symmetrization of f_n w.r.t. the n pairs of variables in $[0,T]\times\mathbb{R}$. For $f_n \in \mathrm{L}_2^n$ and $g_m \in \mathrm{L}_2^m$ we have

$$\mathbb{E}I_n(f_n)I_m(g_m) = \begin{cases} n! \int_{([0,T]\times\mathbb{R})^n} \tilde{f}_n \tilde{g}_n d\mathbb{m}^{\otimes n}, & n = m, \\ 0, & n \neq m. \end{cases}$$

The space $\mathbb{D}_{1,2}$ consists of all random variables $\xi \in \mathrm{L}_2$ such that

$$\|\xi\|_{\mathbb{D}_{1,2}}^2 := \sum_{n=0}^{\infty} (n+1)! \left\|\tilde{f}_n\right\|_{\mathrm{L}_2^n}^2 < \infty.$$

The Malliavin derivative is defined for $\xi \in \mathbb{D}_{1,2}$ by

$$\mathcal{D}_{t,x}\xi := \sum_{n=1}^{\infty} n I_{n-1}\left(\tilde{f}_n\left((t,x),\ \cdot\ \right)\right),$$

for $\mathbb{P}\otimes\mathbb{m}$-a.a. $(\omega,t,x) \in \Omega\times[0,T]\times\mathbb{R}$. Thus $\mathcal{D}\xi \in \mathrm{L}_2(\mathbb{P}\otimes\mathbb{m})$.

We also consider

$$\begin{aligned} \mathbb{D}_{1,2}^0 \quad := \quad & \left\{ \xi = \sum_{n=0}^{\infty} I_n(\tilde{f}_n) \in \mathrm{L}_2 \colon \tilde{f}_n \in \widetilde{\mathrm{L}}_2^n, n \in \mathbb{N}, \right. \\ & \left. \sum_{n=1}^{\infty} (n+1)! \int_0^T \|\tilde{f}_n((t,0),\cdot)\|_{\mathrm{L}_2^{n-1}}^2 dt < \infty \right\} \end{aligned} \tag{3.1}$$

and

$$\mathbb{D}_{1,2}^{\mathbb{R}_0} \;:=\; \Big\{\xi = \sum_{n=0}^{\infty} I_n(\tilde{f}_n) \in \mathrm{L}_2 \colon \tilde{f}_n \in \widetilde{\mathrm{L}}_2^n, n \in \mathbb{N}, \\ \sum_{n=1}^{\infty}(n+1)! \int_{[0,T]\times\mathbb{R}_0} \|\tilde{f}_n((t,x),\cdot)\|^2_{\mathrm{L}_2^{n-1}} \mathrm{m}(dt,dx) < \infty\Big\}.$$

$$\text{If } \sigma > 0 \text{ and } \nu \neq 0 \text{ it holds } \quad \mathbb{D}_{1,2} = \mathbb{D}_{1,2}^0 \cap \mathbb{D}_{1,2}^{\mathbb{R}_0}. \tag{3.2}$$

3.2 From canonical to general probability spaces

Solé et al. introduced in [35] the canonical Lévy space and proved that for $x \neq 0$ the Malliavin derivative $\mathcal{D}_{r,x}\xi$ (defined via chaos expansions) equals in this space an increment quotient. We will discuss here how to transfer results about random variables from the canonical Lévy space to any general probability space carrying a Lévy process provided that the regarded σ -algebra is the completion of the one generated by the Lévy process.

This technique is needed, since key theorems of this section, like Theorem 3.12, will be proven on specific probability spaces. However, the formulation of its assertion is possible also on general probability spaces. The validity of the assertion is then guaranteed by the transfer technique given in Theorem 3.1. Hence, in Section 4, where we apply this section's theorems to BSDEs, we are not restricted to certain specific probability spaces.

Assume $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ to be complete probability spaces with càdlàg Lévy processes $X^i = (X_t^i)_{t\in[0,T]}$, $X_t^i \colon \Omega_i \to \mathbb{R}$, such that X^i corresponds to a given Lévy triplet (γ, σ, ν) for $i = 1, 2$. Furthermore, assume that $\mathcal{F}_i$ is the completion of the σ-algebra generated by X^i. For the processes X^1, X^2, we get the associated independent random measures M^1 and M^2 like in (2.2), and the families of multiple stochastic integrals

$$\left(I_n^1(f_n)\right)_{n\in\mathbb{N}}, \left(I_n^2(f_n)\right)_{n\in\mathbb{N}},$$

respectively. The following assertion is taken from [32, Corollary 4.2], where it is formulated for Lévy processes with paths in $D[0,\infty[$.

Theorem 3.1. *Let $(E, \mathcal{E}, \rho)$ be a σ-finite measure space and let*

$$C^1 \in \mathrm{L}_2\left(\Omega_1 \times E, \mathcal{F}_1 \otimes \mathcal{E}, \mathbb{P}_1 \otimes \rho\right),$$

$$C^2 \in \mathrm{L}_2\left(\Omega_2 \times E, \mathcal{F}_2 \otimes \mathcal{E}, \mathbb{P}_2 \otimes \rho\right)$$

and suppose that these random fields have chaos decompositions

$$C^1 = \sum_{n=0}^{\infty} I_n^1(f_n),\ \mathbb{P}_1 \otimes \rho\text{-a.e.}, \quad C^2 = \sum_{n=0}^{\infty} I_n^2(g_n),\ \mathbb{P}_2 \otimes \rho\text{-a.e.}$$

for f_n, g_n being functions in $\mathrm{L}_2(E, \mathcal{E}, \rho)\hat{\otimes}\mathrm{L}_2^n$ which are symmetric in the last n variables, where '$\hat{\otimes}$' denotes the Hilbert space tensor product.

Assume that for ρ-almost all $e \in E$ there are functionals

$$F_e \colon D\left([0,T]\right) \to \mathbb{R}$$

such that $C^i(e) = F_e\left((X_t^i)_{t\in[0,T]}\right)$, $\mathbb{P}_i$-a.s. for $i = 1, 2$. Then for all $n \in \mathbb{N}$ it holds $f_n = g_n$, $\rho \otimes \mathrm{m}^{\otimes n}$-a.e.

Roughly speaking, if we have the same functionals F_e acting on both Lévy processes X^i defined on the probability spaces $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ for $i = 1, 2$ then the deterministic kernels of their chaos expansions coincide.

The Factorization lemma (see, for instance, [5, Section II.11]) implies that for any $\xi \in \mathrm{L}_2$ there exists a measurable functional $g_\xi \colon D([0,T]) \to \mathbb{R}$ such that

$$\xi(\omega) = g_\xi\left((X_t(\omega))_{0\le t\le T} \right) = g_\xi(X(\omega))$$

for a.a. $\omega \in \Omega$.

The following characterization that $g_\xi(X) \in \mathbb{D}_{1,2}^{\mathbb{R}_0}$ is a consequence from Alòs, León and Vives [1, Corollary 2.3. and Lemma 2.1] (this results hold true for a general Lévy measure since the square integrability of the Lévy process stated at the beginning of [1] is in fact only used from [1, Section 2.4] on) and Theorem 3.1. For details see the proof in [32, Theorem 5.1].

Lemma 3.2. *If $g_\xi(X) \in \mathrm{L}_2$ then $g_\xi(X) \in \mathbb{D}_{1,2}^{\mathbb{R}_0} \iff$*

$$g_\xi(X + x\mathbb{1}_{[t,T]}) - g_\xi(X) \in \mathrm{L}_2(\mathbb{P}\otimes\mathbb{m}) \tag{3.3}$$

and it holds then for $x \neq 0$ $\mathbb{P}\otimes\mathbb{m}$-a.e.

$$\mathcal{D}_{t,x}\xi = g_\xi(X + x\mathbb{1}_{[t,T]}) - g_\xi(X). \tag{3.4}$$

Compared to the approach of [35] which uses the random measure $\sigma dW_t\delta_0(dx) + x\tilde{N}(dt,dx)$, here the according Malliavin derivative for $x \neq 0$ and M from (2.2) is just a *difference* instead of the difference quotient from [35].

Applied on $g_\xi(X(\omega))$ this gives in the canonical space

$$g_\xi(X(\omega_{r,x})) - g_\xi(X(\omega)) = g_\xi(X(\omega) + x\mathbb{1}_{[r,T]}) - g_\xi(X(\omega))$$

for $\mathbb{P}\otimes\mathbb{m}$ a.e. (ω, r, x).

In the situation of the previous lemma, one may ask whether properties of $g_\xi(X)$ that hold $\mathbb{P}$-a.s. are preserved $\mathbb{P}\otimes\mathbb{m}$-a.e. for $g_\xi(X + x\mathbb{1}_{[t,T]})$. The positive answer is given by the following result (the proof can be found in the appendix).

Lemma 3.3. *Let $\Lambda \in \mathcal{G}_T$ be a set with $\mathbb{P}(\{X \in \Lambda\}) = 0$. Then*

$$\mathbb{P}\otimes\mathbb{m}\left(\left\{(\omega, r, v) \in \Omega\times[0,T]\times\mathbb{R}_0 : X(\omega) + v\mathbb{1}_{[r,T]} \in \Lambda\right\}\right) = 0.$$

Corollary 3.4.

(i) Let $f \colon D[0,T]\times\mathbb{R} \to \mathbb{R}$ be a measurable mapping such that $\mathbb{P}$-a.s. $y \mapsto f(X(\omega), y)$ is a Lipschitz function with Lipschitz constant L independent from $\omega \in \Omega$. Then the set

$$\Lambda := \{\mathrm{x} \in D[0,T] : y \mapsto f(\mathrm{x}, y) \text{ is not Lipschitz in } y \text{ with constant } L\}$$

satisfies $\mathbb{P}(X \in \Lambda) = 0$. Lemma 3.3 implies that also

$$y \mapsto f(X(\omega) + v\mathbb{1}_{[r,T]}, y)$$

is a Lipschitz function with constant L for $\mathbb{P}\otimes\mathbb{m}$-a.e. $(\omega, r, v) \in \Omega\times[0,T]\times\mathbb{R}_0$.

(ii) Let $\xi = g_\xi(X) \in L_\infty(\Omega)$. By the same reasoning as in (i)*, it follows from Lemma 3.3 that $\mathbb{P}\otimes\mathbb{m}$-a.e. the random element $g_\xi(X + v\mathbb{1}_{[r,T]})$ is bounded.*

Note that the boundedness of $g_\xi(X + v\mathbb{1}_{[r,T]})$ implies boundedness of the difference in (3.4),

$$g_\xi(X + v\mathbb{1}_{[r,T]}) - g_\xi(X),$$

which – in case of L_2-integrability w.r.t. $\mathbb{P}\otimes\mathbb{m}$ – equals the Malliavin derivative for $v \neq 0$.

3.3 Malliavin calculus for random functions

We want to address the following problem: Let

$$F : \Omega \times \mathbb{R}^d \to \mathbb{R}$$

be jointly measurable, for any $y \in \mathbb{R}^d$ we assume $F(\cdot, y) \in \mathbb{D}_{1,2}$, and for a.a. $\omega \in \Omega$ let $F(\omega, \cdot) \in \mathcal{C}^1(\mathbb{R}^d)$. If $G_1, ..., G_d \in \mathbb{D}_{1,2}$, under which assumption do we get

$$F(\cdot, G_1, ..., G_d) \in \mathbb{D}_{1,2}?$$

We will treat this question in two steps: First we will find conditions on F and $G = (G_1, ..., G_d)$ such that

- $F(\cdot, G) \in \mathbb{D}_{1,2}^{\mathbb{R}_0}$
- $F(\cdot, G) \in \mathbb{D}_{1,2}^{0}$

separately and then use relation (3.2).

3.3.1 The case $F(\cdot, G) \in \mathbb{D}_{1,2}^{\mathbb{R}_0}$

Lemma 3.5. *Assume that $F(\cdot, y) \in \mathbb{D}_{1,2}^{\mathbb{R}_0}$ for all $y \in \mathbb{R}^d$, $F(\cdot, G) \in \mathrm{L}_2$, and $G_1, ..., G_d \in \mathbb{D}_{1,2}^{\mathbb{R}_0}$. Let $F(\omega, \cdot) \in \mathcal{C}(\mathbb{R}^d)$ $\mathbb{P}$-a.s. and let F be represented by the functional $g_F(X, \cdot)$. Then $F(\cdot, G) \in \mathbb{D}_{1,2}^{\mathbb{R}_0} \iff$*

$$\begin{aligned}(\mathcal{D}_{t,x}F)(\cdot, G) + \;\; & g_F(X + x\mathbb{1}_{[t,T]}, G + \mathcal{D}_{t,x}G) - g_F(X + x\mathbb{1}_{[t,T]}, G) \\ & \in \mathrm{L}_2(\Omega \times [0,T] \times \mathbb{R}_0, \mathbb{P} \otimes \mathbb{m}). \end{aligned} \tag{3.5}$$

Proof. By the expression $(\mathcal{D}_{t,x}F)(\cdot, G)$ we mean that we insert the L_2-vector $(G_1, \dots, G_d)$ into the y-variable of $\mathcal{D}_{t,x}F(\cdot, y)$. Furthermore, since by Lemma 3.3, expression $g_F(X(\omega) + x\mathbb{1}_{[t,T]}, y)$ is continuous in y for $\mathbb{P} \otimes \mathbb{m}$-a.e. $(\omega, t, x) \in \Omega \times [0,T] \times \mathbb{R}_0$, taking equivalence classes of

$$g_F(X(\omega) + x\mathbb{1}_{[t,T]}, y) \mid_{y=(G_1(\omega)+\mathcal{D}_{t,x}G_1(\omega),\dots,G_d(\omega)+\mathcal{D}_{t,x}G_d(\omega))}$$

for representatives $(G_1(\omega) + \mathcal{D}_{t,x}G_1(\omega), \dots, G_d(\omega) + \mathcal{D}_{t,x}G_d(\omega))$ leads to a well-defined $\mathrm{L}_0(\mathbb{P} \otimes \mathbb{m})$ object.

For the sufficiency, one can use the same arguments as for [32, Theorem 5.2]. There the proof is carried out only for $d = 1$ but it is easy to see that the multidimensional case can be proved in the same way.

For the necessity we consider $G_1, \dots, G_d$ as given by functionals $g_{G_1}, \dots, g_{G_d}$ and conclude from Lemma 3.2 that

$$\mathcal{D}_{t,x}F(\cdot, y) = g_F(X + x\mathbb{1}_{[t,T]}, y) - g_F(X, y).$$

Hence expression (3.5) equals in fact

$$\begin{aligned} & g_F(X + x\mathbb{1}_{[t,T]}, G_1 + \mathcal{D}_{t,x}G_1, ..., G_d + \mathcal{D}_{t,x}G_d) - g_F(X, G_1, ..., G_d) \\ = \;\; & g_F(X + x\mathbb{1}_{[t,T]}, g_{G_1}(X + x\mathbb{1}_{[t,T]}), ..., g_{G_d}(X + x\mathbb{1}_{[t,T]})) - g_F(X, G_1, ..., G_d) \\ = \;\; & \mathcal{D}_{t,x}F(X, G_1, ..., G_d) \end{aligned}$$

where we have used Lemma 3.2 again. □

3.3.2 The case $F(\cdot, G) \in \mathbb{D}^0_{1,2}$

The Lévy-Itô decomposition implies that the Brownian part and the pure jump part of a Lévy process are independent. Thus we may represent a copy of X on the completion of $(\Omega^W \times \Omega^J, \mathcal{F}^W \otimes \mathcal{F}^J, \mathbb{P}^W \otimes \mathbb{P}^J)$ as

$$X_t(\omega) = \gamma t + \sigma \omega^W_t + J_t(\omega^J), \quad t \in [0,T],$$

where $\omega = (\omega^W, \omega^J)$. Here $(\Omega^W, \mathcal{F}^W, \mathbb{P}^W)$ denotes the completed canonical Wiener space i.e. $\Omega^W := \mathcal{C}_0[0,T]$ is the space of continuous functions starting in 0, and $\mathcal{F}^W$ is the Borel σ-algebra completed with respect to the Wiener measure $\mathbb{P}^W$. The space $(\Omega^J, \mathcal{F}^J, \mathbb{P}^J)$ is a probability space carrying the pure jump process J, where $\mathcal{F}^J$ is generated by J and completed.

To work on the canonical space $(\Omega^W, \mathcal{F}^W, \mathbb{P}^W)$ we continue with a short reminder on Gaussian Hilbert spaces and refer the reader for more information to Janson [20]. Consider the Gaussian Hilbert space $\mathcal{H} := \left\{ \int_0^T h(s) dW_s : h \in \mathrm{L}_2[0,T] \right\}$. Because of Itô's isometry we may identify $\mathcal{H}$ with

$$\mathcal{H}_0 := \mathrm{L}_2[0,T].$$

The space

$$\mathcal{H}_1 := \left\{ \int_0^{\cdot} h(s) ds : h \in \mathrm{L}_2[0,T] \right\}$$

with $\langle \int_0^{\cdot} h_1(s)ds, \int_0^{\cdot} h_2(s)ds \rangle_{\mathcal{H}_1} := \int_0^T h_1(s)h_2(s)ds$ is the Cameron-Martin space. For $h \in \mathcal{H}_0$ we have $g_h \in \mathcal{H}_1$ with

$$g_h(t) := \mathbb{E}\left(\int_0^T h(s) dW_s \, W_t \right) = \int_0^t h(s) ds.$$

The main idea to get sufficient conditions for $F(\cdot, G) \in \mathbb{D}^0_{1,2}$ consists in applying Theorem 3.10 below. We proceed with a collection of definitions and some facts related to this theorem.

In the sequel let E be a separable Hilbert space.

Definition 3.6 ([36], [27]). *Let $1 \leq p < \infty$ and $\mathcal{S} \subseteq \mathbb{D}_{1,p}(\mathbb{P}^W)$ be a dense set of smooth random variables. By $\mathbb{D}^W_{1,p}(E)$ we denote the completion of*

$$\{\xi = \sum_{k=1}^n G_k H_k : G_k \in \mathcal{S}, H_k \in E\}$$

with respect to the norm

$$\|\xi\|_{1;E} := \left(\mathbb{E}\|\xi\|^p_E + \mathbb{E} \int_0^T \| D^W_t \xi\|^p_E dt \right)^{\frac{1}{p}}$$

where $D^W_t \xi := \sum_{k=1}^n (D^W_t G_k) H_k$.

Note that $\mathrm{L}_2(\Omega^J, \mathcal{F}^J, \mathbb{P}^J)$ is a separable Hilbert space, and that the space $\mathbb{D}^W_{1,2}(E)$ for $E := \mathrm{L}_2(\Omega^J, \mathcal{F}^J, \mathbb{P}^J)$ can be identified with $\mathbb{D}^0_{1,2}$ defined in (3.1) (see [1]). This means we may reformulate the question posed in the beginning of this section by asking for sufficient conditions such that

$$F(\cdot, G) \in \mathbb{D}^W_{1,2}(E).$$

The answer will be Theorem 3.12 at the end of this section.

Let E_1 and E_2 be separable Hilbert spaces. A bounded linear operator $A : E_1 \to E_2$ is is called *Hilbert-Schmidt operator* if for some orthonormal basis $\{e_n\}$ in E_1 it holds

$$\|A\|_{HS(E_1,E_2)} := \left(\sum_{n=1}^{\infty} \|Ae_n\|_{E_2}^2 \right)^{\frac{1}{2}} < \infty$$

(see, for example, [9]). We will denote by $HS(\mathcal{H}_0, E)$ the space of Hilbert-Schmidt operators between $\mathcal{H}_0$ and E.

Definition 3.7 ([20],[9])**.** *With* $\mathrm{L}_0(\mathbb{P}^W; E)$ *we denote the space of* E*-valued random variables, equipped with the topology of convergence in probability.*

For $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ *and* $h \in \mathcal{H}_0$ *we define the Cameron-Martin shift by*

$$\rho_h(\xi)(\omega^W) := \xi(\omega^W + g_h).$$

One of the properties of the Cameron-Martin shift is the Cameron-Martin formula. (For an integral of E-valued objects, we always use the Bochner integral.)

Lemma 3.8.

(i) (Cameron-Martin formula). $\mathbb{P}^W \sim \mathbb{P}^W \circ \rho_h^{-1}$ *for* $h \in \mathcal{H}_0$*, and the Radon-Nikodym derivative is given by*

$$\frac{d\mathbb{P}^W \circ \rho_h^{-1}}{d\mathbb{P}^W}(\omega^W) = \exp\left\{ -\frac{1}{2}\int_0^T h(t)^2 dt - \int_0^T h(t) dW_t \right\}.$$

(ii) If $K \in L_p(\mathbb{P}^W; E)$ *for some* $p > 1$ *then for any* $q \in [1, p[$

$$\left\| \int_0^T \rho_{sh} K ds \right\|_{L_q(\mathbb{P}^W;E)} \le \int_0^T \exp\left\{ \frac{s^2}{2(p-q)} \|h\|_{\mathcal{H}_0}^2 \right\} ds \; \|K\|_{L_p(\mathbb{P}^W;E)}.$$

(iii) For every $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ *the map* $\mathcal{H}_0 \to \mathrm{L}_0(\mathbb{P}^W; E) : h \mapsto \rho_h(\xi)$ *is continuous.*

Proof. (i) See Kuo [21, Theorem 1.1].

(ii) Analogously to the proof of Theorem 14.1 (vi) in Janson [20] for $1 \le q < p$ we choose $r = \frac{p}{p-q}$ so that $\frac{1}{r} + \frac{q}{p} = 1$, and by the Cameron-Martin formula and Hölders inequality we get

$$\begin{aligned}
&\left\| \int_0^T \rho_{sh} K ds \right\|_{L_q(\mathbb{P}^W;E)} \\
\le\; & \int_0^T \left(\mathbb{E} \|\rho_{sh} K\|_E^q \right)^{\frac{1}{q}} ds \\
=\; & \int_0^T \left(\mathbb{E} \exp\left\{ s\int_0^T h(t) dW_t - \frac{s^2}{2}\|h\|_{\mathcal{H}_0}^2 \right\} \|K\|_E^q \right)^{\frac{1}{q}} ds \\
\le\; & \|K\|_{L_p(\mathbb{P}^W;E)} \int_0^T \left(\mathbb{E} \exp\left\{ sr\int_0^T h(t) dW_t - \frac{s^2 r}{2}\|h\|_{\mathcal{H}_0}^2 \right\} \right)^{\frac{1}{rq}} ds \\
=\; & \|K\|_{L_p(\mathbb{P}^W;E)} \int_0^T \left(\exp\left\{ \frac{s^2(r^2-r)}{2}\|h\|_{\mathcal{H}_0}^2 \right\} \right)^{\frac{1}{rq}} ds \\
=\; & \|K\|_{L_p(\mathbb{P}^W;E)} \int_0^T \exp\left\{ \frac{s^2}{2(p-q)}\|h\|_{\mathcal{H}_0}^2 \right\} ds.
\end{aligned}$$

(iii) This assertion is formulated for real valued random variables in [20, Theorem 14.1 (viii)] but in [20, Remark 14.6] it is stated that it holds for random variables with values in a separable Banach space. □

Definition 3.9 ([20],[9]). *(i) A random variable $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ is absolutely continuous along $h \in \mathcal{H}_0$ (h-a.c.) if there exists a random variable $\xi^h \in \mathrm{L}_0(\mathbb{P}^W; E)$ such that $\xi^h = \xi$ a.s. and for all $\omega^W \in \Omega^W$ the map*

$$u \mapsto \xi^h(\omega^W + u\, g_h)$$

is absolutely continuous on bounded intervals of $\mathbb{R}$.

(ii) $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ is ray absolutely continuous (r.a.c.) if ξ is h-a.c. for every $h \in \mathcal{H}_0$.

(iii) For $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ and $h \in \mathcal{H}_0$ we say the directional derivative $\partial_h \xi \in \mathrm{L}_0(\Omega^W; E)$ exists if

$$\frac{\rho_{uh}(\xi) - \xi}{u} \overset{\mathbb{P}^W}{\to} \partial_h \xi, \quad u \to 0.$$

(iv) $\xi \in \mathrm{L}_0(\mathbb{P}^W; E)$ is called stochastically Gâteaux differentiable (s.G.d.) if $\partial_h \xi$ exists for every $h \in \mathcal{H}_0$ and there exists an $HS(\mathcal{H}_0, E)$-valued random variable denoted by $\tilde{\mathcal{D}}\xi$ such that for every $h \in \mathcal{H}_0$

$$\partial_h \xi = \langle \tilde{\mathcal{D}}\xi, h \rangle_{\mathcal{H}_0}, \quad \mathbb{P}^W\text{-}a.s.$$

According to Sugita [36], the Malliavin Sobolev spaces $\mathbb{D}^W_{n,p}(E)$ for $n \in \mathbb{N}, 1 < p < \infty$ and the Kusuoka-Stroock Sobolev spaces defined via the properties r.a.c. and s.G.d. coincide. According to Bogachev [9] this holds also for $p = 1$. Here we only use the assertion for $n = 1$:

Theorem 3.10 ([36, Theorem 3.1], [9, Proposition 5.4.6 (iii)]). *Let $p \in [1, \infty[$. Then*

$$\mathbb{D}^W_{1,p}(E) = \{\xi \in \mathrm{L}_p(\mathbb{P}^W; E) : \xi \text{ is r.a.c., s.G.d. and } \tilde{\mathcal{D}}\xi \in \mathrm{L}_p(\mathbb{P}^W; HS(\mathcal{H}_0; E))\},$$

and for $\xi \in \mathbb{D}^W_{1,p}(E)$ it holds $\mathcal{D}^W \xi = \tilde{\mathcal{D}}\xi$ a.s.

We will also need the following result.

Theorem 3.11. *For $h \in \mathcal{H}_0$ and $\xi \in \mathrm{L}_0(\Omega^W; E)$ it holds*

$$\xi \text{ is } h\text{-a.c} \iff \begin{cases} (i)\ \partial_h \xi \text{ exists} \\ (ii) \forall u \in \mathbb{R} : \rho_{uh}\xi(\omega^W) - \xi(\omega^W) = \int_0^u \rho_{sh}(\partial_h \xi)(\omega^W) ds \\ \quad \mathbb{P}^W\text{-a.s.,} \\ \quad \text{where } \int_{-|u|}^{|u|} \|\rho_{sh}(\partial_h \xi)(\omega^W)\|_E \, ds < \infty \ \mathbb{P}^W\text{-a.s.} \\ \quad \text{and } (s, \omega^W) \mapsto \rho_{sh}(\partial_h \xi)(\omega^W) \text{ denotes a jointly} \\ \quad \text{measurable version.} \end{cases}$$

Proof. For $E = \mathbb{R}$ this is Theorem 15.21 of [20]. One can generalize the proof to E-valued random variables since by the Radon-Nikodym property of E (see [16, Corollary IV.1.4]), the fundamental theorem of calculus holds for absolutely continuous functions if Bochner integrals are used. □

With the above preparations we are now able to find sufficient conditions for $F(\cdot, G) \in \mathbb{D}^W_{1,2}(E)$.

Theorem 3.12. *Assume that $E = \mathrm{L}_2(\Omega^J, \mathcal{F}^J, \mathbb{P}^J)$ and*

$$(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega^W \times \Omega^J, \mathcal{F}, \mathbb{P}^W \otimes \mathbb{P}^J),$$

where $\mathcal{F}$ is the completion of $\mathcal{F}^W \otimes \mathcal{F}^J$. Let

$$F : \Omega \times \mathbb{R}^d \to \mathbb{R}$$

be jointly measurable and $G_1, ..., G_d \in \mathbb{D}^W_{1,q}(E)$ for some $q \geq 1$. Suppose that $p > 1$ and *(i)* $F(\omega,\cdot)\in\mathcal{C}^1(\mathbb{R}^d)$ *for a.a.* $\omega\in\Omega$*,*

(ii) *for all* $y\in\mathbb{R}^d: F(\cdot,y)\in\mathbb{D}^W_{1,p}(E)$,

(iii) *for each* $N\in\mathbb{N}$, $\exists K_N\in\bigcup_{r>1}\mathrm{L}_r(\mathbb{P})$ *such that for a.a.* ω *it holds:*

$$\forall y,\tilde y\in B_N(0):=\{x\in\mathbb{R}^d:|x|\le N\}:$$
$$\|(\mathcal{D}^W F(\cdot,y))(\omega)-(\mathcal{D}^W F(\cdot,\tilde y))(\omega)\|_{\mathcal{H}_0}\le K_N(\omega)|y-\tilde y|,$$

(iv) $(\mathcal{D}^W F)(\cdot,G_1,...,G_d)\in\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E))$ *and*

$$\sum_{k=1}^d\frac{\partial}{\partial y_k}F(\cdot,G_1,...,G_d)\mathcal{D}^W G_k\in\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E)).$$

Then

$$F(\cdot,G_1,...,G_d)\in\mathbb{D}^W_{1,p}(E)$$

and

$$\mathcal{D}^W F(\cdot,G_1,...,G_d)=(\mathcal{D}^W F)(\cdot,G_1,...,G_d)+\sum_{k=1}^d\frac{\partial}{\partial y_k}F(\cdot,G_1,...,G_d)\mathcal{D}^W G_k$$

in $\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E))$.

Remark 3.13. In Theorem 3.12 it is possible to use also

$$\text{(iii)'}\quad\forall\varepsilon>0\,\exists\delta_\varepsilon(y)>0:\forall\omega\in\Omega,\forall\tilde y\in B_{\delta_\varepsilon}(y):$$

$$\|(\mathcal{D}^W F(\cdot,y))(\omega)-(\mathcal{D}^W F(\cdot,\tilde y))(\omega)\|_{\mathcal{H}_0}\le\varepsilon.$$

instead of (iii). Neither of both assumptions implies the other one.

Proof. `Step 1.` We will use the characterization of $\mathbb{D}^W_{1,p}(E)$ from Theorem 3.10. In fact, we will prove for any $u\in\mathbb{R}$ and $h\in\mathcal{H}_0$ the relations

$$\begin{aligned}\rho_{uh}F(\omega^W,G(\omega^W))-F(\omega^W,G(\omega^W))&=\int_0^u\rho_{sh}(\partial_h F)(\omega^W,G(\omega^W))ds,\quad\mathbb{P}^W\text{-a.s.,}\\(\partial_h F)(\cdot,G)&=\langle(\mathcal{D}^W F)(\cdot,G)+\nabla_y F(\cdot,G)\cdot\mathcal{D}^W G,h\rangle_{\mathcal{H}_0},\quad\mathbb{P}\text{-a.s.}\end{aligned}\tag{3.6}$$

where the first equation is E-valued with $G=(G_1,\dots,G_d)$, and the second equation is scalar with $\nabla_y=(\frac{\partial}{\partial y_1},\dots,\frac{\partial}{\partial y_d})$.

Since by assumption (iv)

$$(\mathcal{D}^W F)(\cdot,G)+\nabla_y F(\cdot,G)\cdot\mathcal{D}^W G\in\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E))$$

we infer that $\int_{-|u|}^{|u|}\|\rho_{sh}(\partial_h F)(\cdot,G)\|_E ds<\infty$, $\mathbb{P}^W$-a.s and according to Theorem 3.11 it follows from the first line of (3.6) that $F(\cdot,G)$ is r.a.c. From the second line of (3.6) we get that $F(\cdot,G)$ is s.G.d. and

$$\tilde{\mathcal{D}}F(\cdot,G)=(\mathcal{D}^W F)(\cdot,G)+\nabla_y F(\cdot,G)\cdot\mathcal{D}^W G$$

in $\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E))$. Together with Theorem 3.10 this would imply the assertion of the theorem. So it remains to show the relations in (3.6) which will be done in Steps 2 and 3.

`Step 2.` Since $F(\cdot,y)\in\mathbb{D}^W_{1,p}(E)$ we have by Theorem 3.10 that $F(\cdot,y)$ is r.a.c. and

$$\mathcal{D}^W F(\cdot,y)=\tilde{\mathcal{D}}F(\cdot,y)\in\mathrm{L}_p(\mathbb{P}^W;HS(\mathcal{H}_0,E)).$$

Hence by Theorem 3.11 for each $u \in \mathbb{R}$ and $h \in \mathcal{H}_0$ the *E-valued* equation

$$F(\omega^W + ug_h, y) - F(\omega^W, y) = \int_0^u \rho_{sh} \langle (\mathcal{D}^W F(\cdot, y))(\omega^W), h \rangle_{\mathcal{H}_0} ds,$$

holds for all ω^W up to an exception set $C_y \in \mathcal{F}^W$ with $\mathbb{P}^W(C_y) = 0$. Consequently, for each $u \in \mathbb{R}$ and $h \in \mathcal{H}_0$ we have the *real-valued* equation (where we use the notation $\rho_{uh}(\omega) = (\omega^W + ug_h, \omega^J)$)

$$F(\rho_{uh}(\omega), y) - F(\omega, y) = \int_0^u \rho_{sh} \langle (\mathcal{D}^W F(\cdot, y))(\omega), h \rangle_{\mathcal{H}_0} ds, \tag{3.7}$$

for all ω with the exception of a set $\bar{C}_y \in \mathcal{F}$ with $\mathbb{P}(\bar{C}_y) = 0$. Since the LHS is a.s. continuous in y, we can find an exception set $\bar{C} \in \mathcal{F}$ with $\mathbb{P}(\bar{C}) = 0$, which is independent of y, provided that we can show a.s. continuity in y of the RHS. To do this we estimate for $y, \tilde{y} \in B_N(0)$ the expression

$$\begin{aligned}
& \left| \int_0^u \rho_{sh} \langle (\mathcal{D}^W F(\cdot, \tilde{y}))(\omega), h \rangle_{\mathcal{H}_0} ds - \int_0^u \rho_{sh} \langle (\mathcal{D}^W F(\cdot, y))(\omega), h \rangle_{\mathcal{H}_0} ds \right| \\
\leq \quad & \left| \int_0^u \rho_{sh} \langle (\mathcal{D}^W F(\cdot, \tilde{y}))(\omega) - (\mathcal{D}^W F(\cdot, y))(\omega), h \rangle_{\mathcal{H}_0} ds \right| \\
\leq \quad & \|h\|_{\mathcal{H}_0} \int_0^u \|(\mathcal{D}^W F(\cdot, \tilde{y}))(\omega^W + sg_h, \omega^J) - (\mathcal{D}^W F(\cdot, y))(\omega^W + sg_h, \omega^J)\|_{\mathcal{H}_0} ds \\
\leq \quad & \|h\|_{\mathcal{H}_0} |y - \tilde{y}| \int_0^u K_N(\omega^W + sg_h, \omega^J) ds.
\end{aligned}$$

Since by Lemma 3.8 $\int_0^u \rho_{sh} K_N ds < \infty$, $\mathbb{P}$-a.s., it follows that for a.a. ω the RHS of (3.7) is continuous in y. Consequently, on $\Omega \backslash \bar{C} \in \mathcal{F}$ relation (3.7) is true for all $y \in \mathbb{R}^d$. Putting the terms to zero on $\bar{C}$, the right hand side of (3.7) is jointly measurable w.r.t. (ω, y). We may replace y by $G(\omega) := (G_1(\omega), \ldots, G_d(\omega))$ and get

$$F(\rho_{uh}(\omega), G(\omega)) - F(\omega, G(\omega)) = \int_0^u \langle (\mathcal{D}^W F)(\rho_{sh}(\omega), G(\omega)), h \rangle_{\mathcal{H}_0} ds, \ \mathbb{P}\text{- a.s.} \tag{3.8}$$

So far the Cameron-Martin shift ρ_{uh} acts only on the first variable of $F(\omega, G(\omega))$. In the following step we derive the representation for $\rho_{uh} F(\omega, G(\omega))$.

`Step 3.` We show that $F(\cdot, G)$ is r.a.c. For this we choose an interval $[0, t_1]$, $t \in [0, t_1]$, let $0 = s_0 < s_1 < ... < s_n = t_1$ and consider for $s_k^t := s_k \wedge t$ the expression

$$\rho_{th} F(\omega, G(\omega)) - F(\omega, G(\omega)) = \sum_{k=1}^n \rho_{s_k^t h} F(\omega, G(\omega)) - \rho_{s_{k-1}^t h} F(\omega, G(\omega)). \tag{3.9}$$

For any $b := s_k^t$ and $a := s_{k-1}^t$ we derive from (3.8) and the mean-value theorem that a.s.

$$\begin{aligned}
& \rho_{bh} F(\omega, G(\omega)) - \rho_{ah} F(\omega, G(\omega)) \\
= \quad & [F(\rho_{bh}(\omega), G(\rho_{bh}(\omega))) - F(\rho_{ah}(\omega), G(\rho_{bh}(\omega)))] \\
& + [F(\rho_{ah}(\omega), G(\rho_{bh}(\omega))) - F(\rho_{ah}(\omega), G(\rho_{ah}(\omega)))] \\
= \quad & \int_a^b \langle (\mathcal{D}^W F)(\rho_{sh}(\omega), G(\rho_{bh}(\omega))), h \rangle_{\mathcal{H}_0} ds \\
& + \nabla_y F \left(\rho_{ah}(\omega), G(\rho_{ah}(\omega)) + \theta [G(\rho_{bh}(\omega)) - G(\rho_{ah}(\omega))] \right) \\
& \qquad \cdot [G(\rho_{bh}(\omega)) - G(\rho_{ah}(\omega))]
\end{aligned}$$

for some $\theta \in [0,1]$. We may write the last term because $F(\omega, y)$ is $\mathcal{C}^1$ w.r.t. y. Similarly to (3.7) , for each $G_l \in \mathbb{D}^W_{1,q}(E)$, we have for all $t \in \mathbb{R}$ and $h \in \mathcal{H}_0$ that

$$\begin{aligned} \rho_{th} G_l(\omega) - G_l(\omega) &= \int_0^t \langle (\mathcal{D}^W G_l)(\rho_{sh}(\omega)), h\rangle_{\mathcal{H}_0} ds, \quad \mathbb{P}\text{-}a.s. \\ \text{with} \quad m(\omega) &:= \max_l \int_0^{t_1} |\rho_{sh} \langle (\mathcal{D}^W G_l)(\omega), h\rangle_{\mathcal{H}_0}| ds < \infty. \end{aligned} \tag{3.10}$$

To obtain (3.6) we rewrite (3.9) in the following way

$$\begin{aligned} &\rho_{th} F(\omega, G(\omega)) - F(\omega, G(\omega)) \\ =& \int_0^t \rho_{sh} [\langle (\mathcal{D}^W F)(\omega, G(\omega)), h\rangle_{\mathcal{H}_0} + \langle \nabla_y F(\omega, G(\omega)) \cdot (\mathcal{D}^W G)(\omega), h\rangle_{\mathcal{H}_0}] ds \\ &+ \sum_{k=1}^n \text{remainder terms.} \end{aligned}$$

The remainder terms are given by

$$\begin{aligned} & \int_{s^t_{k-1}}^{s^t_k} \rho_{sh} \langle (\mathcal{D}^W F)(\omega, G(\rho_{(s^t_k - s)h}(\omega))) - (\mathcal{D}^W F)(\omega, G(\omega)), h\rangle_{\mathcal{H}_0} ds \\ & + \int_{s^t_{k-1}}^{s^t_k} \{\rho_{s^t_{k-1}h} \nabla_y F((\omega), G(\omega) + \theta[\rho_{(s^t_k - s^t_{k-1})h} G(\omega) - G(\omega)]) \\ & \qquad - \rho_{s^t_{k-1}h} \nabla_y F(\omega, G(\omega))\} \cdot \rho_{sh} \partial_h G(\omega) ds \\ & + \int_{s^t_{k-1}}^{s^t_k} \{\rho_{s^t_{k-1}h} \nabla_y F(\omega, G(\omega)) - \rho_{sh} \nabla_y F(\omega, G(\omega))\} \rho_{sh} \partial_h G(\omega) ds \\ =& \; I_1 + I_2 + I_3, \end{aligned}$$

where we use $\partial_h G(\omega) := \langle (\mathcal{D}^W G)(\omega), h\rangle_{\mathcal{H}_0}$ as an abbreviation. It is sufficient to show that the sum of the remainder terms tends in probability to zero for a fixed sequence of partitions with $\|(s^{t,n}_k)\| := \max_{1 \le k \le n} |s^{t,n}_k - s^{t,n}_{k-1}| \to 0$. Because of (3.10), for arbitrary $\varepsilon_1 > 0$ one can choose $\|(s^{t,n}_k)\|$ sufficiently small such that for all $s \in [s^{t,n}_{k-1}, s^{t,n}_k]$

$$|\rho_{s^{t,n}_k h} G_l(\omega) - \rho_{sh} G_l(\omega))| \le \int_s^{s^{t,n}_k} |\langle \rho_{rh} (\mathcal{D}^W G_l)(\omega), h\rangle| dr < \varepsilon_1.$$

For ω with $\sup_{s^{t,n}_{k-1} \le s \le s^{t,n}_k} |\rho_{sh} G(\omega)| \le N$ assumption (iii) implies

$$\begin{aligned} & |\langle (\mathcal{D}^W F)(\rho_{sh}(\omega), G(\rho_{s^{t,n}_k h}(\omega))) - (\mathcal{D}^W F)(\rho_{sh}(\omega), G(\rho_{sh}(\omega))), h\rangle_{\mathcal{H}_0}| \\ \le& \; \|h\|_{\mathcal{H}_0} K_N(\rho_{sh}(\omega)) \, |G(\rho_{s^{t,n}_k h}(\omega)) - G(\rho_{sh}(\omega))| \\ \le& \; \|h\|_{\mathcal{H}_0} K_N(\rho_{sh}(\omega)) \, \varepsilon_1. \end{aligned}$$

Since Lemma 3.8 (ii) implies that $\int_{s^{t,n}_{k-1}}^{s^{t,n}_k} \rho_{sh} K_N(\omega) ds < \infty$ a.s. we have $I_1 \to 0$ a.s. for $\|(s^{t,n}_k)\| \to 0$.

To estimate I_2 we conclude from assumption (i) that for a.a. ω it holds for all n and $k = 1, ..., n$ that $F(\rho_{s^{t,n}_{k-1} h} \omega, \cdot) \in \mathcal{C}^1(\mathbb{R}^d)$. For any such ω and arbitrary $\varepsilon_2 > 0$ we have

$$|\rho_{s^{t,n}_{k-1} h} \{\nabla_y F(\omega, G(\omega) + \theta[\rho_{(s^{t,n}_k - s^{t,n}_{k-1})h} G(\omega) - G(\omega)]) - \nabla_y F(\omega, G(\omega))\}| < \varepsilon_2$$

if only $\|(s^{t,n}_k)\|$ is small enough.

For the remaining integral I_3 we proceed as follows: By Lemma 3.8 (iii) the map

$$[0,t_1] \ni s \mapsto \rho_{sh} \nabla_y F(\omega^W, G(\omega^W)) \in \mathrm{L}_0(\mathbb{P}^W; E)$$

is uniformly continuous, which then also holds for

$$[0,t_1] \ni s \mapsto \rho_{sh} \nabla_y F(\omega, G(\omega)) \in \mathrm{L}_0(\mathbb{P}).$$

Therefore,

$$\begin{aligned}
&\sum_{k=1}^n \left| \int_{s_{k-1}^{t,n}}^{s_k^{t,n}} \left[\rho_{s_{k-1}^{t,n} h} \nabla_y F(\omega, G(\omega)) - \rho_{sh} \nabla_y F(\omega, G(\omega))\right] \cdot \rho_{sh} (\partial_h G)(\omega) ds \right| \\
&\le \max_{1\le k\le n} \sup_{s_{k-1}^{t,n} \le r \le s_k^{t,n}} \left|\rho_{s_{k-1}^{t,n} h} \nabla_y F(\omega, G(\omega)) - \rho_{rh} \nabla_y F(\omega, G(\omega))\right| \\
&\qquad\qquad \times \int_0^t |\rho_{sh} (\partial_h G)(\omega)| ds \\
&\overset{\mathbb{P}}{\to} 0 \quad \text{for } \|(s_k^{t,n})\| \to 0.
\end{aligned}$$

□

4 Malliavin derivative of solutions to BSDEs

In this section we apply our theorems on Malliavin differentiability of random functions to generators of BSDEs. As a result we state in Theorem 4.4 that under conditions on the smoothness of the data (ξ, f) solutions to BSDEs are Malliavin differentiable. For simplicity, we set $\sigma = 1$ in (2.1). The assertions hold true (with the appropriate modifications) if at least one of them, σ or the Lévy measure ν, are non-zero.

For $0 \le t \le T$ we consider the BSDE

$$\begin{aligned}
Y_t \quad = \quad \xi + \int_t^T f\left(X, s, Y_s, Z_s, \int_{\mathbb{R}_0} g(U_s(x)) g_1(x) \nu(dx)\right) ds - \int_t^T Z_s dW_s \\
- \int_{]t,T]\times\mathbb{R}_0} U_s(x) \tilde{N}(ds, dx),
\end{aligned} \tag{4.1}$$

with $f\colon D[0,T] \times [0,T] \times \mathbb{R}^3 \to \mathbb{R}$. The conditions on g and g_1 are specified in ($\mathbf{A_f}$g) below and ensure that the integral is well-defined. We use the abbreviations

$$[g(u)]_\nu := \int_{\mathbb{R}_0} g(u(x)) g_1(x) \nu(dx) \tag{4.2}$$

where $u : \mathbb{R}_0 \to \mathbb{R}$ denotes a measurable function, and

$$f_g(h, s, y, z, u) := f(h, s, y, z, [g(u)]_\nu),$$

so that

$$\int_t^T f\left(X, s, Y_s, Z_s, \int_{\mathbb{R}_0} g(U_s(x)) g_1(x) \nu(dx)\right) ds = \int_t^T f_g\left(X, s, Y_s, Z_s, U_s\right) ds.$$

The motivation to consider an expression of this form arises from [26] and [7] where BSDEs related to utility maximization have been investigated. However, to show Malliavin differentiability, our expression had to be chosen in a simpler way. For the above expression, when g is the identical map, Malliavin differentiability of (Y, Z, U) has been stated in [13, Theorem 3.5.1].

For shortness of notation, we define

$$\underline{Z}_{s,x} := \begin{cases} Z_s, & x = 0, \\ U_s(x), & x \neq 0 \end{cases}$$

to write

$$\int_t^T Z_s dW_s + \int_{]t,T]\times\mathbb{R}_0} U_s(x)\tilde{N}(ds,dx) = \int_{]t,T]\times\mathbb{R}} \underline{Z}_{s,x} M(ds,dx).$$

For the terminal value ξ and the function f_g we agree upon the following assumptions:

$(\mathbf{A}_\xi)$ $\xi \in \mathbb{D}_{1,2}$.

$(\mathbf{A_f})$ a) $f\colon D[0,T]\times[0,T]\times\mathbb{R}^3 \to \mathbb{R}$ is jointly measurable, adapted to $(\mathcal{G}_t)_{t\in[0,T]}$ defined in (2.4).

b) $\mathbb{E}\int_0^T |f(X,t,0,0,0)|^2\, dt < \infty$.

c) $f(X,.,.) \in \mathcal{C}([0,T]\times\mathbb{R}^3)$ $\mathbb{P}$-a.s. and f satisfies the following Lipschitz condition: There exists a constant L_f such that for all $t\in[0,T], \eta, \tilde{\eta} \in \mathbb{R}^3$

$$|f\left(X,t,\eta\right) - f\left(X,t,\tilde{\eta}\right)| \le L_f|\eta-\tilde{\eta}|,$$

$\mathbb{P}$-a.s.

d) For all $t\in[0,T]$ and $i=1,2,3,$ $\exists\ \partial_{\eta_i} f(X,t,\eta)$ $\mathbb{P}$-a.s. and the functions

$$[0,T]\times\mathbb{R}^3 \ni (t,\eta) \mapsto \partial_{\eta_i} f(X,t,\eta)$$

are $\mathbb{P}$-a.s. continuous.

e) $f(X,t,\eta)\in\mathbb{D}_{1,2}$ for all $(t,\eta)\in[0,T]\times\mathbb{R}^3$, and $\forall t\in[0,T]$, $\forall N\in\mathbb{N}$ $\exists\, K_N^t \in \bigcup_{p>1} \mathrm{L}_p$ such that for a.a. ω

$$\forall \eta,\tilde{\eta} \in B_N(0):$$
$$\|\left(\mathcal{D}_{\cdot,0} f(X,t,\eta)\right)(\omega) - \left(\mathcal{D}_{\cdot,0} f(X,t,\tilde{\eta})\right)(\omega)\|_{\mathcal{H}_0} < K_N^t(\omega)\,|\eta-\tilde{\eta}|\,,$$

where for $\mathcal{D}_{\cdot,0} f(X,t,\eta)$ we always take a progressively measurable version in t.

f) Assume there is a random field $\Gamma\in\mathrm{L}_2(\mathbb{P}\otimes\mathbb{m})$, such that for all random vectors $G\in(\mathrm{L}_2)^3$ and for a.e. t it holds

$$|\left(\mathcal{D}_{s,x} f\right)(t,G)| \le \Gamma_{s,x}, \quad \mathbb{P}\otimes\mathbb{m}\text{-}a.e.$$

where $\left(\mathcal{D}_{s,x} f\right)(t,G) := \mathcal{D}_{s,x} f(X,t,\eta)\,|_{\eta=G}$.

g) $g\in\mathcal{C}^1(\mathbb{R})$ with bounded derivative and $g_1 \in \mathrm{L}_2(\mathbb{R}_0,\mathcal{B}(\mathbb{R}_0),\nu)$.

A triple $(Y,Z,U)\in\mathcal{S}_2\times\mathrm{L}_2(W)\times\mathrm{L}_2(\tilde{N})$ which satisfies (4.1) is called a solution to the BSDE (4.1).

Remark 4.1. 1. For a function $F\colon\Omega\times[0,T]\times\mathbb{R}^3\to\mathbb{R}$ being jointly measurable, adapted to $(\mathcal{F}_t)_{t\in[0,T]}$ one can always find a function f as in ($\mathbf{A_f}$a), such that $\mathbb{P}$-a.s. the equation

$$F(\omega,\cdot,\cdot) = f(X(\omega),\cdot,\cdot)$$

holds. Furthermore, for all $t\in[0,T]$, $\eta\in\mathbb{R}^3$ the equation

$$F(\omega,t,\eta) = f\left(X^t(\omega),t,\eta\right)\ \mathbb{P}\text{-a.s.}$$

is satisfied (for a proof see [15, Theorem 4.9.], [32, Lemma 3.2., Theorem 3.3.], and for the notation $X^t(\omega)$ recall (2.3)). In particular, for functions satisfying ($\mathbf{A_f}$a) it holds

$$f\left(h^t,t,\eta\right) = f\left(h,t,\eta\right)\ \mathbb{P}_X\text{-a.s.}$$

for all $t\in[0,T]$, $\eta\in\mathbb{R}^3$.

2. Assumption ($\mathbf{A}_f$f) is, in fact, stronger than needed in Theorem 4.4 below. It is enough to require that $|(\mathcal{D}_{s,x}f)(t,G)| \le \Gamma_{s,x},\ \mathbb{P}\otimes\mathbb{m}$-$a.e.$ holds for the solution $G=(Y_t,Z_t,U_t)$ and for the members $G=(Y^n_t,Z^n_t,U^n_t)$ of the approximating sequence appearing in the proof of Theorem 4.4. With this more general assumption one can study, for example, BSDEs with linear generators with random coefficients.

3. The assumption ($\mathbf{A}_f$g) on g can be extended to a dependency on t and ω. Also g_1 may be assumed to be time-dependent. To keep the same proof of Theorem 4.4 feasible, we have to impose conditions ($\mathbf{A}_f$a-f) on g (with $\mathbb{R}^3$ replaced by $\mathbb{R}$ as g is then a random process with one parameter). Furthermore, we have to assume that $g_1\colon[0,T]\times\mathbb{R}_0\to\mathbb{R}$ is Borel measurable and that $\|g_1(t,.)\|_{\mathrm{L}_2(\nu)}$ is bounded in $t\in[0,T]$.

To cover the issue of existence of solutions to BSDEs we refer to the following result:

Theorem 4.2 ([37], Lemma 2.4)**.** *Assume (ξ,f) satisfies the assumptions $\xi\in\mathrm{L}_2$ and ($\mathbf{A}_f$a-c). Then the BSDE (4.1) has a unique solution $(Y,Z,U)\in\mathcal{S}_2\times\mathrm{L}_2(W)\times\mathrm{L}_2(\tilde{N})$.*

We cite the stability result of Barles, Buckdahn and Pardoux ([6]) comparing the distance between solutions to the BSDE (4.1) with different terminal conditions and generators.

Theorem 4.3 ([6], Proposition 2.2)**.** *Assume that (ξ,f_g) and (ξ',f'_g) satisfy $\xi,\xi'\in\mathrm{L}_2$ and suppose the generators fulfill ($\mathbf{A}_f$a-c), while g is Lipschitz and $g_1\in\mathrm{L}_2(\mathbb{R}_0,\mathcal{B}(\mathbb{R}_0),\nu)$. Then there exists a constant $C>0$ such that for the corresponding solutions (Y,Z,U) and (Y',Z',U') to (4.1) it holds*

$$\begin{aligned}&\|Y-Y'\|^2_S+\|Z-Z'\|^2_{\mathrm{L}_2(W)}+\|U-U'\|^2_{\mathrm{L}_2(\tilde{N})}\\ \le\ & C\bigg(\|\xi-\xi'\|^2_{\mathrm{L}_2}+\int_0^T\|f_g(X,s,Y_s,Z_s,U_s)-f'_g(X,s,Y_s,Z_s,U_s)\|^2_{\mathrm{L}_2}ds\bigg).\end{aligned}$$

We state now the result about the Malliavin derivative of solutions to BSDEs. For the proof we apply Itô's formula like in the original work due to Pardoux and Peng [28] or in Ankirchner et al. [3]. The benefit is that one does not need any higher moment conditions on the data than L_2. Hence this result is a generalization of El Karoui et al. [18, Theorem 5.3]. It is also more general than [13, Theorem 3.5.1] of Delong: For example, we do not require a canonical Lévy space, the Lévy process does not need to be square integrable, and the generator in (4.1) allows some nonlinear structure w.r.t. $U_s(x)$ thanks to the function g.

Theorem 4.4. *Assume ($\mathbf{A}_\xi$) and ($\mathbf{A}_f$). Then the following assertions hold.*

(i) For $\mathbb{m}$- a.e. $(r,v)\in[0,T]\times\mathbb{R}$ there exists a unique solution $(\mathcal{Y}^{r,v},\mathcal{Z}^{r,v},\mathcal{U}^{r,v})\in\mathcal{S}_2\times\mathrm{L}_2(W)\times\mathrm{L}_2(\tilde{N})$ to the BSDE

$$\begin{aligned}\mathcal{Y}^{r,v}_t &= \mathcal{D}_{r,v}\xi+\int_t^T F_{r,v}\left(s,\mathcal{Y}^{r,v}_s,\mathcal{Z}^{r,v}_s,\mathcal{U}^{r,v}_s\right)ds\\ &\quad-\int_{]t,T]\times\mathbb{R}}\underline{\mathcal{Z}}^{r,v}_{s,x}M(ds,dx),\quad 0\le r\le t\le T\\ \mathcal{Y}^{r,v}_s &= \mathcal{Z}^{r,v}_s=\mathcal{U}^{r,v}_s=0,\quad 0\le s<r\le T,\end{aligned}\tag{4.3}$$

where

$$\underline{\mathcal{Z}}^{r,v}_{s,x}:=\begin{cases}\mathcal{Z}^{r,v}_s, & x=0\\ \mathcal{U}^{r,v}_s(x), & x\neq0,\end{cases}$$

and

$$F_{r,v}(s,y,z,u):=\begin{cases}(\mathcal{D}_{r,0}f_g)(s,y,z,u)\\ +\langle\nabla f(X,s,Y_s,Z_s,[g(U_s)]_\nu),(y,z,[g'(U_s)u]_\nu)\rangle, & v=0\\ \\ f_g\big(X+v\mathbb{1}_{[r,T]},s,Y_s+y,Z_s+z,U_s+u\big)\\ \qquad -f_g(X,s,Y_s,Z_s,U_s), & v\neq 0,\end{cases}$$

with $\nabla=(\partial_{\eta_1},\partial_{\eta_2},\partial_{\eta_3})$.

(ii) *For the solution* (Y,Z,U) *of* (4.1) *it holds*

$$Y,Z\in \mathrm{L}_2([0,T];\mathbb{D}_{1,2}),\quad U\in \mathrm{L}_2([0,T]\times\mathbb{R}_0;\mathbb{D}_{1,2}),\tag{4.4}$$

and $\mathcal{D}_{r,y}Y$ *admits a càdlàg version for* $\mathbb{m}$- *a.e.* $(r,y)\in[0,T]\times\mathbb{R}$.

(iii) $(\mathcal{D}Y,\mathcal{D}Z,\mathcal{D}U)$ *is a version of* $(\mathcal{Y},\mathcal{Z},\mathcal{U})$, *i.e. for* $\mathbb{m}$- *a.e.* (r,v) *it solves*

$$\begin{aligned}\mathcal{D}_{r,v}Y_t=&\mathcal{D}_{r,v}\xi+\int_t^T F_{r,v}\left(s,\mathcal{D}_{r,v}Y_s,\mathcal{D}_{r,v}Z_s,\mathcal{D}_{r,v}U_s\right)ds\\ &-\int_t^T\mathcal{D}_{r,v}Z_s dW_s-\int_{]t,T]\times\mathbb{R}_0}\mathcal{D}_{r,v}U_s(x)\tilde N(ds,dx),\quad 0\le r\le t\le T.\end{aligned}\tag{4.5}$$

(iv) *Setting* $D_{r,v}Y_r(\omega):=\lim_{t\searrow r}\mathcal{D}_{r,v}Y_t(\omega)$ *for all* (r,v,ω) *for which* $\mathcal{D}_{r,v}Y$ *is càdlàg and* $\mathcal{D}_{r,v}Y_r(\omega):=0$ *otherwise, we have*

$${}^p\Big((D_{r,0}Y_r)_{r\in[0,T]}\Big)\text{ is a version of }(Z_r)_{r\in[0,T]},$$
$${}^p\Big((D_{r,v}Y_r)_{r\in[0,T],v\in\mathbb{R}_0}\Big)\text{ is a version of }(U_r(v))_{r\in[0,T],v\in\mathbb{R}_0}.$$

We present an example of a FBSDE where we specify the dependence on ω in the generator by a forward process such that ($\mathbf{A}_f$f) holds.

Example 4.5. Consider the case of a Lévy process X such that $\mathbb{E}|X_t|^2<\infty$ for all $t\in[0,T]$. Assume the generator to be of the type

$$f(s,\omega,y,z,u)=\tilde f(s,\Psi_s(\omega),y,z,u),$$

with $\tilde f$ having a continuous partial derivative in the second variable bounded by K. Assume further that this partial derivative is locally Lipschitz in (y,z,u). Let Ψ denote a forward process given by the SDE

$$d\Psi_s=b(\Psi_s)ds+\sigma(\Psi_s)dW_s+\beta(\Psi_{s-},x)\tilde N(ds,dx)$$

with $\Psi_0\in\mathbb{R}$. Then conditions ($\mathbf{A}_f$e), ($\mathbf{A}_f$f) are satisfied under the requirements

(i) The functions $b\colon\mathbb{R}\to\mathbb{R}$ and $\sigma\colon\mathbb{R}\to\mathbb{R}$ are continuously differentiable with bounded derivative.

(ii) $\beta\colon\mathbb{R}\times\mathbb{R}_0\to\mathbb{R}$ is measurable, satisfies

$$\begin{aligned}&|\beta(\psi,x)|\le C_\beta(1\wedge|x|),\quad(\psi,x)\in\mathbb{R}\times\mathbb{R}_0,\\ &\Big|\beta(\psi,x)-\beta(\hat\psi,x)\Big|\le C_\beta|\psi-\hat\psi|(1\wedge|x|),\quad(\psi,x),(\hat\psi,x)\in\mathbb{R}\times\mathbb{R}_0,\end{aligned}$$

and is continuously differentiable in ψ for fixed $x\in\mathbb{R}_0$.

This follows, since $(\mathcal{D}_{s,x}f)(t,G)$ is given by

$$(\mathcal{D}_{s,x}f)(t,G) = \begin{cases} \tilde{f}_\psi(t,\Psi_t,G)\mathcal{D}_{s,x}\Psi_t, & x=0, \\ \tilde{f}(t,\Psi_t+\mathcal{D}_{s,x}\Psi_t,G)-\tilde{f}(t,\Psi_t,G), & x\neq 0, \end{cases}$$

implying

$$|(\mathcal{D}_{s,x}f)(t,G)| < K\,|\mathcal{D}_{s,x}\Psi_t|\,.$$

Theorem [13, Theorem 4.1.2] states that under the above conditions on b,σ and β,

$$\sup_{r,v}\mathbb{E}\sup_{s\in[0,T]}\left|\frac{\mathcal{D}_{r,v}\Psi_s}{v}\right|^2 < \infty,$$

and refers to [30, Theorem 3] for a proof. Thus, to satisfy ($\mathbf{A}_f$f), we may choose $\Gamma = C\sup_{s\in[0,T]}|\mathcal{D}\Psi_s|$, where C depends on K, C_β and the Lipschitz constants for b,σ and β.

4.1 Proof of Theorem 4.4

Let us start with a lemma providing estimates for the Malliavin derivative of the generator.

Lemma 4.6. *Let $G=(G_1,G_2,G_3)\in(\mathrm{L}_2)^3$ and $\Phi\in(\mathrm{L}_2(\mathbb{P}\otimes\mathbb{m}))^3$. If f satisfies ($\mathbf{A}_f$) it holds for $\mathbb{P}\otimes\mathbb{m}$-a.a. $(\omega,r,v), v\neq 0$, that*

$$\left|f(X+v\mathbb{1}_{[r,T]},t,G+\Phi_{r,v}) - f(X,t,G)\right| \le L_f\,|\Phi_{r,v}| + \Gamma_{r,v}. \tag{4.6}$$

Moreover, for $G\in(\mathbb{D}_{1,2})^3$ it holds $f(X,t,G)\in\mathbb{D}_{1,2}$ and

$$|\mathcal{D}_{r,v}f(X,t,G)| \le L_f\,|D_{r,v}G| + \Gamma_{r,v}, \quad \mathbb{P}\otimes\mathbb{m}\text{-a.e.} \tag{4.7}$$

Proof. According to Corollary 3.4 we may replace X by $X+v\mathbb{1}_{[r,T]}$ and use the Lipschitz property ($\mathbf{A}_f c$) to estimate

$$\left|f(X+v\mathbb{1}_{[r,T]},t,G+\Phi_{r,v}) - f(X+v\mathbb{1}_{[r,T]},t,G)\right| \le L_f\,|\Phi_{r,v}|$$

for $\mathbb{P}\otimes\mathbb{m}$-a.e. (ω,r,v) with $v\neq 0$. From ($\mathbf{A}_f f$) one concludes then (4.6).

For $v\neq 0$ we conclude from Lemma 3.5 that $\mathcal{D}_{r,v}f(X,t,G)\in\mathbb{D}_{1,2}^{\mathbb{R}_0}$ and apply Lemma 3.2 to get

$$\mathcal{D}_{r,v}f(X,t,G) = f(X+v\mathbb{1}_{[r,T]},t,G+\mathcal{D}_{r,v}G) - f(X,t,G)\,,$$

and hence (4.7) follows from (4.6). In the case of $v=0$, by assumption ($\mathbf{A}_f e$) we may apply Theorem 3.12. Thus we get the Malliavin derivative

$$\begin{aligned}\mathcal{D}_{r,0}f(X,t,G) \;&=\; (\mathcal{D}_{r,0}f)(t,G) + \partial_{\eta_1}f(X,t,G)\mathcal{D}_{r,0}G_1 \\ &\quad +\partial_{\eta_2}f(X,t,G)\mathcal{D}_{r,0}G_2 + \partial_{\eta_3}f(X,t,G)\mathcal{D}_{r,0}G_3\end{aligned} \tag{4.8}$$

for $\mathbb{P}\otimes\lambda$ a.a. $(\omega,r)\in\Omega\times[0,T]$. Relation (4.7) follows from conditions ($\mathbf{A}_f c$) and (f) using that the partial derivatives are bounded by L_f. □

Proof of Theorem 4.4. The core of the proof is to conclude assertion (ii) which will be done by an iteration argument. To simplify the notation we do not mention the dependency of f on X in most places.

(i) For those (r,v) such that $\mathcal{D}_{r,v}\xi\in\mathrm{L}_2$ the existence and uniqueness of a solution $(\mathcal{Y}^{r,v},\mathcal{Z}^{r,v},\mathcal{U}^{r,v})$ to (4.3) follows from Theorem 4.2 since $F_{r,v}$ meets the assumptions of the theorem.

(ii) By Theorem 4.3 the solution depends continuously on the terminal condition and $\mathcal{D}\xi$ is measurable w.r.t. (r,v). We infer the measurable dependency $(r,v)\mapsto(\mathcal{Y}^{r,v},\mathcal{Z}^{r,v},\mathcal{U}^{r,v})$ as follows: Since by Theorem 4.3 the mapping

$$\mathrm{L}_2\to\mathcal{S}_2\times\mathrm{L}_2(W)\times\mathrm{L}_2(\tilde{N})\colon \xi\mapsto(Y,Z,U)$$

is continuous one can show the existence of a jointly measurable version of

$$(\mathcal{Y}^{r,v},\mathcal{Z}^{r,v},\mathcal{U}^{r,v}),\quad (r,v)\in[0,T]\times\mathbb{R}$$

by approximating $\mathcal{D}\xi$ with simple functions in $\mathrm{L}_2(\mathbb{P}\otimes\mathbb{m})$. Joint measurability (for example for $\mathcal{Z}$) in all arguments can be gained by identifying the spaces

$$\mathrm{L}_2(\lambda,\mathrm{L}_2(\mathbb{P}\otimes\mathbb{m}))\cong\mathrm{L}_2(\lambda\otimes\mathbb{P}\otimes\mathbb{m}).$$

The quadratic integrability with respect to (r,v) also follows from Theorem 4.3 since $\xi\in\mathbb{D}_{1,2}$.

Using an iteration scheme, starting with $(Y^0,Z^0,U^0)=(0,0,0)$, we get Y^{n+1} by taking the optional projection which implies that

$$Y_t^{n+1}=\mathbb{E}_t\left(\xi+\int_t^T f_g\left(s,Y_s^n,Z_s^n,U_s^n\right)ds\right). \tag{4.9}$$

The process $\underline{Z}^{n+1}$ given by

$$\underline{Z}_{s,x}^{n+1}:=\begin{cases} Z_s^{n+1}, & x=0,\\ U_s^{n+1}(x), & x\neq 0,\end{cases}$$

one gets by the martingale representation theorem w.r.t. M (see, for example, [4]):

$$\begin{aligned}\xi+\int_0^T f_g\left(s,Y_s^n,Z_s^n,U_s^n\right)ds &= \mathbb{E}\left(\xi+\int_0^T f_g\left(s,Y_s^n,Z_s^n,U_s^n\right)ds\right)\\ &\quad+\int_{]0,T]\times\mathbb{R}}\underline{Z}_{s,x}^{n+1}M(ds,dx).\end{aligned} \tag{4.10}$$

Step 1. It is well-known that (Y^n,Z^n,U^n) converges to the solution (Y,Z,U) in $\mathrm{L}_2(W)\times\mathrm{L}_2(W)\times\mathrm{L}_2(\tilde{N})$. Our aim in this step is to show that Y^n, Z^n and U^n are uniformly bounded in n as elements of $\mathrm{L}_2(\lambda;\mathbb{D}_{1,2})$ and $\mathrm{L}_2(\lambda\otimes\nu;\mathbb{D}_{1,2})$, respectively. This will follow from (4.14) below.

Given that $Y^n,Z^n\in\mathrm{L}_2(\lambda;\mathbb{D}_{1,2})$ and $U^n\in\mathrm{L}_2(\lambda\otimes\nu;\mathbb{D}_{1,2})$ one can infer that this also holds for $n+1$: Indeed, $(\mathbf{A}_f g)$ implies that $[g(U_s^n)]_\nu\in\mathbb{D}_{1,2}$ for a.e. s and

$$|\mathcal{D}_{r,v}[g(U_s^n)]_\nu|\le L_g\|g_1\|_{\mathrm{L}_2(\nu)}\|\mathcal{D}_{r,v}U_s^n\|_{\mathrm{L}_2(\nu)}. \tag{4.11}$$

From Lemma 4.6 we get that $f(X,s,Y_s^n,Z_s^n,[g(U_s^n)]_\nu)\in\mathbb{D}_{1,2}$. The above estimate and (4.7) as well as the Malliavin differentiation rules shown by Delong and Imkeller in [14, Lemma 3.1. and Lemma 3.2.] imply that Y^{n+1} as defined in (4.9) is in $\mathrm{L}_2(\lambda;\mathbb{D}_{1,2})$. Then we conclude from (4.10) and [14, Lemma 3.3.] that $Z^{n+1}\in\mathrm{L}_2(\lambda;\mathbb{D}_{1,2})$ and $U^{n+1}\in\mathrm{L}_2(\lambda\otimes\nu;\mathbb{D}_{1,2})$. Especially, we get for $t\in[0,T]$ that $\mathbb{P}$ -a.e.

$$\begin{aligned}\mathcal{D}_{r,v}Y_t^{n+1} &= \mathcal{D}_{r,v}\xi+\int_t^T\mathcal{D}_{r,v}f_g\left(X,s,Y_s^n,Z_s^n,U_s^n\right)ds\\ &\quad-\int_{]t,T]\times\mathbb{R}}\mathcal{D}_{r,v}\underline{Z}_{s,x}^{n+1}M(ds,dx)\text{ for }\mathbb{m}\text{ - }a.a.\,(r,v)\in[0,t]\times\mathbb{R},\end{aligned}$$

$$\begin{aligned}
\mathcal{D}_{r,v} Y_t^{n+1} &= 0 \quad \text{for } \mathbb{m} \text{ - } a.a.\, (r,v) \in (t,T] \times \mathbb{R}, \\
\mathcal{D}_{r,v} \underline{Z}_{t,x}^{n+1} &= 0 \quad \text{for } \mathbb{m} \otimes \mu \text{ - } a.a.\, (r,v,x) \in (t,T] \times \mathbb{R}^2.
\end{aligned} \tag{4.12}$$

Since by [4, Theorem 4.2.12] the process $\left(\int_{]0,t]\times\mathbb{R}} \mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1} M(ds,dx) \right)_{t\in[0,T]}$ admits a càdlàg version, we may take a càdlàg version of both sides.

By Itô's formula (see, for instance, [4]), we conclude that for $0 < r < t$ it holds

$$\begin{aligned}
e^{\beta T} (\mathcal{D}_{r,v}\xi)^2 &= e^{\beta t} (\mathcal{D}_{r,v} Y_t^{n+1})^2 \\
&\quad + \beta \int_t^T e^{\beta s} (\mathcal{D}_{r,v} Y_s^{n+1})^2 ds \\
&\quad - 2 \int_t^T e^{\beta s} \left[\mathcal{D}_{r,v} f_g \left(X, s, Y_s^n, Z_s^n, U_s^n \right) \right] \mathcal{D}_{r,v} Y_s^{n+1} ds \\
&\quad + \int_{]t,T]\times\mathbb{R}} e^{\beta s} [2 (\mathcal{D}_{r,v} Y_{s-}^{n+1}) \mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1} \\
&\qquad\qquad + \mathbb{1}_{\mathbb{R}_0}(x) (\mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1})^2] M(ds,dx) \\
&\quad + \int_{]t,T]\times\mathbb{R}} e^{\beta s} (\mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1})^2 ds \mu(dx) \quad \mathbb{P} \otimes \mathbb{m} \text{ - } a.e.
\end{aligned}$$

One easily checks that the integral w.r.t. M is a uniformly integrable martingale and hence has expectation zero. Therefore, using (4.12), we have for $0 < u < t \leq T$ that

$$\begin{aligned}
&\mathbb{E} e^{\beta t} (\mathcal{D}_{r,v} Y_t^{n+1})^2 + \mathbb{E} \int_{]r,T]\times\mathbb{R}} e^{\beta s} (\mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1})^2 ds \mu(dx) \\
\leq\ & e^{\beta T} \mathbb{E} (\mathcal{D}_{r,v}\xi)^2 + 2 \int_r^T e^{\beta s} \mathbb{E} \left| \left[\mathcal{D}_{r,v} f_g \left(X, s, Y_s^n, Z_s^n, U_s^n \right) \right] \mathcal{D}_{r,v} Y_s^{n+1} \right| ds \\
& - \beta \mathbb{E} \int_r^T e^{\beta s} (\mathcal{D}_{r,v} Y_s^{n+1})^2 ds.
\end{aligned} \tag{4.13}$$

By Young's inequality, (4.11) and Lemma 4.6 we get a constant C_f such that for any $c > 0$,

$$\begin{aligned}
& 2 \left| \left[\mathcal{D}_{r,v} f_g \left(X, s, Y_s^n, Z_s^n, U_s^n \right) \right] \mathcal{D}_{r,v} Y_s^{n+1} \right| \\
\leq\ & c \left| \mathcal{D}_{r,v} Y_s^{n+1} \right|^2 + \frac{C_f}{c} \Big(\left| \Gamma_{r,v} \right|^2 + \left| \mathcal{D}_{r,v} Y_s^n \right|^2 + \left| \mathcal{D}_{r,v} Z_s^n \right|^2 + \| \mathcal{D}_{r,v} U_s^n \|_{\mathrm{L}_2(\nu)}^2 \Big) \\
=\ & c \left| \mathcal{D}_{r,v} Y_s^{n+1} \right|^2 + \frac{C_f}{c} \Big(\left| \Gamma_{r,v} \right|^2 + \left| \mathcal{D}_{r,v} Y_s^n \right|^2 + \int_{\mathbb{R}} \left| \mathcal{D}_{r,v} \underline{Z}_{s,x}^n \right|^2 \mu(dx) \Big).
\end{aligned}$$

Choosing $\beta = c + 1$ and $c = 2C_f$ leads to

$$\begin{aligned}
& \mathbb{E} \int_r^T e^{\beta s} \left| \mathcal{D}_{r,v} Y_s^{n+1} \right|^2 ds + \mathbb{E} \int_{]r,T]\times\mathbb{R}} e^{\beta s} \left| \mathcal{D}_{r,v} \underline{Z}_{s,x}^{n+1} \right|^2 \mathbb{m}(ds,dx) \\
\leq\ & e^{\beta T} \mathbb{E} \left| \mathcal{D}_{r,v}\xi \right|^2 + \frac{1}{2} \int_r^T e^{\beta s} ds\, \mathbb{E} \left| \Gamma_{r,v} \right|^2 \\
& + \frac{1}{2} \left(\mathbb{E} \int_r^T e^{\beta s} \left| \mathcal{D}_{r,v} Y_s^n \right|^2 ds + \mathbb{E} \int_{]r,T]\times\mathbb{R}} e^{\beta s} \left| \mathcal{D}_{r,v} \underline{Z}_{s,x}^n \right|^2 \mathbb{m}(ds,dx) \right).
\end{aligned}$$

Finally, (4.12) and Lemma A.1 imply

$$\int_0^T e^{\beta s} \| \mathcal{D} Y_s^n \|_{L_2(\mathbb{m}\otimes\mathbb{P})}^2 ds + \int_{[0,T]\times\mathbb{R}} e^{\beta s} \| \mathcal{D} \underline{Z}_{s,x}^n \|_{L_2(\mathbb{m}\otimes\mathbb{P})}^2 \mathbb{m}(ds,dx)$$

$$\leq c_\beta \| |\mathcal{D}\xi| + \Gamma \|^2_{\mathrm{L}_2(\mathbb{P}\otimes \mathbb{m})} \text{ for all } n \in \mathbb{N}. \tag{4.14}$$

`Step 2.` We now show that

$$\left\| \mathcal{Y} - \mathcal{D}Y^{n+1} \right\|^2_{\mathrm{L}^2(\mathbb{P}\otimes\lambda\otimes\mathbb{m})} + \left\| \underline{\mathcal{Z}} - \mathcal{D}\underline{Z}^{n+1} \right\|^2_{\mathrm{L}^2(\mathbb{P}\otimes(\mathbb{m})^{\otimes 2})} \to 0, \quad n \to \infty. \tag{4.15}$$

In order to estimate the expressions from (4.15) one can repeat the previous computations for the difference $\mathcal{Y}_t^{r,v} - \mathcal{D}_{r,v}Y_t^{n+1}$ to obtain

$$\begin{aligned} & \mathbb{E}\int_r^T e^{\beta s}(\mathcal{Y}_s^{r,v} - \mathcal{D}_{r,v}Y_s^{n+1})^2 ds + \mathbb{E}\int_{]r,T]\times\mathbb{R}} e^{\beta s}(\underline{\mathcal{Z}}_{s,x}^{r,v} - \mathcal{D}_{r,v}\underline{Z}_{s,x}^{n+1})^2 ds\mu(dx) \\ \leq \quad & \frac{1}{c}\mathbb{E}\int_r^T e^{\beta s}\left| F_{r,v}(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v}) - \mathcal{D}_{r,v}f_g(s, Y_s^n, Z_s^n, U_s^n)\right|^2 ds. \end{aligned} \tag{4.16}$$

for any $c > 0$.

For the case $v = 0$, by using Lipschitz properties of f (which also imply the boundedness of the partial derivatives), we can find a constant C_f' such that

$$\begin{aligned} & \left| F_{r,0}(s, \mathcal{Y}_s^{r,0}, \mathcal{Z}_s^{r,0}, \mathcal{U}_s^{r,0}) - \mathcal{D}_{r,0}f_g(s, Y_s^n, Z_s^n, U_s^n) \right| \\ \leq \quad & C_f'(\left|\mathcal{Y}_s^{r,0} - \mathcal{D}_{r,0}Y_s^n\right| + \left|\mathcal{Z}_s^{r,0} - \mathcal{D}_{r,0}Z_s^n\right| + \left\|\mathcal{U}_s^{r,0} - \mathcal{D}_{r,0}U_s^n\right\|_{\mathrm{L}_2(\nu)}) \\ & +\kappa_n(r,s) \end{aligned} \tag{4.17}$$

where for some $C > 0$

$$\begin{aligned} \kappa_n(r,s) \quad = \quad & C\big(\left|(\mathcal{D}_{r,0}f_g)(s, Y_s, Z_s, U_s) - (\mathcal{D}_{r,0}f_g)(s, Y_s^n, Z_s^n, U_s^n)\right| \wedge \Gamma_{r,0} \\ & +\left|\mathcal{Y}_s^{r,0}\right| \left|\partial_y f_g(s, Y_s, Z_s, U_s) - \partial_y f_g(s, Y_s^n, Z_s^n, U_s^n)\right| \\ & +\left|\mathcal{Z}_s^{r,0}\right| \left|\partial_z f_g(s, Y_s, Z_s, U_s) - \partial_z f_g(s, Y_s^n, Z_s^n, U_s^n)\right| \\ & +\left\|\mathcal{U}_s^{r,0}\right\|_{\mathrm{L}_2(\nu)}(\left|\partial_u f_g(s, Y_s, Z_s, U_s) - \partial_u f_g(s, Y_s^n, Z_s^n, U_s^n)\right| \\ & \quad + \| |g'(U_s) - g'(U_s^n)| g_1 \|_{\mathrm{L}_2(\nu)})\big). \end{aligned} \tag{4.18}$$

Since the sequence (Y^n, Z^n, U^n) converges in $\mathrm{L}_2(W) \times \mathrm{L}_2(W) \times \mathrm{L}_2(\tilde{N})$, condition $(\mathbf{A}_f e)$ holds, and $\partial_y f, \partial_z f, \partial_u f$ as well as g' are bounded and continuous it follows from Vitali's convergence theorem that

$$\delta_n \quad := \quad \mathbb{E}\int_r^T e^{\beta s}\kappa_n(r,s)^2 dr ds \to 0 \text{ for } n \to \infty. \tag{4.19}$$

Now we continue with the case $v \neq 0$. We first realize that for a given $\varepsilon > 0$ we may choose $\alpha > 0$ small enough such that

$$\mathbb{E}\int_r^T \int_{\{|v|<\alpha\}} e^{\beta s}\left| F_{r,v}(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v}) - \mathcal{D}_{r,v}f(s, Y_s^n, Z_s^n, U_s^n)\right|^2 \nu(dv)ds < \varepsilon.$$

This is because from (4.6), (4.7) and (4.2) one gets by a straightforward calculation

$$\begin{aligned} |F_{r,v}(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v})|_{\mathrm{L}_2(\nu)} \quad & \leq \quad \Gamma_{r,v} + L_f(|\mathcal{Y}_s^{r,v}| + |\mathcal{Z}_s^{r,v}| + L_g[|\mathcal{U}_s^{r,v}|]_\nu) \\ & \leq \quad \Gamma_{r,v} + L_{f,g}(|\mathcal{Y}_s^{r,v}| + |\mathcal{Z}_s^{r,v}| + \|\mathcal{U}_s^{r,v}\|_{\mathrm{L}_2(\nu)}) \end{aligned}$$

with $L_{f,g} = L_f(1 + L_g\|g_1\|_{\mathrm{L}_2(\nu)})$ where L_g is the Lipschitz constant of g, and

$$|\mathcal{D}_{r,v}f_g(s, Y_s^n, Z_s^n, U_s^n)| \leq \Gamma_{r,v} + L_{f,g}(|\mathcal{D}_{r,v}Y_s^n| + |\mathcal{D}_{r,v}Z_s^n| + \|\mathcal{D}_{r,v}U_s^n\|_{\mathrm{L}_2(\nu)}).$$

On the set $\{|v| \geq \alpha\}$ we use the Lipschitz properties ($\mathbf{A}_f c$) and ($\mathbf{A}_f g$) to get the estimate

$$\begin{aligned}
&|F_{r,v}(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v}) - \mathcal{D}_{r,v} f_g(s, Y_s^n, Z_s^n, U_s^n)| \\
\leq \quad & \big| f_g\left((X + v\mathbb{1}_{[r,T]}), s, Y_s + \mathcal{Y}_s^{r,v}, Z_s + \mathcal{Z}_s^{r,v}, U_s + \mathcal{U}_s^{r,v}\right) \\
& \quad - f_g\left((X + v\mathbb{1}_{[r,T]}), s, Y_s^n + \mathcal{D}_{r,v} Y_s^n, Z_s^n + \mathcal{D}_{r,v} Z_s^n, U_s^n + \mathcal{D}_{r,v} U_s^n\right) \big| \\
& + \big| f_g\left(X, s, Y_s, Z_s, U_s\right) - f_g\left(X, s, Y_s^n, Z_s^n, U_s^n\right) \big| \\
\leq \quad & L_{f,g} \big[|\mathcal{Y}_s^{r,v} - \mathcal{D}_{r,v} Y_s^n| + |\mathcal{Z}_s^{r,v} - \mathcal{D}_{r,v} Z_s^n| + \|\mathcal{U}_s^{r,v} - \mathcal{D}_{r,v} U_s^n\|_{\mathrm{L}_2(\nu)} \\
& \qquad + 2(|Y_s - Y_s^n| + |Z_s - Z_s^n| + \|U_s - U_s^n\|_{\mathrm{L}_2(\nu)}) \big].
\end{aligned}$$

This gives for any $n \in \mathbb{N}$

$$\begin{aligned}
& \mathbb{E} \int_r^T \int_{[0,T]\times\mathbb{R}} e^{\beta s} |F_{r,v}(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v}) - \mathcal{D}_{r,v} f_g(s, Y_s^n, Z_s^n, U_s^n)|^2 \mathbb{m}(dr, dv) ds \\
\leq \quad & c(L_{f,g}) \mathbb{E} \int_r^T e^{\beta s} \big(\|\mathcal{Y}_s - \mathcal{D} Y_s^n\|_{L_2(\mathbb{m})}^2 + \|\underline{\mathcal{Z}}_{s,.} - \mathcal{D} \underline{Z}_{s,.}^n\|_{L_2(\mathbb{m}\otimes\mu)}^2 \big) ds \\
& + 2c(L_{f,g}) \nu(\{|v| \geq \alpha\}) \mathbb{E} \int_r^T e^{\beta s} \big(|Y_s - Y_s^n|^2 + \|\underline{Z}_{s,.} - \underline{Z}_{s,.}^n\|_{L_2(\mathbb{m}\otimes\mu)}^2 \big) ds \\
& + \delta_n + \varepsilon.
\end{aligned}$$

Choosing c in (4.16) in an appropriate way leads to

$$\begin{aligned}
& \left\| \mathcal{Y} - \mathcal{D} Y^{n+1} \right\|_{\mathrm{L}^2(\mathbb{P}\otimes\lambda\otimes\mathbb{m})}^2 + \left\| \underline{\mathcal{Z}} - \mathcal{D} \underline{Z}^{n+1} \right\|_{\mathrm{L}^2(\mathbb{P}\otimes(\mathbb{m})^{\otimes 2})}^2 \\
\leq \quad & \varepsilon + C_n + \frac{1}{2} \left(\|\mathcal{Y} - \mathcal{D} Y^n\|_{\mathrm{L}^2(\mathbb{P}\otimes\lambda\otimes\mathbb{m})}^2 + \|\underline{\mathcal{Z}} - \mathcal{D} \underline{Z}^n\|_{\mathrm{L}^2(\mathbb{P}\otimes(\mathbb{m})^{\otimes 2})}^2 \right)
\end{aligned}$$

with $C_n = C_n(\alpha)$ tending to zero if $n \to \infty$ for any fixed $\alpha > 0$. We now apply Lemma A.1 and end up with

$$\limsup_{n\to\infty} \left(\|\mathcal{Y} - \mathcal{D} Y^n\|_{\mathrm{L}^2(\mathbb{P}\otimes\lambda\otimes\mathbb{m})}^2 + \|\underline{\mathcal{Z}} - \mathcal{D} \underline{Z}^n\|_{\mathrm{L}^2(\mathbb{P}\otimes(\mathbb{m})^{\otimes 2})}^2 \right) \leq 2\varepsilon.$$

This implies (4.4). Hence we can take the Malliavin derivative of (4.1) and get (4.5) as well as

$$\begin{aligned}
0 \quad = \quad & \mathcal{D}_{r,v} \xi + \int_r^T F_{r,v}\left(s, \mathcal{D}_{r,v} Y_s, \mathcal{D}_{r,v} Z_s, \mathcal{D}_{r,v} U_s\right) ds \\
& - \underline{Z}_{r,v} - \int_{]r,T]\times\mathbb{R}} \mathcal{D}_{r,v} \underline{Z}_{s,x} M(ds, dx), \quad 0 \leq t < r \leq T. \qquad (4.20)
\end{aligned}$$

By the same reasoning as for $\mathcal{D}_{r,v} Y^n$ we may conclude that the RHS of (4.5) has a càdlàg version which we take for $\mathcal{D}_{r,v} Y$.

(iii) This assertion we get comparing (4.3) and (4.5) because of the uniqueness of $(\mathcal{Y}, \mathcal{Z}, \mathcal{U})$.

(iv) We first discuss the measurability of $\lim_{t\searrow r} \mathcal{D}_{r,v} Y_t$ w.r.t. (r, v, ω) which is needed to take the predictable projection. From (4.5) one concludes that for any fixed (r, v) there exists a càdlàg version of $t \mapsto \mathcal{D}_{r,v} Y_t$. By [33, Lemma 1] there exists a jointly in (r, v, t, ω) measurable random map with the following property: for each (r, v) this map has càdlàg paths and is indistinguishable from the above càdlàg version. We assume now that $\mathcal{D}_{r,v} Y_t$ is this measurable random map with càdlàg paths w.r.t. t. Then the pathwise limit $\lim_{t\searrow r} \mathcal{D}_{r,v} Y_t$ is measurable in (r, v, ω) and the assertion follows by comparing the RHS of (4.5) with (4.20). □

4.2 Example: a BSDE related to utility maximization

In [7] and [26] a class of BSDEs is considered which appears in exponential utility maximization. For these BSDEs an additional summand arises in the generator which is only locally Lipschitz and is (in the simplest case) of the form: $[g^\alpha(U_s)]_\nu$ (see (4.2)) with

$$g^\alpha(x) := \frac{e^{\alpha x} - \alpha x - 1}{\alpha} \quad \text{for some } \alpha > 0$$

and $g_1(x) := 1$ for $x \in \mathbb{R}_0$. Consider for $0 \le t \le T$ the following BSDE

$$\begin{aligned} Y_t &= \xi + \int_t^T \left(f_g\left(X, s, Y_s, Z_s, U_s\right) + [g^\alpha(U_s)]_\nu\right) ds - \int_t^T Z_s dW_s \\ &\qquad - \int_{]t,T]\times\mathbb{R}} U_s(x)\tilde{N}(ds, dx), \end{aligned} \tag{4.21}$$

where f_g is defined like in (4.1). Then we have the following assertion:

Corollary 4.7. *Let $\xi \in \mathbb{D}_{1,2}$ and assume that ξ is a.s. bounded and ν is a bounded measure. If $(\mathbf{A}_f)$ is satisfied for f_g and if there exists constants $K_1, K_2 > 0$ such that for all $y, z, u \in \mathbb{R}$*

$$f(X, t, y, z, u) \le K_1 + K_2|y|$$

for $\mathbb{P} \otimes \lambda$ -a.a. $(t, \omega) \in [0, T] \times \Omega$, then the following assertions hold for (4.21).

(i) For $\mathbb{m}$- a.e. $(r, v) \in [0, T] \times \mathbb{R}$ there exists a unique solution $(\mathcal{Y}^{r,v}, \mathcal{Z}^{r,v}, \mathcal{U}^{r,v}) \in \mathcal{S}_2 \times \mathrm{L}_2(W) \times \mathrm{L}_2(\tilde{N})$ to the BSDE

$$\begin{aligned} \mathcal{Y}_t^{r,v} &= \mathcal{D}_{r,v}\xi + \int_t^T \left(F_{r,v}\left(s, \mathcal{Y}_s^{r,v}, \mathcal{Z}_s^{r,v}, \mathcal{U}_s^{r,v}\right) + G_{r,v}(s, \mathcal{U}_s^{r,v})\right) ds \\ &\qquad - \int_{]t,T]\times\mathbb{R}} \underline{\mathcal{Z}}_{s,x}^{r,v} M(ds, dx), \quad 0 \le r \le t \le T \\ \mathcal{Y}_s^{r,v} &= \mathcal{Z}_s^{r,v} = \mathcal{U}_s^{r,v} = 0, \quad 0 \le s < r \le T, \end{aligned} \tag{4.22}$$

with $F_{r,v}$ and $\underline{\mathcal{Z}}_{s,x}^{r,v}$ given in Theorem 4.4 and

$$G_{r,v}(s, u) := \begin{cases} \left[(e^{\alpha U_s} - 1)u\right]_\nu, & v = 0, \\ \left[e^{\alpha U_s} g^\alpha(u) + \frac{e^{\alpha U_s} - 1}{\alpha} u\right]_\nu, & v \ne 0. \end{cases}$$

(ii) For the solution (Y, Z, U) of (4.21) it holds

$$Y, Z \in \mathrm{L}_2([0, T]; \mathbb{D}_{1,2}), \quad U \in \mathrm{L}_2([0, T] \times \mathbb{R}_0; \mathbb{D}_{1,2}),$$

and $\mathcal{D}_{r,y}Y$ admits a càdlàg version for $\mathbb{m}$- a.e. $(r, y) \in [0, T] \times \mathbb{R}$.

(iii) $(\mathcal{D}Y, \mathcal{D}Z, \mathcal{D}U)$ is a version of $(\mathcal{Y}, \mathcal{Z}, \mathcal{U})$, i.e. for $\mathbb{m}$- a.e. (r, v) it solves (4.22).

(iv) Setting $D_{r,v}Y_r(\omega) := \lim_{t\searrow r} \mathcal{D}_{r,v}Y_t(\omega)$ for all (r, v, ω) for which $\mathcal{D}_{r,v}Y$ is càdlàg and $\mathcal{D}_{r,v}Y_r(\omega) := 0$ otherwise, we have

$$^p\left((D_{r,0}Y_r)_{r\in[0,T]}\right) \text{ is a version of } (Z_r)_{r\in[0,T]},$$

$$^p\left((D_{r,v}Y_r)_{r\in[0,T],v\in\mathbb{R}_0}\right) \text{ is a version of } (U_r(v))_{r\in[0,T],v\in\mathbb{R}_0}.$$

Proof. Since ξ is a.s. bounded, the Lévy measure ν is finite and the generator satisfies the conditions of [7, Theorem 3.5.] it follows that $\|Y\|_{\mathcal{S}_\infty} < \infty$ and

$$|U_s(x)| \le 2\|Y\|_{\mathcal{S}_\infty} \quad \text{for } \mathbb{P}\otimes\lambda\otimes\nu\text{- a.e. } (\omega, s, x). \tag{4.23}$$

From the fact that g^α is locally Lipschitz and U is a.e. bounded it follows that $(\mathbf{A}_f)$ (especially the Lipschitz condition) can be seen as satisfied also for $[g^\alpha(U_s)]_\nu$:

We find a $\mathcal{C}^1$ function $\widehat{g^\alpha}$ such that $g^\alpha = \widehat{g^\alpha}$ on $[-2\|Y\|_{\mathcal{S}_\infty}, 2\|Y\|_{\mathcal{S}_\infty}]$ and

$$\operatorname{supp}(\widehat{g^\alpha}) \subseteq [-3\|Y\|_{\mathcal{S}_\infty}, 3\|Y\|_{\mathcal{S}_\infty}] \,.$$

Since by (4.23), $g^\alpha(U_s(x)) = \widehat{g^\alpha}(U_s(x))$, $\mathbb{P}\otimes\lambda\otimes\nu$- a.e., it follows that for all $t\in[0,T]$

$$\int_t^T \!\left(f_g\left(X,s,Y_s,Z_s,U_s\right) + [g^\alpha(U_s)]_\nu\right) ds = \int_t^T \!\left(f_g\left(X,s,Y_s,Z_s,U_s\right) + [\widehat{g^\alpha}(U_s)]_\nu\right) ds,$$

$\mathbb{P}$-a.s. So the solution of (4.21) also satisfies the BSDE with g^α replaced by $\widehat{g^\alpha}$ which satisfies the assumptions of Theorem 4.4. □

A Appendix

Proof of Lemma 3.3

`Step 1.` We have the a.s. representation of the Lévy process X as

$$X_t = \gamma t + \sigma W_t + J_t.$$

We denote $B_t := \gamma t + \sigma W_t$. Because of

$$\mathbb{P}\left(\{X\in\Lambda\}\right) = \int_{D[0,T]} \mathbb{P}_J\left(\Lambda - h\right)\mathbb{P}_B(dh),$$

we may restrict ourselves to 'pure jump processes' (i.e. $X = J$).

`Step 2.` Assume that X is a compound Poisson process. Then $\nu(\mathbb{R}_0) < \infty$. We define $\hat{\mathbb{P}} := \mathbb{P}\otimes\frac{\lambda\otimes\nu}{T\nu(\mathbb{R}_0)}$ on $\big(\Omega\times[0,T]\times\mathbb{R}_0, \mathcal{F}\otimes\mathcal{B}([0,T]\times\mathbb{R}_0)\big)$ and

$$\hat{X}_t(\omega,r,v) := X_t(\omega) + \beta_t(r,v)$$

where $\beta_t(r,v) := v\mathbb{1}_{[r,T]}(t)$. By the law of total probability we get

$$\hat{\mathbb{P}}\left(\hat{X}\in\Lambda\right) = \sum_{k=0}^{\infty}\hat{\mathbb{P}}\left(\hat{X}\in\Lambda\middle| N(]0,T]\times\mathbb{R}_0) = k\right)\hat{\mathbb{P}}\left(N(]0,T]\times\mathbb{R}_0) = k\right). \tag{A.1}$$

The conditional probabilities

$$\hat{\mathbb{P}}\left(\hat{X}\in\Lambda\middle| N(]0,T]\times\mathbb{R}_0) = k\right), \quad k\in\mathbb{N},$$

are the distributions of an independent sum of β and the compound Poisson process X, conditioned on the event that the process X jumps k times in $]0,T]$. The probability law of this conditioned compound Poisson process is the same as the law of a piecewise constant process which has exactly k independent, uniformly distributed jumps in $[0,T]$ whose jump sizes are independently identically distributed according to $\frac{\nu}{\nu(\mathbb{R}_0)}$ and independent from the jump times. Therefore it holds that

$$\hat{\mathbb{P}}\left(\hat{X}\in\Lambda\middle| N(]0,T]\times\mathbb{R}_0) = k\right)$$

$$= \frac{(\lambda\otimes\nu)^{\otimes(k+1)}}{T^{k+1}\nu\left(\mathbb{R}_0\right)^{k+1}}\left(\left\{((t_1,x_1),\dots,(t_k,x_k),(r,v)) : \sum_{l=1}^{k} x_l\mathbb{1}_{[t_l,T]} + v\mathbb{1}_{[r,T]} \in \Lambda\right\}\right)$$
$$= \mathbb{P}\left(X\in\Lambda | N(]0,T]\times\mathbb{R}_0) = k+1\right) = 0,$$

where we used the argument concerning the distribution of a conditioned Poisson process again to come to the last line. Hence, all summands of (A.1) are zero, which shows the assertion for the special case of this step.

`Step 3.` To extend the second step to the case of a general pure-jump Lévy process X we split up $\mathbb{R}_0$ into sets S_p, $p\geq 1$ such that $0<\nu(S_p)<\infty$. Without loss of generality set $S_1 := \{x\in\mathbb{R} : |x|>1\}$. We may assume that the sequence $(S_p)_{p\geq 1}$ is infinite, else we would be in the compound Poisson case again. From the proof of the Lévy-Itô decomposition it follows that

$$X_t = \lim_{n\to\infty}\sum_{p=2}^{n}\left(X_t^{(p)} - t\int_{S_p} x\nu(dx)\right) + X_t^{(1)},$$

where the convergence is $\mathbb{P}$-a.s., uniformly in $t\in[0,T]$ and the $(X^{(p)})$ given by

$$X_t^{(p)} = \int_{[0,t]\times S_p} xN(ds,dx),$$

are independent compound Poisson processes which have jumps distributed by $\frac{\nu|_{S_p}}{\nu(S_p)}$. Since for $k,p\in\mathbb{N}$ with $p\geq 1$,

$$0 < (\mathbb{P}\otimes\lambda\otimes\nu)\big(\{N(]0,T]\times S_p)=k\}\times[0,T]\times S_p\big) < \infty,$$

we can proceed in a similar way as in (A.1) for σ-finite measures: Let

$$\overline{X}_t^{(p)} := X_t - X_t^{(p)}, \quad 0\leq t\leq T,$$

an notice that $\overline{X}^{(p)}$ and $X^{(p)}$ are independent. Then

$$\begin{aligned}
&(\mathbb{P}\otimes\lambda\otimes\nu)\left(\hat{X}\in\Lambda\right)\\
&= \sum_{\substack{p=1,\\k=0}}^{\infty}(\mathbb{P}\otimes\lambda\otimes\nu)\left(X+\beta\in\Lambda\middle|\big\{N(]0,T]\times S_p)=k\big\}\times[0,T]\times S_p\right)\\
&\qquad\times(\mathbb{P}\otimes\lambda\otimes\nu)\left(\big\{N(]0,T]\times S_p)=k\big\}\times[0,T]\times S_p\right). \qquad \text{(A.2)}
\end{aligned}$$

From Steps 1 and 2 we conclude that the summands on the RHS of (A.2) are zero again by

$$\begin{aligned}
&(\mathbb{P}\otimes\lambda\otimes\nu)\left(\overline{X}^{(p)}+X^{(p)}+\beta\in\Lambda\middle|\big\{N(]0,T]\times S_p)=k\big\}\times[0,T]\times S_p\right)\\
&=(\mathbb{P}\otimes\lambda\otimes\nu)\left(\overline{X}^{(p)}+X^{(p)}\in\Lambda\middle|\big\{N(]0,T]\times S_p)=k+1\big\}\times[0,T]\times S_p\right)\\
&=0,
\end{aligned}$$

which proves Step 3. □

Lemma A.1. *Let $(g_n)_{n\geq 0}$ be a sequence of nonnegative numbers satisfying $g_0=0$ and*

$$g_{n+1}\leq\varepsilon + C_n + \frac{1}{2}g_n,$$

where $\varepsilon>0$ and $\lim_{n\to\infty}C_n=0$. Then it holds that

$$\limsup_{n\to\infty} g_n \leq 2\varepsilon.$$

Especially, if $C_n=0$ for all $n\in\mathbb{N}$, then $g_n\leq 2\varepsilon$ for all $n\in\mathbb{N}$.

References

[1] E. Alòs, J. A. León and J. Vives, *An anticipating Itô formula for Lévy processes*, ALEA Lat. Am. J. Probab. Math. Stat. 4, 285–305, 2008. MR-2456970

[2] S. Ankirchner, P. Imkeller, *Quadratic hedging of weather and catastrophe risk by using short term climate predictions*, HU Berlin, Preprint, 2008.

[3] S. Ankirchner, P. Imkeller and G. Dos Reis, *Classical and Variational Differentiability of BSDEs with Quadratic Growth*, Electron. J. Probab. 12, no. 53, 1418–1453, 2007. MR-2354164

[4] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, 2004. MR-2072890

[5] H. Bauer, *Measure and Integration Theory*, de Gruyter, 2001. MR-1897176

[6] G. Barles, R. Buckdahn and É. Pardoux, *Backward stochastic differential equations and integral-partial differential equations*, Stochastics Stochastics Rep. 60, no. 1–2, 57–83, 1997. MR-1436432

[7] D. Becherer, *Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging*, Ann. Appl. Probab. 16, no. 4, 1733–2275, 2006. MR-2288712

[8] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons, New York, 1968. MR-0233396

[9] V. Bogachev, *Gaussian Measures*, AMS, 1998. MR-1642391

[10] B. Bouchard and R. Elie, *Discrete time approximation of decoupled Forward-Backward SDE with jumps*, Stochastic Process. Appl. 118, 53–75, 2008. MR-2376252

[11] B. Bouchard, R. Elie and N. Touzi, *Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs*, In Advanced Financial Modelling. Radon Ser. Comput. Appl. Math. 8 91–124. de Gruyter, Berlin, 2009. MR-2648459

[12] B. Bouchard and N. Touzi, *Discrete time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl. 111, 175–206, 2004. MR-2056536

[13] Ł. Delong, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, Springer, 2013. MR-3089193

[14] Ł. Delong, P. Imkeller, *On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures*, Stochastic Process. Appl. 120, 1748–1775, 2010. MR-2673973

[15] O. Delzeith, *On Skorohod spaces as universal sample path spaces*, arXiv:math/0412092v1, 2004.

[16] J. Diestel, and J. J. Uhl Jr., *Vector Measures*, AMS 1977. MR-0453964

[17] G. Di Nunno, B. Øksendal, F. Proske, *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, 2009. MR-2460554

[18] N. El Karoui, S. Peng, and M.C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance 7, 1–71, 1997. MR-1434407

[19] K. Itô, *Spectral type of the shift transformation of differential process with stationary increments*, Trans. Amer. Math. Soc. 81, 253–263, 1956. MR-0077017

[20] S. Janson, *Gaussian Hilbert Spaces*, Cambridge, 1997. MR-1474726

[21] H.-H. Kuo, *Gaussian measures in Banach spaces*, Bulletin of the American Mathematical Society 82, no. 5, 695–700, 1976. MR-1566896

[22] Y. Lee and H. Shih, *Analysis of geralized Lévy functionals*, J. Funct. Analysis, 211, 1–70, 2004. MR-2054618

[23] A. Løkka, *Martingale representation of functionals of Lévy processes*, Stoch. Anal. Appl., 22, 867–892, 2004. MR-2062949

[24] T. Mastrolia, D. Possamaï and A. Réveillac, *On the Malliavin differentiability of BSDEs,* arXiv:1404.1026, 2014.

[25] P.A. Meyer, *Une remarque sur le calcul stochastique dépendant d'un paramètre*, Séminaire de probabilités (Strasbourg), tome 13, 199–203, 1979. MR-0544791

[26] M.-A. Morlais, *A new existence result for quadratic BSDEs with jumps with application to the utility maximization problem*, Stochastic Process. Appl. 120, 1966–1995, 2010. MR-2673984

[27] D. Nualart, *The Malliavin calculus and related topics. Second edition*, Probability and its Applications (New York). Springer-Verlag, Berlin, 2006. MR-2200233

[28] É. Pardoux, S. Peng, *Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations*, Stochastic partial differential equations and their applications (Charlotte, NC, 1991), 200–217, Lecture Notes in Control and Inform. Sci., 176, Springer, Berlin, 1992. MR-1176785

[29] É. Pardoux, S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. 14, no. 1, 55–61, 1990. MR-1037747

[30] E. Petrou, *Malliavin calculus in Lévy spaces and applications to finance*, Electron. J. Probab., 13, no. 27, 852–879, 2008. MR-2399298

[31] J. Picard, *On the existence of smooth densities for jump processes*, P.T.R.F. 105, 481–511, 1996. MR-1402654

[32] A. Steinicke, *Functionals of a Lévy process on canonical and generic probability spaces*, J. Theoret. Probab., DOI: 10.1007/s10959-014-0583-7, 2014.

[33] C. Stricker and M. Yor, *Calcul stochastique dépendant d'un paramètre*, Z. Wahrsch. Verw. Gebiete, 45, no. 2, 109–133, 1978. MR-0510530

[34] J. Solé, F. Utzet, J Vives, *Chaos expansions and Malliavin calculus for Lévy processes*, Stoch. Anal. Appl., 595–612, Abel Symp., 2, Springer, Berlin, 2007. MR-2397807

[35] J. Solé, F. Utzet, J. Vives, *Canonical Lévy process and Malliavin calculus*, Stochastic Process. Appl. 117, pp. 165–187, 2007. MR-2290191

[36] H. Sugita, *On a characterization of the Sobolev spaces over an abstract Wiener space*, J. Math. Kyoto Univ. 25, no. 4, 717–725, 1985. MR-0810975

[37] S. Tang and X. Li, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM J. Control and Optim. 32, 1447–1475, 1994. MR-1288257

Acknowledgments. We would like to thank S. Geiss for his helpful comments concerning the measurability needed in the proof of Theorem 4.4 (iv).