ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

On the Malliavin differentiability of BSDEs

Thibaut Mastrolia[a], Dylan Possamaï[a] and Anthony Réveillac[b]

[a]*Université Paris-Dauphine, CEREMADE UMR CNRS 7534, Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France.*
E-mail: mastrolia@ceremade.dauphine.fr; possamai@ceremade.dauphine.fr
[b]*IMT UMR CNRS 5219, Université de Toulouse, INSA de Toulouse, 135 avenue de Rangueil, 31077 Toulouse Cedex 4, France.*
E-mail: anthony.reveillac@insa-toulouse.fr

Received 4 November 2014; revised 24 August 2015; accepted 14 October 2015

Abstract. In this paper we provide new conditions for the Malliavin differentiability of solutions of Lipschitz or quadratic BSDEs. Our results rely on the interpretation of the Malliavin derivative as a Gâteaux derivative in the directions of the Cameron–Martin space. Incidentally, we provide a new formulation for the characterization of the Malliavin–Sobolev type spaces $\mathbb{D}^{1,p}$.

Résumé. Dans cet article, nous donnons de nouvelles conditions nous assurant que les solutions d'EDSR à générateurs lipschitziens ou à croissance quadratique sont différentiables au sens de Malliavin, en utilisant l'interprétation de la dérivée de Malliavin comme dérivée de Gâteaux directionnelle par rapport à l'espace de Cameron–Martin. Ce résultat est en outre basé sur une nouvelle caractérisation des espaces de Malliavin–Sobolev $\mathbb{D}^{1,p}$ que nous fournissons.

1. Introduction

Backward Stochastic Differential Equations (BSDEs) have been studied extensively in the last two decades as they naturally arise in the context of stochastic control problems (for instance in Finance see [9]), and as they provide a probabilistic representation for solution to semi-linear parabolic PDEs, via a non-linear Feynman–Kac formula (see [20]). Before going further let us recall that this class of equations has been introduced in [4,19,20] and that a BSDE can be formulated as:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\,ds - \int_t^T Z_s\,dW_s, \quad t \in [0, T], \tag{1.1}$$

where T is a fixed positive number, $W := (W_t)_{t \in [0,T]}$ is a one-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$ with natural filtration $(\mathcal{F}_t)_{t \in [0,T]}$. The data of the equation are the $\mathcal{F}_T$-measurable r.v. ξ, called the terminal condition, and the mapping $f : [0, T] \times \Omega \times \mathbb{R}^2 \to \mathbb{R}$ which is a progressively measurable process and where according to the notations used in the literature we write $f(t, y, z)$ for $f(t, \omega, y, z)$. A solution to the BSDE (1.1) is then a pair of predictable processes (Y, Z), with appropriate integrability properties, such that Relation (1.1) holds $\mathbb{P}$-a.s.

When dealing with applications, one needs to obtain regularity properties on the solution (Y, Z), such as the Malliavin differentiability of the random variables Y_t, Z_t at a given time t in $[0, T]$. Note that for the Z component

this question needs to be clarified a little bit because of the definition of Z, cf. Theorem 5.1 for a precise statement. More precisely, one needs to answer the following question:

Which conditions on the data ξ and f in (1.1) ensure that Y_t, Z_t are Malliavin differentiable?

This question was first addressed in the paper [20] in a Markovian setting, that is when $\xi := g(X_T)$ and $f(t, \omega, y, z) := h(t, X_t(\omega), y, z)$ where $g : \mathbb{R} \to \mathbb{R}$ and $h : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ are regular enough deterministic functions and $X := (X_t)_{t \in [0, T]}$ is the unique solution to a SDE of the form:

$$X_t = X_0 + \int_0^t \sigma(s, X_s)\, dW_s + \int_0^t b(s, X_s)\, ds, \quad t \in [0, T],$$

with regular enough coefficients $\sigma, b : [0, T] \times \mathbb{R} \to \mathbb{R}$. In that framework, Pardoux and Peng proved in [20, Proposition 2.2] that, under (essentially) the following conditions:

(PP1) g is continuously differentiable with bounded derivative.
(PP2) h is continuously differentiable in (x, y, z) with bounded derivatives uniformly in time,

Y_t is Malliavin differentiable at any time t, with a similar statement for Z, and the Malliavin derivatives of Y and Z provide a solution to an explicit linear BSDE. To be more precise, in [20] the authors make one assumption for the whole paper which is stronger than (PP1)–(PP2) above. However a careful reading of the proof of [20, Proposition 2.2] enables one to conclude that Conditions (PP1)–(PP2) are sufficient to obtain the Malliavin differentiability of the solution. Assumptions (PP1)–(PP2) look pretty intuitive since they basically require the Malliavin differentiability of the terminal condition ξ and of the generator f once the component (y, z) are frozen, i.e., of the process $(t, \omega) \mapsto f(t, \omega, y, z)$ for given (y, z). Hence, it is natural to expect that the latter conditions can be easily generalized to the non-Markovian framework. Unfortunately, the first result in that direction, given by El Karoui, Peng and Quenez in [9], requires more stringent conditions than the aforementioned intuitive ones. More explicitly, the main result in [9] concerning the Malliavin differentiability of the solution to the BSDE (1.1) (essentially) involves the following conditions (see [9, Proposition 5.3] for a precise statement):

(EPQ1) ξ is Malliavin differentiable[1] and $\mathbb{E}[|\xi|^4] < +\infty$.
(EPQ2) At any time $t \in [0, T]$, the r.v. $\omega \mapsto f(t, \omega, Y_t, Z_t)$ is Malliavin differentiable[2] with Malliavin derivative denoted by $D.f(t, Y_t, Z_t)$ such that there exists a predictable process $K^\theta := (K_t^\theta)_{t \in [0, T]}$ with $\int_0^T \mathbb{E}[(\int_0^T |K_s^\theta|^2\, ds)^2]\, d\theta < +\infty$, and such that for any $(y_1, y_2, z_1, z_2) \in \mathbb{R}^4$ it holds for a.e. $\theta \in [0, T]$ that:

$$\left| D_\theta f(t, \omega, y_1, z_1) - D_\theta f(t, \omega, y_2, z_2) \right| \leq K_t^\theta(\omega)\big(|y_1 - y_2| + |z_1 - z_2|\big).$$

Roughly speaking, this means that ξ and $\omega \mapsto f(t, \omega, y, z)$ have to be Malliavin differentiable, but in order to prove that Y and Z are Malliavin differentiable, one needs to enforce an extra regularity conditions on each of the data: that is ξ has a finite moment of order 4, and the Malliavin derivative of the driver f is Lipschitz continuous in (y, z) with a sufficiently integrable stochastic Lipschitz constant K. Note that a careful reading of the proof allows one to conclude that the moment conditions on ξ and Df can actually be relaxed to hold only in $L^{2+\varepsilon}$ for some $\varepsilon > 0$. Besides, as noted in [9, Remark at the bottom of p. 59], if K is bounded then the proof can be modified so that the extra integrability condition on ξ (i.e. $\mathbb{E}[|\xi|^4] < +\infty$) can be dropped. However, even in this case, one can check that in the Markovian framework, Conditions (EPQ1)–(EPQ2) are strictly stronger than Conditions (PP1)–(PP2).

Since these two seminal papers, the most notable extension was concerned with the study of the Malliavin differentiability of (Y, Z) in a quadratic setting, that is to say when the generator f has quadratic growth in the z variable, a problem addressed in [2,8,11,12]. Notice nonetheless that the proofs in these references are strongly influenced by the ones in the Lipschitz setting of [9,20], as they all start by approximating the quadratic generators by Lipschitz ones, to which they apply the results of [9,20]. The applications of the Malliavin differentiability of BSDEs also received a lot of attention in the literature. Hence, it was used in the context of numerical schemes for BSDEs in, among others,

---

[1]That is, ξ is in $\mathbb{D}^{1,2}$.
[2]In fact as an adapted process it belongs to $\mathbb{D}^{1,2}$, we refer to the space $\mathbb{L}_{1,2}^a$ whose precise definition is recalled in [9, p. 58].

[6,10], or to study the existence and regularity of densities for the marginal laws of (Y, Z) in [1,3,17]. However, in all the above references, the authors always refer to either [9,20] in a Lipschitz context or [2] in a quadratic context, when stating differentiability results in the Malliavin sense (see for instance the sentence before Theorem 2.2 in [3], or Step 2 in the proof of Theorem 3.3 in [1], which refers to [3], or the proof of Part a) of Theorem 2.6 in [10], or Proposition 3.2 in [17]).

The aim of this paper is to provide an alternative sufficient condition to (EPQ1)–(EPQ2) for the Malliavin differentiability of the solution to a BSDE of the form (1.1) in the general non-Markovian setting. Our main result in that direction is Theorem 5.1 below, using a fundamentally different approach from [9,20], as well as different type of assumptions. Since they involve some notations concerning the analysis on the Wiener space, we refrain from detailing them immediately, and rather explain informally what are the main differences between our approach and the one of [9]. For the sake of simplicity in this introduction, we expose the problem under interest in the one dimensional case, however the present paper handles also the higher dimensional case. A natural way to solve a BSDE of the form (1.1) when the driver f is Lipschitz in (y, z) is to make use of a Picard iteration, that is to say a family (Y^n, Z^n) of solutions to BSDEs satisfying

$$Y_t^n = \xi + \int_t^T f\left(s, Y_s^{n-1}, Z_s^{n-1}\right) ds - \int_t^T Z_s^n \, dW_s, \quad t \in [0, T], \tag{1.2}$$

where $Y^0 \equiv Z^0 \equiv 0$. Then, a fixed point argument allows one to construct, in appropriate spaces, a solution (Y, Z) to Equation (1.1). If ξ and $f(t, y, z)$ are Malliavin differentiable, then so is (Y^n, Z^n). Then, it just remains to prove that this property extends to the limits Y and Z of respectively Y^n and Z^n, in appropriate spaces. More precisely this is done by a uniform (in n) control of the Sobolev norms of Y^n, Z^n or equivalently by proving that the Malliavin derivatives (DY^n, DZ^n) of (Y^n, Z^n) converge to the solution of a linear BSDE whose solution will be the Malliavin derivatives (DY, DZ) of Y and Z. This last step is exactly where the extra regularity (EPQ1)–(EPQ2) is needed. It appears quite clearly that for this approach, the conditions of [9] cannot be optimized in the general case. Even though this idea seems pretty natural, it is based on a choice somehow arbitrary. Indeed, a necessary condition for DY_t to be well defined at a given time t, is that there exists a sequence of random variables $(F^n)_n$ converging to Y_t in L^2 such that each variable F^n is Malliavin differentiable with derivative DF^n and such that DF^n converges, with respect to a suitable norm, to DY_t. As a consequence, in the approach described above, one believes that this sequence $(F^n)_n$ can be chosen to be the Picard iteration $(Y^n)_n$. Once again, this idea looks very natural, according to the same type of proofs for SDEs, but then one sees that in the BSDE framework this intuitive idea leads to pretty heavy assumptions. We elaborate a little bit more on this point in Section 6.3.

Regarding the discussion above, one could think of trying to find a sequence of processes known to approximate the Malliavin derivative of Y (and Z) when Y is Malliavin differentiable. This approximation is provided by the well-known interpretation of the Malliavin derivative as a Gâteaux derivative in the directions of the Cameron–Martin space. More precisely, a necessary condition for Y_t to belong to $\mathbb{D}^{1,2}$, is that for any absolutely continuous function h starting from 0 at 0 with derivative denoted $\dot{h}$, the difference quotient $\varepsilon^{-1}(Y_t(\omega + \varepsilon h) - Y_t(\omega))$ converges, in a sense to be made precise, as ε goes to 0 to $\langle DY_t, \dot{h} \rangle_{L^2([0,T])}$. This fact was initially given by Malliavin and then extended by Stroock, Shigekawa, Kusuoka and Sugita in a series of papers [15,16,22–24]. In addition, Sugita proved in [24] that a r.v. F is Malliavin differentiable if it is *ray absolutely continuous*[3] and if it is stochastically Gâteaux differentiable. Using the main ideas of [24] we provide incidentally a new formulation of the characterization of the Malliavin–Sobolev type spaces $\mathbb{D}^{1,p}$ in Theorem 4.1. Since we did not find explicitly this characterization in the literature, we believe that this result is new and maybe interesting by itself. The main point is that this formulation is especially handy when dealing with stochastic equations like BSDEs. With this result at hand, we obtain new conditions (see Assumptions (D), (H_1) and (H_2) at the beginning of Section 5) for Y, Z to be Malliavin differentiable, see Theorem 5.1. Our assumptions refine those of [9,20] in the Markovian case, and our approach is directly applicable to quadratic growth BSDEs since we do not rely on any approximation procedure. We refer the reader to Section 6 for some examples and a discussion on the differences between our approach and the one of [9,20].

The rest of the paper is organised as follows. We start below with some preliminaries. Then we turn in Section 3 to some elements of analysis on the Wiener space. Our characterization of the sets $\mathbb{D}^{1,p}$ is given in Section 4, and the

---

[3]We refer to Section 4 where this notion is recalled.

material on the Malliavin differentiability of BSDEs itself is contained in Section 5. We provide applications and a comparison of the results in Section 6. Finally, we extend our approach to quadratic growth BSDEs in Section 7.

2. Preliminaries

2.1. *Notations*

We fix throughout the paper a time horizon $T > 0$ and d a positive integer. For any positive integer k, we denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^k$ and by $\cdot$ the inner product, without mention of k which will be clear in the context. For any positive integers n and m, we identify $\mathbb{R}^{n \times m}$ with the space of real matrices with n rows and m columns, endowed with the Euclidean norm on $\mathbb{R}^{n \times m}$. Let M be in $\mathbb{R}^{n \times m}$, $1 \leq j \leq n$ and $1 \leq \ell \leq m$, we denote by $M^{j,:} \in \mathbb{R}^{1 \times m}$ (resp. $M^{:,\ell} \in \mathbb{R}^{n,1}$) its jth row (resp. its ℓth column). We set $M^{\top} \in \mathbb{R}^{m \times n}$ to be the transpose of M. We also identify $\mathbb{R}^k$ with $\mathbb{R}^{1,k}$. Let now $\Omega := C_0([0, T], \mathbb{R}^d)$ be the canonical Wiener space of continuous function $\omega := (\omega^1, \ldots, \omega^d)^{\top}$ from $[0, T]$ to $\mathbb{R}^d$ such that $\omega(0) = (0, \ldots, 0)^{\top}$. Let $W := (W_t^1, \ldots, W_t^d)_{t \in [0,T]}^{\top}$ be the canonical Wiener process, that is, for any time t in $[0, T]$, W_t denotes the evaluation mapping: $W_t^i(\omega) := \omega_t^i$ for any element ω in Ω and i in $\{1, \ldots, d\}$. We set $\mathbb{F}^o$ the natural filtration of W. Under the Wiener measure $\mathbb{P}_0$, the process W is a standard Brownian motion and we denote by $\mathbb{F} := (\mathcal{F}_t)_{t \in [0,T]}$ the usual augmentation (which is right-continuous and complete) of $\mathbb{F}^o$ under $\mathbb{P}_0$. Unless otherwise stated, all the expectations considered in this paper will have to be understood as expectations under $\mathbb{P}_0$, and all notions of measurability for elements of Ω will be with respect to the filtration $\mathbb{F}$ or the σ-field $\mathcal{F}_T$.

For any Hilbert space $\mathcal{K}$, for any $p \geq 1$ and for any $t \in [0, T]$, we set $L^p([t, T]; \mathcal{K})$ to be following space

$$L^p\big([t, T]; \mathcal{K}\big) := \left\{ f : [t, T] \longrightarrow \mathcal{K}, \text{Borel-measurable, s.t. } \int_t^T \|f(s)\|_{\mathcal{K}}^p \, ds < +\infty \right\},$$

where the norm $\|\cdot\|_{\mathcal{K}}$ is the one canonically induced by the inner product on $\mathcal{K}$. We denote, for simplicity, by $\mathfrak{H} := L^2([0, T]; \mathbb{R}^d)$ and by $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$ its canonical inner product, that is to say

$$\langle f, g \rangle_{\mathfrak{H}} := \int_0^T f(s) \cdot g(s) \, ds = \sum_{i=1}^d \int_0^T f^i(s) g^i(s) \, ds, \quad (f, g) \in \mathfrak{H}^2.$$

Let now H be the Cameron–Martin space that is the space of functions in Ω which are absolutely continuous with square-integrable derivative and which start from 0 at 0:

$$H := \left\{ h : [0, T] \longrightarrow \mathbb{R}^d, \exists \dot{h} \in \mathfrak{H}, h(t) = \int_0^t \dot{h}(x) \, dx, \forall t \in [0, T] \right\}.$$

For any h in H, we will always denote by $\dot{h}$ a version of its Radon–Nykodym density with respect to the Lebesgue measure. Then, H is an Hilbert space equipped with the inner product $\langle h_1, h_2 \rangle_H := \langle \dot{h_1}, \dot{h_2} \rangle_{\mathfrak{H}}$, for any $(h_1, h_2) \in H \times H$, and with associated norm $\|h\|_H^2 := \langle \dot{h}, \dot{h} \rangle_{\mathfrak{H}}$. Define next $L^p(\mathcal{K})$ as the set of all $\mathcal{F}_T$-measurable random variables F which are valued in an Hilbert space $\mathcal{K}$, and such that $\|F\|_{L^p(\mathcal{K})}^p < +\infty$, where

$$\|F\|_{L^p(\mathcal{K})} := \big(\mathbb{E}[\|F\|_{\mathcal{K}}^p]\big)^{1/p}.$$

Let now $\mathcal{S}$ be the set of cylindrical functionals, that is the set of $\mathbb{R}$-valued random variables F of the form

$$F = f\big(W(h_1), \ldots, W(h_n)\big), \quad (h_1, \ldots, h_n) \in H^n, f \in C_b^{\infty}(\mathbb{R}^n), \text{ for some } n \geq 1, \tag{2.1}$$

where $W(h) := \int_0^T \dot{h}_s \cdot dW_s := \sum_{i=1}^d \int_0^T \dot{h}_s^i \, dW_s^i$ for any h in H and where $C_b^{\infty}(\mathbb{R}^n)$ denotes the space of bounded mapping which are infinitely continuously differentiable with bounded derivatives. For any F in $\mathcal{S}$ of the form (2.1), the Malliavin derivative ∇F of F is defined as the following H-valued random variable:

$$\nabla F := \sum_{i=1}^n f_{x_i}\big(W(h_1), \ldots, W(h_n)\big) h_i, \tag{2.2}$$

where $f_{x_i} := \frac{\partial f}{\partial x_i}$. It is then customary to identify ∇F with the stochastic process $(\nabla_t F)_{t \in [0,T]}$, where $\nabla_t F \in \mathbb{R}^d$ for any $t \in [0, T]$. More precisely, we define for any $(t, \omega) \in [0, T] \times \Omega$,

$$\nabla_t F(\omega) := \sum_{i=1}^n f_{x_i}\big(W(h_1)(\omega), \ldots, W(h_n)(\omega)\big) h_i(t).$$

Denote then by $\mathbb{D}^{1,p}$ the closure of $\mathcal{S}$ with respect to the Malliavin–Sobolev semi-norm $\|\cdot\|_{1,p}$, defined as:

$$\|F\|_{1,p} := \big(\mathbb{E}\big[|F|^p\big] + \mathbb{E}\big[\|\nabla F\|_H^p\big]\big)^{1/p}.$$

We set $\mathbb{D}^{1,\infty} := \bigcap_{p \geq 2} \mathbb{D}^{1,p}$. In order to link our notations with the ones of the related papers [9,20] we make use of the notation DF to represent the derivative of ∇F as:

$$\nabla_t F = \int_0^t D_s F \, ds, \quad t \in [0, T].$$

We denote by $\delta : L^p(H) \to L^p(\mathbb{R})$ the adjoint operator of ∇ by the following duality relationship:

$$\mathbb{E}\big[F\delta(u)\big] = \mathbb{E}\big[\langle \nabla F, u \rangle_H\big], \quad \forall u \in \mathrm{dom}(\delta), \text{ where}$$

$$\mathrm{dom}(\delta) := \big\{ u \in L^p(H), \exists c_u > 0, \big|\mathbb{E}\big[\langle \nabla F, u \rangle_H\big]\big| \leq c_u \|F\|_{L^p(\mathbb{R})}, \forall F \in \mathbb{D}^{1,p} \big\}.$$

δ is also known under the name of Skorohod (or divergence) operator. Recall that any element u of the form $u := Gh$ with G in $\mathcal{S}$ and h in H belongs to $\mathrm{dom}(\delta)$ and that

$$\delta(Gh) = GW(h) - \langle \nabla G, h \rangle_H, \tag{2.3}$$

see for example [18, Relation (1.46)]. Note that for any h in H, $\delta(h) = W(h)$.

Notice that in [24] the cylindrical space, that we will denote by $\mathcal{P}$ in the following, is the space of functionals F of the form (2.1) with f a polynomial. More precisely let $\mathcal{P}$ be the set of polynomial cylindrical functionals, that is the set of random variables F of the form

$$F = f\big(W(h_1), \ldots, W(h_n)\big), \quad (h_1, \ldots, h_n) \in H^n, f \in \mathbb{R}^n[X], \text{ for some } n \geq 1, \tag{2.4}$$

where $\mathbb{R}^n[X]$ denotes the set of polynomials of degree less or equal to n. However, the closures of both $\mathcal{S}$ and $\mathcal{P}$ with respect to any $\|\cdot\|_{1,p}$ coincide, as any polynomial together with its derivative can be approximated in $L^p(\mathbb{R}^n)$ (see Lemma 2.1 below).

Lemma 2.1. *Let G be in $\mathcal{P}$. There exists a sequence $(G^N)_{N \geq 1} \subset \mathcal{S}$ such that $\lim_{N \to +\infty} G^N = G$ in $\mathbb{D}^{1,r}$ for any $r \geq 1$.*

Proof. Let $G := f(W(h_1), \ldots, W(h_n))$ with $n \geq 1$, h_i in H and f in $\mathbb{R}^n[X]$. Without loss of generality, we assume that the family $(h_1, \ldots, h_n)$ is orthonormal in H. Let θ be a cutoff function, that is a mapping $\theta : \mathbb{R}^n \to \mathbb{R}^+$ such that $\theta(x) = 1$ if $\|x\| < 1$, $\theta(x) = 0$ for $\|x\| \geq 2$, and such that $\theta \in C_b^\infty(\mathbb{R}^n)$. For $N \geq 1$, we set:

$$G^N := f^N\big(W(h_1), \ldots, W(h_n)\big), \qquad f^N(x) := f(x) \times \theta(x/N), \quad x \in \mathbb{R}^n.$$

Note that each random variable G^N belongs to $\mathcal{S}$. Fix $r \geq 1$. We aim at proving that $\lim_{N \to +\infty} \|G^N - G\|_{1,r} = 0$. On the one hand,

$$\mathbb{E}\big[|G^N - G|^r\big] = \mathbb{E}\left[|G|^r \left|\theta\left(\frac{W(h_1)}{N}, \ldots, \frac{W(h_n)}{N}\right) - 1\right|^r\right]$$

$$\leq \mathbb{E}\big[|G|^{2r}\big]^{1/2} \mathbb{E}\left[\left|\theta\left(\frac{W(h_1)}{N}, \ldots, \frac{W(h_n)}{N}\right) - 1\right|^{2r}\right]^{1/2} \leq C \int_{\mathbb{R}^n \setminus B^n(0,N)} e^{-\|x\|^2/2} \, dx,$$

where C is a positive constant. Hence, $\lim_{N \to +\infty} \mathbb{E}[|G^N - G|^r] = 0$. We now turn to the proof of the convergence of the derivatives. We have

$$\nabla G^N = \sum_{i=1}^n \frac{\partial f^N}{\partial x_i} \big(W(h_1), \ldots, W(h_n) \big) h_i,$$

with $\frac{\partial f^N}{\partial x_i}(x) = \frac{\partial f}{\partial x_i} \theta(x/N) + N^{-1} f(x) \frac{\partial \theta}{\partial x_i}(x/N)$. Hence:

$$
\begin{aligned}
\mathbb{E}\big[\big\|\nabla(G^N - G)\big\|_H^{2r}\big] &= \sum_{i=1}^n \mathbb{E}\left[\left|\frac{\partial f^N}{\partial x_i} - \frac{\partial f}{\partial x_i}\right|^{2r} \big(W(h_1), \ldots, W(h_n) \big)\right] \\
&\leq C\left(\sum_{i=1}^n \mathbb{E}\left[\left|\frac{\partial f}{\partial x_i}\big(W(h_1), \ldots, W(h_n)\big)\right|^{2r} \left|\theta\left(\frac{W(h_1)}{N}, \ldots, \frac{W(h_n)}{N}\right) - 1\right|^{2r}\right] \right. \\
&\qquad\left. + N^{-2r} \mathbb{E}\left[\left|\frac{\partial \theta}{\partial x_i}\left(\frac{W(h_1)}{N}, \ldots, \frac{W(h_n)}{N}\right) \times f\big(W(h_1), \ldots, W(h_n)\big)\right|^{2r}\right]\right) \\
&\underset{N \to +\infty}{\longrightarrow} 0. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square
\end{aligned}
$$

We conclude this section by introducing the following norms and spaces which are of interest when studying BSDEs. For any positive integers p, n, we set $\mathbb{S}_n^p$ the space of $\mathbb{R}^n$-valued, continuous and $\mathbb{F}$-progressively measurable processes Y s.t.

$$|Y|_{\mathbb{S}_n^p}^p := \mathbb{E}\left[\sup_{0 \leq t \leq T} \|Y_t\|^p\right] < +\infty.$$

We denote by $\mathbb{H}_{n,d}^p$ the space of $\mathbb{R}^{n \times d}$-valued and $\mathbb{F}$-predictable processes Z such that

$$|Z|_{\mathbb{H}_{n,d}^p}^p := \mathbb{E}\left[\left(\int_0^T \sum_{j=1}^n \|Z_t^j\|^2 \, dt\right)^{p/2}\right] < +\infty.$$

We set $\mathbb{S}^p := \mathbb{S}_1^p$ and $\mathbb{H}_d^p := \mathbb{H}_{1,d}^p$.

3. Some elements of analysis on the Wiener space

One of the main tool that we will use throughout this paper is the shift operator along directions in the Cameron–Martin space. More precisely, for any $h \in H$, we define the following shift operator $\tau_h : \Omega \to \Omega$ by

$$\tau_h(\omega) := \omega + h := \big(\omega^1 + h^1, \ldots, \omega^d + h^d\big)^\top.$$

Note that the fact that h belongs to H ensures that τ_h is a measurable shift on the Wiener space. In fact, one can be a bit more precise, since according to [25, Lemma B.2.1] for any $\mathcal{F}_T$-measurable r.v. F the mapping $h \mapsto F \circ \tau_h$ is continuous in probability from H to $L^0(\mathbb{R}^d)$, the space of real-valued and $\mathcal{F}_T$-measurable random variables, see Lemma 3.2 below. Taking $F = \mathrm{Id}$, one gets that τ_h is a continuous mapping on Ω for any h in H. We list below some other properties of such shifts.

Lemma 3.1 ([25, Appendix B.2]). *Let X and Y be two $\mathcal{F}_T$-measurable random variables. If $X = Y$, $\mathbb{P}_0$-a.s., then for any h in H,*

$$X \circ \tau_h = Y \circ \tau_h, \quad \mathbb{P}_0\text{-a.s.}$$

We recall, the quite surprising result that any r.v. is continuous in probability in the directions of the Cameron–Martin space. More precisely:

Lemma 3.2 ([25, Lemma B.2.1]). *Let F be a $\mathcal{F}_T$-measurable random variable. The mapping $h \mapsto F \circ \tau_h$ is continuous from H to $L^0(\mathbb{R}^d)$ where the convergence is in probability.*

One of the main technique when working with shifts on the path space is the famous Cameron–Martin formula.

Proposition 3.1 (Cameron–Martin formula, see e.g. [25, Appendix B.1]). *Let F be a $\mathcal{F}_T$-measurable random variable and let h be in H. Then, when both sides are well defined*

$$\mathbb{E}[F \circ \tau_h] = \mathbb{E}\left[F \exp\left(\int_0^T \dot{h}(s) \cdot dW_s - \frac{1}{2} \int_0^T \|\dot{h}(s)\|^2 \, ds \right) \right].$$

For further reference, we also emphasize that for any $h \in H$ and for any $p \geq 1$, the stochastic exponential $\mathcal{E}(\int_0^\cdot \dot{h}(s) \cdot dW_s) := \exp(\int_0^\cdot \dot{h}(s) \cdot dW_s - \frac{1}{2} \int_0^\cdot \|\dot{h}(s)\|^2 \, ds)$ verifies

$$\mathcal{E}\left(\int_0^\cdot \dot{h}(s) \cdot dW_s \right) \in \mathbb{S}^p, \quad \forall p \geq 1. \tag{3.1}$$

Lemma 3.3. *Let t in $[0, T]$ and let F be a $\mathcal{F}_t$-measurable random variable. For any h in H, it holds that*

$$F \circ \tau_h = F \circ \tau_{\tilde{h}^t}, \quad \mathbb{P}_0\text{-a.s.},$$

where

$$\tilde{h}^t(s) := \left(\int_0^s \dot{h}^1(u) \mathbf{1}_{0 \leq u \leq t} \, du, \ldots, \int_0^s \dot{h}^d(u) \mathbf{1}_{0 \leq u \leq t} \, du \right)^\top.$$

In particular, $F \circ \tau_h$ is $\mathcal{F}_t$-measurable.

Proof. It is well known that by definition of $\mathbb{P}_0$, any $\mathcal{F}_t$-measurable random variable admits a $\mathcal{F}_t^o$-measurable version. Therefore, there exists some measurable map $\varphi : \Omega \to \mathbb{R}$, such that

$$F = \varphi(W_{\cdot \wedge t}), \quad \mathbb{P}_0\text{-a.s.}$$

Hence, we deduce by Lemma 3.1 that for $\mathbb{P}_0$-a.e. $\omega \in \Omega$

$$F \circ \tau_h(\omega) = \varphi\big(W_{\cdot \wedge t}(\omega)\big) \circ \tau_h = \varphi\big(W_{\cdot \wedge t} \circ \tau_h(\omega)\big) = \varphi\big(\omega(\cdot \wedge t) + h(\cdot \wedge t)\big) = F \circ \tau_{\tilde{h}^t}(\omega). \qquad \square$$

We conclude this section with the following lemma which might be known. However since we did not find it in the literature we provide a proof in order to make this paper self-contained.

Lemma 3.4. *Let $Z \in \mathbb{H}_d^2$ and h in H. It holds that*

$$\int_0^T Z_s \cdot dW_s \circ \tau_h = \int_0^T Z_s \circ \tau_h \cdot dW_s + \int_0^T Z_s \circ \tau_h \cdot \dot{h}(s) \, ds, \quad \mathbb{P}_0\text{-a.s.}$$

Proof. Let $\mathfrak{S}$ be the class of simple processes X of the form

$$X_t := \left(X_t^1, \ldots, X_t^d \right)^\top, \qquad X_t^j = \sum_{i=0}^{n_j} \lambda_i^j \mathbf{1}_{(t_i^j, t_{i+1}^j]}(t), \quad j \in \{1, \ldots, d\},$$

where for any $j \in \{1, \ldots, d\}$, $n_j \in \mathbb{N}^*$, $t_0^j = 0 < t_1^j < \cdots < t_n^j = T$ and where for any $0 \le i \le n_j$, $(\lambda_i^j)_{i=1,\ldots,n_j}$ are $\mathcal{F}_{t_i^j}$-measurable and in $L^2(\mathbb{R})$.

We start by proving the result for Z in $\mathfrak{S}$ and then we prove the result for any element Z in $\mathbb{H}_d^2$ using a density argument. Let $Z \in \mathfrak{S}$ with the decomposition

$$Z_s = (Z_s^1, \ldots, Z_s^d)^\top, \qquad Z_s^j := \sum_{i=0}^{n_j} \lambda_i^j \mathbf{1}_{(t_i^j, t_{i+1}^j]}(s), \quad s \in [0, T], j \in \{1, \ldots, d\}.$$

Then, for any $h \in H$ and for every $\omega \in \Omega$,

$$\begin{aligned}
\left(\int_0^T Z_s \cdot dW_s \circ \tau_h \right)(\omega) &= \left(\sum_{j=1}^d \sum_{i=0}^{n_j} \lambda_i^j \left(W_{t_{i+1}^j}^j - W_{t_i^j}^j \right) \right) \circ \tau_h(\omega) \\
&= \sum_{j=1}^d \sum_{i=0}^{n_j} \lambda_i^j (\omega + h) \left(W_{t_{i+1}^j}^j - W_{t_i^j}^j \right)(\omega + h) \\
&= \sum_{j=1}^d \sum_{i=0}^{n_j} \lambda_i^j \circ \tau_h(\omega) \left(\omega^j(t_{i+1}^j) - \omega(t_i^j) + h^j(t_{i+1}^j) - h^j(t_i^j) \right) \\
&= \int_0^T Z_s \circ \tau_h \cdot dW_s(\omega) + \int_0^T Z_s \circ \tau_h(\omega) \cdot dh_s,
\end{aligned}$$

which gives the desired result since h is absolutely continuous. We extend this result to processes Z in $\mathbb{H}_d^2$. Let $Z \in \mathbb{H}_d^2$, then there exists a sequence $(Z^n)_{n \in \mathbb{N}}$ in $\mathfrak{S}$ which converges to Z in $\mathbb{H}_d^2$. Hence,

$$\begin{aligned}
&\mathbb{E}\left[\left| \int_0^T Z_s \cdot dW_s \circ \tau_h - \int_0^T Z_s \circ \tau_h \cdot dW_s - \int_0^T Z_s \circ \tau_h \cdot dh_s \right| \right] \\
&\le \mathbb{E}\left[\left| \int_0^T Z_s \cdot dW_s \circ \tau_h - \int_0^T Z_s^n \cdot dW_s \circ \tau_h \right| \right] + \mathbb{E}\left[\left| \int_0^T Z_s^n \circ \tau_h \cdot dW_s - \int_0^T Z_s \circ \tau_h \cdot dW_s \right| \right] \\
&\quad + \mathbb{E}\left[\left| \int_0^T Z_s^n \circ \tau_h \cdot dh_s - \int_0^T Z_s \circ \tau_h \cdot dh_s \right| \right] \\
&\le \underbrace{\mathbb{E}\left[\left| \int_0^T (Z_s - Z_s^n) \cdot dW_s \right| \circ \tau_h \right]}_{=:A^n} + \underbrace{\mathbb{E}\left[\left| \int_0^T (Z_s^n - Z_s) \circ \tau_h \cdot dW_s \right| \right]}_{=:B^n} \\
&\quad + \underbrace{\mathbb{E}\left[\left| \int_0^T (Z_s^n - Z_s) \circ \tau_h \cdot dh_s \right| \right]}_{=:C^n}.
\end{aligned}$$

Let us estimate these three terms. First, using Proposition 3.1, Cauchy–Schwarz inequality, then Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned}
A^n &= \mathbb{E}\left[\left| \int_0^T (Z_s - Z_s^n) \cdot dW_s \right| e^{\int_0^T \dot{h}(s) \cdot dW_s - (1/2) \int_0^T \|\dot{h}(s)\|^2 \, ds} \right] \\
&\le \mathbb{E}\left[\int_0^T \|Z_s - Z_s^n\|^2 \, ds \right]^{1/2} \mathbb{E}\left[\mathcal{E}\left(\int_0^\cdot \dot{h}(s) \cdot dW_s \right)_T^2 \right]^{1/2}.
\end{aligned}$$

By (3.1), this clearly goes to 0 as n goes to infinity. Similarly, using Burkholder–Davis–Gundy Inequality, we have

$$B^n \leq \mathbb{E}\left[\left(\int_0^T \|(Z_s^n - Z_s) \circ \tau_h\|^2 ds\right)^{1/2}\right] = \mathbb{E}\left[\left(\int_0^T \|Z_s^n - Z_s\|^2 ds\right)^{1/2} \circ \tau_h\right].$$

Therefore, we can use Proposition 3.1 and Cauchy–Schwarz inequality, to also deduce that $B^n \to_{n\to+\infty} 0$. Finally, we have

$$
\begin{aligned}
C^n &= \mathbb{E}\left[\mathcal{E}\left(\int_0^T \dot{h}(s) \cdot dW_s\right)\left|\int_0^T (Z_s^n - Z_s) \cdot \dot{h}(s) \, ds\right|\right] \\
&\leq \mathbb{E}\left[\mathcal{E}\left(\int_0^T \dot{h}(s) \cdot dW_s\right)^2\right]^{1/2} \mathbb{E}\left[\left(\int_0^T \|Z_s^n - Z_s\|\|\dot{h}(s)\| \, ds\right)^2\right]^{1/2} \\
&\leq \mathbb{E}\left[\mathcal{E}\left(\int_0^T \dot{h}(s) \cdot dW_s\right)^2\right]^{1/2} \mathbb{E}\left[\int_0^T \|Z_s^n - Z_s\|^2 ds\right]^{1/2}\left(\int_0^T \|\dot{h}(s)\|^2 ds\right)^{1/2},
\end{aligned}
$$

which also goes to 0 as n goes to infinity. Therefore the proof is complete. $\qquad\square$

This result entails the following useful consequence. Let t in $(0, T]$ and h in H such that $\dot{h}_s = (0, \ldots, 0)$ for $s \geq t$. Then for any Z in $\mathbb{H}_d^2$, it holds that:

$$\int_t^T Z_s \cdot dW_s \circ \tau_h = \int_t^T Z_s \circ \tau_h \cdot dW_s, \quad \mathbb{P}_0\text{-a.s.,} \tag{3.2}$$

since $\int_t^T Z_s \circ \tau_h \cdot \dot{h}(s) \, ds = 0$.

4. A characterization of Malliavin differentiability

Before going further, we would like to recall the main finding of [24]. Any Malliavin–Sobolev type space $\mathbb{D}^{1,p}$ as defined in Section 2 (originally defined by Malliavin [16] and Shigekawa [22]) agrees with the Sobolev space (due to Stroock [23] and Kusuoka [15]) $\tilde{\mathbb{D}}^{1,p}$ consisting in the set of *Ray Absolutely Continuous* (RAC) and *Stochastically Gâteaux Differentiable* (SGD) r.v. F in $L^p(\mathbb{R})$, where these notions are defined as follows:

(RAC) For any h in H, there exists a r.v. $\tilde{F}_h$ such that $\tilde{F}_h = F$, $\mathbb{P}_0$-a.s., and such that for any ω in Ω, $t \in \mathbb{R} \mapsto \tilde{F}_h(\omega + th)$ is absolutely continuous, where $th := (th^1, \ldots, th^d)$.
(SGD) There exists $\mathcal{D}F$ in $L^p(H)$ such that for any h in H,

$$\frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} \xrightarrow[\varepsilon\to 0]{} \langle \mathcal{D}F, h\rangle_H, \quad \text{in probability.} \tag{4.1}$$

In addition, for any F in $\mathbb{D}^{1,p}$, $\nabla F = \mathcal{D}F$, $\mathbb{P}_0$-a.s. Note that according to the statement of Step 1 in the proof of [24, Theorem 3.1], if F is (RAC) and (SGD) then for any h in H and any $\varepsilon > 0$ it holds that

$$\varepsilon^{-1}(\tilde{F}_h \circ \tau_{\varepsilon h} - \tilde{F}_h) = \varepsilon^{-1}\int_0^\varepsilon \langle \nabla F \circ \tau_{sh}, h\rangle_H \, ds, \quad \mathbb{P}_0\text{-a.s.}$$

Furthermore, by Lemma 3.1, we have for any ε there exists a set A^ε such that $\mathbb{P}_0[A^\varepsilon] = 0$ and $F \circ \tau_{\varepsilon h} = \tilde{F}_h \circ \tau_{\varepsilon h}$ and $F = \tilde{F}_h$ outside A^ε. Hence, for any ε in $(0, 1)$, the relation above rewrites as:

$$\varepsilon^{-1}(F \circ \tau_{\varepsilon h} - F) = \varepsilon^{-1}\int_0^\varepsilon \langle \nabla F \circ \tau_{sh}, h\rangle_H \, ds, \quad \mathbb{P}_0\text{-a.s.} \tag{4.2}$$

Remark 4.1. *It has actually been proved by Janson* [14] *that* (4.2) *is equivalent to* (RAC) *and* (SGD), *for any* $p > 1$, *see Lemma* 15.89. *Notice that* [14] *also obtained a similar characterization for* $p = 1$ (*see Lemma* 15.71). *However, as stated in Remark* 4 *of* [24], *the identification of the Kusuoka–Stroock and Shigekawa spaces when* $p = 1$ *is still an open result, so that we never consider the case* $p = 1$ *in this paper.*

The main result of this section is the following theorem whose proof is postponed to the end of the section.

Theorem 4.1. *Let* $p > 1$ *and* $F \in L^p(\mathbb{R})$. *The following properties are equivalent.*

(i) *F belongs to* $\mathbb{D}^{1,p}$.

(ii) *There exists $\mathcal{D}F$ in $L^p(H)$ such that for any h in H and any $q \in [1, p)$*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left| \frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} - \langle \mathcal{D}F, h \rangle_H \right|^q \right] = 0.$$

(iii) *There exists $\mathcal{D}F$ in $L^p(H)$ and there exists $q \in [1, p)$ such that for any h in H*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left| \frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} - \langle \mathcal{D}F, h \rangle_H \right|^q \right] = 0.$$

(iv) *There exists $\mathcal{D}F$ in $L^p(H)$ such that for any h in H*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left| \frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} - \langle \mathcal{D}F, h \rangle_H \right| \right] = 0.$$

In that case, $\mathcal{D}F = \nabla F$.

Remark 4.2. *The implication* (ii) $\Rightarrow$ (i) *when* $q = p = 2$ *already appears in* [5] (*see Proposition* 8.11.3). *This is of course contained in our result.*

We now give the following lemma which characterizes the Malliavin derivative using the duality formula involving the Skorohod operator (also called divergence operator).

Lemma 4.1. *Let $\varepsilon > 0$ and $1 < p < +\infty$. Suppose that $F \in L^{1+\varepsilon}(\mathbb{R})$ and assume that there exists $\mathcal{D}F$ in $L^p(H)$ such that:*

$$\mathbb{E}\big[F \delta(Gh) \big] = \mathbb{E}\big[G \langle \mathcal{D}F, h \rangle_H \big],$$

for every $G \in \mathcal{S}$ and $h \in H$. Then, it holds that $F \in \mathbb{D}^{1,p}$, and $\mathcal{D}F = \nabla F$, $\mathbb{P}_0$-a.s.

Proof. We know (see e.g. [24, Corollary 2.1]) that the result is true if $\mathcal{S}$ is replaced by $\mathcal{P}$. Let G be in $\mathcal{P}$. By Lemma 2.1 there exists (G^N) in $\mathcal{S}$ such that G^N approximates G in $\mathbb{D}^{1,p}$. Let h in H. For any $N \geq 1$, we have

$$\mathbb{E}\big[F \delta(Gh) \big] = \mathbb{E}\big[F \big(G W(h) - \langle \nabla G, h \rangle_H \big) \big]$$
$$= \mathbb{E}\big[F \big(G^N W(h) - \langle \nabla G^N, h \rangle_H \big) \big] - \mathbb{E}\big[(G^N - G) F W(h) - F \langle \nabla (G^N - G), h \rangle_H \big]$$
$$= \mathbb{E}\big[F \delta(G^N h) \big] - \mathbb{E}\big[(G^N - G) F W(h) - F \langle \nabla (G^N - G), h \rangle_H \big]$$
$$= \mathbb{E}\big[G^N \langle \mathcal{D}F, h \rangle_H \big] - \mathbb{E}\big[(G^N - G) F W(h) - F \langle \nabla (G^N - G), h \rangle_H \big].$$

Furthermore, by Lemma 2.1

$$\left| \mathbb{E}\big[(G^N - G) F W(h) \big] \right| \leq \mathbb{E}\big[|F W(h)|^p \big]^{1/p} \mathbb{E}\big[|G^N - G|^{\bar{p}} \big]^{1/\bar{p}} \underset{N \to +\infty}{\longrightarrow} 0,$$

with $1 < p < 1 + \varepsilon$ and where $\bar{p}$ is the conjugate of p, and

$$\left| \mathbb{E}\big[F \langle \nabla (G^N - G), h \rangle_H \big] \right| \leq \mathbb{E}\big[|F|^p \big]^{1/p} \mathbb{E}\big[\| \nabla (G^N - G) \|_H^{\bar{p}} \big]^{1/\bar{p}} \| h \|_H \underset{N \to +\infty}{\longrightarrow} 0,$$

by Lemma 2.1 again. Hence,

$$\mathbb{E}\big[F \delta(Gh) \big] = \lim_{N \to +\infty} \mathbb{E}\big[G^N \langle \mathcal{D}F, h \rangle_H \big]$$

$$= \mathbb{E}\big[G \langle \mathcal{D}F, h \rangle_H \big] + \lim_{N \to +\infty} \mathbb{E}\big[(G^N - G) \langle \mathcal{D}F, h \rangle_H \big],$$

and

$$\lim_{N \to +\infty} \left| \mathbb{E}\big[(G^N - G) \langle \mathcal{D}F, h \rangle_H \big] \right| \leq \lim_{N \to +\infty} \mathbb{E}\big[|G^N - G|^p \big]^{1/p} \mathbb{E}\big[\| \mathcal{D}F \|_H^{\bar{p}} \big]^{1/\bar{p}} \| h \|_H \underset{N \to +\infty}{\longrightarrow} 0.$$

Thus we have proved that for any G in $\mathcal{P}$ and for any h in H,

$$\mathbb{E}\big[F \delta(Gh) \big] = \mathbb{E}\big[G \langle \mathcal{D}F, h \rangle_H \big],$$

which gives the result by [24, Corollary 2.1]. □

We now prove the following lemma for the Malliavin differentiability of a given random variable.

Lemma 4.2. *Let $p > 1$. Let F be in $\mathbb{D}^{1,p}$. Then, for any q in $[1, p)$ and for any h in H,*

$$\frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} \underset{\varepsilon \to 0}{\longrightarrow} \langle \nabla F, h \rangle_H \quad \text{in } L^q(\mathbb{R}).$$

Proof. Fix q in $[1, p)$, h in H and $\eta > 0$ such that $q + \eta < p$. We know from [24, Theorem 3.1] that since F is in $\mathbb{D}^{1,p}$, F is (SGD), (RAC), and Relation (4.2) holds true. We thus have using Jensen's inequality

$$\mathbb{E}\big[|\varepsilon^{-1}(F \circ \tau_{\varepsilon h} - F)|^{q+\eta} \big] = \mathbb{E}\left[\varepsilon^{-(q+\eta)} \left| \int_0^\varepsilon \langle \nabla F \circ \tau_{sh}, h \rangle_H \, ds \right|^{q+\eta} \right]$$

$$\leq \varepsilon^{-1} \mathbb{E}\left[\int_0^\varepsilon \big| \langle \nabla F \circ \tau_{sh}, h \rangle_H \big|^{q+\eta} \, ds \right]$$

$$= \varepsilon^{-1} \int_0^\varepsilon \mathbb{E}\big[|\langle \nabla F, h \rangle_H|^{q+\eta} \circ \tau_{sh} \big] \, ds$$

$$= \varepsilon^{-1} \int_0^\varepsilon \mathbb{E}\left[|\langle \nabla F, h \rangle_H|^{q+\eta} \mathcal{E}\left(s \int_0^T \dot{h}_r \cdot dW_r \right) \right] ds$$

$$\leq \mathbb{E}\big[|\langle \nabla F, h \rangle_H|^p \big]^{(q+\eta)/p} \sup_{t \in (0,1)} \mathbb{E}\left[\left| \mathcal{E}\left(t \int_0^T \dot{h}_r \cdot dW_r \right) \right|^{p/(p-q-\eta)} \right]^{(p-q-\eta)/p}$$

$$< +\infty.$$

Hence by de La Vallée Poussin Criterion, we deduce that the family of random variables $(|\varepsilon^{-1}(F \circ \tau_{\varepsilon h} - F)|^q)_{\varepsilon \in (0,1)}$ is uniformly integrable which together with the convergence in probability (4.1) gives the result. □

Remark 4.3. *Note that the conclusion of the previous lemma may fail for $q = p$.*[4]

---

[4]After the first version of this paper, a counter example has been given in [13].

We can now proceed with the proof of Theorem 4.1.

Proof of Theorem 4.1. From Lemma 4.2 we have (i) $\Rightarrow$ (ii) and of course (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). We turn to (iv) $\Rightarrow$ (i). Let F be such that there exists $\mathcal{D}F$ in $L^p(H)$ such that

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left|\frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} - \langle \mathcal{D}F, h \rangle_H \right|\right] = 0.$$

The proof consists in applying Lemma 4.1 by proving the duality relationship

$$\mathbb{E}\big[F\delta(Gh)\big] = \mathbb{E}\big[G\langle \mathcal{D}F, h \rangle_H\big], \quad G \in \mathcal{S}, h \in H. \tag{4.3}$$

By Lemma A.1 (in the Appendix) with $\varepsilon = 0$,

$$\begin{aligned}
\mathbb{E}\big[F\delta(Gh)\big] &= \frac{d}{d\varepsilon}\mathbb{E}[F \circ \tau_{\varepsilon h} G]\bigg|_{\varepsilon=0} \\
&= \lim_{\eta \to 0} \frac{1}{\eta}\mathbb{E}\big[(F \circ \tau_{\eta h} - F)G\big] \\
&= \lim_{\eta \to 0} \mathbb{E}\left[\left(\frac{F \circ \tau_{\eta h} - F}{\eta} - \langle \mathcal{D}F, h \rangle_H\right)G\right] \\
&\quad + \mathbb{E}\big[\langle \mathcal{D}F, h \rangle_H G\big] \\
&= \mathbb{E}\big[\langle \mathcal{D}F, h \rangle_H G\big], \tag{4.4}
\end{aligned}$$

where the proof that the first term on the right-hand side goes to 0 is reported below.

Note that $\mathbb{E}[|\langle \mathcal{D}F, h \rangle_H G|] < +\infty$ since G is bounded and $\mathcal{D}F$ belongs to $L^p(H)$. Equality (4.4) is justified by Hölder's inequality since

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(\frac{F \circ \tau_{\eta h} - F}{\eta} - \langle \mathcal{D}F, h \rangle_H\right)G\right|\right] \\
&\quad \leq \|G\|_\infty \mathbb{E}\left[\left|\frac{F \circ \tau_{\eta h} - F}{\eta} - \langle \mathcal{D}F, h \rangle_H\right|\right] \underset{\varepsilon \to 0}{\longrightarrow} 0. \qquad \square
\end{aligned}$$

Corollary 4.1. *Let F be in $\mathbb{D}^{1,p}$. For any $\varepsilon > 0$ and any h in H, $F \circ \tau_{\varepsilon h}$ belongs to $\mathbb{D}^{1,p}$ and $\nabla(F \circ \tau_{\varepsilon h}) = (\nabla F) \circ \tau_{\varepsilon h}$.*

Proof. Let F be in $\mathbb{D}^{1,p}$. Using Theorem 4.1, we know that for any h in H and any $q \in [1, p)$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left|\frac{F \circ \tau_{\varepsilon h} - F}{\varepsilon} - \langle \nabla F, h \rangle_H \right|^q\right] = 0.$$

By Lemma A.1 (in the Appendix) it holds that

$$\begin{aligned}
\mathbb{E}\big[F \circ \tau_{\varepsilon h}\delta(Gh)\big] &= \frac{d}{d\varepsilon}\mathbb{E}[F \circ \tau_{\varepsilon h} G] \\
&= \lim_{\eta \to 0} \frac{1}{\eta}\mathbb{E}\big[(F \circ \tau_{(\varepsilon+\eta)h} - F \circ \tau_{\varepsilon h})G\big] \\
&= \lim_{\eta \to 0} \mathbb{E}\left[\left(\frac{F \circ \tau_{(\varepsilon+\eta)h} - F \circ \tau_{\varepsilon h}}{\eta} - \langle (\nabla F) \circ \tau_{\varepsilon h}, h \rangle_H\right)G\right] \\
&\quad + \mathbb{E}\big[\langle (\nabla F) \circ \tau_{\varepsilon h}, h \rangle_H G\big] \\
&= \mathbb{E}\big[\langle (\nabla F) \circ \tau_{\varepsilon h}, h \rangle_H G\big], \tag{4.5}
\end{aligned}$$

where the proof that the first term on the right-hand side goes to 0 is reported below.

Note that $\mathbb{E}[|\langle(\nabla F)\circ\tau_{\varepsilon h},h\rangle_H G|]<+\infty$ since $\langle(\nabla F)\circ\tau_{\varepsilon h},h\rangle_H=\langle\nabla F,h\rangle_H\circ\tau_{\varepsilon h}$, $\mathbb{P}_0$-a.s., G belongs to all the spaces $L^r(\mathbb{R})$ for $r\geq 1$ and

$$\mathbb{E}\big[\big|\langle\nabla F,h\rangle_H\big|^p\big]\leq\|h\|_H^p\mathbb{E}\big[\|\nabla F\|_H^p\big]<+\infty.$$

Equality (4.5) is justified by Hölder's inequality since

$$\mathbb{E}\bigg[\bigg|\bigg(\frac{F\circ\tau_{(\varepsilon+\eta)h}-F\circ\tau_{\varepsilon h}}{\eta}-\langle\nabla F\circ\tau_{\varepsilon h},h\rangle_H\bigg)G\bigg|\bigg]$$

$$\leq\mathbb{E}\bigg[\bigg|\frac{F\circ\tau_{(\varepsilon+\eta)h}-F\circ\tau_{\varepsilon h}}{\eta}-\langle\nabla F\circ\tau_{\varepsilon h},h\rangle_H\bigg|^r\bigg]^{1/r}\mathbb{E}\big[|G|^{\bar{r}}\big]^{1/\bar{r}}$$

$$=\mathbb{E}\bigg[\bigg|\frac{F\circ\tau_{\eta h}-F}{\eta}-\langle\nabla F,h\rangle_H\bigg|^r\circ\tau_{\varepsilon h}\bigg]^{1/r}\mathbb{E}\big[|G|^{\bar{r}}\big]^{1/\bar{r}}$$

$$\leq\mathbb{E}\bigg[\bigg|\frac{F\circ\tau_{\eta h}-F}{\eta}-\langle\nabla F,h\rangle_H\bigg|^q\bigg]^{1/q}\mathbb{E}\bigg[\mathcal{E}\bigg(\varepsilon\int_0^T\dot{h}(s)\cdot dW_s\bigg)^{\bar{\alpha}}\bigg]^{1/(r\bar{\alpha})}\mathbb{E}\big[|G|^{\bar{r}}\big]^{1/\bar{r}},$$

where $1<r<q$ and $\alpha:=\frac{q}{r}$ and where $\bar{r}$ (resp. $\bar{\alpha}$) is the Hölder conjugate of r (resp. α). Consequently, $\mathbb{E}[F\circ\tau_{\varepsilon h}\delta(Gh)]=\mathbb{E}[\langle\nabla F\circ\tau_{\varepsilon h},h\rangle_H G]$, and from Lemma 4.1 $\nabla(F\circ\tau_{\varepsilon h})=(\nabla F)\circ\tau_{\varepsilon h}$. □

5. Malliavin's differentiability of BSDEs

In this section we derive a sufficient condition ensuring that the solution to a BSDE is Malliavin differentiable. To simplify the comparison of the results with the companion papers [9,20] we adopt the notations used in these papers concerning the Malliavin calculus. More precisely, for any F in $\mathbb{D}^{1,p}$ (for $p>1$) we have defined the Malliavin derivative ∇F as an H-valued random variable. Recall that denoting DF the derivative of ∇F that is $\nabla_t F=\int_0^t D_r F\,dr$, DF coincides with the Malliavin derivative introduced in [9,18,20]. In particular $\langle\nabla F,h\rangle_H=\langle DF,\dot{h}\rangle_{\mathfrak{H}}$ for any h in H.

Let n be a positive integer, we consider now the following BSDE:

$$Y_t=\xi+\int_t^T f(r,Y_r,Z_r)\,dr-\int_t^T Z_r\,dW_r,\quad t\in[0,T],\mathbb{P}_0\text{-a.s.,}\tag{5.1}$$

where $\xi:=(\xi^1,\ldots,\xi^n)^\top$ is a $\mathcal{F}_T$-measurable $\mathbb{R}^n$-valued r.v. and $f:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}^n$ is a $\mathbb{F}$-progressively measurable process where as usual the ω-dependence is omitted.

The aim of this section is to show that for any $t\in[0,T]$, we can apply Theorem 4.1 under the following assumptions:

(L) The map $(y,z)\mapsto f(\cdot,y,z)$ is differentiable with uniformly bounded and continuous partial derivatives. We denote by $f_y:=(\frac{\partial f^j}{\partial y_k})_{j\in\{1,\ldots,n\},k\in\{1,\ldots,n\}}$ the Jacobian matrix of f with respect to y, where j (resp. k) indexes the columns (resp. the rows) of f_y and f_y^j denotes the gradient of f^j. We denote by f_z^j, for any $j\in\{1,\ldots,n\}$ the Jacobian matrix of f^j with respect to z, that is $f_z^j=(\frac{\partial f^j}{\partial z_{k,l}})_{k\in\{1,\ldots,n\},l\in\{1,\ldots,d\}}$.

(D) ξ belongs to $(\mathbb{D}^{1,2})^n$, for any $(y,z)\in\mathbb{R}^n\times\mathbb{R}^{n\times d}$, the map $(t,\omega)\mapsto f(t,\omega,y,z)$ is in $L^2([0,T];(\mathbb{D}^{1,2})^n)$, $f(\cdot,y,z)$ and $Df(\cdot,y,z)$ are $\mathbb{F}$-progressively measurable, and

$$\mathbb{E}\bigg[\int_0^T\sum_{j=1}^n|D.f^j(s,Y_s,Z_s)|_{\mathfrak{H}}^2\,ds\bigg]<+\infty.$$

(H_1) There exists $p \in (1, 2)$ such that for any $h \in H$ and for any $j \in \{1, \ldots, n\}$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left(\int_0^T \left| \frac{f^j(s, \cdot + \varepsilon h, Y_s, Z_s) - f^j(s, \cdot, Y_s, Z_s)}{\varepsilon} - \langle Df^j(s, \cdot, Y_s, Z_s), \dot{h} \rangle_{\mathfrak{H}} \right| ds \right)^p \right] = 0.$$

(H_2) Let $(\varepsilon_k)_{k \in \mathbb{N}}$ be a sequence in $(0, 1]$ such that $\lim_{k \to +\infty} \varepsilon_k = 0$, and let $(Y^k, Z^k)_k$ be a sequence of random variables which converges in $\mathbb{S}_n^p \times \mathbb{H}_{n,d}^p$ for any $p \in [1, 2)$ to some (Y, Z). Then for all $h \in H$ and for all $j \in \{1, \ldots, n\}$, the following convergences hold in probability

$$\left| f_y^j(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot) - f_y^j(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T]; \mathbb{R}^n)} \xrightarrow[k \to +\infty]{} 0,$$

$$\left| f_z^j(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot^k) - f_z^j(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T]; \mathbb{R}^{n \times d})} \xrightarrow[k \to +\infty]{} 0, \tag{5.2}$$

or

$$\left| f_y^j(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot^k) - f_y^j(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T]; \mathbb{R}^n)} \xrightarrow[k \to +\infty]{} 0,$$

$$\left| f_z^j(\cdot, \omega + \varepsilon_k h, Y_\cdot, Z_\cdot^k) - f_z^j(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T]; \mathbb{R}^{n \times d})} \xrightarrow[k \to +\infty]{} 0. \tag{5.3}$$

Before turning to the main result of this section, we would like to comment on Assumption (H_2). In this explanation we set $n = 1$ for the sake of simplicity. On the one hand, by Lemma 3.2, at given (s, y, z), $f_y(s, \omega + \varepsilon_k h, y, z)$ converges in probability to $f_y(s, \omega, y, z)$ as n goes to infinity. On the other hand, $f_y(s, \omega, \cdot)$ is continuous by assumption. Thus, Condition (H_2) is just requiring joint continuity of f_y in $L^2([0, T])$. The same comment holds for f_z. Note finally, that since f_y is assumed to be bounded, a sufficient condition for (H_2) to hold true is that $f_y(t, Y_t^k, Z_t)$ converges in probability to $f_y(t, Y_t, Z_t)$ for dt-almost every t (and the same for f_z).

Recall that for any M in $\mathbb{R}^{n \times d}$, $1 \leq j \leq n$ and $1 \leq \ell \leq d$, we denote by $M^{j,:} \in \mathbb{R}^{1 \times d}$ (resp. $M^{:,\ell} \in \mathbb{R}^{n,1}$) its jth row (resp. its ℓth column). We thus have

Theorem 5.1. *Let t be in $[0, T]$. Under Assumptions* (L), (D), (H_1) *and* (H_2), Y_t *belongs to* $(\mathbb{D}^{1,2})^n$ *and* $(Z^{j,:})^\top \in L^2([t, T]; (\mathbb{D}^{1,2})^d)$, $j \in \{1, \ldots, n\}$.

Proof. We only consider the case where (5.2) holds in Assumption (H_2), since the other one can be treated similarly. We prove first that Y_t^j belongs to $\mathbb{D}^{1,p}$ for any j in $\{1, \ldots, n\}$ where $p \in (1, 2)$ is the exponent appearing in Assumption (H_1), and then we extend the result to $\mathbb{D}^{1,2}$. To this end we aim at applying Theorem 4.1. Fix j in $\{1, \ldots, n\}$. Let h in H. Since Y^j is $\mathbb{F}$-progressively measurable, by Lemma 3.3, we can assume without loss of generality that $\dot{h}_s = 0$ for $s > t$. Let $\varepsilon > 0$. By Lemmas 3.1 and 3.4, it holds that

$$Y_s^j \circ \tau_{\varepsilon h} = \xi^j \circ \tau_{\varepsilon h} + \int_s^T f^j(r, Y_r, Z_r) \circ \tau_{\varepsilon h} \, dr - \int_s^T \left(Z_r^{j,:}\right)^\top \circ \tau_{\varepsilon h} \cdot dW_r, \quad \forall s \in [t, T], \mathbb{P}_0\text{-a.s.}$$

As a consequence, setting for the sake of simplicity

$$Y_s^\varepsilon := \frac{1}{\varepsilon}(Y_s \circ \tau_{\varepsilon h} - Y_s), \qquad Z_s^\varepsilon := \frac{1}{\varepsilon}(Z_s \circ \tau_{\varepsilon h} - Z_s), \qquad \xi^\varepsilon := \frac{1}{\varepsilon}(\xi \circ \tau_{\varepsilon h} - \xi), \quad s \in [t, T],$$

we have for any $j \in \{1, \ldots, n\}$

$$(Y_s^\varepsilon)^j = (\xi^\varepsilon)^j + \int_s^T \left(\tilde{A}_r^\varepsilon + \tilde{A}_r^{y,\varepsilon} \cdot Y_r^\varepsilon + \sum_{k=1}^d (f_z^j)^{:,k}(r, \cdot + \varepsilon h, Y_r \circ \tau_{\varepsilon h}, \tilde{Z}_r^k) \cdot (Z_r^\varepsilon)^{:,k} \right) dr$$

$$- \int_s^T \left((Z_r^\varepsilon)^{j,:}\right)^\top \cdot dW_r, \tag{5.4}$$

with

$$\tilde{A}_r^{y,\varepsilon} := f_y^j(r, \cdot + \varepsilon h, \bar{Y}_r^{\varepsilon,h}, Z_r), \qquad \tilde{A}_r^\varepsilon := \frac{1}{\varepsilon}\big(f^j(r, \cdot + \varepsilon h, Y_r, Z_r) - f^j(r, \cdot, Y_r, Z_r)\big),$$

where $\bar{Y}_r^{\varepsilon,h}$ is a convex combination of Y_r and $Y_r \circ \tau_{\varepsilon h}$ and where for any $k \in \{1, \ldots, d\}$, we have $\tilde{Z}_r^k := (Z_r^{:,1} \circ \tau_{\varepsilon h}, \ldots, Z_r^{:,k-1} \circ \tau_{\varepsilon h}, \bar{Z}_r^{:,k}, Z_r^{:,k+1}, \ldots, Z_r^{:,d})$ where $\bar{Z}_r^{:,k}$ is a convex combination of $Z_r^{:,k} \circ \tau_{\varepsilon h}$ and $Z_r^{:,k}$.

Under Assumptions (D) and (L), the following linear BSDE on $[t, T]$ has a unique solution $(\hat{Y}^h, \hat{Z}^h)$ in $\mathbb{S}_n^2 \times \mathbb{H}_{n,d}^2$ with $\hat{Y}^h := ((\hat{Y}^h)^j)_{j \in \{1, \ldots, n\}}$, $\hat{Z}^h := ((\hat{Z}^h)^{j,:})_{j \in \{1, \ldots, n\}}$ and for any $j \in \{1, \ldots, n\}$

$$(\hat{Y}_s^h)^j = \langle \nabla(\xi^j), h \rangle_H + \int_s^T \bigg(\langle \nabla(f^j)(r, \cdot, Y_r, Z_r), h \rangle_H + f_y^j(r, \cdot, Y_r, Z_r) \cdot \hat{Y}_r^h $$
$$+ \sum_{k=1}^d (f_z^j)^{:,k}(r, \cdot + \varepsilon h, Y_r, Z_r) \cdot (\hat{Z}_r^h)^{:,k} \bigg) dr - \int_s^T ((\hat{Z}_r^h)^{j,:})^\top \cdot dW_r. \tag{5.5}$$

Using a priori estimates (see Proposition 3.2 in [7]) in L^p, we have for some constant C_p, independent of ε

$$\mathbb{E}\bigg[\sup_{s \in [t,T]} |(Y_s^\varepsilon)^j - (\hat{Y}_s^h)^j|^p \bigg] + \mathbb{E}\bigg[\bigg(\int_t^T \|((Z_s^\varepsilon)^{j,:})^\top - ((\hat{Z}_s^h)^{j,:})^\top\|^2 \, ds \bigg)^{p/2} \bigg]$$
$$\leq C_p \bigg(\mathbb{E}\big[|(\xi^\varepsilon)^j - \langle \nabla(\xi^j), h \rangle_H|^p \big] + \mathbb{E}\bigg[\bigg(\int_0^T |\tilde{A}_s^\varepsilon - \langle \nabla(f^j)(s, \cdot, Y_s, Z_s), h \rangle_H| \, ds \bigg)^p \bigg] \bigg)$$
$$+ C_p \mathbb{E}\bigg[\bigg(\int_0^T \|\tilde{A}_s^{y,\varepsilon} - f_y^j(s, \cdot, Y_s, Z_s)\| \|\hat{Y}_s^h\| \, ds \bigg)^p \bigg]$$
$$+ C_p \sum_{k=1}^d \mathbb{E}\bigg[\bigg(\int_0^T \|(f_z^j)^{:,k}(s, \cdot + \varepsilon h, Y_s \circ \tau_{\varepsilon h}, \tilde{Z}_s^k) - (f_z^j)^{:,k}(s, \cdot + \varepsilon h, Y_s, Z_s)\| \|(\hat{Z}_s^h)^{:,k}\| \, ds \bigg)^p \bigg]. \tag{5.6}$$

Since ξ^j is in $\mathbb{D}^{1,2}$, $\lim_{\varepsilon \to 0} \mathbb{E}[|(\xi^\varepsilon)^j - \langle \nabla(\xi^j), h \rangle_H|^p] = 0$ by Lemma 4.2. By Assumption (H_1), the second term in the right-hand side of (5.6) goes to 0 as ε goes to 0. For the last two terms, we will use Assumption (H_2). First, the above estimate implies directly that $(Y^j \circ \tau_{\varepsilon h} - Y^j, (Z^{j,:})^\top \circ \tau_{\varepsilon h} - (Z^{j,:})^\top)_\varepsilon$ goes to 0 in $\mathbb{S}^q \times \mathbb{H}_d^q$ as ε goes to 0 for any $q \in (1, 2)$. We can therefore conclude with Assumption (H_2), together with the fact that $\|f_y^j\|$ is bounded, that by the dominated convergence theorem:

$$\mathbb{E}\bigg[\bigg(\int_0^T \|\tilde{A}_s^{y,\varepsilon} - f_y^j(s, \cdot, Y_s, Z_s)\| \|\hat{Y}_s^h\| \, ds \bigg)^p \bigg]$$
$$\leq C \mathbb{E}\bigg[\bigg(\int_0^T \|\tilde{A}_s^{y,\varepsilon} - f_y^j(s, \cdot, Y_s, Z_s)\|^2 \, ds \bigg)^p \bigg]^{1/2} \mathbb{E}\bigg[\bigg(\int_0^T \|\hat{Y}_s^h\|^2 \, ds \bigg)^p \bigg]^{1/2} \xrightarrow[\varepsilon \to 0]{} 0.$$

We can show similarly that the last term on the right-hand side of (5.6) also goes to 0, by using the fact that for any $j \in \{1, \ldots, n\}$, $((\hat{Z}^h)^{j,:})^\top \in \mathbb{H}_d^2$. It just remains to prove that for any $j \in \{1, \ldots, n\}$, $(\hat{Y}_t^h)^j$ is a random operator on H or equivalently that there exists $\mathcal{D}Y_t^j$ an H-valued r.v. such that $(\hat{Y}_t^h)^j = \langle \mathcal{D}Y_t^j, h \rangle_H$ for any h in H. To this end, let $(h_k)_k$ be an orthonormal system in H, we set for any $j \in \{1, \ldots, n\}$, $\ell \in \{1, \ldots, d\}$:

$$\mathcal{D}Y_t^j := \sum_{k \geq 1} (\hat{Y}_t^{h_k})^j h_k, \qquad \mathcal{D}Z_s^{j,l} := \sum_{k \geq 1} ((\hat{Z}_s^{h_k})^{j,l}) h_k.$$

Note that these elements are well defined, since one can prove that $\mathcal{D}Y_t^j \in L^2(H)$ and that $\mathcal{D}Z^{j,\ell} \in L^2([t, T]; H)$. Indeed using once again a priori estimates for affine BSDEs, there exists $C > 0$ (which may differ from line to line)

such that:

$$\mathbb{E}\left[\sum_{j=1}^{n}\|\mathcal{D}Y_t^j\|_H^2 + \sum_{j=1}^{n}\sum_{\ell=1}^{d}\int_t^T\|\mathcal{D}Z_s^{j,\ell}\|_H^2\,ds\right]$$

$$= \sum_{k\geq 1}\mathbb{E}\left[\sum_{j=1}^{n}|(\hat{Y}_t^{h_k})^j|^2 + \sum_{j=1}^{n}\sum_{\ell=1}^{d}\int_t^T|(\hat{Z}_s^{h_k})^{j,\ell}|^2\,ds\right]$$

$$\leq C\sum_{k\geq 1}\mathbb{E}\left[\sum_{j=1}^{n}|\langle\nabla(\xi^j),h_k\rangle_H|^2 + \sum_{j=1}^{n}\int_t^T|\langle\nabla(f^j)(s,Y_s,Z_s),h_k\rangle_H|^2\,ds\right]$$

$$\leq C\sum_{j=1}^{n}\mathbb{E}\left[\|\nabla(\xi^j)\|_H^2 + \int_t^T\|\nabla(f^j)(s,Y_s,Z_s)\|_H^2\,ds\right] < +\infty, \tag{5.7}$$

by our assumptions on ξ and f. We now identify $(\hat{Y}_t^h)^j$ (respectively $(\hat{Z}_t^h)^{j,\ell}$) with the inner product $\langle\mathcal{D}Y_t^j,h\rangle_H$ (respectively $\langle\mathcal{D}Z_t^{j,\ell},h\rangle_H$). For any $s\geq t$, it holds that:

$$\langle\mathcal{D}Y_s^j,h\rangle_H = \sum_{k\geq 1}\langle\nabla(\xi^j),h_k\rangle_H\langle h_k,h\rangle_H + \sum_{k\geq 1}\langle h_k,h\rangle_H\int_s^T\langle\nabla(f^j)(r,Y_r,Z_r),h_k\rangle_H\,dr$$

$$+ \sum_{k\geq 1}\langle h_k,h\rangle_H\int_s^T\left(f_y^j(r,Y_r,Z_r)\cdot\hat{Y}_r^{h_k} + \sum_{l=1}^{d}(f_z^j)^{:,l}(r,\cdot+\varepsilon h,Y_r,Z_r)\cdot(\hat{Z}_r^{h_k})^{:,l}\right)dr$$

$$+ \sum_{k\geq 1}\langle h_k,h\rangle_H\int_s^T((\hat{Z}_r^{h_k})^{j,:})^\top\cdot dW_r$$

$$= \langle\nabla(\xi^j),h\rangle_H + \int_s^T\left(\langle\nabla(f^j)(r,Y_r,Z_r),h\rangle_H + f_y^j(r,Y_r,Z_r)\cdot\sum_{k\geq 1}\hat{Y}_r^{h_k}\langle h_k,h\rangle_H\right.$$

$$\left. + \sum_{k\geq 1}\langle h_k,h\rangle_H\sum_{l=1}^{d}(f_z^j)^{:,l}(r,\cdot+\varepsilon h,Y_r,Z_r)\cdot(\hat{Z}_r^{h_k})^{:,l}\right)dr$$

$$+ \int_s^T\sum_{k\geq 1}\langle h_k,h\rangle_H((\hat{Z}_r^{h_k})^{j,:})^\top\cdot dW_r,$$

where we justify the exchange between the series and the Riemann integrals by Fubini's theorem. Concerning the Wiener integral we make use of the stochastic Fubini's theorem (see e.g. [26]) since by a priori estimates:

$$\sum_{k\geq 1}\mathbb{E}\left[\int_0^T|\langle h_k,h\rangle_H|\times\|(\hat{Z}^{h_k})_t^{j,:}\|^2\,dt\right]^{1/2}$$

$$\leq C\sum_{k\geq 1}|\langle h_k,h\rangle_H|\mathbb{E}\left[|\langle\nabla(\xi^j),h_k\rangle_H|^2 + \int_0^T|\langle\nabla(f^j)(r,Y_r,Z_r),h_k\rangle_H|^2\,dr\right]^{1/2}$$

$$\leq C\mathbb{E}\left[\|\nabla(\xi^j)\|_H^2 + \int_0^T\|\nabla(f^j)(r,Y_r,Z_r)\|_H^2\,dr\right] < +\infty,$$

where C is a constant which may vary from line to line. By setting

$$\mathcal{D}Y_s^h := (\langle\mathcal{D}Y_s^j,h\rangle_H)_{j\in\{1,\dots,n\}}, \qquad \mathcal{D}Z_s^h := (\langle\mathcal{D}Z_s^{j,l},h\rangle_H)_{1\leq j\leq n,1\leq\ell\leq d},$$

we obtain

$$\langle \mathcal{D}Y_s^j, h \rangle_H = \langle \nabla(\xi^j), h \rangle_H + \int_s^T \left(\langle \nabla(f^j)(r, Y_r, Z_r), h \rangle_H + f_y^j(r, Y_r, Z_r) \cdot \mathcal{D}Y_s^h \right.$$

$$\left. + \sum_{\ell=1}^d (f_z^j)^{:,\ell}(r, \cdot + \varepsilon h, Y_r, Z_r) \cdot (\mathcal{D}Z_r^h)^{:,\ell} \right) dr + \int_s^T ((\mathcal{D}Z_r^h)^{j,:})^\top \cdot dW_r.$$

Thus, by uniqueness of the solution to affine BSDEs with square integrable data, it holds that $(\hat{Y}_t^h)^j = \langle \mathcal{D}Y_t^j, h \rangle_H$ in $L^2(\mathbb{R})$ and $(\hat{Z}^h)^{j,:}\mathbf{1}_{[t,T]} = (\mathcal{D}Z^h)^{j,:}$ in $\mathbb{H}_d^2$ for any h in H. Thus, using Estimate (5.6) we have proved that for any h in H,

$$\lim_{\varepsilon \to 0} \mathbb{E}\big[|(Y_t^\varepsilon)^j - \langle \mathcal{D}Y_t^j, h \rangle_H|^p \big] = 0.$$

Hence by Theorem 4.1, Y_t^j belongs to $\mathbb{D}^{1,p}$ and $\nabla Y_t^j = \mathcal{D}Y_t^j$.

If we set

$$\mathcal{D} \int_t^T (Z_s^{j,:})^\top \cdot dW_s := \sum_{k \geq 1} \int_t^T ((\hat{Z}_s^{h_k})^{j,:})^\top \cdot dW_s h_k,$$

the stochastic Fubini theorem implies that:

$$\left\langle \mathcal{D} \int_t^T (Z_s^{j,:})^\top \cdot dW_s, h \right\rangle_H = \int_t^T ((\mathcal{D}Z_s^h)^{j,:})^\top \cdot dW_s.$$

Moreover, Burkholder–Davis–Gundy's inequality implies that there exists $\tilde{C}_p > 0$ such that

$$\mathbb{E}\bigg[\bigg| \varepsilon^{-1}\bigg(\int_t^T (Z_r^{j,:})^\top \cdot dW_r \circ \tau_{\varepsilon h} - \int_t^T (Z_r^{j,:})^\top \cdot dW_r \bigg) - \bigg\langle \mathcal{D} \int_t^T (Z_r^{j,:})^\top \cdot dW_r, h \bigg\rangle_H \bigg|^p \bigg]$$

$$= \mathbb{E}\bigg[\bigg| \varepsilon^{-1}\bigg(\int_t^T (Z_r^{j,:})^\top \cdot dW_r \circ \tau_{\varepsilon h} - \int_t^T (Z_r^{j,:})^\top \cdot dW_r \bigg) - \int_t^T ((\mathcal{D}Z_r^h)^{j,:})^\top \cdot dW_r \bigg|^p \bigg]$$

$$\leq \mathbb{E}\bigg[\sup_{t \leq s \leq T} \bigg| \int_t^s \big(\varepsilon^{-1}((Z_r^{j,:})^\top \circ \tau_{\varepsilon h} - (Z_r^{j,:})^\top) - ((\mathcal{D}Z_r^h)^{j,:})^\top \big) \cdot dW_r \bigg|^p \bigg]$$

$$\leq \tilde{C}_p \mathbb{E}\bigg[\bigg(\int_t^T \|((Z_r^\varepsilon)^{j,:})^\top - ((\mathcal{D}Z_r^h)^{j,:})^\top\|^2 dr \bigg)^{p/2} \bigg]$$

$$= \tilde{C}_p \mathbb{E}\bigg[\bigg(\int_t^T \|((Z_r^\varepsilon)^{j,:})^\top - ((\hat{Z}_r^h)^{j,:})^\top\|^2 dr \bigg)^{p/2} \bigg],$$

where we have used in the last inequality the fact that $(\hat{Z}^h)^{j,:}\mathbf{1}_{[t,T]} = (\mathcal{D}Z^h)^{j,:}$ in $\mathbb{H}_d^2$ for any h in H.

The right-hand side above tends to 0 as ε goes to 0, once again by (5.6). Therefore,

$$\int_t^T (Z_s^{j,:})^\top \cdot dW_s \in \mathbb{D}^{1,p} \quad \text{and} \quad \nabla \int_t^T (Z_s^{j,:})^\top \cdot dW_s = \mathcal{D} \int_t^T (Z_s^{j,:})^\top \cdot dW_s.$$

Furthermore, by the computations (5.7) we deduce that Y_t^j belongs to $\mathbb{D}^{1,2}$ and that $\int_t^T (Z_s^{j,:})^\top \cdot dW_s$ belong to $\mathbb{D}^{1,2}$ which, by [20, Lemma 2.3] implies that $Z^{j,:}$ belongs to $L^2([t,T]; (\mathbb{D}^{1,2})^d)$. Finally, to match with the notations of the papers [9,20] let us denote by DY^j and $DZ^{j,:}$ the derivatives of respectively ∇Y^j and $\nabla Z^{j,:}$. We also define $D_s Y \in \mathbb{R}^{n \times d}$ by

$$(D_s Y)^{j,\ell} := (D_s Y^j)^\ell, \quad 1 \leq j \leq n, 1 \leq \ell \leq d,$$

and similarly for $D_s \xi$, $D_s f$ and $D_s Z^{\cdot,k}$, $1 \leq k \leq d$. Let us also define for any $1 \leq j \leq n$, $D_s Z^{j,\cdot} \in \mathbb{R}^{d \times n}$ by

$$\left(D_s Z^{j,\cdot}\right)^{k,\ell} := \left(D_s Z^{j,k}\right)^{\ell}, \quad 1 \leq k \leq d, 1 \leq \ell \leq n.$$

We thus obtain using the chain rule formula

$$(D_s Y_t)^{j,\ell} = (D_s \xi)^{j,\ell} + \int_t^T \left((D_s f)^{j,\ell} (r, Y_r, Z_r) + f_y^j (r, Y_r, Z_r) \cdot \left((D_s Y_r)^{j,\cdot} \right)^\top \right) dr$$

$$+ \int_t^T \sum_{k=1}^d (f_z^j)^{\cdot,k} (r, Y_r, Z_r) \cdot \left((D_s Z_r^{k,\cdot})^{j,\cdot} \right)^\top dr - \int_t^T \left(D_s Z_r^{j,\cdot} \right)^{\cdot,\ell} \cdot dW_r, \tag{5.8}$$

which can be interpreted as an affine BSDE. $\qquad\square$

Remark 5.1. *Let $j \in \{1, \dots, n\}$. We would like to point out that since each process $(Z^{j,\cdot})^\top$ is defined as a $\mathfrak{H}$-valued r.v., one may be careful not to study Z directly at a given time, as Z_t is not well defined for a given t. Hence, in the proof we rather study at any time t the random variable $\int_t^T (Z_s^{j,\cdot})^\top \cdot dW_s$ and prove that it belongs to $\mathbb{D}^{1,2}$. Then by* [20, *Lemma* 2.3] *the latter result is equivalent to the fact that $(Z^{j,\cdot})^\top$ belongs to $L^2([t, T]; (\mathbb{D}^{1,2})^d)$.*

Remark 5.2. *We emphasize that our criterion can also be used to study higher-order differentiability properties for (Y, Z). For instance, the pair (DY, DZ) is itself the solution of a (linear) BSDE. Therefore, as long as one is able to derive appropriate a priori estimates for this BSDE, the methodology above can then be applied to obtain conditions ensuring second-order Malliavin differentiability of (Y, Z). Notice nonetheless that when handling higher order derivatives, products of lower order derivatives appear. One may then need to add conditions on the coefficients ensuring strong integrability properties of (Y, Z) and their Malliavin derivatives.*

6. Applications and discussion of the results

For simplicity, in this section we will enforce that $n = d = 1$.

6.1. *Application to FBSDEs*

We consider in this section a FBSDE of the form

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) \, ds + \int_0^t \sigma(s, X_s) \, dW_s, & t \in [0, T], \mathbb{P}_0\text{-a.s.,} \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s, & t \in [0, T], \mathbb{P}_0\text{-a.s.,} \end{cases} \tag{6.1}$$

where $X_0 \in \mathbb{R}$. We make the following assumptions:

(A_1) $b, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ are continuous in time and continuously differentiable in space for any fixed time t and such that there exist $k_b, k_\sigma > 0$ with

$$\left| b_x(t, x) \right| \leq k_b, \qquad \left| \sigma_x(t, x) \right| \leq k_\sigma, \quad \text{for all } x \in \mathbb{R}.$$

Besides $b(t, 0), \sigma(t, 0)$ are bounded functions of t.

(A_2) (i) g is continuously differentiable with polynomial growth.

(ii) $f : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ is continuously differentiable in (x, y, z) with bounded first partial derivatives in y, z uniformly in t, such that $\mathbb{E}[\int_0^T |f(s, 0, 0, 0)|^2 \, ds] < +\infty$ and satisfying for some $C > 0$

$$\exists (q, \kappa) \in \mathbb{R}_+ \times [0, 2), \quad \left| f_x(t, x, y, z) \right| \leq C\left(1 + |y|^\kappa + |z|^\kappa + |x|^q\right), \quad \forall (t, x, y, z) \in [0, T] \times \mathbb{R}^3.$$

Notice that under (A_1) and (A_2), the FBSDE (6.1) admits a unique solution (X, Y, Z) (see [19]). The well-known following lemma provides the existence of a Malliavin derivative for X_t for all $t \in [0, T]$ under Assumption (A_1) (see e.g. [18, Theorem 2.2.1]).

Lemma 6.1. *Under Assumption* (A_1), *for any* $p \geq 1$, $X_t \in \mathbb{D}^{1,p}$ *for all* $t \in [0, T]$, *and* $X \in \mathbb{S}^p$.

The following proposition shows that Assumptions (A_1) and (A_2), which are actually weaker than (PP1) and (PP2), imply our new Assumptions (H_1) and (H_2). As a corollary, using Theorem 5.1, we recover the original result [20, Proposition 2.2].

Proposition 6.1. *Let* (X, Y, Z) *be the unique solution of FBSDE* (6.1). *Under Assumptions* (A_1), (A_2), *Assumptions* (L), (D), (H_1) *and* (H_2) *hold.*

Remark 6.1. *We insist on the fact that in the Markovian case, the original Assumptions* (PP1) *and* (PP2) *of* [20] *imply directly our new Assumptions* (H_1) *and* (H_2) *while Assumptions* (EPQ1) *and* (EPQ2) *of* [9] *are strictly stronger than* (PP1) *and* (PP2). *In other words, in the Markovian case our assumptions are enough to recover the original result* [20, *Proposition* 2.2], *without any additional conditions.*

Proof of Proposition 6.1. From Lemma 6.1, Property (D) holds by the chain rule formula and (L) follows from our assumptions. It remains to prove (H_1) and (H_2). We start with (H_1). Let $1 < p < 2$ and h in H. Below C denotes a positive constant which can differ from line to line. Recall that from our assumptions,

$$\mathbb{E}\left[\sup_{t \in [0,T]} |Y_t|^r + \left(\int_0^T |Z_t|^2 \, dt \right)^{r/2} \right] < \infty, \quad \forall r \leq 2. \tag{6.2}$$

Denoting by $\bar{X}_t$ a random point between X_t and $X_t \circ \tau_{\varepsilon h}$, where we suppressed the dependence on ε for notational simplicity. We have for any t in $[0, T]$, that

$$\mathbb{E}\left[\left| \varepsilon^{-1} \left(f(t, X_t \circ \tau_{\varepsilon h}, Y_t, Z_t) - f(t, X_t, Y_t, Z_t) \right) - f_x(t, X_t, Y_t, Z_t) \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^p \right]$$

$$= \mathbb{E}\left[\left| \frac{X_t \circ \tau_{\varepsilon h} - X_t}{\varepsilon} f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t) \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^p \right]$$

$$\leq C \mathbb{E}\left[\left| \varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^p \left(1 + |Y_t|^{\kappa p} + |Z_t|^{\kappa p} + |X_t|^{pq} + |X_t \circ \tau_{\varepsilon h}|^{pq} \right) \right]$$

$$\quad + C \mathbb{E}\left[\left| f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t) \right|^p \left| \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^p \right]$$

$$\leq C \mathbb{E}\left[\left| \varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^{pr} \right]^{1/r} \mathbb{E}\left[\left(1 + |Y_t|^{\kappa p} + |Z_t|^{\kappa p} + |X_t|^{pq} + |X_t \circ \tau_{\varepsilon h}|^{pq} \right)^{\bar{r}} \right]^{1/\bar{r}}$$

$$\quad + C \mathbb{E}\left[\left| f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t) \right|^p \left| \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^p \right]$$

$$=: A_t^{1,\varepsilon} + A_t^{2,\varepsilon},$$

where $\bar{r} > 1$ and p are chosen so that $p\kappa\bar{r} < 2$ and r denotes the Hölder conjugate of $\bar{r}$. Using the above estimates, we deduce

$$\mathbb{E}\left[\left(\int_0^T \left| \varepsilon^{-1} \left(f(t, X_t \circ \tau_{\varepsilon h}, Y_t, Z_t) - f(t, X_t, Y_t, Z_t) \right) - f_x(t, X_t, Y_t, Z_t) \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right| dt \right)^p \right]$$

$$\leq \int_0^T \left(A_t^{1,\varepsilon} + A_t^{2,\varepsilon} \right) dt.$$

Then, we have

$$\int_0^T A_t^{1,\varepsilon} \, dt \leq C \left(\int_0^T \mathbb{E}\left[\left| \varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h} \rangle_{\mathfrak{H}} \right|^{pr} \right]^{2/r} dt \right)^{1/2}$$

$$\times \left(\int_0^T \mathbb{E}\left[\left(1 + |Y_t|^{\kappa p} + |Z_t|^{\kappa p} + |X_t|^{pq} + |X_t \circ \tau_{\varepsilon h}|^{pq} \right)^{\bar{r}} \right]^{2/\bar{r}} dt \right)^{1/2}. \tag{6.3}$$

In addition by Lemma 3.4, we have that $M^{\varepsilon,h} := X \circ \tau_{\varepsilon h} - X$ is solution to the linear SDE:

$$dM_t^{\varepsilon,h} = M_t^{\varepsilon,h}\big(b_x(t, \underline{X}_t)\, dt + \sigma_x(t, \underline{X}_t)\, dW_t\big) + \varepsilon\sigma(t, X_t \circ \tau_{\varepsilon h})\dot{h}_t\, dt,$$

where $\underline{X}_s$ denotes once again a random point between X_s and $X_s \circ \tau_{\varepsilon h}$. Hence using Assumption (A_1) and standard estimates for SDEs, we get that for any $q \geq 1$,

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\sup_{t \in [0,T]} |X_t \circ \tau_{\varepsilon h} - X_t|^q\Big] = 0.$$

Following the same lines as above, and recalling that $N^h := \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}$ is solution to the SDE:

$$dN_t^h = N_t^h\big(b_x(t, X_t)\, dt + \sigma_x(t, X_t)\, dW_t\big) + \sigma(t, X_t)\dot{h}_t\, dt,$$

we get that the process $P^{\varepsilon,h} := \varepsilon^{-1}(X \circ \tau_{\varepsilon h} - X) - \langle DX, \dot{h}\rangle_{\mathfrak{H}}$ is solution to the affine SDE:

$$dP_t^{\varepsilon,h} = dH_t^{\varepsilon} + P_t^{\varepsilon,h}\big(b_x(t, \underline{X}_t)\, dt + \sigma_x(t, \underline{X}_t)\, dW_t\big),$$

with

$$dH_t^{\varepsilon} := \big(\langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big(b_x(t, \underline{X}_t) - b_x(t, X_t)\big) + \dot{h}_t\big(\sigma(t, X_t \circ \tau_{\varepsilon h}) - \sigma(t, X_t)\big)\big)\, dt$$
$$+ \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big(\sigma_x(t, \underline{X}_t) - \sigma_x(t, X_t)\big)\, dW_t.$$

Using the fact that σ_x, b_x are bounded, σ has linear growth and is continuous, we get by similar computations than those done several times in this paper that:

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\sup_{t \in [0,T]} |H_t^{\varepsilon}|^q\Big] = 0, \quad \forall q \geq 1,$$

from which we deduce using the explicit representation of solutions to affine SDEs (see e.g. [21, Theorem V.9.53]) that

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\sup_{t \in [0,T]} \big|\varepsilon^{-1}(X \circ \tau_{\varepsilon h} - X) - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big|^q\Big] = 0, \quad \forall q \geq 1.$$

As a consequence, combining this estimate with (6.3), we get that:

$$\int_0^T A_t^{1,\varepsilon}\, dt \leq C\Big(\mathbb{E}\Big[\sup_{t \in [0,T]} \big|\varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big|^{pr}\Big]^{2/r}\Big)^{1/2},$$

which goes to 0 as ε goes to 0, since we recall that we have chosen $p, \bar{r} > 1$ so that $\kappa p \bar{r} < 2$, which implies by (6.2), Lemma 6.1 and the Cameron–Martin formula that

$$\int_0^T \mathbb{E}\big[\big(1 + |Y_t|^{\kappa p} + |Z_t|^{\kappa p} + |X_t|^{pq} + |X_t \circ \tau_{\varepsilon h}|^{pq}\big)^{\bar{r}}\big]^{2/\bar{r}}\, dt < \infty.$$

Concerning the term $\int_0^T A_t^{2,\varepsilon}\, dt$, choosing $\tilde{p} > 1$ so that $p\tilde{p} < 2$, it holds by Hölder and by Jensen inequalities that

$$\int_0^T A_t^{2,\varepsilon}\, dt \leq C\Big(\int_0^T \mathbb{E}\big[\big|f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t)\big|^{p\tilde{p}}\big]\, dt\Big)^{1/\tilde{p}},$$

since

$$\mathbb{E}\Big[\sup_{t \in [0,T]} \big|\langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big|^q\Big] < \infty, \quad \forall q > 1.$$

As

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\sup_{t \in [0,T]} |X_t \circ \tau_{\varepsilon h} - X_t|^q \Big] = 0, \quad \forall q \geq 1$$

it holds that

$$\lim_{\varepsilon \to 0} \big| f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t) \big|^{p\tilde{p}} = 0, \quad \mathbb{P}_0 \otimes dt\text{-a.e.}$$

Furthermore, for any $2 > \rho > 1$,

$$\sup_{\varepsilon \in (0,1)} \int_0^T \mathbb{E}\big[\big| f_x(t, \bar{X}_t, Y_t, Z_t) - f_x(t, X_t, Y_t, Z_t) \big|^{\rho p \tilde{p}} \big] dt$$

$$\leq C \sup_{\varepsilon \in (0,1)} \int_0^T \mathbb{E}\big[\big(1 + |X_t|^q + |X_t \circ \tau_{\varepsilon h}|^q + |Y_t|^\kappa + |Z_t|^\kappa \big)^{\rho p \tilde{p}} \big] dt < \infty,$$

by choosing p small enough so that $\kappa \rho p \tilde{p} \leq 2$. So by Lebesgue's dominated convergence theorem,

$$\lim_{\varepsilon \to 0} \int_0^T A_t^{2,\varepsilon} \, dt = 0,$$

which proves (H_1). Concerning, (H_2) we just mention that f_y (respectively f_z) is bounded, jointly continuous in (x, y, z) and we make use of Lemma 3.2. □

6.2. *Affine BSDEs*

The aim of this section is to prove that with our condition, we can provide weaker conditions compared to [9] for affine BSDEs. We take a driver of the form

$$f(t, \omega, y, z) := \alpha_t(\omega) + \beta_t(\omega)y + \gamma_t(\omega)z$$

with bounded $\mathbb{F}$-progressively measurable processes such that $\alpha, \beta, \gamma \in L^2([0,T]; \mathbb{D}^{1,2})$, and ξ in $\mathbb{D}^{1,2}$. The conditions given in [9, Proposition 5.3] for proving that the associated solution (Y, Z) is Malliavin differentiable read as follows (together with some measurability conditions):

$$\exists \eta > 0 \text{ such that } \quad \mathbb{E}\big[|\xi|^{2+\eta} \big] < +\infty \quad and \quad \int_0^T \mathbb{E}\Bigg[\bigg(\int_\theta^T \big| K_\theta(s) \big|^2 ds \bigg)^{2+\eta} \Bigg]^{1/(2+\eta)} d\theta < +\infty, \tag{6.4}$$

with $K_\theta(s) := |D_\theta \beta(s)| + |D_\theta \gamma(s)|$.

In our setting, one needs to check Assumptions (L), (D), (H_1) and (H_2). As mentioned below by Lemma 3.2 Condition (H_2) comes for free, and Assumptions (D) and (L) are also trivially satisfied. The interesting point is that (H_1) is true as soon as (6.4) is replaced with:

$$\exists \eta > 0 \text{ such that } \quad \lim_{\varepsilon \to 0} \int_0^T \mathbb{E}\big[\big| \varepsilon^{-1}(\mu_t \circ \tau_{\varepsilon h} - \mu_t) - \langle D\mu_t, \dot{h} \rangle_{\mathfrak{H}} \big|^{2+\eta} \big] dt = 0, \quad \text{for } \mu \in \{\beta, \gamma\}. \tag{6.5}$$

Hence our condition only involves a condition on γ and β and not on ξ. For instance if β and γ are given as:

$$\beta_t = \varphi_1(X_t), \qquad \gamma_t := \varphi_2(X_t), \quad t \in [0,T],$$

with φ_1, φ_2 two smooth functions with polynomial growth and X is the solution to an SDE of the form of the one considered in Section 6.1, then the requirements of Conditions (6.4) and (6.5) are satisfied for β and γ, however in contradistinction to Condition (6.4), Assumption (6.5) does not put extra regularity on the terminal condition ξ.

We make precise our result.

Proposition 6.2. *Let ξ in $\mathbb{D}^{1,2}$, and α, β, γ bounded $\mathbb{F}$-progressively processes in $L^2([0, T]; \mathbb{D}^{1,2})$ such that $D\alpha$, $D\beta$ and $D\gamma$ are $\mathbb{F}$-progressively measurable. Assume that Assumption (6.5) is in force. Then for any t in $[0, T]$, Y_t belongs to $\mathbb{D}^{1,2}$, $Z \in L^2([t, T]; \mathbb{D}^{1,2})$ where (Y, Z) is the unique solution in $\mathbb{S}^2 \times \mathbb{H}^2$ to the affine BSDE:*

$$Y_t = \xi + \int_t^T (\alpha_s + \beta_s Y_s + \gamma_s Z_s) \, ds - \int_t^T Z_s \, dW_s, \quad t \in [0, T].$$

Proof. Once again we check that assumptions of Theorem 4.1 are in force. Properties (D) and (L) are immediately satisfied. Let $f(t, \omega, y, z) := \alpha_t(\omega) + \beta_t(\omega)y + \gamma_t(\omega)z$. Since $f_y(t, \omega, y, z) = \beta_t(\omega)$, and $f_z(t, \omega, y, z) = \gamma_t(\omega)$ we immediately get by Lemma 3.2 and since β, γ are bounded that (H_2) is satisfied. Concerning (H_1), we have for any $1 < p < 2$ and h in H, that

$$\mathbb{E}\left[\left(\int_0^T \left|\varepsilon^{-1}\left(f(t, \cdot + \varepsilon h, Y_t, Z_t) - f(t, \cdot, Y_t, Z_t)\right) - \langle Df(t, \cdot, Y_t, Z_t), \dot{h}\rangle_{\mathfrak{H}}\right| dt\right)^p\right]$$

$$\leq C\mathbb{E}\left[\left(\int_0^T \left|\varepsilon^{-1}(\alpha_t \circ \tau_{\varepsilon h} - \alpha_t) - \langle D\alpha_t, \dot{h}\rangle_{\mathfrak{H}}\right| dt\right)^p\right]$$

$$+ C\mathbb{E}\left[\left(\int_0^T \left|Y_t\left(\varepsilon^{-1}(\beta_t \circ \tau_{\varepsilon h} - \beta_t) - \langle D\beta_t, \dot{h}\rangle_{\mathfrak{H}}\right)\right| dt\right)^p\right]$$

$$+ C\mathbb{E}\left[\left(\int_0^T \left|Z_t\left(\varepsilon^{-1}(\gamma_t \circ \tau_{\varepsilon h} - \gamma_t) - \langle D\gamma_t, \dot{h}\rangle_{\mathfrak{H}}\right)\right| dt\right)^p\right]$$

$$=: A_1^\varepsilon + A_2^\varepsilon + A_3^\varepsilon, \tag{6.6}$$

where C is a constant. By Lemma A.2 we have that $\lim_{\varepsilon \to 0} A_1^\varepsilon = 0$. We consider the term A_3^ε. We have that:

$$A_3^\varepsilon \leq \mathbb{E}\left[\int_0^T \left|Z_t\left(\varepsilon^{-1}(\gamma_t \circ \tau_{\varepsilon h} - \gamma_t) - \langle D\gamma_t, \dot{h}\rangle_{\mathfrak{H}}\right)\right|^p dt\right]$$

$$\leq C\int_0^T \mathbb{E}[|Z_t|^2]^{p/2}\mathbb{E}\left[\left|\varepsilon^{-1}(\gamma_t \circ \tau_{\varepsilon h} - \gamma_t) - \langle D\gamma_t, \dot{h}\rangle_{\mathfrak{H}}\right|^{2p/(2-p)}\right]^{(2-p)/2} dt$$

$$\leq C\left(\int_0^T \mathbb{E}[|Z_t|^2] \, dt\right)^{p/2}\left(\int_0^T \mathbb{E}\left[\left|\varepsilon^{-1}(\gamma_t \circ \tau_{\varepsilon h} - \gamma_t) - \langle D\gamma_t, \dot{h}\rangle_{\mathfrak{H}}\right|^{2p/(2-p)}\right] dt\right)^{(2-p)/2}.$$

Choosing p such that $\frac{2p}{2-p} = 2 + \eta$ we get that A_3^ε converges to 0 as ε tends to 0 by (6.5). Similarly, $\lim_{\varepsilon \to 0} A_2^\varepsilon = 0$ for this choice of p. □

Remark 6.2. *Note that, since the BSDE is affine, Y_t can be expressed explicitly as*:

$$Y_t = \mathbb{E}\left[M_{t,T}\xi - \int_t^T M_{t,s}\alpha_s \, ds \,\Big|\, \mathcal{F}_t\right],$$

where

$$M_{t,s} := \exp\left(\int_t^s \gamma_u \, dW_u - \frac{1}{2}\int_t^s |\gamma_u|^2 \, du + \int_t^s \beta_u \, du\right), \quad s \in [t, T].$$

Hence, on the one hand, Y_t belongs to $\mathbb{D}^{1,2}$ if and only if the coefficients α, β, γ belong to $L^2([0, T]; \mathbb{D}^{1,2})$ and ξ is in $\mathbb{D}^{1,2}$. The same conclusion follows for the Z component. Hence, neither our Condition (6.5) nor the one of [9], namely (6.4), are sharp. However, both are sharp in the case where $\beta = \gamma = 0$. On the other hand, Condition (6.4) or (6.5) give more information than the simple fact that Y, Z are Malliavin differentiable, since they imply that the BSDE solved by (DY, DZ) is limit in $\mathbb{S}^2 \times \mathbb{H}^2$ of respectively (DY^n, DZ^n) (where (Y^n, Z^n) is the solution to the Picard iteration equation at order n approximating (Y, Z)) for (6.4), and of the difference quotient $(\varepsilon^{-1}(Y \circ \tau_{\varepsilon h} - Y), \varepsilon^{-1}(Z \circ \tau_{\varepsilon h} - Z))$ in our case (6.5).

6.3. Discussion and comparison of the results

We would like before going to the quadratic BSDE case to make a comment about the difference between our approach and the one of [9,20] and our approach. In these references, the authors consider the sequence of BSDEs:

$$Y_t^n = \xi + \int_t^T f\left(s, Y_s^{n-1}, Z_s^{n-1}\right) ds - \int_t^T Z_s^n \, dW_s, \quad t \in [0, T],$$

which approximate in $\mathbb{S}^2 \times \mathbb{H}^2$ the solution to the original BSDE:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s, \quad t \in [0, T].$$

Now, under mild assumptions on f, the processes (Y^n, Z^n) are Malliavin differentiable and it holds that a version of $(D_r Y_t^n, D_r Z_t^n)$ satisfies for $t \in [0, T], r \le t$:

$$D_r Y_t^n = D_r \xi + \int_t^T \left[D_r f\left(s, \Theta_s^{n-1}\right) + \partial_y f\left(s, \Theta_s^{n-1}\right) D_r Y_s^{n-1} + \partial_z f\left(s, \Theta_s^{n-1}\right) D_r Z_s^{n-1} \right] ds$$

$$- \int_t^T D_r Z_s^n \, dW_s,$$

with $\Theta_s^{n-1} := (Y_s^{n-1}, Z_s^{n-1})$. On the other if (Y, Z) where Malliavin differentiable we would have that a version of $(D_r Y_t, D_r Z_t)$ would satisfy for $t \in [0, T], r \le t$:

$$D_r Y_t = D_r \xi + \int_t^T \left[D_r f(s, Y_s, Z_s) + \partial_y f(s, Y_s, Z_s) D_r Y_s + \partial_z f(s, Y_s, Z_s) D_r Z_s \right] ds - \int_t^T D_r Z_s \, dW_s.$$

In other words, assuming that $\partial_y f$ and $\partial_z f$ to be continuous, we would get formally that (DY^n, DZ^n) converges to (DY, DZ) (in $\mathbb{S}^2 \times \mathbb{H}^2$) as n goes to infinity provided that at the limit one can replace $D_r f(s, Y_s^{n-1}, Z_s^{n-1})$ by $D_r f(s, Y_s, Z_s)$ which is exactly where comes the main assumption in [9,20] which impose $D_r f$ to be (stochastic) Lipschitz continuous in (y, z) with integrability conditions on the Lipschitz constant to make the aforementioned argument rigorous. However, it is not a necessary condition for (Y, Z) to be Malliavin differentiable that (DY^n, DZ^n) to converge to (DY, DZ). However, for Y_t to be in $\mathbb{D}^{1,2}$, it is necessary (and sufficient) that $\varepsilon^{-1}(Y_t \circ \tau_{\varepsilon h} - Y_t)$ converges in L^p for some $p < 2$ to $\langle DY_t, \dot{h} \rangle_{\mathfrak{H}}$ for any h in H (according to Theorem 4.1). Hence, this is an advantage of our conditions.

7. Extension to quadratic growth BSDEs

The aim of this section is to extend our previous results to so-called quadratic growth BSDEs. Some results for these equations already exist in the literature, see in particular [2,12] or the thesis [8], however they are generally limited to specific forms of the generators or to a Markovian setting. We will show that our approach to the Malliavin differentiability is flexible enough to be able to treat this problem without major modifications to our proofs. Since the wellposedness theory for multidimensional quadratic BSDEs is still an open problem, we enforce $n = 1$ throughout this section.

We will now list our assumptions in this quadratic setting

(D$_\infty$) ξ is bounded, belongs to $\mathbb{D}^{1,\infty}$ and its Malliavin derivative $D\xi$ is bounded, for any $(y, z) \in \mathbb{R} \times \mathbb{R}^d$, $(t, \omega) \mapsto f(t, \omega, y, z)$ is in $L^2([0, T]; \mathbb{D}^{1,\infty})$, $f(\cdot, y, z)$ and $Df(\cdot, y, z)$ are $\mathbb{F}$-progressively measurable, $Df(\cdot, y, z)$ is uniformly bounded in y, z.

(Q) The map $(y, z) \mapsto f(\cdot, y, z)$ is continuously differentiable and there exists some constant $C > 0$ such that for any $(s, \omega, y, z, z') \in [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$

$$\left| f\left(s, \omega, y, z\right) - f\left(s, \omega, y, z'\right) \right| \le C\left(1 + \|z\| + \|z'\|\right)\|z - z'\|, \qquad \left| f(s, \omega, 0, 0) \right| \le C,$$

$$\left| f_y(s, \omega, y, z) \right| \le C, \qquad \left\| f_z(s, \omega, y, z) \right\| \le C\left(1 + \|z\|\right),$$

where $f_z = (\frac{\partial f}{\partial z_l})_{l \in \{1,...,d\}}$ denotes the gradient of f with respect to the z variable.

$(H_{1,\infty})$ For any $p > 1$ and for any $h \in H$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left(\int_0^T \left| \frac{f(s, \cdot + \varepsilon h, Y_s, Z_s) - f(s, \cdot, Y_s, Z_s)}{\varepsilon} - \langle Df(s, \cdot, Y_s, Z_s), \dot{h}\rangle_{\mathfrak{H}} \right| ds\right)^p\right] = 0.$$

$(H_{2,\infty})$ Let $(\varepsilon_k)_{k \in \mathbb{N}}$ be a sequence in $(0, 1]$ such that $\lim_{k \to +\infty} \varepsilon_k = 0$, and let $(Y^k, Z^k)_k$ be a sequence of random variables which converges in $\mathbb{S}^p \times \mathbb{H}_d^p$ for any $p > 1$ to some (Y, Z). Then for all $h \in H$, the following convergences hold in probability

$$\left| f_y(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot) - f_y(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T];\mathbb{R})} \underset{k \to +\infty}{\longrightarrow} 0,$$

$$\left| f_z(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot^k) - f_z(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T];\mathbb{R}^d)} \underset{n \to +\infty}{\longrightarrow} 0,$$

(7.1)

or

$$\left| f_y(\cdot, \omega + \varepsilon_k h, Y_\cdot^k, Z_\cdot^k) - f_y(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T];\mathbb{R})} \underset{k \to +\infty}{\longrightarrow} 0,$$

$$\left| f_z(\cdot, \omega + \varepsilon_k h, Y_\cdot, Z_\cdot^k) - f_z(\cdot, \omega, Y_\cdot, Z_\cdot) \right|_{L^2([0,T];\mathbb{R}^d)} \underset{k \to +\infty}{\longrightarrow} 0.$$

(7.2)

Let $\mathbb{S}^\infty$ be the set of $\mathbb{F}$-progressively measurable processes Y such that $\sup_{t \in [0,T]} |Y_t|$ is bounded and $\mathbb{H}_{\text{BMO}}^2$ the set of $\mathbb{R}^d$-valued predictable processes Z such that:

$$\underset{\tau \in \mathcal{T}}{\text{essup}}\, \mathbb{E}\left[\int_\tau^T \|Z_s\|^2\, ds \,\Big|\, \mathcal{F}_\tau\right] < +\infty, \quad \mathbb{P}_0\text{-a.s.},$$

where $\mathcal{T}$ denotes the set of $\mathbb{F}$-stopping times with values in $[0, T]$. We start by recalling the following by now classical results on quadratic growth BSDEs and stochastic Lipschitz BSDEs, which can be found among others in [12].

Proposition 7.1. *Under Assumptions* (D_∞) *and* (Q), *the BSDEs* (5.1) *and* (5.5) *both admit a unique solution in* $\mathbb{S}^\infty \times \mathbb{H}_{\text{BMO}}^2$.

We have the following extension of Theorem 5.1.

Theorem 7.1. *Let t be in* $[0, T]$. *Under Assumptions* (D_∞), (Q), $(H_{1,\infty})$ *and* $(H_{2,\infty})$, Y_t *belongs to* $\mathbb{D}^{1,\infty}$ *and* $Z \in L^\infty([t, T]; (\mathbb{D}^{1,2})^d)$.

Proof. We follow the proof of Theorem (5.1), using the same notations. Since the BSDEs are now quadratic, we can use the a priori estimates of Lemma A.1 in [12] to obtain that for any $p > 1$, there exists some $q > 1$ such that

$$\mathbb{E}\left[\sup_{s \in [t,T]} |Y_s^\varepsilon - \hat{Y}_s^h|^{2p}\right] + \mathbb{E}\left[\left(\int_t^T \|Z_s^\varepsilon - \hat{Z}_s^h\|^2\, ds\right)^p\right]$$

$$\leq C_p\left(\mathbb{E}[|\xi^\varepsilon - \langle D\xi, \dot{h}\rangle_{\mathfrak{H}}|^{pq}]^{1/q} + \mathbb{E}\left[\left(\int_0^T \tilde{A}_s^\varepsilon - \langle Df(s, \cdot, Y_s, Z_s), \dot{h}\rangle_{\mathfrak{H}}\, ds\right)^{pq}\right]^{1/q}\right)$$

$$+ C_p \mathbb{E}\left[\left(\int_0^T |\tilde{A}_s^{y,\varepsilon} - f_y(s, \cdot, Y_s, Z_s)| |\hat{Y}_s^h|\, ds\right)^{pq}\right]^{1/q}$$

$$+ C_p \mathbb{E}\left[\left(\int_0^T \|\tilde{A}_s^{z,\varepsilon} - f_z(s, \cdot, Y_s, Z_s)\| \|\hat{Z}_s^h\|\, ds\right)^{pq}\right]^{1/q},$$

(7.3)

where we set

$$\tilde{A}_s^{y,\varepsilon} := f_y\big(s, \cdot + \varepsilon h, \bar{Y}_s^{\varepsilon,h}, Z_s\big),$$

$$\tilde{A}_s^{\varepsilon} := \frac{1}{\varepsilon}\big(f(r, \cdot + \varepsilon h, Y_s, Z_s) - f(s, \cdot, Y_s, Z_s)\big),$$

where $\bar{Y}_r^{\varepsilon,h}$ is a convex combination of Y_r and $Y_r \circ \tau_{\varepsilon h}$ and

$$\tilde{A}_s^{z,\varepsilon} := \big((\tilde{A}_s^{z,\varepsilon})^k\big)_{k \in \{1,\dots,d\}}$$

with $(\tilde{A}_s^{z,\varepsilon})^k := f_z(s, \cdot + \varepsilon h, Y_s \circ \tau_{\varepsilon h}, \tilde{Z}_r^k)$, and $\tilde{Z}_s^k$ is as in the proof of Theorem 5.1.

Since $\xi \in \mathbb{D}^{1,\infty}$, the first term on the right-hand side above goes to 0 thanks to Theorem 4.1. Moreover, the second term also goes to 0 thanks to Assumption (H_∞). Then, since f_y is bounded by Assumption (Q) and since $\tilde{Y}^h \in \mathbb{S}^\infty$ by Proposition 7.1, we can easily conclude with Assumption (H_2) and the dominated convergence theorem that the third term on the right-hand side also goes to 0. Let us now concentrate on the fourth term involving the control variable. By Cauchy–Schwarz inequality we have that

$$\mathbb{E}\left[\left(\int_0^T \big\| \tilde{A}_s^{z,\varepsilon} - f_z(s, \cdot, Y_s, Z_s)\big\| \big\|\hat{Z}_s^h\big\| \, ds\right)^{pq}\right]$$

$$\leq \mathbb{E}\left[\left(\int_0^T \big\| \tilde{A}_s^{z,\varepsilon} - f_z(s, \cdot, Y_s, Z_s)\big\|^2 \, ds\right)^{pq}\right]^{1/2} \mathbb{E}\left[\left(\int_0^T \big\|\tilde{Z}_s^h\big\|^2 \, ds\right)^{pq}\right]^{1/2}. \tag{7.4}$$

Since $(\hat{Y}^h, \hat{Z}^h)$ is the solution to the stochastic linear BSDE (5.5) with bounded coefficients Df and f_y (by (D_∞)) and $f_z(s, Y_s, Z_s)$ is in $\mathbb{H}^2_{\mathrm{BMO}}$ since $\|f_z(s, Y_s, Z_s)\| \leq C(1 + \|Z_s\|)$ (by Assumption (Q)), we deduce that $\hat{Z}^h \in \mathbb{H}^2_{\mathrm{BMO}}$ which implies that $\hat{Z}^h \in \mathbb{H}^m_d$ for any $m > 1$ by the energy inequalities. Furthermore, for any $\eta > 0$ it holds that

$$\mathbb{E}\left[\left(\int_0^T \big\| \tilde{A}_s^{z,\varepsilon} - f_z(s, \cdot, Y_s, Z_s)\big\| \big\|\hat{Z}_s^h\big\| \, ds\right)^{pq+\eta}\right]$$

$$\leq C\mathbb{E}\left[\left(\int_0^T \big(1 + \|Z_s\| + \|Z_s \circ \tau_{\varepsilon h}\|\big)\big\|\hat{Z}_s^h\big\| \, ds\right)^{pq+\eta}\right]$$

$$\leq C\mathbb{E}\left[\left(\int_0^T \big(1 + \|Z_s\| + \|Z_s \circ \tau_{\varepsilon h}\|\big)^2 \, ds\right)^{(pq+\eta)/2}\left(\int_0^T \big\|\hat{Z}_s^h\big\|^2 \, ds\right)^{(pq+\eta)/2}\right]$$

$$\leq C\mathbb{E}\left[\left(\int_0^T \big(1 + \|Z_s\| + \|Z_s \circ \tau_{\varepsilon h}\|\big)^2 \, ds\right)^{pq+\eta}\right]^{1/2}\mathbb{E}\left[\left(\int_0^T \big\|\hat{Z}_s^h\big\|^2 \, ds\right)^{pq+\eta}\right]^{1/2}$$

$$\leq C\left(1 + \mathbb{E}\left[\left(\int_0^T \|Z_s\|^2 \, ds\right)^{p'}\right]^{1/q'}\right)\mathbb{E}\left[\left(\int_0^T \big\|\hat{Z}_s^h\big\|^2 \, ds\right)^{pq+\eta}\right]^{1/2} < +\infty,$$

where $p', q' > 1$ using Hölder inequality and Proposition 3.1. Hence, taking limit as ε goes to 0 in (7.4) we get that $\lim_{\varepsilon \to 0} \mathbb{E}[\sup_{s \in [t,T]} |Y_s^\varepsilon - \hat{Y}_s^h|^{2p}] + \mathbb{E}[(\int_t^T \|Z_s^\varepsilon - \hat{Z}_s^h\|^2 \, ds)^p] = 0$. Following the same lines as in the proof of Theorem 5.1, one can use a priori estimates for quadratic growth BSDEs to obtain that $\hat{Y}^h$ and $\hat{Z}^h$ are linear operators-valued r.v. This proves that Y_t and $\int_t^T Z_s^\top \cdot dW_s$ belongs to $\mathbb{D}^{1,\infty}$ by Theorem 4.1. In particular, $Z\mathbf{1}_{[t,T]}$ belongs to $L^2([t,T]; (\mathbb{D}^{1,2})^d)$ (see [20]). Moreover, since $(D_t Y, D_t Z)$ is the solution of the stochastic linear BSDE (5.8) for any $t \in [0,T]$ and Assumptions (D_∞) and (Q) hold, from the relation $(D_t Y_t)^j = (Z_t)^j$ for any j in $\{1,\dots,d\}$, and for all $t \in [0,T]$ we obtain $Z\mathbf{1}_{[t,T]} \in L^\infty([t,T]; (\mathbb{D}^{1,2})^d)$. $\square$

Remark 7.1. *We would like to point out that our conditions cover the case of Markovian quadratic BSDEs presented in* [11, *Theorem* 2.9]. *Indeed, assume that we consider a forward-backward system of the form* (6.1) *where the solution*

process X to the forward SDE is m-dimensional with m a positive integer (*so that we match with the notations and assumptions of* [11, *Theorem* 2.9]) *under assumptions,* (D_∞), (Q), (A_1), (A_2)(i) *and where* (A_2)(ii) *is replaced by the following assumption*:

(A_2)(ii$'$) $f : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ *is continuously differentiable in* (x, y, z) *and satisfying for some* $C > 0$

$$\exists q \in \mathbb{R}_+, \quad \|f_x(t, x, y, z)\| \le C\big(1 + |y| + \|z\|^2 + \|x\|^q\big), \quad \forall (t, x, y, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d,$$

where $f_x := (\frac{\partial f}{\partial x_l})_{l \in \{1,\dots,m\}}$ *denotes the gradient of* f *with respect to the variable* x. *Under these assumptions, we can check that* $(H_{1,\infty})$ *and* $(H_{2,\infty})$ *are in force. To see this we just make a comment about how the proof of Proposition* 6.1 *has to be modified to obtain* $(H_{1,\infty})$, *whereas* $(H_{2,\infty})$ *is met trivially. Using the notations of this proof one can manage a term of the form*:

$$\mathbb{E}\left[\left(\int_0^T \big\|\varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big\| \big\|f_x(t, \bar{X}_t, Y_t, Z_t)\big\| dt\right)^p\right]$$

as follows:

$$\mathbb{E}\left[\left(\int_0^T \big\|\varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big\| \big\|f_x(t, \bar{X}_t, Y_t, Z_t)\big\| dt\right)^p\right]$$

$$\le C\mathbb{E}\left[\sup_{t \in [0,T]}\left\|\frac{X_t \circ \tau_{\varepsilon h} - X_t}{\varepsilon} - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\right\|^p\right.$$

$$\left. \times \left(\int_0^T \big(1 + \|X_t\|^q + \|X_t \circ \tau_{\varepsilon h}\|^q + |Y_t| + \|Z_t\|^2\big) dt\right)^p\right]$$

$$\le C\mathbb{E}\left[\sup_{t \in [0,T]} \big\|\varepsilon^{-1}(X_t \circ \tau_{\varepsilon h} - X_t) - \langle DX_t, \dot{h}\rangle_{\mathfrak{H}}\big\|^{2p}\right]^{1/2}$$

$$\times \mathbb{E}\left[\left(\int_0^T \big(1 + \|X_t\|^q + \|X_t \circ \tau_{\varepsilon h}\|^q + |Y_t| + \|Z_t\|^2\big) dt\right)^{2p}\right]^{1/2},$$

which goes to 0 *as* ε *goes to* 0 *since* Z *belongs to* $\mathbb{H}^2_{\text{BMO}}$ *and since* Y *is bounded. The term involving* $A^{2,\varepsilon}$ *can be treated similarly.*

Appendix

The following lemma was remarked in [24, Remark 2] with the set of polynomial cylindrical functions $\mathcal{P}$, we provide a proof of it with the set of cylindrical functions $\mathcal{S}$.

Lemma A.1. *Let* $p > 1$ *and* F *be in* $L^p(\mathbb{R})$, $G \in \mathcal{S}$ *and* $h \in H$. *The mapping* $\varepsilon \mapsto \mathbb{E}[F \circ \tau_{\varepsilon h}G]$ *is differentiable in* ε *and*

$$\frac{d}{d\varepsilon}\mathbb{E}[F \circ \tau_{\varepsilon h}G] = \mathbb{E}\big[F \circ \tau_{\varepsilon h}\delta(Gh)\big]. \tag{A.1}$$

Proof. Let $\eta > 0$, by the Cameron–Martin formula, we have that

$$\eta^{-1}\big(\mathbb{E}[F \circ \tau_{(\eta+\varepsilon)h}G] - \mathbb{E}[F \circ \tau_{\varepsilon h}G]\big)$$

$$= \mathbb{E}\left[F \circ \tau_{\varepsilon h}\frac{G \circ \tau_{-\eta h}\exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2)\int_0^T \|\dot{h}(u)\|^2 du) - G}{\eta}\right].$$

Hence

$$\lim_{\eta \to 0} \eta^{-1} \big(\mathbb{E}[F \circ \tau_{(\eta+\varepsilon)h} G] - \mathbb{E}[F \circ \tau_{\varepsilon h} G] \big)$$

$$= \mathbb{E}\bigg[F \circ \tau_{\varepsilon h} \lim_{\eta \to 0} \frac{G \circ \tau_{-\eta h} \exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2) \int_0^T \|\dot{h}(u)\|^2 \, du) - G}{s} \bigg]$$

$$= \mathbb{E}\bigg[F \circ \tau_{\varepsilon h} \lim_{\eta \to 0} \bigg(\frac{G \circ \tau_{-\eta h} - G}{\eta} + G \circ \tau_{-\eta h} \frac{\exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2) \int_0^T \|\dot{h}(u)\|^2 \, du) - 1}{\eta} \bigg) \bigg],$$

where the exchange between the limit and the expectation is justified by the fact that

$$\sup_{\eta \in (0,1]} \eta^{-q} \mathbb{E}\bigg[\bigg| G \circ \tau_{-\eta h} \exp\bigg(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - \frac{|\eta|^2}{2} \int_0^T \|\dot{h}(u)\|^2 \, du \bigg) - G \bigg|^q \bigg] < +\infty \tag{A.2}$$

for any $q > 1$ and by the Cameron–Martin formula. Indeed for any r in $(1, p)$ we have by Hölder inequality:

$$\mathbb{E}\bigg[\bigg| F \circ \tau_{\varepsilon h} \frac{G \circ \tau_{-\eta h} \exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2) \int_0^T \|\dot{h}(u)\|^2 \, du) - G}{\eta} \bigg|^r \bigg]$$

$$\leq \underbrace{\mathbb{E}\big[|F \circ \tau_{\varepsilon h}|^{p_1} \big]^{r/p_1}}_{=:E_1} \underbrace{\mathbb{E}\bigg[\bigg| \frac{G \circ \tau_{-\eta h} \exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2) \int_0^T \|\dot{h}(u)\|^2 \, du) - G}{\eta} \bigg|^{r p_2} \bigg]^{1/p_2}}_{=:E_2},$$

where $r < p_1 < p$ and p_2 is the Hölder conjugate of p_1/r. Using Cameron–Martin formula for E_1, Relation (3.1) and Hölder inequality with $r_1 = \frac{p}{p_1}$ and r_2 such that $\frac{1}{r_1} + \frac{1}{r_2} = 1$, we deduce that:

$$E_1 \leq \mathbb{E}\big[|F|^p \big]^{r/p} \mathbb{E}\bigg[\bigg| \mathcal{E}\bigg(\int_0^T \dot{h}_s^\top \cdot dW_s \bigg) \bigg|^{r_2} \bigg]^{1/r_2} < +\infty.$$

We now turn to E_2, for any $q > 1$

$$\sup_{\eta \in (0,1]} \mathbb{E}\bigg[\bigg| \frac{G \circ \tau_{-\eta h} \mathcal{E}(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u) - G}{\eta} \bigg|^q \bigg]$$

$$\leq \underbrace{\sup_{\eta \in (0,1]} \mathbb{E}\bigg[\bigg| \frac{G \circ \tau_{-\eta h} - G}{\eta} \mathcal{E}\bigg(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u \bigg) \bigg|^q \bigg]}_{=:A_1} + \underbrace{\sup_{\eta \in (0,1]} \mathbb{E}\bigg[\bigg| \frac{\mathcal{E}(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u) - 1}{\eta} G \bigg|^q \bigg]}_{=:A_2},$$

hence, on the one hand there exists $\alpha_1, \alpha_2 > 1$ such that:

$$A_1 \leq \sup_{\eta \in (0,1]} \eta^{-q} \mathbb{E}\big[|G \circ \tau_{-\eta h} - G|^{q\alpha_1} \big]^{1/\alpha_1} \mathbb{E}\bigg[\bigg| \mathcal{E}\bigg(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u \bigg) \bigg|^{q\alpha_2} \bigg]^{1/\alpha_2} < +\infty,$$

using the fact that G is polynomial, so G is locally Lipschitz and we conclude by Relation (3.1). On the other hand, using the mean value theorem and Relation (3.1), we obtain also $A_2 < +\infty$. We deduce that Relation (A.2) holds. Moreover, given that $G \in \mathcal{P}$ is polynomial, we deduce that $\frac{G \circ \tau_{-\eta h} - G}{\eta} \to_{\eta \to 0} -\langle \nabla G, h \rangle_H$ a.s. Hence,

$$\lim_{\eta \to 0} \eta^{-1} \big(\mathbb{E}[F \circ \tau_{(\eta+\varepsilon)h} G] - \mathbb{E}[F \circ \tau_{\varepsilon h} G] \big)$$

$$= \mathbb{E}\bigg[F \circ \tau_{\varepsilon h} \lim_{\eta \to 0} \bigg(\frac{G \circ \tau_{-\eta h} - G}{\eta} + G \circ \tau_{-\eta h} \frac{\exp(\eta \int_0^T \dot{h}(u)^\top \cdot dW_u - (|\eta|^2/2) \int_0^T \|\dot{h}(u)\|^2 \, du) - 1}{\eta} \bigg) \bigg]$$

$$= \mathbb{E}\big[F \circ \tau_{\varepsilon h} \big(-\langle \nabla G, h \rangle_H + G \delta(h) \big) \big] = \mathbb{E}\big[F \circ \tau_{\varepsilon h} \delta(Gh) \big],$$

by (2.3), so (A.1) holds. $\qquad\square$

Lemma A.2. *Let α in $L^2([0, T]; \mathbb{D}^{1,2})$. Then for any p in $(1, 2)$,*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\int_0^T \left|\frac{\alpha_s \circ \tau_{\varepsilon h} - \alpha_s}{\varepsilon} - \langle \nabla \alpha_s, h\rangle_H\right|^p ds\right] = 0.$$

Proof. Note first that the space $L^2([0, T]; \mathbb{D}^{1,2})$ can be identified with the space $\mathbb{D}^{1,2}(\mathfrak{H})$ which is the completion of the set of $\mathfrak{H}$-valued r.v. of the form:

$$\sum_{i=1}^n F_i u_i, \quad F_i \in \mathcal{S}, u_i \in L^2([0, T]), n \geq 1,$$

with respect to the norm $\|\cdot\|_{1,2,2}$ defined as:

$$\|u\|_{1,2,2}^2 := \mathbb{E}\big[\|u\|_{L^2([0,T])}^2\big] + \mathbb{E}\big[\|\nabla u\|_{H \otimes L^2([0,T])}^2\big].$$

Alternatively, an element u in $\mathbb{D}^{1,2}(\mathfrak{H})$ is identified with a stochastic process such that for almost avery t in $[0, T]$, u_t belongs to $\mathbb{D}^{1,2}$ and such that

$$\mathbb{E}\big[\|\nabla u\|_{H \otimes L^2([0,T])}^2\big] = \mathbb{E}\left[\int_0^T \int_0^T |D_s u_t|^2 ds\, dt\right] < +\infty.$$

Hence we can assume that α belongs to $\mathbb{D}^{1,2}(\mathfrak{H})$. Thus by [24, Theorem 3.1], α satisfies (RAC) and (SGD), which entails in this setting that for any h in H, there exists a $\mathfrak{H}$-valued r.v. $\tilde{\alpha}_h$ such that $\tilde{\alpha}_h = \alpha$ in $\mathfrak{H}$, $\mathbb{P}_0$-a.s., and for any $\varepsilon > 0$

$$\frac{\tilde{\alpha}_h \circ \tau_{\varepsilon h} - \tilde{\alpha}_h}{\varepsilon} = \varepsilon^{-1} \int_0^\varepsilon \langle \nabla \alpha \circ \tau_{sh}, h\rangle_H ds, \quad \text{in } \mathfrak{H}, \mathbb{P}_0\text{-a.s.}$$

Using Lemma 3.1 we thus get that for any $r \in (p, 2)$, it holds that:

$$\mathbb{E}\left[\int_0^T \left|\varepsilon^{-1}(\alpha_s \circ \tau_{\varepsilon h} - \alpha_s)\right|^r ds\right] = \mathbb{E}\left[\int_0^T \left|\varepsilon^{-1}\big((\tilde{\alpha}_h)(s) \circ \tau_{\varepsilon h} - (\tilde{\alpha}_h)(s)\big)\right|^r ds\right]$$

$$\leq \mathbb{E}\left[\int_0^T \varepsilon^{-1} \int_0^\varepsilon \left|\langle \nabla \alpha_s \circ \tau_{uh}, h\rangle_H\right|^r du\, ds\right]$$

$$\leq C \int_0^T \mathbb{E}\big[\left|\langle \nabla \alpha_s, h\rangle_H\right|^p\big]^{r/p} ds$$

$$\leq C \mathbb{E}\left[\int_0^T \left|\langle \nabla \alpha_s, h\rangle_H\right|^p ds\right]^{r/p}$$

$$\leq C \|h\|_H^r \mathbb{E}\big[\|\nabla \alpha\|_{H \otimes \mathfrak{H}}^p\big]^{r/p} < +\infty,$$

where we have used Cameron–Martin formula and similar computations to those of the proof of Lemma 4.2, and C denotes a positive constant which can differ from line to line. Hence, the family $(\int_0^T |\varepsilon^{-1}(\alpha_s \circ \tau_{\varepsilon h} - \alpha_s) - \langle \nabla \alpha_s, h\rangle_H|^p ds)_{\varepsilon \in (0,1)}$ is uniformly integrable. In addition, by Property (SGD), $\varepsilon^{-1}(\alpha \circ \tau_{\varepsilon h} - \alpha)$ converges in probability to $\langle \nabla \alpha, h\rangle_H$ (with respect to the norm $L^2([0, T])$) which implies that $\int_0^T |\varepsilon^{-1}(\alpha_s \circ \tau_{\varepsilon h} - \alpha_s) - \langle \nabla \alpha_s, h\rangle_H|^p ds$ converges in probability to 0 as ε goes to 0, which provides the result. $\qquad\square$

Acknowledgements

References

[1] O. Aboura and S. Bourguin. Density estimates for solutions to one dimensional backward SDE's. *Potential Anal.* **38** (2) (2013) 573–587. MR3015365

[2] S. Ankirchner, P. Imkeller and G. dos Reis. Classical and variational differentiability of bsdes with quadratic growth. *Electron. J. Probab.* **12** (53) (2007) 1418–1453. MR2354164

[3] F. Antonelli and A. Kohatsu-Higa. Densities of one-dimensional backward SDEs. *Potential Anal.* **22** (3) (2005) 263–287. MR2134722

[4] J.-M. Bismut. Contrôle des systèmes linéaires quadratiques: Applications de l'intégrale stochastique. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)* 180–264. *Lecture Notes in Math.* **649**. Springer, Berlin, 1978. MR0520007

[5] V. I. Bogachev. *Differentiable Measures and the Malliavin Calculus. Mathematical Surveys and Monographs* **164**, Am. Math. Soc., Providence, 2010. MR2663405

[6] P. Briand and C. Labart. Simulation of BSDEs by Wiener chaos expansion. *Ann. Appl. Probab.* **24** (3) (2014) 1129–1171. MR3199982

[7] P. Briand, B. Delyon, Y. Hu, E. Pardoux and L. Stoica. L^p solutions of backward stochastic differential equations. *Stochastic Process. Appl.* **108** (1) (2003) 109–129. MR2008603

[8] G. dos Reis. On some properties of solutions of quadratic growth BSDE and applications in finance and insurance. Ph.D. thesis, Humboldt University, 2010.

[9] N. El Karoui, S. Peng and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance* **7** (1) (1997) 1–71. MR1434407

[10] Y. Hu, D. Nualart and X. Zhong. Malliavin calculus for backward stochastic differential equations and application to numerical solutions. *Ann. Appl. Probab.* **21** (6) (2011) 2379–2423. MR2895419

[11] P. Imkeller and G. dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Process. Appl.* **120** (3) (2010) 348–379. MR2584898

[12] P. Imkeller, A. Réveillac and A. Richter. Differentiability of quadratic BSDEs generated by continuous martingales. *Ann. Appl. Probab.* **22** (1) (2012) 285–336. MR2932548

[13] P. Imkeller, T. Mastrolia, D. Possamaï and A. Réveillac. A note on the Malliavin–Sobolev spaces. *Statist. Probab. Lett.* **19** (2016) 45–53. MR3434959

[14] S. Janson. *Gaussian Hilbert Spaces*. Cambridge University Press, Cambridge, 1997. MR1474726

[15] S. Kusuoka. Dirichlet forms and diffusion processes on Banach spaces. *J. Fac. Sci., Univ. Tokyo, Sect. 1A, Math.* **29** (1) (1982) 79–95. MR0657873

[16] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)* 195–263. Wiley, New York–Chichester–Brisbane, 1978. MR0536013

[17] T. Mastrolia, D. Possamaï and A. Réveillac. Density analysis of BSDEs. *Ann. Probab.* To appear, 2014. Available at arXiv:1402.4416.

[18] D. Nualart. *The Malliavin Calculus and Related Topics*, 2nd edition. *Probability and Its Applications*. Springer, Berlin, 2006. MR2200233

[19] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** (1) (1990) 55–61. MR1037747

[20] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications* 200–217. Springer, Berlin, 1992. MR1176785

[21] P. E. Protter. *Stochastic Integration and Differential Equations*. 2nd edition. *Stochastic Modelling and Applied Probability* **21**. Springer, Berlin, 2005. Version 2.1, Corrected third printing. MR2273672

[22] I. Shigekawa. Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.* **20** (2) (1980) 263–289. MR0582167

[23] D. W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations. I. *Math. Syst. Theory* **14** (1) (1981) 25–65. MR0603973

[24] H. Sugita. On a characterization of the Sobolev spaces over an abstract Wiener space. *J. Math. Kyoto Univ.* **25** (4) (1985) 717–725. MR0810975

[25] A. S. Üstünel and M. Zakai. *Transformation of Measure on Wiener Space*. Springer, Berlin, 2000. MR1736980

[26] M. Veraar. The stochastic Fubini theorem revisited. *Stochastics* **84** (4) (2012) 543–551. MR2966093