ON A CLASS OF MARKOV BSDES WITH GENERALISED DRIVER

ELENA ISSOGLIO[1] AND FRANCESCO RUSSO[2]

Abstract. We are concerned with BSDEs where the driver contains a distributional term (in the sense of generalised functions). We introduce an integral operator to give sense to the equation and then we show the existence of a strong solution. Because of the irregularity of the driver, the Y-component of a couple (Y, Z) solving the BSDE is not necessarily a semimartingale but a weak Dirichlet process.

Key words and phrases. Backward stochastic differential equations (BSDEs); distributional driver; weak Dirichlet process; pointwise product; generalised and rough coefficients.
2010 MSC. 60G20; 60H30; 60J60.

1. Introduction

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

$$Y_t = \xi + \int_t^T Z_r b(r, W_r)\mathrm{d}r + \int_t^T f(r, W_r, Y_r, Z_r)\mathrm{d}r - \int_t^T Z_r \mathrm{d}W_r. \quad (1)$$

We are interested in a class of coefficients b of distributional type of the form $b \in C([0,T]; H_q^{-\beta}(\mathbb{R}^d;\mathbb{R}^d))$ for some $\beta \in (0, 1/2)$. The objects appearing in (1) take values in the following sets: $t \in [0,T]$, $\xi, W, Y \in \mathbb{R}^d$, $Z \in \mathbb{R}^{d\times d}$ and $f(t, W, Y, Z) \in \mathbb{R}^d$ (all vectors being column vectors). Here $\xi = \Phi(W_T)$ for some deterministic function Φ.

The classical notion of Brownian BSDE was introduced in 1990 by E. Pardoux and S. Peng in [16], after an early work of J. M. Bismut in 1973 in [2]. It is a stochastic differential equation with prescribed

[1]Department of Mathematics, University of Leeds, Leeds, LS2 9JT, UK E.Issoglio@leeds.ac.uk
[2]ENSTA ParisTech, Unité de Mathématiques appliquées, 828, boulevard des Maréchaux, F-91120 Palaiseau, France francesco.russo@ensta-paristech.fr

Date: May 4, 2018.

terminal condition ξ and driver $\hat{f}$ expressed by

$$Y_t = \xi + \int_t^T \hat{f}\,(r, \omega, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r \mathrm{d}W_r. \tag{2}$$

The unknown is a couple (Y, Z) of adapted processes. Of particular interest is the case where the randomness of the driver is expressed through a forward diffusion process X and the terminal condition only depends on X_T. Existence and uniqueness of the solution for the above equation was established first supposing (essentially) only Lipschitz conditions on the driver $\hat{f}$ with respect to the y and z variables. In subsequent works those conditions were considerably relaxed, see [17] and references therein for recent contributions on the topic. When $b = 0$, formally speaking (1) can be obtained as a special case of (2) setting $\hat{f}(r, \omega, y, z) = f(r, W_r(\omega), y, z)$. However, the driver in (1) includes the term $r \mapsto Z_r b(r, W_r(\omega))$, where b is a distribution in the second variable, hence this cannot be reduced to the classical case. As an example of generalised functions b which are allowed here, one can think of the derivative of a Hölder continuous function with Hölder parameter larger than $\frac{1}{2}$ (plus some growth condition at infinity).

One of the main applications of BSDEs is their use in providing probabilistic representations to the solution of certain non-linear PDEs. It is known (at least in the classical case) that when $\xi = \Phi(W_T)$, then BSDE (1) is linked to a PDE of the form

$$\begin{cases} \partial_t u + \frac{1}{2}\Delta u = -\nabla u^* \, b - f(u, \nabla u) \\ u(T) = \Phi, \end{cases}$$

see Section 2 for details about the notation. If u is the solution of the PDE, then one has the probabilistic representation $u(t, W_t) = Y_t$ and $\nabla u^*(t, W_t) = Z_t$, also known as non-linear Feynman-Kac formula. In this paper we obtain indeed a probabilistic representation for the solution u of the PDE using the BSDE. The novelty here is the rough driver b and the fact that the analysis is done entirely in terms of the original Brownian filtration even when the driver has the singular term Zb.

The topic of stochastic equations involving distributional coefficients has attracted a lot of interest, in particular for (forward) SDEs. See for example [6, 10, 9] in the case where the solution is not a semimartingale. See also [20] and more recently [8, 4]. For what concerns the case of backward SDEs involving a distribution we mention the works [7] on (reflected) BSDEs with distribution as terminal condition, and [23] whose authors studied a one-dimensional BSDE (with random terminal time) involving distributional coefficients via a forward stochastic process. Recently [1] considered a BSDE with a general forward process that covers various singular situations, including the case when the drift is distributional. [5] considered BSDEs where the driver is a Young integral. The work that is mostly related to our paper though, is [13]. There the authors study a BSDE like (2) where the driver has a component $Z_r b(r, W_r)$, and b is a distribution like in the present paper. It is clear that the integral $\int_t^T Z_r b(r, W_r)\mathrm{d}r$ needs to be carefully defined because the distribution $b(r, \cdot)$ cannot be evaluated at the point W_r and moreover a distribution in general cannot be multiplied by a function (in this case by the stochastic process Z_r). The way that this is carried out in [13] is by means of the Itô trick: they effectively replace the (not well-defined) distributional term $Z_r b(r, W_r)$ with known quantities, and get a new stochastic equation. For the latter, they are able to show existence and uniqueness of a solution under Lipschitz continuity conditions on the remaining part of the driver f (and some growth condition at 0).

In this paper we make a substantial step towards a deeper understanding of equations with distributional drivers. The main difference with [13] is that here we give a meaning to the distributional term $Z_r b(r, W_r)$ rather than replacing it with known objects, and effectively we solve the original BSDE rather than a different one. In this paper the underlying forward process is the Brownian motion itself. This means that the BSDE is constructed on the probability space where the Brownian motion lives, rather than the space where a (weak) solution of a forward SDE lies. Thus we provide a genuine *strong* solution to the BSDE.

We start by introducing an equivalent formulation of the BSDE (see Definition 3.3) that makes use of an integral operator $A^{W,Y}(b)$ introduced in Definition 3.1 to replace the term $\int_t^T Z_r b(r, W_r)\mathrm{d}r$. The operator is well-defined for smooth bs and coincides with the classical integral. The idea is to extend such operator to a class of drivers b that includes the original distributional driver. We carry out this study in Section 5.1 in the Markovian setting, where the component Y and the terminal condition ξ are deterministic functions of the Brownian motion W. In Proposition 5.5 we show that the integral operator satisfies certain continuity properties which are essential to prove the existence of the extension in the first place. Furthermore, a representation property is needed to effectively employ this operator and solve the original BSDE. For instance we show a chain rule in Proposition 5.7 which provides an explicit representation for the integral $A^{W,W}(b)$ even in the case when b is a distribution. Note that when b is a distribution then

$A^{W,W}(b)$ is defined as a limit, so having an explicit representation (chain rule) is very useful. The chain rule we just mentioned is shown for the special case $A^{W,W}$ and it is then linked to the general case $A^{Y,W}$ via the deterministic transformation γ, where $Y_r = \gamma(r, W_r)$. In Section 5.2 we state and proof the main results of existence (and uniqueness in a special class) of a solution to the BSDE (12) under the Markovian framework. The Markovian analysis is based on analytical properties of the PDE associated to the BSDE, which is introduced and investigated in Section 4.

Note. When we talk about *smooth drivers* we mean drivers for which the *classical* BSDE theory can be applied. For example smooth drivers are elements of $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. On the other hand, when we talk about *rough drivers* we mean elements in $C([0,T]; H_q^{-\beta})$, for which a new framework within BSDEs theory is developed here.

Throughout the paper c and C denote positive constants whose specific value is not important and may change from line to line.

2. Preliminaries and notation

Function spaces - notation. We denote by $C^{0,1}([0,T] \times \mathbb{R}^d)$ the space of real-valued continuous functions on $[0,T] \times \mathbb{R}^d$ which are continuously differentiable in the variable $x \in \mathbb{R}^d$. By $\varphi_n \to 0$ in $C^{0,1}$ we mean that φ_n and $\nabla \varphi_n$ (the gradient taken w.r.t. the x-variable) converge to 0 uniformly on compacts. The space $C^{0,1}$ is then endowed with the topology related to this convergence. For a vector $\varphi = (\varphi_1, \ldots, \varphi_d)$ such that $\varphi_i \in C^{0,1}([0,T] \times \mathbb{R}^d)$ for all i, we write $\varphi \in C^{0,1}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ or $\varphi \in C^{0,1}$ for shortness. Similarly we denote by $C^{1,2}([0,T] \times \mathbb{R}^d)$ the space of real-valued functions on $[0,T] \times \mathbb{R}^d$ which are continuously differentiable once in t and twice in x, and by $C^{1,2} := C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. The topology is similar to the one for $C^{0,1}$. Moreover we use $C_c(\mathbb{R}^d)$ to denote the space of continuous functions of x with compact support and $C_c^\infty(\mathbb{R}^d)$ to denote the space of infinitely differentiable functions with compact support. Again the short-hand notation for $\mathbb{R}^d$-valued functions is $C_c := C_c(\mathbb{R}^d; \mathbb{R}^d)$ and $C_c := C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$. The Euclidean norm in $\mathbb{R}$ and $\mathbb{R}^d$, and the Frobenius norm in $\mathbb{R}^{d \times d}$ will be denoted by $|\cdot|$. For a vector v, its transpose is denoted by v^*. If v is a real-valued function of $x \in \mathbb{R}$ then ∇v^* denotes the transpose of the column vector ∇v. Moreover is u is a vector-valued function of x then ∇u is a matrix where the j-th column is given by ∇u_j so that $(\nabla u)_{i,j} = \frac{\partial}{\partial x_i} u_j$. For the matrix ∇u, we denote its transposed by ∇u^*.

Stochastic analysis tools. Throughout the paper $(\Omega, \mathcal{G}, P)$ is a probability space on which a d-dimensional Brownian motion $W := (W_t)_t$ is defined, with Brownian filtration $\mathcal{F} := (\mathcal{F}_t)_t$.

We denote by $\mathcal{C}$ the space of continuous stochastic processes indexed by $[0,T]$ with values in $\mathbb{R}^d$. In this space we will consider u.c.p. convergence (uniform convergence in probability) for stochastic processes. More precisely, we say that a family of stochastic processes X^n indexed by $[0,T]$ converges u.c.p. to X in $\mathcal{C}$ if

$$\sup_{s\in[0,T]} |X^n_s - X_s| \to 0 \text{ in probability.}$$

The following definitions of covariation process and weak-Dirichlet process are taken from [11], see also [21] for more details.
Given two stochastic processes $Y := (Y_t)_t$ and $X := (X_t)_t$, we denote by $[Y, X]$ the *covariation process* of Y and X which is defined by

$$[Y,X]_t := \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_0^t (Y_{s+\varepsilon} - Y_s)(X_{s+\varepsilon} - X_s)\mathrm{d}s,$$

if the limit exists in the u.c.p. sense in t. If X, Y are d-dimensional processes then $[Y,X] \in \mathbb{R}^{d\times d}$ is the tensor covariation and it is defined component by component by $([Y,X])_{i,j} = [Y_i, X_j]$, if it exists. Note that the covariation is not symmetric because the matrix does not need to be squared and in particular we have $[Y,X] = [X,Y]^*$. This concept extends the classical covariation of continuous semimartingales. We remark that the covariation of a bounded variation process and a continuous process is always zero.
Given a filtration $\mathcal{F} := (\mathcal{F}_t)_t$, a real process D is said to be an $\mathcal{F}$*-weak Dirichlet* process if it can be written as $D = M + A$ where

(i) $M := (M_t)_t$ is an $\mathcal{F}$-local martingale;
(ii) $(A_t)_t$ is an *orthogonal martingale process*, namely a process such that $[A, N] = 0$ for every $\mathcal{F}$-continuous local martingale N. For convenience we also set $A_0 = 0$.

It was shown that the decomposition $D = M + A$ is unique and every $\mathcal{F}$-semimartingale is an $\mathcal{F}$-weak Dirichlet process. A vector $D = (D^1, \dots, D^d)$ is an $\mathcal{F}$-weak Dirichlet process if every component D^i is an $\mathcal{F}$-weak Dirichlet process. We will drop the $\mathcal{F}$ and simply write weak Dirichlet process when it is clear what filtration $\mathcal{F}$ we are considering.

Proposition 2.1. *Let $v \in C^{0,1}([0,T]\times\mathbb{R}^d)$ and S^1 (resp. S^2) be an $\mathbb{R}^d$-valued (resp. $\mathbb{R}$-valued) continuous $\mathcal{F}$-semimartingale with martingale* *component M^1 (resp. M^2). Then*

$$[v(\cdot, S^1), S^2]_t = \int_0^t \nabla v^*(r, S^1_r)\mathrm{d}[M^1, M^2]_r. \tag{3}$$

Proof. Let us denote by $M^v_t := \int_0^t \nabla v^*(r, S^1_r)\mathrm{d}M^1_r$. By [11, Corollary 3.11] we have that $v(\cdot, S^1)$ is a weak Dirichlet process with martingale component M^v. If A^v is the related orthogonal martingale process, we know that $[A^v, N] = 0$ for any $\mathcal{F}$-continuous local martingale N, see [22, Proposition 1.7.(b)]. Consequently the left-hand side of (3) gives

$$\begin{aligned}[v(\cdot, S^1), S^2]_t &= [M^v, M^2]_t \\ &= \left[\int_0^\cdot \nabla v^*(r, S^1_r)\mathrm{d}M^1_r, M^2\right] \\ &= \int_0^t \nabla v^*(r, S^1_r)\mathrm{d}[M^1, M^2]_r,\end{aligned}$$

where the last equality holds true because the covariation $[\cdot,\cdot]$ extends the one of semimartingales. □

When v is a vector-valued function (say u), the covariation becomes a matrix and an analogous result holds, as stated in the corollary below (in the special case when u is a function of Brownian motion).

Corollary 2.2. *Let $\phi \in C^{0,1}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$, W be an $\mathbb{R}^d$-valued $\mathcal{F}$-Brownian motion and N an $\mathcal{F}$-continuous local martingale with values in $\mathbb{R}^d$. Then*

$$[\phi(\cdot, W), N]_t = \int_0^t \nabla \phi^*(r, W_r)\mathrm{d}[W, N]_r.$$

Heat semigroup and fractional Sobolev spaces. We denote by $\mathcal{S}(\mathbb{R}^d)$ the space of $\mathbb{R}^d$-valued Schwartz functions and by $\mathcal{S}'(\mathbb{R}^d)$ the space of Schwartz distributions. The heat semigroup $(P(t), t \geq 0)$ with kernel $p_t(x) = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{2t}\right)$ acts on $\mathcal{S}(\mathbb{R}^d)$ and can be extended to $\mathcal{S}'(\mathbb{R}^d)$ by duality. One can also consider the restriction of the semigroup mapping any $L^r(\mathbb{R}^d)$ into itself for $1 < r < \infty$, which is denoted again by $(P(t), t \geq 0)$. This restriction is a bounded analytic semigroup, see [3, Theorems 1.4.1, 1.4.2]. It is known that one can define fractional Sobolev spaces using such semigroup and its generator $\frac{1}{2}\Delta$, more precisely if we define $A := I - \frac{1}{2}\Delta$, then its fractional powers are well-defined for any power $s \in \mathbb{R}$ (see [18]) and $H^s_r(\mathbb{R}^d) := A^{-s/2}(L^r(\mathbb{R}^d))$. These are Banach spaces endowed with the norm $\|u\|_{H^s_r} := \|A^{s/2}u\|_{L^r}$. Using the fact that $D(A^{s/2}) = D((-\frac{1}{2}\Delta)^{s/2}) = H^s_r(\mathbb{R}^d)$ and that $A^{-\alpha/2}$ is an isomorphism between

$H^s_r(\mathbb{R}^d)$ and $H^{s+\alpha}_r(\mathbb{R}^d)$, for each $\alpha \in \mathbb{R}$, one has for $\delta > \beta > 0, \delta+\beta < 1$ and $0 < t \le T$ that $P(t) : H^{-\beta}_r(\mathbb{R}^d) \to H^{1+\delta}_r(\mathbb{R}^d)$ for all $1 < r < \infty$ and

$$\|P(t)w\|_{H^{1+\delta}_r(\mathbb{R}^d)} \le Ce^t t^{-\frac{1+\delta+\beta}{2}} \|w\|_{H^{-\beta}_r(\mathbb{R}^d)}, \tag{4}$$

for $w \in H^{-\beta}_r(\mathbb{R}^d), t > 0$. This follows from a similar property for the bounded analytic semigroup $(e^{-t}P(t), t \ge 0)$ generated by $-A$ which is stated in [8, Lemma 10], see also [12, Proposition 3.2] for the analogous on domains $D \subset \mathbb{R}^d$.

As done already before in this paper, we denote by H^s_r the spaces $H^s_r(\mathbb{R}^d;\mathbb{R}^d)$, whose definition is as above for each component. Note that by slight abuse of notation the same H^s_r might be the space $H^s_r(\mathbb{R}^d;\mathbb{R}^{d\times d})$, especially when considering functions like ∇u. When we write $u \in H^s_r$ we mean that each component u_i is in $H^s_r(\mathbb{R}^d)$. The norm will be denoted with the same notation for simplicity.

Pointwise product. Here we recall the definition of the *pointwise product* between a function and a distribution, for more details see [19]. Let $g \in \mathcal{S}'(\mathbb{R}^d)$. We choose a function $\psi \in \mathcal{S}(\mathbb{R}^d)$ such that $0 \le \psi(x) \le 1$, for every $x \in \mathbb{R}^d$ and

$$\psi(x) = \begin{cases} 1, & |x| < 1, \\ 0, & |x| \ge 2. \end{cases}$$

For every $j \in \mathbb{N}$, we consider the approximation $S^j g$ of g as follows:

$$S^j g(x) := \mathscr{F}^{-1}\left(\psi\left(\frac{\xi}{2^j}\right)\mathscr{F}(g)\right)(x),$$

where $\mathscr{F}(g)$ and $\mathscr{F}^{-1}(g)$ are the Fourier transform and the inverse Fourier transform of g, respectively. The product gh of $g, h \in \mathcal{S}'(\mathbb{R}^d)$ is defined as

$$gh := \lim_{j\to\infty} S^j g S^j h, \tag{5}$$

if the limit exists in $\mathcal{S}'(\mathbb{R}^d)$.

Lemma 2.3. [19, Theorem 4.4.3/1] *Let $g \in H^{-\beta}_q(\mathbb{R}^d)$, $h \in H^\delta_p(\mathbb{R}^d)$ for $1 < p, q < \infty$, $q > \max(p, \frac{d}{\delta})$, $0 < \beta < \frac{1}{2}$ and $\beta < \delta$. Then the pointwise product gh is well-defined, it belongs to the space $H^{-\beta}_p(\mathbb{R}^d)$ and we have the following bound:*

$$\|gh\|_{H^{-\beta}_p(\mathbb{R}^d)} \le c\|g\|_{H^{-\beta}_q(\mathbb{R}^d)} \cdot \|h\|_{H^\delta_p(\mathbb{R}^d)}. \tag{6}$$

In this paper we will always use this product in such fractional Sobolev spaces.

More on function spaces. We repeat that when we talk about smooth drivers we consider elements of $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ or of $C_c^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$, which is defined to be the space of all $f \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ such that $\frac{\partial^\alpha f}{\partial x^\alpha}$ exists for all multi-indexes α and $\frac{\partial^\alpha f}{\partial x^\alpha} \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. It is clear that each function in $C_c^\infty([0,T] \times \mathbb{R}^d)$ is an element of $L^r(\mathbb{R}^d)$ for any fixed time $t \in [0,T]$ and for $2 \leq r \leq \infty$, and moreover it is continuous with respect to the topology in $L^r(\mathbb{R}^d)$. Since $L^r(\mathbb{R}^d) \subset H_r^s(\mathbb{R}^d)$ for $s \leq 0$ we have the inclusion $C_c^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d) \subset C([0,T]; H_r^s)$.

For the following, see [24, Section 2.7.1]. The closures of $\mathcal{S}(\mathbb{R}^d)$ with respect to the norms

$$\|h\|_{C_b^{0,0}(\mathbb{R}^d)} := \|h\|_{L^\infty(\mathbb{R}^d)}$$

and

$$\|h\|_{C_b^{1,0}(\mathbb{R}^d)} := \|h\|_{L^\infty(\mathbb{R}^d)} + \|\nabla h\|_{L^\infty(\mathbb{R}^d)}$$

respectively, are denoted by $C_b^{0,0}(\mathbb{R}^d)$ and $C_b^{1,0}(\mathbb{R}^d)$. For any $\alpha > 0$, we consider the Banach spaces

$$C^{0+\alpha}(\mathbb{R}^d) = \{h \in C_b^{0,0}(\mathbb{R}^d) : \|h\|_{C^{0+\alpha}(\mathbb{R}^d)} < \infty\}$$
$$C^{1+\alpha}(\mathbb{R}^d) = \{h \in C_b^{1,0}(\mathbb{R}^d) : \|h\|_{C^{1+\alpha}(\mathbb{R}^d)} < \infty\},$$

endowed with the norms

$$\|h\|_{C^{0+\alpha}(\mathbb{R}^d)} := \|h\|_{L^\infty(\mathbb{R}^d)} + \sup_{x \neq y \in \mathbb{R}^d} \frac{|h(x) - h(y)|}{|x-y|^\alpha}$$
$$\|h\|_{C^{1+\alpha}(\mathbb{R}^d)} := \|h\|_{L^\infty(\mathbb{R}^d)} + \|\nabla h\|_{L^\infty(\mathbb{R}^d)} + \sup_{x \neq y \in \mathbb{R}^d} \frac{|\nabla h(x) - \nabla h(y)|}{|x-y|^\alpha},$$

respectively. We denote by $C^{0+\alpha}$ and $C^{1+\alpha}$ the analogous spaces for $\mathbb{R}^d$-valued functions and the corresponding norms by $\|\cdot\|_{C^{0+\alpha}}$ and $\|\cdot\|_{C^{1+\alpha}}$.

Let B be a Banach space. We denote by $C([0,T]; B)$ the Banach space of B-valued continuous functions and its sup norm by $\|\cdot\|_{C([0,T];B)}$. For $h \in C([0,T]; B)$ and $\rho \geq 1$ we also use the family of equivalent norms $\{\|\cdot\|^{(\rho)}_{C([0,T];B)}, \rho \geq 1\}$, defined by

$$\|h\|^{(\rho)}_{C([0,T];B)} := \sup_{0 \leq t \leq T} e^{-\rho t} \|h(t)\|_B. \tag{7}$$

The following lemma is a fractional Sobolev embedding theorem which will be used several times in this paper. It is a generalisation of the Morrey inequality to fractional Sobolev spaces. For the proof we refer to [24, Theorem 2.8.1, Remark 2].

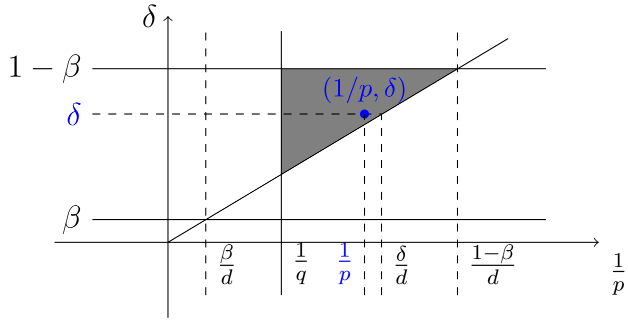

FIGURE 1. The set $K(\beta,q)$ for $d>1$. Given any couple β,q that satisfies the assumptions, the grey region shows all possible δ,p.

Lemma 2.4 (Fractional Morrey inequality). *Let $0<\delta<1$ and $d/\delta<r<\infty$. If $h\in H_r^{1+\delta}(\mathbb{R}^d)$ then there exists a unique version of h (which we denote again by h) such that h is differentiable. Moreover $h\in C^{1+\alpha}(\mathbb{R}^d)$ with $\alpha=\delta-d/r$ and*

$$\|h\|_{C^{1+\alpha}(\mathbb{R}^d)} \le c\|h\|_{H_r^{1+\delta}(\mathbb{R}^d)}, \quad \|\nabla h\|_{C^{0+\alpha}(\mathbb{R}^d)} \le c\|\nabla h\|_{H_r^{\delta}(\mathbb{R}^d)}, \tag{8}$$

where $c=c(\delta,r,d)$ is a universal constant. In particular h and ∇h are bounded.

Assumptions. Later in the paper we will use the following assumptions about the parameters and the functions involved.

Assumption 2.5. *We always choose $(\delta,p)\in K(\beta,q)$, where the latter set is defined below in two different cases.*

Case $d\ge 2$: *Let $\beta\in\left(0,\frac{1}{2}\right)$ and $q\in\left(\frac{d}{1-\beta},\frac{d}{\beta}\right)$. For given β and q as above we define the set*

$$K(\beta,q) := \left\{(\delta,p)\in\mathbb{R}^2:\ \beta<\delta<1-\beta,\ \frac{d}{\delta}<p<q\right\}, \tag{9}$$

which is drawn in Figure 1.

Case $d=1$: *In this case we let $\beta\in\left(0,\frac{1}{2}\right)$ and $q\in\left(2,\frac{1}{\beta}\right)$. For given β and q as above we define the set*

$$K(\beta,q) := \left\{(\delta,p)\in\mathbb{R}^2:\ \beta<\delta<1-\beta,\ \frac{1}{\delta}<p<q, 2\le p\right\}, \tag{10}$$

which is drawn in Figure 2.

Note that $K(\beta,q)$ is non-empty since $\beta<\frac{1}{2}$ and $\frac{d}{1-\beta}<q<\frac{d}{\beta}$. The set $K(\beta,q)$ was first introduced in [8] without the restriction $q,p\ge 2$.

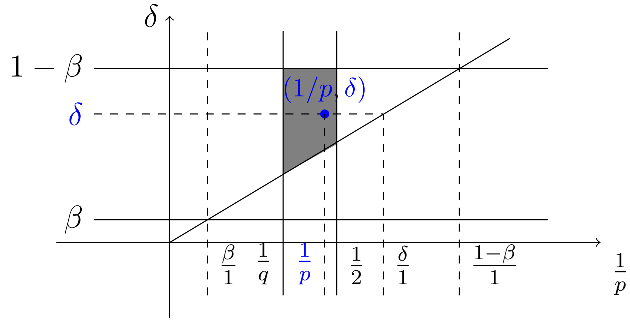

FIGURE 2. The set $K(\beta, q)$ for $d = 1$. Given any couple β, q that satisfies the assumptions, the grey region shows all possible δ, p.

This is satisfied anyway if $d > 1$. If $d = 1$ then the set of admissible couples (δ, p) is shown in Figure 2.

The following assumption concerns the driver f and the terminal condition Φ (note that the terminal condition ξ in the BSDE will be replaced by $\Phi(W)$ in later sections).

Assumption 2.6. *We assume the following.*

- $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ *is an element of* $H^{1+\delta+2\gamma}(\mathbb{R}^d)$ *for some* $0 < \gamma < \frac{1-\delta-\beta}{2}$*;*
- $f : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \to \mathbb{R}^d$ *is continuous in* (x, y, z) *uniformly in* t*, and is Lipschitz continuous in* (y, z) *uniformly in* t *and* x*, i.e.* $|f(t, x, y, z) - f(t, x, y', z')| \le L(|y - y'| + |z - z'|)$ *for any* $y, y' \in \mathbb{R}^d$ *and* $z, z' \in \mathbb{R}^{d\times d}$*. Moreover* $f(t, x, 0, 0)$ *is continuous in* (t, x)*;*
- $\sup_{t,x} |f(t, x, 0, 0)| < \infty$ *a.s. and* $\sup_{t\in[0,T]} \|f(t, \cdot, 0, 0)\|_{L^p} < \infty$*.*

3. Alternative representation for the BSDE

In this section we propose an alternative representation for the BSDE (1) which turns out to be well-suited for BSDEs with rough drivers and it is equivalent to the one above if the driver is smooth, see Proposition 3.5 below.

To be able to consider rough drivers, the main term in (1) that needs to be (re)defined is the integral $\int_t^T Z_r b(r, W_r)\mathrm{d}r$. Here we recall that b is a column $\mathbb{R}^d$-vector and $Z \in \mathbb{R}^{d\times d}$ so that the integral is a column vector. We introduce the following integral operator.

Definition 3.1. *Let* $W = (W_t)_t$ *be a* d*-dimensional Brownian motion with filtration* $\mathcal{F} = (\mathcal{F}_t)_t$ *and* $Y = (Y_t)_t$ *be a continuous* $\mathbb{R}^d$*-valued* *stochastic process such that the* $(d \times d)$*-covariation matrix* $[W, Y]$ *exists and all the components have finite variation.*

The integral operator $A^{W,Y}$ *is defined on the space* $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ *by*

$$\begin{array}{rrcl} A^{W,Y}: & C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d) & \to & \mathcal{C} \\ & l & \mapsto & A^{W,Y}_\cdot(l), \end{array}$$

where

$$A^{W,Y}_t(l) := \left(\int_0^t l^*(r, W_r) \mathrm{d}[W, Y]_r \right)^* \tag{11}$$

for all $t \in [0,T]$*. Here* l *and* $A^{W,Y}_t(l)$ *are* d*-dimensional column vectors.*

We observe that in the special case when $Y = W$ the operator $A^{W,W}$ applied to l is nothing but $\int_0^\cdot l(r, W_r)\mathrm{d}r$ (see the introduction of Section 5.1 for more details).
Moreover, for this smooth class of drivers $l \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ the integral in (11) is well-defined because $[W, Y]$ is a matrix with finite variation components by assumption. Our aim is to define such integral operator $A^{W,Y}$ for rough driver, as specified in the next definition.

Definition 3.2. *Let* E *be a Polish space which contains* $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ *as a dense subset. We define the integral operator* $A^{W,Y} : E \to \mathcal{C}$ *as the continuous extension of the operator defined in Definition 3.1, provided that the extension exists.*

In Section 5 we will prove the existence of such extension for $E = C([0,T]; H^s_r)$ with $2 \le r < \infty$ and $-\frac{1}{2} < s \le 0$. Using this extension we can reformulate BSDE (1) for a rough driver and give a precise meaning to its solution.

Definition 3.3. *Let* $b \in C([0,T]; \mathcal{S}')$*. Let* E *be a Polish space of* $\mathcal{S}'$*-valued functions including* $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ *as a dense subset and such that* $b \in E$*. We say that a continuous* $\mathbb{R}^d$*-valued stochastic process* Y *is a solution of BSDE* (1) *if:*

(i) $A^{W,Y}$ *exists as an operator according to Definition 3.2;*
(ii) $A^{W,Y}_\cdot(b)$ *is an orthogonal martingale process;*
(iii) $Y_T = \xi$*;*
(iv) *the process* $M = (M_t)_t$ *given by*

$$M_t := Y_t - Y_0 + A^{W,Y}_t(b) + \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y, W]_r}{\mathrm{d}r}\right) \mathrm{d}r \tag{12}$$

is a square-integrable $\mathcal{F}$*-martingale, where* $\mathcal{F}$ *is the Brownian filtration.*

Remark 3.4.
- *Such solution Y is a weak-Dirichlet process with orthogonal martingale component given by*

$$A_t^{W,Y}(b) + \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y,W]_r}{\mathrm{d}r}\right) \mathrm{d}r.$$

- *We have $[Y,W] = [M,W]$, thus the covariation process is absolutely continuous with respect to* $\mathrm{d}r$ *component by component and hence all terms appearing in the driver f in* (12) *are well-defined.*
- *Definition 3.3 above makes sense also in the case when ξ is a generic square integrable random variable and the random dependence in the driver f is allowed to be on the whole past $\{W_s; s \le r\}$ instead of only on W_r.*

In the next proposition we see how the classical formulation of the BSDE is equivalent to the one introduced above if $b \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. In this case clearly $A^{W,Y}$ is itself the trivial extension to $E = C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ of the operator introduced in Definition 3.1.

Proposition 3.5. *Let Y be a d-dimensional adapted process and $b \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. Then Y is a solution of* (1) *according to Definition 3.3 with respect to some E if and only if there exists a predictable $(d \times d)$-dimensional process Z such that (Y,Z) is a solution of BSDE* (1) *in the classical sense.*

Proof. Suppose that (Y,Z) is a classical solution of (1). We set $E = C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. This ensures that point (i) of Definition 3.3 is satisfied and $A_\cdot^{W,Y}(b) = \left(\int_0^\cdot b^*(r, W_r) \mathrm{d}[W,Y]_r\right)^*$. Using (1) we have

$$[W,Y]_t = \left[W, Y_0 - \int_0^\cdot Z_r b(r, W_r)\mathrm{d}r - \int_0^\cdot f(r, W_r, Y_r, Z_r)\mathrm{d}r + \int_0^\cdot Z_r \mathrm{d}W_r\right]_t,$$

where the covariation is a matrix and it is calculated component by component. Clearly the only non-zero term is given by the stochastic integral and so we get

$$[W,Y]_t = \left[W, \int_0^\cdot Z_r \mathrm{d}W_r\right]_t = \int_0^t Z_r^* \mathrm{d}r,$$

hence $\mathrm{d}[W,Y]_r = Z_r^* \mathrm{d}r$, and in particular

$$A_\cdot^{W,Y}(b) = \left(\int_0^\cdot b^*(r, W_r) Z_r^* \mathrm{d}r\right)^* = \int_0^\cdot Z_r b(r, W_r)\mathrm{d}r.$$

Being of bounded variation, the latter is clearly an orthogonal martingale process, which is point (ii) in Definition 3.3. Point (iii) is trivial. Point (iv) is also satisfied because

$$Y_t - Y_0 + A_t^{W,Y}(b) + \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y,W]_r}{\mathrm{d}r}\right)\mathrm{d}r = \int_0^t Z_r \mathrm{d}W_r$$

and the right-hand side is a square integrable $\mathcal{F}$-martingale.

Conversely, let Y be a solution of (1) according to Definition 3.3 with respect to E. We know that

$$M_t := Y_t - Y_0 + A_t^{W,Y}(b) + \int_0^t f(r, W_r, Y_r, \frac{\mathrm{d}[Y,W]_r}{\mathrm{d}r})\mathrm{d}r$$

is a square integrable martingale by point (iv) in Definition 3.3, hence by the martingale representation theorem there exists a square-integrable process Z such that $M_t = \int_0^t Z_r \mathrm{d}W_r$. Moreover $A^{W,Y}$ is an orthogonal martingale process by point (ii), thus $[W,Y]_t = [W,M]_t = \int_0^t Z_r^* \mathrm{d}r$. Therefore $A_t^{W,Y}(b) = (\int_0^t b^*(s, W_s)\mathrm{d}[W,Y]_s)^* = \int_0^t Z_s b(s, W_s)\mathrm{d}s$ and this concludes the proof. □

Remark 3.6. *We observe that, in the classical formulation of BSDEs, Z is always directly determined by Y since $\frac{\mathrm{d}}{\mathrm{d}t}[Y,W]_t = Z_t$.*

To conclude this section we point out that the new setting and formulation introduced in Definition 3.3 in fact coincide with the classical ones even in the case when $b \in L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. This can be seen by observing two facts. The first one is that a BSDE with a driver $b \in L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ makes sense without the introduction of the operator A and can be studied with classical methods (*à la* Pardoux-Peng). On the other hand we will show (see Theorem 5.11) that the operator $A^{W,W}$ applied to a driver in $C([0,T]; H^s_r) \cap L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ is compatible with integrals of drivers in $L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ defined classically. Hence the framework presented here coincides with the classical one not only for $b \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ (as shown in Proposition 3.5) but also for $b \in L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$.

4. Analytical PDE results

In this section we collect and prove some results about several PDEs that will be used in Section 5. In particular, a key point in the subsequent analysis will be to show that the integral operator $A^{Y,W}$ appearing in (12) is well-defined for rough drivers and this will be done with the aid of the following auxiliary PDEs and relative results.

The parameters β and q are fixed and chosen according to Assumption 2.5. These are directly linked to the regularity of the rough driver

b. Moreover the parameters (δ, p) are chosen in $K(\beta, q)$ and in particular $\frac{d}{\delta} < p < q$.

The first auxiliary PDE is

$$\begin{cases} \partial_t \phi + \frac{1}{2}\Delta\phi = l \\ \phi(T) = \Psi, \end{cases} \tag{13}$$

where $\Psi \in H_r^{1+\delta}$ and $l \in C([0,T]; H_p^{-\beta})$. Here the Laplacian Δ acts on ϕ componentwise and the resulting object is a vector with i-th component given by $\Delta\phi_i$. With a slight abuse of notation we use $\Delta\phi$ for the whole vector. We consider the mild formulation of (13) which is given by

$$\phi(t) = P(T-t)\Psi + \int_t^T P(r-t)l(r)\mathrm{d}r, \tag{14}$$

where $\{P(t), t \geq 0\}$ is the semigroup generated by $\frac{1}{2}\Delta$.
It is known that if a classical solution exists then it coincides with the solution of (14) (mild formulation) and it has certain regularity properties as recalled in the lemma below for smooth Φ and l. For more details and a proof see for example [15, Theorem 5.1.4, part (iv)].

Lemma 4.1. *Let $l \in C_c^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ and $\Psi \in C^{2+\epsilon}(\mathbb{R}^d; \mathbb{R}^d)$ for some $0 < \epsilon < 1$. The solution ϕ to (13) is at least of class $C^{1,2+\epsilon}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$.*

In the general case that suits our framework (i.e. for rough ls and Ψ in fractional Sobolev spaces) we have the following results.

Lemma 4.2. *Let β, δ, p and q be chosen according to Assumption 2.5.*

(i) *If $\Psi \in H_p^{1+\delta}$ then $t \mapsto P(T-t)\Psi$ is a continuous function with values in $H_p^{1+\delta}$.*

(ii) *If $l \in C\left([0,T]; H_p^{-\beta}\right)$ then the function $t \mapsto \int_t^T P(s-t) l(s)\,\mathrm{d}s$ is in $C^\gamma\left([0,T]; H_p^{2-2\epsilon-\beta}\right)$ for every $\epsilon > 0$ and $\gamma \in (0,\epsilon)$.*
In particular one can always choose ϵ such that $2-2\epsilon-\beta = 1+\delta$.

Proof. Item (i) is well-known and can be shown using the theory of semigroups and the fact that $P(t)$ is the heat semigroup (contractive semigroup).
Item (ii) follows by first applying [8, Proposition 11] with the time t replaced by $T-t$ and then making a change of time to the resulting integral to get a backward integral, namely transforming the integrator variable r into $s = t - r$. □

Lemma 4.3. *Let Assumption 2.5 hold, and let $\Psi \in H_p^{1+\delta}$ and $l \in C([0,T]; H_p^{-\beta})$. The expression ϕ given in (14) is well-defined and* *belongs to* $C([0,T];H_p^{1+\delta}) \subset C([0,T];C^{1+\alpha})$ *and to* $C^{0,1}$*, where* $\alpha = \delta - d/p$*. Moreover we have*

$$\|\phi(t)\|_{H_p^{1+\delta}} \leq ce^T \|\Psi\|_{H_p^{1+\delta}} + (T-t)^{\frac{1-\delta-\beta}{2}} \|l\|_{C([0,T];H_p^{-\beta})}$$

and

$$\|\phi\|_{C([0,T];C^{1+\alpha})} \leq c\|\phi\|_{C([0,T];H_p^{1+\delta})}.$$

Proof. For the first term in (14) we have that $t \mapsto P(T-t)\Psi \in H_p^{1+\delta}$ is continuous by Lemma 4.2, item (i). Moreover by the mapping property (4) of the semigroup in Sobolev spaces, its norm is bounded by $\sup_{t\in[0,T]} \|P(T-t)\Psi\|_{H_p^{1+\delta}} \leq ce^T\|\Psi\|_{H_p^{1+\delta}}$. For the second term in (14) we have continuity as a function of time by Lemma 4.2, item (ii) and again by the mapping property of semigroups we get the bound

$$\begin{aligned} \left\| \int_t^T P(s-t)l(s)\mathrm{d}s \right\|_{H_p^{1+\delta}} &\leq ce^T \int_t^T (s-t)^{-\frac{1+\delta+\beta}{2}} \|l(s)\|_{H_p^{-\beta}} \mathrm{d}s \\ &\leq c(T-t)^{\frac{1-\delta-\beta}{2}} \|l\|_{C([0,T];H_p^{-\beta})}, \end{aligned}$$

which ensures that $\phi \in C([0,T];H_p^{1+\delta})$ since $1-\delta-\beta > 0$ by assumption on the parameters. Moreover $\delta > d/p$ again by Assumption 2.5 and so by the fractional Morrey inequality (Lemma 2.4) we have

$$\|\phi(t)\|_{C^{1+\alpha}} \leq c\|\phi(t)\|_{H_p^{1+\delta}}.$$

Hence taking the supremum over $t \in [0,T]$ we get

$$\|\phi\|_{C([0,T];C^{1+\alpha})} \leq c\|\phi\|_{C([0,T];H_p^{1+\delta})}.$$

From this it follows that the solution ϕ is jointly continuous in t and x and once differentiable in x, namely $\phi \in C^{0,1}$ as wanted (for a proof of a result similar to the last claim see [8, Lemma 21]). □

The following corollary follows from Lemma 4.3 by the linearity of the PDE.

Corollary 4.4. *Let Assumption 2.5 hold. Let* $(l_n)_n \subset C([0,T];H_p^{-\beta})$ *be a sequence such that* $l_n \to l$ *in this space and let* $\Psi_n \to \Psi$ *in* $H_p^{1+\delta}$ *with* $(\Psi_n)_n \subset H_p^{1+\delta}$*. Let* ϕ_n *denote the solution of* (13) *with* l_n *in place of* l*. Then* $\phi_n \to \phi$ *in* $C^{0,1}$*.*

Another important PDE that will appear in the next section is the PDE associated to BSDE (1) in the Markovian case, which will be used to construct the solution to the BSDE, namely

$$\begin{cases} \partial_t u(t) + \frac{1}{2}\Delta u(t) = -\nabla u^*(t)\, b(t) - f(t,\cdot,u(t),\nabla u(t)) \\ u(T) = \Phi. \end{cases} \tag{15}$$

We note that the term Δu (as in PDE (13) above) and the term $\nabla u^* b$ are defined componentwise, in particular the i-th component of $\nabla u^* b$ is given by $\nabla u_i^* b$. A mild solution to PDE (15) is a function u that satisfies

$$u(t) = P(T-t)\Phi - \int_t^T P(r-t)\left(\nabla u^*(r) b(r)\right) \mathrm{d}r \qquad (16)$$
$$- \int_t^T P(r-t) f(r, \cdot, u(r), \nabla u(r)) \mathrm{d}r$$

in an appropriate function space (specified below). Each component in the term $\nabla u^*(r)b(r)$ is defined by means of the pointwise product (recalled in Section 2) and it is well-defined as an element of $H_p^{-\beta}$ when $b(t) \in H_q^{-\beta}$ and $\nabla u^*(t) \in H_p^{\delta}$.

Equation (15) was first studied in [12] on a bounded domain $D \subset \mathbb{R}^d$ and with $f \equiv 0$. It was then solved in $\mathbb{R}^d$ in [8] with $f = 0$, and in [13] with f non zero. In particular in [13] the authors obtain an existence and uniqueness result for a function $\tilde{f} : [0,T] \times H_p^{1+\delta} \times H_p^{\delta} \to H_p^0$ with some Lipschitz regularity and boundedness at 0. We want to apply this result later on, but we will need to consider $\tilde{f}$ to be the *same* function f appearing in BSDE (1). Clearly some care is needed because the f appearing in the BSDE is a function of t, x, y and z and its regularity stated in Assumption 2.6 is given pointwise, unlike $\tilde{f}$. On the other hand, to get a fixpoint for the PDE we need some Lipschitz regularity in terms of the function spaces. The way to merge these two settings is to consider a function $\tilde{f}$ (which will have the appropriate Lipschitz regularity) by setting $\tilde{f}(t,u,v) = f(t, \cdot u(t), \nabla u(t))$ for any $u \in H_p^{1+\delta}$ and $v \in H_p^{\delta}$, with f from Assumption 2.6 (we will abuse the notation and write f for both). Then $\tilde{f}$ satisfies the required conditions, as explained in [13, Remark 2.5], in particular $\tilde{f}$ is Lipschitz continuous in the Sobolev spaces

$$\|\tilde{f}(t,u,v) - \tilde{f}(t,u',v')\|_{H_p^0} \leq c(\|u-u'\|_{H_p^{1+\delta}} + \|v-v'\|_{H_p^{\delta}}). \qquad (17)$$

Theorem 5, and Lemmata 5 and 8 in [13] give the following existence, uniqueness and regularity result.

Theorem 4.5 (Issoglio, Jing)**.** *Under Assumption 2.5 and Assumption 2.6 there exists a unique mild solution u to (15) in $C([0,T]; H_p^{1+\delta})$. Moreover $u(t) \in C^{1+\alpha}$ for all $t \in [0,T]$, where $\alpha = \delta - d/p$, and $u \in C^{0,1}([0,T] \times \mathbb{R}^d)$.*

A small note: in [13] the result is valid even if $b \in L^\infty([0,T]; H_q^{-\beta})$.

5. The Markovian case with distributional driver

In this section we carry out the analysis of BSDE (1) when $b \in C([0,T]; H_q^{-\beta})$ in the *Markovian setting*. The Markovian setting means that the process Y and the r.v. ξ are deterministic functions of W, namely $\xi = \Phi(W_T)$ and $Y_t = \gamma(t, W_t)$ for some deterministic functions Φ and γ, the regularity of which is specified below.

As already mentioned previously, one of the main issues when dealing with rough drivers is to show that the integral operator $A^{W,Y}$ can be extended to $C([0,T]; H_q^{-\beta})$. This extension is performed in Subsection 5.1 below. In Subsection 5.2 we will show existence (and uniqueness) of a solution to BSDE (1) according to Definition 3.3 when b is a rough driver.

5.1. **Preliminary properties.** In this section we show how to extend the operator $A^{W,Y}$ to rough drivers. Let us focus on the smooth case for a moment. The first key observation is that in the Markovian setting we can rewrite $A^{W,Y}$ in terms of $A^{W,W}$, where we recall that

$$A^{W,W} : C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d) \to \mathcal{C} \tag{18}$$

is the integral operator from Definition 3.1 when $Y = W$ and $\mathcal{C}$ is the space of continuous paths on $[0,T]$ with values in $\mathbb{R}^d$. In the special case when $Y = W$, the covariation is a multiple of the d-dimensional identity matrix I_d, so that $\mathrm{d}[W,W]_r = I_d \mathrm{d}r$. In particular this means that for any $l \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ we have

$$A_t^{W,W}(l) = \left(\int_0^t l^*(r, W_r) I_d \mathrm{d}r \right)^* = \int_0^t l(r, W_r) \mathrm{d}r, \tag{19}$$

for all $t \in [0,T]$. To see that $A^{W,Y}$ can be written in terms of $A^{W,W}$ in the Markovian case, suppose that there exists a function $\gamma \in C^{0,1}$ such that $Y_t = \gamma(t, W_t)$, hence by Corollary 2.2 we have $[\gamma(\cdot, W), W]_t = \int_0^t \nabla\gamma^*(r, W_r)\mathrm{d}r$ and so $[W, Y]_t = [W, \gamma(\cdot, W)]_t = \int_0^t \nabla\gamma(r, W_r)\mathrm{d}r$. Thus for any smooth driver l in $C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ we have the following representation for the integral operator:

$$\begin{aligned} A_t^{W,Y}(l) &= \left(\int_0^t l^*(r, W_r) \mathrm{d}[W, Y]_r \right)^* \\ &= \left(\int_0^t l^*(r, W_r) \nabla\gamma(r, W_r) \mathrm{d}r \right)^* \\ &= \int_0^t \nabla\gamma^*(r, W_r) l(r, W_r) \mathrm{d}r \\ &= A_t^{W,W}(\nabla\gamma^* \, l). \end{aligned} \tag{20}$$

By Theorem 4.5 $u \in C^{0,1}$ and so equation (20) holds true also in the case where γ is replaced by the solution u of PDE (15).

Before going into details on the extension of $A^{W,Y}$ we state a useful density result, the proof of which is postponed to the Appendix.

Lemma 5.1. *We have $C_c^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d) \subset C([0,T];H_r^s)$ for any $-\frac{1}{2} < s \leq 0$ and $2 \leq r < \infty$, and the inclusion is dense.*

Remark 5.2. *In Lemma 5.1 one can replace C_c^∞ with the larger space C_c and therefore obtain that also the space $C_c([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ is dense in $C([0,T];H_r^s)$.*

The next result provides us with an explicit representation (chain rule) of $A^{W,W}$ for smooth l, and this representation will still hold in the rough case.

Proposition 5.3 (Chain rule - smooth case)**.** *Let Assumption 2.5 hold, let $l \in C_c([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ and $\Psi \in H_p^{1+\delta}$. Let us denote by ϕ the function given by the expression* (14)*. Then for the integral operator $A^{W,W}$ given in* (18) *we have the representation*

$$A_t^{W,W}(l) = \phi(t,W_t) - \phi(0,W_0) - \int_0^t \nabla\phi^*(r,W_r)\mathrm{d}W_r, \qquad (21)$$

for all $t \in [0,T]$.

We note that the structure of the representation (21) does not change when Ψ changes (although obviously the actual function ϕ changes when Ψ changes).

Proof. Let $(l_n)_n$ be a sequence in $C_c^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ such that $l_n \to l$ in $C([0,T];H_p^{-\beta})$, which can be constructed by Lemma 5.1 since $-\frac{1}{2} < -\beta \leq 0$ and $2 \leq p \leq \infty$ by Assumption 2.5. Moreover a similar approximation can be done for Ψ, namely since C_c^∞ is dense in $H_p^{1+\delta}$ (see Step 1 of the proof of Lemma 5.1) we can also construct a sequence $(\Psi_n) \subset C_c^\infty$ such that $\Psi_n \to \Psi$ in $H_p^{1+\delta}$. Let ϕ_n denote the expression (14), where l is replaced by l_n and Ψ by Ψ_n. Then ϕ_n is at least of class $C^{1,2}$ on $[0,T]\times\mathbb{R}^d$ by Lemma 4.1. Given the expression (19), the

PDE (13) and Itô's formula we get

$$
\begin{aligned}
A_t^{W,W}(l_n) &= \int_0^t l_n(r, W_r)\mathrm{d}r \\
&= \int_0^t \left(\partial_t\phi_n(r, W_r) + \frac{1}{2}\Delta\phi_n(r, W_r)\right)\mathrm{d}r \\
&= \phi_n(t, W_t) - \phi_n(0, W_0) - \int_0^t \nabla\phi_n^*(r, W_r)\mathrm{d}W_r
\end{aligned}
\tag{22}
$$

for $0 \le t \le T$. By Corollary 4.4 we have that $\phi_n \to \phi$ in $C^{0,1}$, thus applying Lemma 5.4 below we conclude that

$$
\begin{aligned}
A_\cdot^{W,W}(l) &= \lim_{n\to\infty} A_\cdot^{W,W}(l_n) \\
&= \lim_{n\to\infty} \left(\phi_n(\cdot, W_\cdot) - \phi_n(0, W_0) - \int_0^\cdot \nabla\phi_n^*(r, W_r)\mathrm{d}W_r\right) \\
&= \phi(\cdot, W_t) - \phi(0, W_0) - \int_0^\cdot \nabla\phi^*(r, W_r)\mathrm{d}W_r
\end{aligned}
$$

and the proof is complete. □

Lemma 5.4. *Let $g_n, g \in C^{0,1}$ such that $g_n \to g$ in the same space. Then*

$$
g_n(\cdot, W_\cdot) - g_n(0, W_0) - \int_0^\cdot \nabla g_n^*(r, W_r)\mathrm{d}W_r \tag{23}
$$

converges to

$$
g(\cdot, W_\cdot) - g(0, W_0) - \int_0^\cdot \nabla g^*(r, W_r)\mathrm{d}W_r
$$

u.c.p. in $\mathcal{C}$.

Proof. Obviously it is enough to consider $g = 0$. Clearly $g_n(\cdot, W)$ converges uniformly to 0 a.s., and in particular uniformly in probability. Setting $f_n = \nabla g_n^*$ it remains to show that

$$
\int_0^\cdot f_n(r, W_r)\mathrm{d}W_r \to 0 \text{ u.c.p.}
$$

According to [14, Proposition 2.26] it is enough to show that

$$
\int_0^T |f_n(r, W_r)|^2\mathrm{d}r \to 0 \tag{24}
$$

in probability. Now $f_n \to 0$ uniformly on each compact by assumption, which implies that (24) holds a.s. □

The following proposition will be used to extend the integral operator $A^{W,W}$ to the space of rough drivers, see Remark 5.6, part (1).

Proposition 5.5. *The operator $A^{W,W}$ (defined in Definition 3.1 in the special case $Y = W$) is continuous with respect to the topology $C([0,T]; H_p^{-\beta})$.*

Proof. Let $(l_n)_n \subset C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ be a sequence such that $l_n \to 0$ in $C([0,T]; H_p^{-\beta})$. Let ϕ_n be given by (14) with l replaced by l_n. By Corollary 4.4 we get $\phi_n \to 0$ in $C^{0,1}$. Using the chain rule (Proposition 5.3) and taking the u.c.p.-limit in $\mathcal{C}$ as $n \to \infty$ we get by Lemma 5.4

$$\begin{aligned}\lim_{n\to\infty} A_\cdot^{W,W}(l_n) &= \lim_{n\to\infty} \left(\phi_n(\cdot, W) - \phi_n(0, W_0) - \int_0^\cdot \nabla\phi_n^*(r, W_r) \mathrm{d}W_r \right) \\ &= 0.\end{aligned}$$

The continuity of the integral operator $A^{W,W}$ at 0 implies the continuity everywhere by linearity. □

In what follows we are interested in drivers $b \in C([0,T]; H_q^{-\beta})$, so we would like to extend the operator $A^{W,Y}$ to $b \in C([0,T]; H_q^{-\beta})$. This will be done by using the operator $A^{W,W}$, which will be calculated in $\nabla\gamma^* b$ for some appropriate function γ, and $\nabla\gamma^* b$ will belong to $C([0,T]; H_p^{-\beta})$. For this reason we start by extending the operator $A^{W,W}$ to the space $E = C([0,T]; H_p^{-\beta})$, as explained below.

Remark 5.6.

1. *By Lemma 5.1 and Proposition 5.5 we can extend the operator $A^{W,W}$ continuously to $E = C([0,T]; H_p^{-\beta})$, where the parameters p and $-\beta$ are chosen according to Assumption 2.5. So $A^{W,W}$ is well-defined according to Definition 3.2.*
2. *Clearly the extended operator $A^{W,W}$ defined in Remark 5.6 part (1) is continuous, i.e. we have*
$$A_\cdot^{W,W}(l) = \lim_{n\to\infty} A_\cdot^{W,W}(l_n)$$
in $\mathcal{C}$ for any sequence $(l_n)_n$ such that $l_n \to l$ in $C([0,T]; H_p^{-\beta})$.

We can now easily prove the chain rule in the rough case, thus we get an explicit representation of $A_t^{W,W}(l)$ in terms of the solution ϕ of equation (13) when $l \in C([0,T]; H_p^{-\beta})$.

Proposition 5.7 (Chain rule - rough case)**.** *Let Assumption 2.5 hold, $l \in C([0,T]; H_p^{-\beta})$ and ϕ be given by (14) for a terminal condition $\Psi \in H_p^{1+\delta}$. Then for all $t \in [0,T]$ we have the representation*

$$A_t^{W,W}(l) = \phi(t, W_t) - \phi(0, W_0) - \int_0^t \nabla\phi^*(r, W_r) \mathrm{d}W_r. \qquad (25)$$

Moreover $A^{W,W}(l)$ is an orthogonal martingale process.

Note that this chain rule does not depend on the actual Ψ chosen, in particular we can pick $\Psi = 0$ or $\Psi = \Phi$.

Proof. By Lemma 5.1 we can take a sequence $l_n \to l$ in $C([0,T]; H_p^{-\beta})$ such that $(l_n)_n \subset C_c^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$. By Remark 5.6 part (2) and the chain rule for the smooth case (Proposition 5.3) we get

$$\begin{aligned} A_\cdot^{W,W}(l) &= \lim_{n\to\infty} A_\cdot^{W,W}(l_n) \\ &= \lim_{n\to\infty} \left(\phi_n(\cdot, W) - \phi_n(0, W_0) - \int_0^\cdot \nabla\phi_n^*(r, W_r) \mathrm{d}W_r \right). \end{aligned}$$

Moreover we can apply Corollary 4.4 to ϕ_n because indeed $l_n \to l$ in $C([0,T]; H_p^{-\beta})$ and thus $\phi_n \to \phi$ in $C^{0,1}$. Finally by Lemma 5.4 we can take the u.c.p. limit in $\mathcal{C}$ when $n \to \infty$ and we get

$$A_\cdot^{W,W}(l) = \phi(\cdot, W) - \phi(0, W_0) - \int_0^\cdot \nabla\phi^*(r, W_r) \mathrm{d}W_r.$$

To show that $A^{W,W}(l)$ is an orthogonal martingale process we use the representation (25) and calculate the covariation of each term on the right-hand side with an arbitrary continuous $\mathcal{F}$-local martingale N with values in $\mathbb{R}^d$. By Corollary 2.2

$$[\phi(\cdot, W) - \phi(0, W_0), N]_t = [\phi(\cdot, W), N]_t = \int_0^t \nabla\phi^*(r, W_r) \mathrm{d}[W, N]_r,$$

having used the fact that $\phi \in C^{0,1}$. Since the covariation operator extends the one of semimartingales, the covariation of N and the last term on the right-hand side of (25) gives

$$[-\int_0^\cdot \nabla\phi^*(r, W_r) \mathrm{d}W_r, N]_t = -\int_0^t \nabla\phi^*(r, W_r) \mathrm{d}[W, N]_r,$$

thus $[A^{W,W}(l), N]_t = 0$ as required. □

The next lemma is a continuity result that will be used in Proposition 5.9 to show the extension of the operator $A^{W,Y}$ to $C([0,T]; H_q^{-\beta})$.

Lemma 5.8. *Let $\gamma \in C([0,T]; H_p^{1+\delta})$. For any sequence $(l_n)_n \subset C([0,T]; H_q^{-\beta})$ such that $l_n \to l$ in $C([0,T]; H_q^{-\beta})$, then $\nabla\gamma^* \, l$ is an element of $C([0,T]; H_p^{-\beta})$ and $\nabla\gamma^* \, l_n \to \nabla\gamma^* \, l$ in the same space.*

Proof. In the space $H_p^{-\beta}$ the norm of the pointwise product for each t

$$\|\nabla\gamma^*(t) l_n(t) - \nabla\gamma^*(t) l(t)\|_{H_p^{-\beta}} = \|\nabla\gamma^*(t)(l_n(t) - l(t))\|_{H_p^{-\beta}}$$

is bounded by $c\|\nabla\gamma^*(t)\|_{H_p^{1+\delta}}\|l_n(t)-l(t)\|_{H_q^{-\beta}}$ thanks to Lemma 2.3 applied to each component. Taking the supremum over time $t\in[0,T]$ we get

$$\sup_{t\in[0,T]}\|\nabla\gamma^*(t)(l_n-l)(t)\|_{H_p^{-\beta}}\leq c\|\gamma^*\|_{C([0,T];H_p^{1+\delta})}\|l_n-l\|_{C([0,T];H_q^{-\beta})}$$

and the right-hand side goes to zero as $n\to\infty$ by assumption. This concludes the proof. □

Proposition 5.9. *Let Assumption 2.5 hold. Suppose $Y_t=\gamma(t,W_t)$ for some $\gamma\in C([0,T];H_p^{1+\delta})$. Then the map $A^{W,Y}$ is well-defined in the sense of Definition 3.2 with $E=C([0,T];H_q^{-\beta})$ and*

$$A^{W,Y}(l)=A^{W,W}(\nabla\gamma^*\,l), \tag{26}$$

for all $l\in E$.

Proof. We start by observing that $C_c([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ is dense in $E=C([0,T];H_q^{-\beta})$ by Lemma 5.1. Moreover $A^{W,W}$ is well-defined in $C([0,T];H_p^{-\beta})$ by Remark 5.6 part (1) and it is continuous. Let $l_n\to l$ in E. We want to prove that $A^{W,Y}(l_n)$ converges to the RHS of (26). Taking into account (20) and the fact that $l_n\in C_c([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ we have

$$A^{W,Y}(l_n)=A^{W,W}(\nabla\gamma^*\,l_n).$$

Note that the map $l\mapsto\nabla\gamma^*\,l$ is continuous from $C([0,T];H_q^{-\beta})$ to $C([0,T];H_p^{-\beta})$ thus $A^{W,W}(\nabla\gamma^*\,l_n)\to A^{W,W}(\nabla\gamma^*\,l)$ in $\mathcal{C}$ because of compositions of continuous maps. This concludes the proof. □

Remark 5.10. *We observe that in* [13] *the authors deal with the singular integral term $\int_0^t Z_s b(s,W_s)\mathrm{d}s$ by replacing it with known terms. In particular, they* define it using the chain rule (25) *with $l=\nabla u^*\,b$ but without proving it. Obviously their virtual solution coincide with the one constructed here.*

Finally we end this section with a result on classical drivers g. We show that for a function $g\in C([0,T];H_r^s)\cap L^\infty_{\mathrm{loc}}([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ with $-\frac{1}{2}<s\leq 0$ and $2\leq r<\infty$, then the operator $A^{W,W}$ defined as an extension to $E=C([0,T];H_r^s)$ and evaluated in g coincides with the classical integral $\int_0^\cdot g(s,W_s)\mathrm{d}s$.

Theorem 5.11. *Let $g\in C([0,T];H_r^s)\cap L^\infty_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ with $-\frac{1}{2}<s\leq 0$ and $2\leq r<\infty$, with g column vector. Suppose that $A^{W,W}$ is well-defined in the sense of Definition 3.2 with $E=C([0,T];H_r^s)$. Then*

$$A^{W,W}_\cdot(g)=\int_0^\cdot g(s,W_s)\mathrm{d}s. \tag{27}$$

Note that the operator $A^{W,W}$ is well-defined for example if $s=-\beta$ and $r=p$ see Remark 5.6.

Proof. The proof is split in two steps. In Step 1 we show that (27) holds for $g\in C([0,T];H^s_r)\cap L^\infty_{\text{loc}}([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ and bounded functions with compact support. In Step 2 we treat the general case.

The proof is written for real-valued functions, and can be applied component by component.

Step 1. *g bounded function with compact support.*
We consider a sequence $\phi_N:\mathbb{R}^d\to\mathbb{R}$ of mollifiers converging to the Dirac measure and for each N we define an operator P_N acting on $h\in H^s_r(\mathbb{R}^d)$ by

$$P_N h := (h*\phi_N).$$

It is easy to show that for every $h\in H^s_r(\mathbb{R}^d)$ then P_N and $A^{-s/2}:=(I-\frac{1}{2}\Delta)^{-s/2}$ commute, that is

$$P_N(A^{-s/2}h)=A^{-s/2}(P_N h). \tag{28}$$

Indeed by the definition of the norm in the $H^s_r(\mathbb{R}^d)$-spaces, we have $A^{-s/2}h\in L^r(\mathbb{R}^d)$. Denoting by $\mathcal{F}$ the Fourier transform in $L^r(\mathbb{R}^d)$ we have

$$\begin{aligned}\mathcal{F}\left(A^{-s/2}(P_N h)\right)(\xi)&=\left(1+\frac{\xi^2}{2}\right)^{-s/2}\mathcal{F}(P_N h)(\xi)\\&=\left(1+\frac{\xi^2}{2}\right)^{-s/2}\mathcal{F}(h)(\xi)\mathcal{F}(\phi_N)(\xi)\\&=\mathcal{F}\left(A^{-s/2}h\right)(\xi)\mathcal{F}(\phi_N)(\xi).\end{aligned}$$

Taking the inverse Fourier transform on both sides we obtain the commutation property as stated in (28). Now it easily follows that

$$P_N h\to h \text{ in } H^s_r(\mathbb{R}^d), \text{ as } N\to\infty \tag{29}$$

for every $h\in H^s_r(\mathbb{R}^d)$, using the definition of the norm in the fractional Sobolev spaces, the property that $P_N f\to f$ in $L^r(\mathbb{R}^d)$ for f in the latter space (in particular for $f=A^{-s/2}h$) and the commutation property (28). Moreover P_N is a contraction in the same spaces, namely

$$\|P_N h\|_{H^s_r(\mathbb{R}^d)}\le\|h\|_{H^s_r(\mathbb{R}^d)}. \tag{30}$$

This can be seen by observing that

$$\|P_N h\|_{H^s_r(\mathbb{R}^d)}=\|A^{-s/2}(P_N h)\|_{H^0_r(\mathbb{R}^d)}=\|P_N(A^{-s/2}h)\|_{H^0_r(\mathbb{R}^d)},$$

where we have used (28), and the latter is bounded by $\|A^{-s/2}h\|_{H^0_r(\mathbb{R}^d)}=\|h\|_{H^s_r(\mathbb{R}^d)}$ because P_N is a contraction operator in $H^0_r(\mathbb{R}^d)=L^r(\mathbb{R}^d)$. Property (30) is applied to $h=g(t,\cdot)$ for all $t\in[0,T]$ to show that the function $t \mapsto P_N g(t,\cdot)$ is continuous from $[0,T]$ to $H^s_r(\mathbb{R}^d)$. Indeed for any sequence $t_k \to t$ we have

$$\begin{aligned}\|P_N g(t_k,\cdot) - P_N g(t,\cdot)\|_{H^s_r(\mathbb{R}^d)} &= \|P_N(g(t_k,\cdot) - g(t,\cdot))\|_{H^s_r(\mathbb{R}^d)} \\ &\le \|g(t_k,\cdot) - g(t,\cdot)\|_{H^s_r(\mathbb{R}^d)},\end{aligned}$$

which goes to zero by assumption on g. To show that

$$P_N g \to g \tag{31}$$

in $C([0,T]; H^s_r(\mathbb{R}^d))$, we use Lemma A.1 with $H = H^s_r(\mathbb{R}^d)$. We can do so since the family of operators $(P_N)_N$ is linear and equibounded in $H^s_r(\mathbb{R}^d)$ by (30), and it fulfils (29). Thus defining the compact K in $H^s_r(\mathbb{R}^d)$ by $K := \{g(t) : t \in [0,T]\}$ we have

$$\sup_{a\in K} \|P_N a - a\|_{H^s_r(\mathbb{R}^d)} = \sup_{0\le t\le T} \|P_N g(t,\cdot) - g(t,\cdot)\|_{H^s_r(\mathbb{R}^d)}$$

and by Lemma A.1 the quantity above converges to 0 as $N \to \infty$. At this point we observe that $P_N g \in C_c([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ because both g and ϕ_N have compact support. Therefore $A^{W,W}(P_N g)$ is well-defined and (27) holds for g replaced by $P_N g$ thanks to (19). Moreover by (31) we can apply Remark 5.6, part (2) and get

$$\lim_{N\to\infty} A^{W,W}(P_N g) = A^{W,W}(g) \text{ in } \mathcal{C}. \tag{32}$$

Finally we can see that

$$\int_0^{\cdot} P_N g(s, W_s)\mathrm{d}s \to \int_0^{\cdot} g(s, W_s)\mathrm{d}s \tag{33}$$

$u.c.p$ when $N \to \infty$. Indeed

$$\begin{aligned} &E\left[\sup_{0\le t\le T} |\int_0^t (P_N g - g)(s, W_s)\mathrm{d}s|\right] \\ &\le E\left[\sup_{0\le t\le T} \int_0^t |P_N g - g|(s, W_s)\mathrm{d}s\right] \\ &\le \int_0^T \int_{\mathbb{R}^d} |P_N g - g|(s,y) p_s(y)\mathrm{d}y\mathrm{d}s, \end{aligned} \tag{34}$$

where $p_s(y)$ is the mean-zero Gaussian density in $\mathbb{R}^d$ with variance s. Now for almost all $(s,x) \in [0,T] \times \mathbb{R}^d$ we have $|P_N g(s,x)| \le \|g\|_{L^\infty([0,T]\times\mathbb{R}^d)} \le C$, because g is bounded by assumption. This, together with the fact that

$$\int_{[0,T]\times\mathbb{R}^d} p_s(y)\mathrm{d}s\mathrm{d}y = T$$

implies that (34) is bounded by $2CT$. Moreover for almost all $(s,x) \in [0,T] \times \mathbb{R}^d$ we also have

$$(P_N g - g)(s,x) \to 0.$$

By Lebesgue dominated convergence theorem the RHS of (34) converges to 0. This implies (33) and with (32) we conclude.
Step 2. *General case $g \in C([0,T]; H^s_r(\mathbb{R}^d)) \cap L^\infty_{loc}([0,T] \times \mathbb{R}^d)$.*
Let us define $\tau_M := \inf\{t \geq 0 \text{ such that } |W_t| > M\}$. Clearly $\tau_M \to \infty$ a.s. as $M \to \infty$. Moreover we define a family of smooth functions

$$\chi_M(x) = \begin{cases} 1 & \text{if } |x| \leq M \\ 0 & \text{if } |x| \geq M+1 \end{cases}$$

and with $0 \leq \chi_M(x) \leq 1$. Then we set $g_M(s,x) := g(s,x)\chi_M(x)$. It is clear that $g_M(s,W_s) = g(s,W_s)$ for all ω and for all $s \leq t \wedge \tau_M$ for any arbitrary t, hence

$$\int_0^{t\wedge\tau_M} g(s,W_s)\mathrm{d}s = \int_0^{t\wedge\tau_M} g_M(s,W_s)\mathrm{d}s. \tag{35}$$

On the other hand we know that g_M is bounded and has compact support by definition, and that $g_M \in C([0,T]; H^s_r(\mathbb{R}^d))$ because g is in the same space and χ_M is smooth (using the pointwise multipliers property, see [25, Section 2.2.2]). So Step 1 applies to g_M

$$A^{W,W}(g_M) = \int_0^{\cdot} g_M(s,W_s)\mathrm{d}s.$$

and in particular it holds for the time $t \wedge \tau_M$, that is

$$A^{W,W}_{t\wedge\tau_M}(g_M) = \int_0^{t\wedge\tau_M} g_M(s,W_s)\mathrm{d}s. \tag{36}$$

Now we want to show that

$$A^{W,W}_{\cdot\wedge\tau_M}(g_M) = A^{W,W}_{\cdot\wedge\tau_M}(g). \tag{37}$$

To this aim, let us consider an approximating sequence $(g^n)_n$ of g in $C_c([0,T] \times \mathbb{R}^d)$, which exists due to Lemma 5.1. Then we set $g^n_M := g^n\chi_M$ for each n, and this is an approximating sequence for g_M in $C([0,T]; H^s_r(\mathbb{R}^d))$. Indeed the linear map $\phi \mapsto \phi\chi_M$ is continuous in $C([0,T]; H^s_r(\mathbb{R}^d))$ by [25, equation (2.50)], namely there exists a constant $c(M)$ only dependent on χ_M such that

$$\|\phi\chi_M\|_{H^s_r(\mathbb{R}^d)} \leq c(M)\|\phi\|_{H^s_r(\mathbb{R}^d)}.$$

Then

$$\begin{aligned}\|g^n_M - g_M\|_{C([0,T];H^s_r(\mathbb{R}^d))} &= \|g^n\chi_M - g\chi_M\|_{C([0,T];H^s_r(\mathbb{R}^d))}\\ &= \sup_{0\le t\le T}\|(g^n(t,\cdot)-g(t,\cdot))\chi_M\|_{H^s_r(\mathbb{R}^d)}\\ &\le c(M)\sup_{0\le t\le T}\|g^n(t,\cdot)-g(t,\cdot)\|_{H^s_r(\mathbb{R}^d)}\\ &= c(M)\|g^n-g\|_{C([0,T];H^s_r(\mathbb{R}^d))},\end{aligned}$$

and since g^n converges to g in $C([0,T];H^s_r(\mathbb{R}^d))$ then so does g^n_M to g_M. For each n we have

$$A^{W,W}_{\cdot\wedge\tau_M}(g^n_M) = A^{W,W}_{\cdot\wedge\tau_M}(g^n) \tag{38}$$

because both sides are defined explicitly and the two functions coincide before τ_M. We note that $A^{W,W}_\cdot(g^n_M)$ (resp. $A^{W,W}_\cdot(g^n)$) converges u.c.p. to $A^{W,W}_\cdot(g_M)$ (resp. $A^{W,W}_\cdot(g)$) as $n\to\infty$. The truncated processes, which are the left-hand side and the right hand-side of (38) also converge u.c.p., hence we get (37). This, together with (35) and (36) gives

$$\int_0^{\cdot\wedge\tau_M} g(s,W_s)\mathrm{d}s = A^{W,W}_{\cdot\wedge\tau_M}(g). \tag{39}$$

For almost all ω there exists $n_0(\omega)$ such that for all $M>n_0(\omega)$ we have $\tau_M(\omega)\ge T$, then taking the limit as $M\to\infty$ of (39) we conclude. □

Corollary 5.12 (chain rule for L^∞_{loc}). *If $g\in L^\infty_{loc}([0,T]\times\mathbb{R}^d;\mathbb{R}^d)\cap C([0,T];H^{-\beta}_p)$ then*

$$\int_0^t g(s,W_s)\mathrm{d}s = \phi(t,W_t)-\phi(0,W_0)-\int_0^t \nabla\phi^*(s,W_s)\mathrm{d}W_r,$$

where ϕ is the solution of (13) *with $\Psi\in H^{1+\delta}_p$, given by* (14).

Proof. This follows by Theorem 5.11 and Proposition 5.7 with $l=g$. □

5.2. **Existence and Uniqueness.** Here we show that in the Markovian case we have existence and uniqueness of a solution to (1) when $b\in C([0,T];H^{-\beta}_q)$. In particular, in Theorem 5.13 we construct a solution to BSDE (1) with $\xi=\Phi(W_T)$ using the solution to the associated PDE (15), and we show that such solution is unique in Theorem 5.14 in the class of solutions Y that can be written as $Y_t=\gamma(t,W_t)$.

For ease of reading, we rewrite the formal meaning of the BSDE (1) under consideration:

$$Y_t=\Phi(W_T)+\int_t^T Z_r b(r,W_r)\mathrm{d}r+\int_t^T f(r,W_r,Y_r,Z_r)\mathrm{d}r-\int_t^T Z_r\mathrm{d}W_r.$$

Theorem 5.13 (Existence)**.** *Let Assumption 2.5 and Assumption 2.6 hold and let* $b \in C([0,T]; H_q^{-\beta})$*. We denote by* u *be the unique mild solution to* (15)*. Then* $Y_t = u(t, W_t)$ *is a solution of* (1) *according to Definition 3.3 with* $E = C([0,T]; H_q^{-\beta})$*.*

Proof. First we observe that thanks to Theorem 4.5 we have $u \in C([0,T]; H_p^{1+\delta})$. Thus by Proposition 5.9 the operator $A^{Y,W}$ appearing in Definition 3.3 is well-defined in $E = C([0,T]; H_q^{-\beta})$ and we have

$$A_t^{W,Y}(b) = A_t^{W,W}(\nabla u^* \, b). \tag{40}$$

This is an orthogonal martingale process by Proposition 5.7 with $l = \nabla u^* \, b$. The latter is an element of $C([0,T]; H_p^{-\beta})$, and this is shown by Lemma 5.8. Moreover $u(T) = \Phi$ implies that $Y_T = u(T, W_T) = \Phi(W_T)$ so that parts (i)-(iii) of Definition 3.3 are verified. The last point to check is part (iv) in the same Definition, namely that

$$M_t := Y_t - Y_0 + A_t^{W,Y}(b) + \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y,W]_r}{\mathrm{d}r}\right) \mathrm{d}r$$

is a square integrable martingale. The term with the driver f becomes

$$\begin{aligned} \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y,W]_r}{\mathrm{d}r}\right) \mathrm{d}r &= \int_0^t f\left(r, W_r, u(r, W_r), \nabla u(r, W_r)\right) \mathrm{d}r \\ &= \int_0^t \tilde{f}(r, W_r) \mathrm{d}r, \end{aligned} \tag{41}$$

where $\tilde{f}(t,x) = f(t, x, u(t,x), \nabla u(t,x))$. Since $u \in C^{0,1}$ and f is continuous then $\tilde{f} \in L^\infty_{\text{loc}}([0,T] \times \mathbb{R}^d)$. We also have that $\tilde{f} \in C([0,T]; L^p)$ since f is Lipschitz in (y,z) uniformly in t, x, and $x \mapsto f(t,x,0,0)$ is an element of L^p uniformly in $t \in [0,T]$ by Assumption 2.6 and $u(t), \nabla u(t)$ are in L^p uniformly in t since $u \in C([0,T]; H_p^{1+\delta})$. So in particular $\tilde{f} \in C([0,T]; H_p^0)$ and hence by Theorem 5.11 we have

$$\int_0^t \tilde{f}(r, W_r) \mathrm{d}r = A_t^{W,W}(\tilde{f}).$$

Moreover by (40) and the linearity of $A^{W,W}$ one gets

$$\begin{aligned} M_t &= Y_t - Y_0 + A_t^{W,W}(\nabla u^* \, b) + A_t^{W,W}(\tilde{f}) \\ &= Y_t - Y_0 + A_t^{W,W}(\nabla u^* \, b + \tilde{f}) \\ &= Y_t - Y_0 - A_t^{W,W}(-\nabla u^* \, b - \tilde{f}). \end{aligned}$$

Now we apply the chain rule to $A_t^{W,W}(-\nabla u^* \, b - \tilde{f})$, namely Proposition 5.7 with $l = -\nabla u^* \, b - \tilde{f}$ on the RHS of (13). Note that in this case (25) holds for $\phi = u$ because the function u verifies (14) with $l = -\nabla u^* b - \tilde{f}$, see indeed (16). Thus we get

$$\begin{aligned} M_t =& Y_t - Y_0 - A_t^{W,W}(-\nabla u^* b - \tilde{f}) \\ =& u(t, W_t) - u(0, W_0) \\ & - u(t, W_t) + u(0, W_0) + \int_0^t \nabla u^*(r, W_r) \mathrm{d}W_r \end{aligned}$$

so that

$$M_t = \int_0^t \nabla u^*(r, W_r) \mathrm{d}W_r,$$

which is clearly a square integrable $\mathcal{F}$-martingale because ∇u^* is uniformly bounded since $u \in C^{1+\alpha}$ by Theorem 4.5. □

Theorem 5.14 (Uniqueness in the class $Y_t = \gamma(t, W_t)$)**.** *Let Assumption 2.5 and Assumption 2.6 hold and let $b \in C([0,T]; H_q^{-\beta})$. If the solution of (1) according to Definition 3.3 with $E = C([0,T]; H_q^{-\beta})$ can be written as $Y_t = \gamma(t, W_t)$ for some $\gamma \in C([0,T]; H_p^{1+\delta})$, then it is unique.*

Proof. Suppose that $Y_t^i = \gamma^i(t, W_t)$, $i = 1, 2$ are solutions to (1) according to Definition 3.3 and let us denote by

$$M_t^i := Y_t^i - Y_0^i + A_t^{W,Y^i}(b) + \int_0^t f\left(r, W_r, Y_r^i, \frac{\mathrm{d}[Y^i, W]_r}{\mathrm{d}r}\right) \mathrm{d}r, \quad (42)$$

which is a martingale by part (iv) of Definition 3.3. Moreover

$$(\nabla \gamma^i)^* b \in C([0,T]; H_p^{-\beta}) \quad (43)$$

by Lemma 5.8. By assumption on Y^i we can apply Proposition 5.9 and write

$$A_t^{Y^i,W}(b) = A_t^{W,W}((\nabla \gamma^i)^* b). \quad (44)$$

Furthermore by Corollary 2.2 we have

$$\begin{aligned} &\int_0^t f\left(r, W_r, Y_r^i, \frac{\mathrm{d}[Y^i, W]_r}{\mathrm{d}r}\right) \mathrm{d}r \\ &= \int_0^t f\left(r, W_r, \gamma^i(r, W_r), \nabla \gamma^i(r, W_r)\right) \mathrm{d}r \\ &= \int_0^t \tilde{f}^i(r, W_r) \mathrm{d}r, \end{aligned} \quad (45)$$

where $\tilde{f}^i(t,x) := f(t, x, \gamma^i(t,x), \nabla \gamma^i(t,x))$. We note that

$$\tilde{f}^i \in L^\infty_{\mathrm{loc}} \cap C([0,T]; L^p), \quad (46)$$

which can be proven similarly to the considerations below (41) in the proof of the previous existence theorem. Thus we can apply Theorem 5.11, so (45) $= A_t^{W,W}(\tilde{f}^i)$. By (44) and the additivity of $A^{W,W}$ we have

$$M_t^i = Y_t^i - Y_0^i + A_t^{W,W}((\nabla\gamma^i)^* \, b + \tilde{f}^i). \tag{47}$$

Let us consider the PDE

$$\begin{cases} \partial_t h^i(t) + \frac{1}{2}\Delta h^i(t) = (\nabla\gamma^i)^*(t)\, b(t) + f(t, \cdot, \gamma^i, \nabla\gamma^i) \\ h^i(T) = 0, \end{cases} \tag{48}$$

which is PDE (13) with $(\nabla\gamma^i)^*(t)\, b(t) + f(t,\cdot,\gamma^i,\nabla\gamma^i) = (\nabla\gamma^i)^*(t)\, b(t) + \tilde{f}^i(t,\cdot) \in C([0,T]; H_p^{-\beta})$ (by (46) and (43)) on the right-hand side in place of l. We denote by h^i, $i = 1, 2$ the corresponding (mild solution) expression (14), which belongs to $C([0,T]; C^{1+\alpha})$ by Lemma 4.3. Then $(\nabla h^i)^*$ is bounded. By the chain rule (Proposition (5.7)) we get

$$A_t^{W,W}((\nabla\gamma^i)^* \, b + \tilde{f}^i) = h^i(t, W_t) - h^i(0, W_0) - \int_0^t (\nabla h^i)^*(r, W_r)\mathrm{d}W_r. \tag{49}$$

Plugging (49) into (47) we get

$$\begin{aligned} M_t^i =& \gamma^i(t, W_t) - \gamma^i(0, W_0) + h^i(t, W_t) - h^i(0, W_0) \\ & - \int_0^t (\nabla h^i)^*(r, W_r)\mathrm{d}W_r. \end{aligned}$$

Subtracting M_T^i from both sides and rearranging the terms we obtain

$$\begin{aligned} \gamma^i(t, W_t) + h^i(t, W_t) =& -(M_T^i - M_t^i) - \int_t^T (\nabla h^i)^*(r, W_r)\mathrm{d}W_r \\ & + \gamma^i(T, W_T) + h^i(T, W_T) \\ =& \Phi(W_T) - (\tilde{M}_T^i - \tilde{M}_t^i), \end{aligned} \tag{50}$$

where we have set $\tilde{M}_t^i := M_t^i + \int_0^t \nabla h^i(r, W_r)\mathrm{d}W_r$ and we have used the fact that $h^i(T, W_T) = 0$ by (48) and that $\gamma^i(T, W_T) = \Phi(W_T)$ by item (iii) of Definition 3.3. Clearly $\tilde{M}^i$ is another martingale since $(\nabla h^i)^*$ is bounded. So the left-hand side of equality (50) can be represented by

$$\gamma^i(t, W_t) + h^i(t, W_t) = E\left[\Phi(W_T)|\mathcal{F}_t\right].$$

The above equality holds for $i = 1, 2$ and since the right-hand side is the same, we get

$$\gamma^1(t, W_t) + h^1(t, W_t) = \gamma^2(t, W_t) + h^2(t, W_t)$$

almost surely. From this we can infer that

$$\gamma^1(t, x) + h^1(t, x) = \gamma^2(t, x) + h^2(t, x), \tag{51}$$

for every $t \in [0,T]$ and $x \in \mathbb{R}^d$ in the following way: suppose that we have a continuous function η such that $\eta(t, W_t) = 0$ almost surely. Then

$$0 = E[|\eta(t, W_t)|] = \int_{[0,T]\times\mathbb{R}^d} |\eta(t,x)| p_t(x) \mathrm{d}t\mathrm{d}x$$

and since $p_t(x) > 0$ we get that $\eta(t,x) = 0$ almost everywhere. In fact this holds everywhere because η is continuous. Setting $\gamma^i(t) := \gamma^i(t,\cdot)$ for all $t \in [0,T]$ and $i = 1,2$, it remains to show that $\gamma^1 = \gamma^2$. We know that $\gamma^1(t) - \gamma^2(t) = h^2(t) - h^1(t)$ by (51). The idea is to bound the difference $h^2 - h^1$ in the ρ-equivalent norm for the space $C([0,T]; H_p^{1+\delta})$. To do so we work with the time reversed functions $\hat{h}^i(s) := h^i(T-s)$, which clearly have the same regularity as h^i and also the same norm in $C([0,T]; H_p^{1+\delta})$ and ρ-equivalent norm. Setting $\hat{b}(s) := b(T-s), \hat{\gamma}^i(s) := \gamma^i(T-s)$ and $\hat{f}(s,y,z) := f(T-s,y,z)$ we have

$$\begin{aligned}\hat{h}^2(t) - \hat{h}^1(t) = &\int_0^t P(t-r)\left((\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r))^*\hat{b}(r)\right)\mathrm{d}r \\ &+ \int_0^t P(t-r)\left(\hat{f}(r, \hat{\gamma}^2(r), \nabla\hat{\gamma}^2(r) - \hat{f}(r, \hat{\gamma}^1(r), \nabla\hat{\gamma}^1(r))\right)\mathrm{d}r.\end{aligned}$$

Taking the ρ-equivalent norm (see (7)) of the difference above, we have

$$\begin{aligned}&\|\hat{h}^2 - \hat{h}^1\|^{(\rho)}_{C([0,T];H_p^{1+\delta})} \\ &= \sup_{0\le t\le T} e^{-\rho t}\|\hat{h}^2(t) - \hat{h}^1(t)\|_{H_p^{1+\delta}} \\ &\le \sup_{0\le t\le T} e^{-\rho t}\|\int_0^t P(t-r)\left((\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r))^*\hat{b}(r)\right)\mathrm{d}r\|_{H_p^{1+\delta}} \\ &\quad + \sup_{0\le t\le T} e^{-\rho t}\|\int_0^t P(t-r)\Big(\hat{f}(r, \hat{\gamma}^2(r), \nabla\hat{\gamma}^2(r) \\ &\quad - \hat{f}(r, \hat{\gamma}^1(r), \nabla\hat{\gamma}^1(r))\Big)\mathrm{d}r\|_{H_p^{1+\delta}} \\ &=: (A) + (B).\end{aligned}$$

To bound the first term we use the pointwise product estimate for fixed time $r \in [0,T]$ (Lemma 2.3), the mapping property (4) of the semigroup, and the definition of the ρ-equivalent norm (7). We get

$$
\begin{aligned}
(A) &\leq c \sup_{0\leq t\leq T} \int_0^t e^{-\rho t}(t-r)^{-\frac{1+\delta+\beta}{2}} \|\hat{b}(r)\|_{H_q^{-\beta}} \|\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r)\|_{H_p^\delta} \mathrm{d}r \\
&\leq c\|\hat{b}\|_{C([0,T];H_q^{-\beta})} \sup_{0\leq t\leq T} \int_0^t e^{-\rho(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}} \cdot \\
&\qquad\qquad e^{-\rho r}\|\hat{\gamma}^2(r) - \hat{\gamma}^1(r)\|_{H_p^{1+\delta}} \mathrm{d}r \qquad (52) \\
&\leq c\|\hat{\gamma}^2 - \hat{\gamma}^1\|^{(\rho)}_{C([0,T];H_p^{1+\delta})} \sup_{0\leq t\leq T} \int_0^t e^{-\rho(t-r)}(t-r)^{-\frac{1+\delta+\beta}{2}} \mathrm{d}r \\
&\leq c\rho^{\frac{\delta+\beta-1}{2}} \|\hat{\gamma}^2 - \hat{\gamma}^1\|^{(\rho)}_{C([0,T];H_p^{1+\delta})},
\end{aligned}
$$

having used the Gamma function and the bound

$$\int_0^t e^{-\rho r} r^\alpha \mathrm{d}r \leq \Gamma(\alpha+1)\rho^{-(\alpha+1)}$$

in the latter inequality, with $\alpha = -\frac{1+\delta+\beta}{2}$. Note that $-(\alpha+1) = \frac{\delta+\beta-1}{2} < 0$ so we have $\rho^{\frac{\delta+\beta-1}{2}} \to 0$ as $\rho \to \infty$.
To bound term (B) we do similarly but use the mapping property of the semigroup from H_p^0 to $H_p^{1+\delta}$ and the Lipschitz regularity (17) of $\hat{f}$ so we get

$$
\begin{aligned}
(B) &\leq c \sup_{0\leq t\leq T} \int_0^t e^{-\rho(t-r)}(t-r)^{-\frac{1+\delta}{2}} \cdot \\
&\qquad e^{-\rho r}(c\|\hat{\gamma}^2(r) - \hat{\gamma}^1(r)\|_{H_p^{1+\delta}} + \|\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r)\|_{H_p^\delta}) \mathrm{d}r \qquad (53) \\
&\leq c\rho^{\frac{\delta-1}{2}} \|\hat{\gamma}^2 - \hat{\gamma}^1\|^{(\rho)}_{C([0,T];H_p^{1+\delta})}.
\end{aligned}
$$

Collecting the estimates (52) and (53), we get

$$\|\gamma^1 - \gamma^2\|^{(\rho)}_{C([0,T];H_p^{1+\delta})} \leq c(\rho^{\frac{\delta-1}{2}} + \rho^{\frac{\delta+\beta-1}{2}})\|\gamma^1 - \gamma^2\|^{(\rho)}_{C([0,T];H_p^{1+\delta})},$$

so

$$\|\gamma^1 - \gamma^2\|^{(\rho)}_{C([0,T];H_p^{1+\delta})}(1 - c(\rho^{\frac{\delta-1}{2}} + \rho^{\frac{\delta+\beta-1}{2}})) \leq 0,$$

where c depends on b and T but not on γ^i or ρ. We choose ρ large enough such that $1 - c(\rho^{\frac{\delta-1}{2}} + \rho^{\frac{\delta+\beta-1}{2}}) > 0$, which implies $\gamma^1 = \gamma^2$ and shows shows that $Y^1 = Y^2$. □

Appendix A. A technical lemma and proof of Lemma 5.1

We first state and prove a technical Lemma that is used in the proof of the density below and that has been used in the proof of Theorem 5.11.

Lemma A.1. *Let $(H, \|\cdot\|)$ be a normed space and $(P_N)_N$ be a family of linear equibounded operators on H such that for each $a \in H$ we have $P_N a \to a$ in H. Then for any compact $K \subset H$ we have*

$$\sup_{a\in K} \|P_N a - a\| \to 0,$$

as $N \to \infty$.

Proof. Let $\delta > 0$. Since K is compact, we can construct a finite cover of size δ, for example $K \subseteq \cup_{i=1}^m B(a_i, \delta)$. For a given $a \in H$ there exists $j \in \{1, \dots, m\}$ such that $a \in B(a_j, \delta)$. Then we write

$$\begin{aligned} \|P_N a - a\| &\le \|P_N(a - a_j)\| + \|P_N a_j - a_j\| + \|a_j - a\| \\ &\le (1+c)\|a - a_j\| + \max_{i=1,\dots,m} \|P_N a_i - a_i\| \\ &\le (1+c)\delta + \max_{i=1,\dots,m} \|P_N a_i - a_i\|, \end{aligned}$$

where c is the bound of the operator norms related to P_N. Then $\sup_{a\in K} \|P_N a - a\| \le (1+c)\delta + \max_{i=1,\dots,m} \|P_N a_i - a_i\|$ and so taking the lim sup on both sides we get

$$\limsup_{N\to\infty} \sup_{a\in K} \|P_N a - a\| \le (1+c)\delta$$

since $\lim_{N\to\infty} \|P_N a_i - a_i\| = 0$ for all $i \in \{1, \dots, m\}$. By the fact that δ is arbitrary we get

$$\lim_{N\to\infty} \sup_{a\in K} \|P_N a - a\| = 0$$

as wanted. □

Before proving Lemma 5.1 we introduce the Haar wavelet functions and illustrate their use within the context of fractional Sobolev spaces H^s_r. For simplicity of notation we recall only the case of the Haar wavelets on $\mathbb{R}$ (see [26], Section 2.2, eqn (2.93)–(2.96)) and leave to the reader the extension to $\mathbb{R}^d$ which can be found in Section 2.3 of the same book. We define

$$h_M(x) := \begin{cases} 1 & \text{if } 0 \le x < \frac{1}{2}, \\ -1 & \text{if } \frac{1}{2} \le x < 1, \\ 0 & \text{if } x \notin [0,1), \end{cases}$$

$$h_F(x) := |h_M(x)|, \quad h_{-1,m}(x) := \sqrt{2} h_F(x-m), \; m \in \mathbb{Z},$$

and

$$h_{j,m}(x) := h_M(2^j x - m),\ j \in N_0, m \in \mathbb{Z}.$$

Then the family

$$\{h_{j,m}, j \in \mathbb{N}_0 \cup \{-1\}, m \in \mathbb{Z}\} \tag{54}$$

is an unconditional basis of $H^s_r(\mathbb{R})$ for $2 \leq r \leq \infty$ and $-\frac{1}{2} < s < \frac{1}{r}$ by [26, Theorem 2.9, (ii)]. Note that $r = \infty$ is included here but is not included in Lemma 5.1 because of Step 1 in the proof below. The analogous result in dimension $d \geq 1$ is given in Theorem 2.21, (ii). Moreover for any $h \in H^s_r(\mathbb{R})$ we have the unique representation

$$h = \sum_{j=-1}^{\infty} \sum_{m \in \mathbb{Z}} \mu_{j,m} 2^{-j(s-\frac{1}{r})} h_{j,m}$$

where

$$\mu_{j,m} := 2^{j(s-\frac{1}{r}+1)} \int_{\mathbb{R}} h(x) h_{j,m}(x) \mathrm{d}x, \tag{55}$$

and the integral has to be interpreted as a dual pairing as mentioned in [26, Theorem 2.9], see also [25, Remark 1.14]. Rewriting the same series with a different notation $\bar{\mu}_{j,m} := 2^j \int_{\mathbb{R}} h(x) h_{j,m}(x) \mathrm{d}x$ we get another equivalent representation for h given by

$$h = \sum_{j=-1}^{\infty} \sum_{m \in \mathbb{Z}} \bar{\mu}_{j,m} h_{j,m}. \tag{56}$$

Defining the projector P_N as

$$P_N h := \sum_{j=-1}^{N} \sum_{m=-N}^{N} \bar{\mu}_{j,m} h_{j,m}, \tag{57}$$

for h of the form (56), then clearly $P_N h \in H^s_r(\mathbb{R})$ and

$$\|h - P_N h\|_{H^s_r(\mathbb{R}^d)} \to 0, \tag{58}$$

as $N \to \infty$.

Remark A.2. *We observe that the projector P_N enjoys the bound*

$$\|P_N h\|_{H^s_r(\mathbb{R})} \leq \|h\|_{H^s_r(\mathbb{R})}.$$

This can be seen as follows. We denote by $\mu(h)$ the collection of $\mu_{j,m}$ given by (55) *for some h. Then for $2 \leq r \leq \infty$ the map $h \mapsto \mu(h)$ is an isomorphism between H^s_r and f^-_{r2}, where the latter is a space of sequences. For a precise definition of f^-_{r2}, its norm and the statement of this isomorphism property, see* [26] *in particular, see Section*

2.2.3, Theorem 2.9 for the 1-dimensional case and Section 2.3.2, Theorem 2.21 for the d-dimensional one. Moreover the sequence of coefficients $\mu(P_N h)$ coincide with $\mu(h)$ for all $j, |m| > N$ and is zero otherwise. Thus by definition of the norm of f^-_{r2}, we have $\|\mu_{j,m}(P_N h)\|_{f^-_{r2}} \leq \|\mu_{j,m}(h)\|_{f^-_{r2}}$ and this together with the isomorphism implies $\|P_N h\|_{H^s_r(\mathbb{R})} \leq \|h\|_{H^s_r(\mathbb{R})}$ as stated.

Proof of Lemma 5.1. We will show that the dense inclusion holds for real-valued functions, namely that $C^\infty_c([0,T]\times\mathbb{R}^d) \subset C([0,T]; H^s_r(\mathbb{R}^d))$. To get the full statement it is then enough to apply this result to each component of functions in $C([0,T]; H^s_r)$.

Step 1: *Density of $C^\infty_c(\mathbb{R}^d)$ in $H^s_r(\mathbb{R}^d)$.* It is a known result that $C^\infty_c(\mathbb{R}^d)$ is dense in $H^s_r(\mathbb{R}^d)$ for all $1 < r < \infty$ and $-\infty < s < \infty$. For a proof see for example [24, Theorem in Section 2.3.2, part (b)].

Step 2: *Non-smooth approximating sequence for $l \in C([0,T]; H^s_r(\mathbb{R}^d))$.* We consider $d = 1$ in the proof for simplicity of notation and explanation, but the same methodology extends to the case $d \geq 1$, see for example [26, Section 2.3.1]. We will use here the notation of Section 2.2.2 in the same book, which deals with the case $d = 1$, in particular let $\{h_{j,m}, j \in \mathbb{N}_0 \cup \{-1\}, m \in \mathbb{Z}\}$ be the Haar basis on $L^2(\mathbb{R})$ defined in (54). Now let $l \in C([0,T]; H^s_r(\mathbb{R}))$ and let $t \in [0,T]$. We recall that $(P_N)_N$ defined by (57) is a family of linear operators acting on $H^s_r(\mathbb{R})$. The coefficients $\bar\mu$ of $P_N l(t)$ are now parametrized by time, namely $\bar\mu_{j,m}(t) = 2^j \int_{\mathbb{R}} l(t,x) h_{j,m}(x)\mathrm{d}x$. By (58) we have that $P_N l(t) \to l(t)$ in $H^s_r(\mathbb{R})$ as $N \to \infty$, for all $t \in [0,T]$. It is clear by definition of the coefficients that $t \mapsto \bar\mu_{j,m}(t)$ is continuous and each term $t \mapsto \bar\mu_{j,m}(t)h_{j,m}$ in the finite sum belongs to $C([0,T]; H^s_r(\mathbb{R}))$ hence $P_N l \in C([0,T]; H^s_r(\mathbb{R}))$. We will now show that $P_N l \to l$ in $C([0,T]; H^s_r(\mathbb{R}))$, namely that

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \|l(t) - P_N l(t)\|_{H^s_r(\mathbb{R})} = 0. \tag{59}$$

To prove this, we want to use Lemma A.1 with the compact $K := \{l(t) : t \in [0,T]\} \subset H^s_r(\mathbb{R})$ and the projection P_N defined by (57). The family of functions $t \mapsto P_N l(t)$ is bounded in N in the space $C([0,T]; H^s_r(\mathbb{R}))$ by Remark A.2. Since $\{l(t), t \in [0,T]\}$ is a compact set in H^s_r then we can apply Lemma A.1 and we get (59).

Step 3: *Smoothing of non-smooth approximating sequence.* The last step consists in showing that for any $l_N(t) := P_N l(t)$ from Step 2 and for any $\varepsilon > 0$, we can find an element $t \mapsto \tilde l_N(t)$ which is an element of $C^\infty_c([0,T] \times \mathbb{R})$ and such that $\sup_{t\in[0,T]} \|l_N(t) - \tilde l_N(t)\|_{H^s_r(\mathbb{R})} < \varepsilon$. Then this would conclude the argument and show the density of $C_c^\infty([0,T]\times\mathbb{R})$ in $C([0,T];H_r^s(\mathbb{R}))$.

To find $\tilde{l}_N(\cdot)$ we observe that $l_N(t)$ is a finite sum of terms of the type $\mu_{j,m}(t)h_{j,m}$, where the μs are continuous in time and $h_{j,m}$ is an element of the Haar basis. For each of this terms using Step 1 we can find $\tilde{h}_{j,m}\in C_c^\infty(\mathbb{R})$ such that

$$\|h_{j,m}-\tilde{h}_{j,m}\|_{H_r^s(\mathbb{R})}<\frac{\varepsilon}{\max_{t\in[0,T]}\sum_{j,m}|\mu_{j,m}(t)|},$$

where the sum appearing in the denominator is over the finite set of indices $j\in\{-1,0,\dots,N\}$ and $m\in\{-N,\dots,0,\dots,N\}$. Then we set

$$\tilde{l}_N(t):=\sum_{j,m}\mu_{j,m}(t)\tilde{h}_{j,m},$$

where again the sum over j,m is a finite sum. Then for any $t\in[0,T]$ we have

$$\begin{aligned}\|\tilde{l}_N(t)-l_N(t)\|_{H_r^s(\mathbb{R})}&=\|\sum_{j,m}\mu_{j,m}(t)(h_{j,m}-\tilde{h}_{j,m})\|_{H_r^s(\mathbb{R})}\\&\le\max_{t\in[0,T]}\sum_{j,m}|\mu_{j,m}(t)|\|(h_{j,m}-\tilde{h}_{j,m})\|_{H_r^s(\mathbb{R})}\\&<\varepsilon.\end{aligned}$$

□

ACKNOWLEDGEMENTS. The contribution of the second named author was partially supported by the grant 346300 for IMPAN from the Simons Foundation and the matching 2015-2019 Polish MNiSW fund.

References

[1] A. Barrasso and F. Russo. Backward Stochastic Differential Equations with no driving martingale, Markov processes and associated Pseudo Partial Differential Equations. part II: Decoupled mild solutions and examples. 2017. Preprint, hal-01505974.

[2] J.M. Bismut. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.*, 44:384–404, 1973.

[3] E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[4] F. Delarue and R. Diel. Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields*, 165(1-2):1–63, 2016.

[5] J. Diehl and J. Zhang. Backward stochastic differential equations with Young drift. *Probab. Uncertain. Quant. Risk*, 2:Paper No. 5, 17, 2017.

[6] H.-J. Engelbert and J. Wolf. Strong Markov local Dirichlet processes and stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 43(2):331–348, 1998.

[7] M. Erraoui, Y. Ouknine, and A. Sbi. Reflected solutions of backward stochastic differential equations with distribution as terminal condition. *Random Operators and Stochastic Equations*, 6(1):1–16, 1998.

[8] F. Flandoli, E. Issoglio, and F. Russo. Multidimensional SDEs with distributional coefficients. *T. Am. Math. Soc.*, 369:1665–1688, 2017.

[9] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.*, 40(2):493–542, 2003.

[10] F. Flandoli, F. Russo, and J. Wolf. Some SDEs with distributional drift. II. Lyons-Zheng structure, Itô's formula and semimartingale characterization. *Random Oper. Stochastic Equations*, 12(2):145–184, 2004.

[11] F. Gozzi and F. Russo. Weak Dirichlet processes with a stochastic control perspective. *Stochastic Processes and their Applications*, 116(11):1563 – 1583, 2006.

[12] E. Issoglio. Transport equations with fractal noise - existence, uniqueness and regularity of the solution. *J. Analysis and its App.*, 32(1):37–53, 2013.

[13] E. Issoglio and S. Jing. Forward-Backward SDEs with distributional coefficients. preprint - ArXiv 2016 (arXiv:1605.01558).

[14] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.

[15] A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems*. Progress in Nonlinear Differential Equations and their Applications, 16. Birkhäuser Verlag, Basel, 1995.

[16] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[17] E. Pardoux and A. Răşcanu. *Stochastic differential equations, backward SDEs, partial differential equations*, volume 69 of *Stochastic Modelling and Applied Probability*. Springer, Cham, 2014.

[18] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[19] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[20] F. Russo and G. Trutnau. Some parabolic PDEs whose drift is an irregular random noise in space. *Ann. Probab.*, 35(6):2213–2262, 2007.

[21] F. Russo and P. Vallois. The generalized covariation process and Itô formula. *Stochastic Processes and their Applications*, 59(1):81 – 104, 1995.

[22] F. Russo and P. Vallois. Elements of stochastic calculus via regularization. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 147–185. Springer, Berlin, 2007.

[23] F. Russo and L. Wurzer. Elliptic PDEs with distributional drift and backward SDEs driven by a càdlàg martingale with random terminal time. *Stoch. Dyn.*, 17(4):1750030 (36 pages), 2017.

[24] H. Triebel. *Interpolation theory, function spaces, differential operators*. North-Holland Amsterdam, Heidelberg, first edition, 1978.

[25] H. Triebel. *Function Spaces and Wavelets on Domains*, volume 7 of *EMS Tracts in Mathematics*. European Mathematical Society, 2008.

[26] H. Triebel. *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, volume 11 of *EMS Tracts in Mathematics*. European Mathematical Society, 2010.