A Feynman–Kac result via Markov BSDEs with generalised drivers

ELENA ISSOGLIO[1] and FRANCESCO RUSSO[2]

[1]*Department of Mathematics, University of Leeds, Leeds, LS2 9JT, UK. E-mail: e.issoglio@leeds.ac.uk*
[2]*ENSTA Paris, Institut Polytechnique de Paris, Unité de Mathématiques appliquées, 828, bd. des Maréchaux, F-91120 Palaiseau, France. E-mail: francesco.russo@ensta-paris.fr*

In this paper, we investigate BSDEs where the driver contains a distributional term (in the sense of generalised functions) and derive general Feynman–Kac formulae related to these BSDEs. We introduce an integral operator to give sense to the equation and then we show the existence of a strong solution employing results on a related PDE. Due to 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.

Keywords: backward stochastic differential equations (BSDEs); distributional driver; Feynman–Kac formula; generalised and rough coefficients; pointwise product; weak Dirichlet process

1. Introduction

In this paper we consider Markov backward stochastic differential equations (BSDEs) where the driver is a generalised function (Schwarz distribution), and investigate Feynman–Kac type formulae in this general setting. The classical notion of Brownian BSDE was introduced in 1990 by E. Pardoux and S. Peng in [25], after an early work of J. M. Bismut in 1973 in [2]. It is a stochastic differential equation with prescribed terminal condition ξ and driver $\hat{f}$ expressed by

$$Y_t = \xi + \int_t^T \hat{f}(r, \omega, Y_r, Z_r) \, dr - \int_t^T Z_r \, dW_r. \tag{1}$$

The unknown is a couple (Y, Z) of adapted processes. 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 [26] and references therein for recent contributions on the topic. Ever since the earliest papers, the field of BSDEs has attracted the interest of a wide number of mathematicians. This is due to the fact that BSDEs turned out to be powerful tools that allowed new and unexpected applications.

Of particular interest is the case where the randomness of the driver in (1) is expressed through a forward diffusion process X and the terminal condition only depends on X_T. We denominate this situation as the *Markov case*. In the present paper, we consider the Markov case where the randomness of the driver $\hat{f}$ depends only on the Brownian motion $W(\omega)$. The key novelty is that $\hat{f}$ has a linear part in Z of the form $Z_r b(r, W_r(\omega))$ where b is a suitable generalised function. In

particular, we consider BSDEs of the form

$$Y_t = \xi + \int_t^T Z_r b(r, W_r) \, dr + \int_t^T f(r, W_r, Y_r, Z_r) \, dr - \int_t^T Z_r \, dW_r. \tag{2}$$

We are interested in a class of coefficients b of distributional type, namely

$$b \in C\big([0, T]; H_q^{-\beta}(\mathbb{R}^d; \mathbb{R}^d)\big),$$

for some $\beta \in (0, 1/2)$. The objects appearing in (2) 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. As an example of generalised function b which is 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).

Our motivation for looking at these very irregular coefficients comes both from applications and from theoretical issues. Indeed, BSDEs like (2) and variations of those equations with the same low regularity of coefficients, arise from vastly different contexts from pricing and hedging problems, to stochastic control, to probabilistic representation of PDEs. Below we illustrate some examples of applications of the BSDE (2) with distributional driver.

- BSDEs intervene classically in financial modelling, see, for example, [10]. If ξ is a contingent claim based on some asset price X (already discounted), then the *price* and the *self-financing strategy* at time t are provided by the couple (Y_t, Z_t) which fulfills

$$Y_t = \xi - \int_t^T Z_r \, dX_r. \tag{3}$$

An interesting case concerns the hedging problem when the underlying X is not a semi-martingale, even though Delbaen & Schachermayer's fundamental theory imposes that X is a semimartingale if no arbitrage is to be excluded. However, these no-arbitrage issues can be solved by imposing extra constraints on the class of admissible strategies. For example, [4] considered a model driven by fractional Brownian motion (which is not a semimartingale): there arbitrage was prevented by not allowing continuous trading. In that context, the integral in (3) obviously exists because the strategy processes are of bounded variation. However in general, a fundamental issue is that the integral in (3) has to be suitably defined. For instance in [6], where X is a finite quadratic variation process (but non necessarily a semimartingale), the integral in (3) is a forward integral, and no-arbitrage is guaranteed by appropriately restricting the class of admissible strategies. Suppose now that the asset price is modelled by the *rough process*

$$X_t = W_t - \int_0^t b(s, W_s) \, ds,$$

where $b(s, \cdot)$ is a Schwarz distribution. Then in this case (3) reduces to BSDE (2) with $f \equiv 0$, that is

$$Y_t = \xi - \int_t^T Z_r \, dW_r + \int_t^T Z_r b(r, W_r) \, dr.$$

Note that the latter integral has still to be defined.

- BSDEs are also powerful tools that help to solve stochastic control problems. For example, suppose that X follows a stochastic controlled dynamics

$$dX_t = \mu(t, X_t, \alpha_t)\, dt + \sigma(t, X_t, \alpha_t)\, dW_t, \tag{4}$$

where α is the control process that acts on the drift and the volatility. Let $d = 1$ for simplicity. Suppose that we are interested in maximising the functional $J(\alpha) = \mathbb{E}[\Phi(X_T)]$ as a function of the control α. It is known that this stochastic control problem can be solved with the help of the stochastic maximum principle (Pontryagin maximum principle), see, for example, [28], Section 6.4.2. In this setting, one needs to solve a BSDE, called *adjoint equation*, where the driver is the derivative of the Hamiltonian $\mathcal{H}(t, x, a, y, z) := \mu(t, x, a)y + \sigma(t, x, a)z$, that is, the BSDE takes the form

$$-dY_t = \mathcal{D}_x \mathcal{H}(t, X_t, \alpha_t, Y_t, Z_t)\, dt - Z_t\, dW_t, \tag{5}$$

with terminal condition $Y_T = \mathcal{D}_x \Phi(X_T)$. Here $\mathcal{D}_x \mathcal{H}$ denotes the derivative of $\mathcal{H}$ with respect to the variable x. It is clear that if $x \mapsto \sigma(t, x, a)$ is a continuous function which is not differentiable, then the driver of the BSDE will contain some *singular* elements. More specifically, consider for instance the case when $\sigma(t, x, a) = \sigma_0(t, x)\sigma_1(a)$ where $\sigma_1(\cdot)$ is bounded and $\sigma_0(t, \cdot) \in H_p^s(\mathbb{R})$ where $s < 1$. Then $\mathcal{D}_x \sigma_0(t, \cdot) \in H_p^{s-1}(\mathbb{R})$ and $s - 1 < 0$, that is a generalised function like b in (2). Indeed, in this case we recover a BSDE where there is a rough part linear in z, namely $\mathcal{D}_x \sigma(t, x, \alpha(t, x))z =: b(t, x)z$, much like (2) with $-\beta = s - 1$. We remark that any s'-Hölder continuous function $\sigma_0(t, \cdot)$ with compact support belongs to the fractional Sobolev space $H_p^s(\mathbb{R})$ for any $s < s'$ and $p \geq 2$, see, for example, [18], Proposition 4.1. Moreover, if the diffusion coefficient is s'-Hölder continuous with $s' \geq \frac{1}{2}$ and if the drift μ is Lipschitz, one can show pathwise uniqueness of a solution X to (4) for every given control α following the proof of [22], Proposition 2.13, Ch 5. In finance, such kind of non-smooth volatility σ can be obtained if one looks for example at CIR models with *uncertain volatility*, where $\mu(t, x, a) = bx + c$ and $\sigma(t, x, a) = \sqrt{x}a$. Here the control a is a scaling parameter that represents the uncertainty of the volatility and varies between two given values a_1, a_2.

- As we mentioned earlier, another main application 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 (2) is linked to a PDE of the form

$$\begin{cases} \partial_t u + \dfrac{1}{2}\Delta u = -\nabla u^* b - f(\cdot, u, \nabla u), \\ u(T) = \Phi, \end{cases} \tag{6}$$

see Section 2 for details about the notation. If u is the solution of the PDE, then $Y_t := u(t, W_t)$ and $Z_t := \nabla u^*(t, W_t)$ is a solution to the BSDE (2). We emphasize that (Y, Z) is a strong solution to the BSDE related to the Brownian filtration related to W, which is then used to represent the solution u to the PDE via non-linear Feynman–Kac type formulae. Note that if we were to work with SDEs with distributional coefficients, we would have representation of the (linear) PDE via weak solutions and not strong solutions, because in this case the solution to the SDE is weak, see [13].

Motivated by these examples, we study BSDE (2) and the Feynman–Kac representation of the solution to PDE (6) from a theoretical perspective. PDEs with distributional coefficients appear naturally as Fokker–Planck type equations for diffusions in irregular medium or polymers, see, for example, [8,24,34]. The topic of stochastic differential equations involving distributional coefficients has attracted a lot of interest, in particular for (forward) SDEs. See, for example, [11, 14,15] in the case where the solution is not a semimartingale. See also [30] and more recently [8, 13]. For what concerns the case of backward SDEs involving a distribution we mention the works [12] on (reflected) BSDEs with distribution as terminal condition, and [33] whose authors studied a one-dimensional BSDE (with random terminal time) involving distributional coefficients via a forward stochastic process. In [9], they considered BSDEs, where the driver is a Young integral. Recently, [1,20] studied Markov BSDEs with special forward process with distributional drift, using different techniques than ours.

In this paper, we make a substantial step towards a deeper understanding of backward equations with distributional drivers and their link to rough non-linear PDEs expressed via Feynman–Kac type formulae. It is worth noticing that even though one expects that BSDE (1) is somehow equivalent to PDE (6), this is a priori not clear in the singular case when b is a distribution. We rigorously prove this fact in the present paper. Our idea is to give an intrinsic meaning to the distributional term $Z_r b(r, W_r)$ in order to define and solve the BSDE. We start by introducing an integral operator $A^{Y,W}$ (see Definitions 3.1 and 3.2) that will provide a proper mathematical meaning to the term $\int_t^T Z_r b(r, W_r) \, dr$ when evaluated in b. This operator is defined in terms of the Brownian motion W and a process Y. In the special case when $Y = W$, $A^{W,W}$ will be denominated as the *occupation time operator*, since $A^{W,W}(g)$ can be linked to the occupation time formula, see Remark 5.7. Using the integral operator $A^{Y,W}$ we introduce an equivalent formulation of the BSDE (1) (see Definition 3.3) and show that it extends the classical notion of solution from Pardoux–Peng, see Proposition 3.5. In Proposition 5.4, we show that the occupation time operator $A^{W,W}$ is well-defined for bs in a specific class of distributions, namely in the fractional Sobolev space $H_q^{-\beta}$ where the parameters satisfy Assumption 2.6. In Proposition 5.6, we show a chain rule for $\phi(t, W_t)$ for a certain class of $\phi \in C^{0,1}$ (related to the heat equation (18)), and the remainder in the chain rule is expressed in terms of the occupation time operator $A^{W,W}$. Our main results are Theorem 5.13, where we prove the existence of a solution to the BSDE (17) in the Markovian framework given in terms of the solution of PDE (6), and Corollary 5.14, which is the Feynman–Kac formula for the probabilistic representation of the solution of the PDE. We also investigate uniqueness of the solution of the BSDE in a particular class (Proposition 5.15).

The paper is organised as follows. In Section 2, we recall useful results, set the notation and state the assumptions needed later on. In Section 3, we define the integral operator $A^{Y,W}$ and introduce the equivalent formulation of the BSDE. Section 4 collects important analytical properties of the PDE associated to the BSDE in the Markovian case. In Section 5.1, we investigate the properties of the occupation time operator and in Section 5.2, we state and prove the main results of existence of a solution to the BSDE and the corresponding Feynman–Kac formula. Finally in the Appendix, we state and prove a technical result needed in the paper, as well as two technical proofs which have been moved here for ease of reading.

2. Preliminaries and notation

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

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

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

The following definitions of covariation process and weak-Dirichlet process are taken from [16], see also [31] 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.

Definition 2.1. 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 a *martingale-orthogonal process*, namely a process such that $[A, N] = 0$ for every $\mathcal{F}$-continuous local martingale N. For convenience we also set $A_0 = 0$.

Note that in [16] they use the name *weak zero energy process* for the *martingale-orthogonal process*. 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 $D = (D^1, \ldots, D^d)$ 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 2.2. *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*

$$\left[v(\cdot, S^1), S^2 \right]_t = \int_0^t \nabla v^* (r, S_r^1) \, \mathrm{d} \left[M^1, M^2 \right]_r. \tag{7}$$

Proof. Let us denote by $M_t^v := \int_0^t \nabla v^*(r, S_r^1) \, \mathrm{d} M_r^1$. By [16], 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 martingale-orthogonal process, we know that $[A^v, N] = 0$ for any $\mathcal{F}$-continuous local martingale N, see [32], Proposition 1.7.(b). Consequently the left-hand side of (7) gives

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

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

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 2.3. *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*

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

Heat semigroup and fractional Sobolev spaces

We denote by $\mathcal{S}(\mathbb{R}^d)$ the space of $\mathbb{R}$-valued Schwartz functions and by $\mathcal{S}'(\mathbb{R}^d)$ the space of Schwartz distributions. Setting $A := I - \frac{1}{2}\Delta$, we can view this as an operator in $\mathcal{S}'(\mathbb{R}^d)$. Then its fractional powers A^α are well-defined on the same space for any power $\alpha \in \mathbb{R}$ by means of Fourier transform (see, e.g., [37], Remark 1.2(iii)). One can define the classical fractional Sobolev spaces via these fractional powers, that is $H_r^s(\mathbb{R}^d) := A^{-s/2}(L^r(\mathbb{R}^d))$. These are Banach spaces endowed with the norm $\|u\|_{H_r^s} := \|A^{s/2}u\|_{L^r}$. It is also known that $A^{-\alpha/2}$ is an isomorphism between $H_r^s(\mathbb{R}^d)$ and $H_r^{s+\alpha}(\mathbb{R}^d)$, for each $\alpha \in \mathbb{R}$, (see again [37], Remark 1.2(iii)). We denote by $(P(t), t \geq 0)$ the heat semigroup which acts on any $L^r(\mathbb{R}^d)$ for $1 < r < \infty$, with kernel $p_t(x) = \frac{1}{(2\pi t)^{d/2}} \exp(-\frac{|x|^2}{2t})$. This is a bounded analytic semigroup generated by $\frac{1}{2}\Delta$, see [7], Theorems 1.4.1, 1.4.2. We denote by $(S(t), t \geq 0)$ the semigroup given by $S(t) := e^{-t}P(t)$. If we consider A as an unbounded operator on $L^r(\mathbb{R}^d)$, then it is well known that the semigroup S is generated by $-A$ and $D(A) = H_r^2(\mathbb{R}^d)$. Fractional powers of A, as unbounded operator on $L^r(\mathbb{R}^d)$, where $-A$ is the generator of an analytic semigroup can also be defined (see [27], Section 2.6) and a key fact that links these operators with fractional Sobolev spaces is that $D(A^{s/2}) = H_r^s(\mathbb{R}^d)$, which follows from interpolation theory.[1] Using this and the isomorphism property one has for $\delta > \beta > 0$, $\delta + \beta < 1$ and $0 < t \leq T$ that $P(t) : H_r^{-\beta}(\mathbb{R}^d) \to H_r^{1+\delta}(\mathbb{R}^d)$ for all $1 < r < \infty$ and

$$\|P(t)w\|_{H_r^{1+\delta}(\mathbb{R}^d)} \leq Ce^t t^{-\frac{1+\delta+\beta}{2}} \|w\|_{H_r^{-\beta}(\mathbb{R}^d)}, \tag{8}$$

for $w \in H_r^{-\beta}(\mathbb{R}^d)$, $t > 0$. This follows from a similar property for the bounded analytic semigroup S which is stated in [13], Lemma 10, see also [18], Proposition 3.2, for the analogous on domains $D \subset \mathbb{R}^d$. Moreover it is easy to show[2] that $P(t)$ is a contraction on $H_r^s(\mathbb{R}^d)$ for all $1 < r < \infty$ and $s \geq 0$, that is for all $w \in H_r^s(\mathbb{R}^d)$ we have

$$\|P(t)w\|_{H_r^s(\mathbb{R}^d)} \leq \|w\|_{H_r^s(\mathbb{R}^d)}. \tag{9}$$

---

[1] This can be seen by applying [35], Theorem 1.15.3, with $\alpha = 0$, $\beta = 1$ and $0 < \theta < 1$ to get $D(A^\theta) = [D(A^0), D(A^1)]_\theta$. The latter is equal to $[L^p, H_p^2]_\theta$ because of known result on the operator A and its integer powers. Finally using [35], Theorem 2.4.2/1, with $q_0 = q_1 = q = 2$, $p_0 = p_1 = p$ and $s_0 = 1$, $s_1 = 2$ (so that $s = 2\theta$) one gets $[L^p, H_p^2]_\theta = [H_p^0, H_p^2]_\theta = H_p^{2\theta}$.

[2] This can be seen by writing $P(t) = e^t e^{-t} P(t) = e^t S(t)$. Since $-A$ is the generator of S we have that $S(t) : L^r \to D(A^{s/2})$ by [27], Chapter 2, Thm 6.13(a). Moreover $D(A^{s/2}) = H_r^s$ as recalled above. Let $w \in H_r^s$, so we also have $w \in L^r$ thus $S(t)w \in H_s^r$. Then by the definition of norm in H_r^r we get $\|P(t)w\|_{H_r^s} = e^t \|S(t)w\|_{H_r^s} = e^t \|A^{s/2} S(t)w\|_{L^r}$. Now applying [27], Chapter 2, Thm 6.13(b), we know that $A^{s/2}$ and $S(t)$ commute and using the contractivity of $P(t)$ on L^r we get $\|P(t)w\|_{H_r^s} \leq \|e^t S(t) A^{s/2} w\|_{L^r} \leq \|A^{s/2} w\|_{L^r}$ and the latter is equal to $\|w\|_{H_r^s}$ by definition of the norm.

As done already before in this paper, we denote by H_r^s the spaces $H_r^s(\mathbb{R}^d; \mathbb{R}^d)$, whose definition is as above for each component. Note that by slight abuse of notation the same H_r^s might be the space $H_r^s(\mathbb{R}^d; \mathbb{R}^{d \times d})$, especially when considering functions like ∇u. When we write $u \in H_r^s$ we mean that each component u_i is in $H_r^s(\mathbb{R}^d)$. 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 [29]. Let $g \in \mathcal{S}'(\mathbb{R}^d)$. We choose a function $\psi \in \mathcal{S}(\mathbb{R}^d)$ such that $0 \leq \psi(x) \leq 1$, for every $x \in \mathbb{R}^d$ and

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

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

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

where $\mathcal{F}(g)$ and $\mathcal{F}^{-1}(g)$ 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

$$gh := \lim_{j \to \infty} S^j g S^j h, \tag{10}$$

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

Lemma 2.4 ([29], Theorem 4.4.3/1). *Let* $g \in H_q^{-\beta}(\mathbb{R}^d)$, $h \in H_p^\delta(\mathbb{R}^d)$ *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* $H_p^{-\beta}(\mathbb{R}^d)$ *and we have the bound*

$$\|gh\|_{H_p^{-\beta}(\mathbb{R}^d)} \leq c\|g\|_{H_q^{-\beta}(\mathbb{R}^d)} \cdot \|h\|_{H_p^\delta(\mathbb{R}^d)}. \tag{11}$$

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

More on function spaces

We observe 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 [35], 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 $\{\|\cdot\|_{C([0,T];B)}^{(\rho)}, \rho \geq 1\}$, defined by

$$\|h\|_{C([0,T];B)}^{(\rho)} := \sup_{0 \leq t \leq T} e^{-\rho t} \|h(t)\|_B. \tag{12}$$

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 [35], Theorem 2.8.1, Remark 2.

Lemma 2.5 (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*
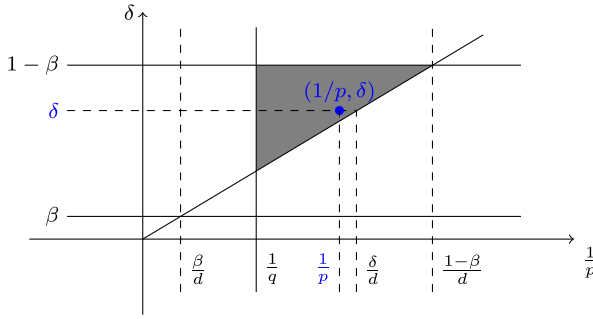
$$\|h\|_{C^{1+\alpha}(\mathbb{R}^d)} \leq c\|h\|_{H_r^{1+\delta}(\mathbb{R}^d)}, \qquad \|\nabla h\|_{C^{0+\alpha}(\mathbb{R}^d)} \leq c\|\nabla h\|_{H_r^\delta(\mathbb{R}^d)}, \tag{13}$$

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

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.

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

$$K(\beta, q) := \left\{ (\delta, p) \in \mathbb{R}^2 : \beta < \delta < 1 - \beta, \frac{d}{\delta} < p < q \right\}, \tag{14}$$
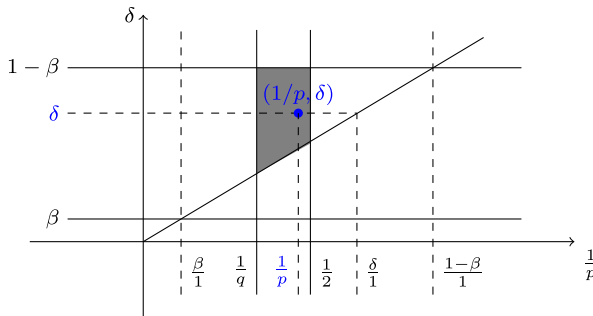
which is drawn in Figure 1.

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

$$K(\beta, q) := \left\{ (\delta, p) \in \mathbb{R}^2 : \beta < \delta < 1 - \beta, \frac{1}{\delta} < p < q, 2 \leq p \right\}, \tag{15}$$

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 [13] without the restriction $q, p \geq 2$. This is satisfied anyway if $d > 1$. If $d = 1$ then the set of admissible couples (δ, p) is shown in Figure 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.

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

Assumption 2.7. 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, that is, $|f(t, x, y, z) - f(t, x, y', z')| \leq L(|y - y'| + |z - z'|)$ 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 (2) 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.

Let $W = (W_t)_t$ be a d-dimensional Brownian motion equipped with its canonical filtration $\mathcal{F} = (\mathcal{F}_t)_t$. To be able to consider rough drivers, the main term in (2) that needs to be (re)defined is the integral $\int_t^T Z_r b(r, W_r) \, dr$. 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 $Y = (Y_t)_t$ 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

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

$$l \mapsto A_{\cdot}^{W,Y}(l),$$

where

$$A_t^{W,Y}(l) := \left(\int_0^t l^*(r, W_r) \, d[W, Y]_r \right)^* \tag{16}$$

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

We observe that in the special case when $Y = W$ the occupation time operator $A^{W,W}$ applied to l is nothing but $\int_0^{\cdot} l(r, W_r) \, dr$ (see the introduction of Section 5.1 for more details). Moreover, for the class of functions $l \in C_c([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ the integral in (16) 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 generalised functions, 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 $E = C([0, T]; H_r^s)$ with $2 \leq r < \infty$ and $-\frac{1}{2} < s \leq 0$. Using this extension, we can reformulate BSDE (2) 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 (2) if:

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

$$M_t := Y_t - Y_0 + A_t^{W,Y}(b) + \int_0^t f\left(r, W_r, Y_r, \frac{\mathrm{d}[Y, W]_r}{\mathrm{d}r}\right) \mathrm{d}r \qquad (17)$$

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 in the sense of Definition 2.1 with martingale-orthogonal process A 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 (17) 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 $\{W_s; s \leq r\}$ instead of only on W_r.
- Another generalization of Pardoux–Peng BSDEs that allows the solution Y not to be a semimartingale appeared for example in [5], where the authors introduce and study generalised backward differential equations. In their formulation of BSDE (see [5], Definitions 2.1 and 2.2) they consider a functional $F_t(Y, M)$ for every adapted cadlag process Y and L^p-martingale M, which in general may not be a semimartingale. In our setting, the object corresponding to this functional would be $(\int_0^t b^*(r, W_r) \, \mathrm{d}[Y, W]_r)^*$. However this integral is not defined for every cadlag adapted process Y, since b is a distribution and the covariation $[Y, W]$ is not well-defined a priori, hence that setting cannot be used here.

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* (2) *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* (2) *in the classical sense.*

Proof. Suppose that (Y, Z) is a classical solution of (2). 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 $A^{W,Y}_\cdot(b) = (\int_0^\cdot b^*(r, W_r) \, d[W, Y]_r)^*$. Using (2) we have

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

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

$$[W, Y]_t = \left[W, \int_0^\cdot Z_r \, dW_r \right]_t = \int_0^t Z_r^* \, dr,$$

hence $d[W, Y]_r = Z_r^* \, dr$, and in particular

$$A^{W,Y}_\cdot(b) = \left(\int_0^\cdot b^*(r, W_r) Z_r^* \, dr \right)^* = \int_0^\cdot Z_r b(r, W_r) \, dr.$$

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

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

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

Conversely, let Y be a solution of (2) according to Definition 3.3 with respect to E. We know that

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

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 $M_t = \int_0^t Z_r \, dW_r$. Moreover, $A^{W,Y}$ is a martingale-orthogonal process by point (ii), thus $[W, Y]_t = [W, M]_t = \int_0^t Z_r^* \, dr$. Therefore, $A^{W,Y}_t(b) = (\int_0^t b^*(s, W_s) \, d[W, Y]_s)^* = \int_0^t Z_s b(s, W_s) \, ds$ and this concludes the proof. $\qquad \square$

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

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 $b \in L^\infty_{loc}([0, T] \times$

$\mathbb{R}^d; \mathbb{R}^d$). 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 $C([0, T]; H^s_r) \cap L^\infty_{\text{loc}}([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ for $-1/2 < s \leq 0$ and $2 \leq r < \infty$ 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 (17) is well-defined for suitable generalised functions 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 2.6. 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

$$\begin{cases} \partial_t \phi + \dfrac{1}{2}\Delta \phi = l, \\ \phi(T) = \Psi, \end{cases} \tag{18}$$

where $\Psi \in H^{1+\delta}_p$ and $l \in C([0, T]; H^{-\beta}_p)$. 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 (18) which is given by

$$\phi(t) = P(T - t)\Psi + \int_t^T P(r - t)l(r)\,dr, \tag{19}$$

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 (19) (mild formulation) and it has certain regularity properties as recalled in the lemma below for smooth Ψ and l. For more details and a proof see, for example, [23], Theorem 5.1.4, part (iv).

Lemma 4.1. *Let* $l \in C^\infty_c([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ *and* $\Psi \in C^{2+\epsilon}(\mathbb{R}^d; \mathbb{R}^d)$ *for some* $0 < \epsilon < 1$. *The solution* ϕ *to* (18) *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* 2.6.

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

(ii) If $l \in C([0, T]; H_p^{-\beta})$ then the function $t \mapsto \int_t^T P(s - t)l(s)\,ds$ is in $C^\gamma([0, T]; H_p^{2-2\epsilon-\beta})$ 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) follows from three facts: 1. well-known continuity of the heat semigroup $S(t) = e^{-t}P(t)$ in L^p; 2. continuity of $S(t)$ in H_p^s for all $s \geq 0$, which follows from the fact that $A^{s/2}$ commutes with $S(t)$ (see [27], Chapter 2, Thm 6.13(b)) so that one has $\|S(t)w - S(t_0)w\|_{H_p^s} = \|S(t)A^{s/2}w - S(t_0)A^{s/2}w\|_{L^p}$ for each $t_0 \in [0, T]$; 3. the link between $S(t)$ and $P(t)$ via the continuous scalar function e^t so that $P(t) = e^t S(t)$ is still continuous in H_p^s as a function of t.

Item (ii) follows by first applying [13], 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 2.6 hold, and let $\Psi \in H_p^{1+\delta}$ and $l \in C([0, T]; H_p^{-\beta})$. The expression ϕ given in (19) 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

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

and

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

Proof. For the first term in (19), we have that $t \mapsto P(T - t)\Psi \in H_p^{1+\delta}$ is continuous by Lemma 4.2, item (i). Moreover by (9), we have $\|P(T - t)\Psi\|_{H_p^{1+\delta}} \leq \|\Psi\|_{H_p^{1+\delta}}$ for all $t \in [0, T]$. For the second term in (19), we have continuity as a function of time by Lemma 4.2, item (ii) and again by the mapping property (8) of the semigroup in Sobolev spaces we get the bound

$$\left\| \int_t^T P(s - t)l(s)\,ds \right\|_{H_p^{1+\delta}} \leq ce^T \int_t^T (s - t)^{-\frac{1+\delta+\beta}{2}} \|l(s)\|_{H_p^{-\beta}}\,ds$$

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

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 2.6 and so by the fractional Morrey inequality (Lemma 2.5) 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 [13], Lemma 21). $\qquad\square$

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

Corollary 4.4. *Let Assumption* 2.6 *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* (18) *with* l_n *in place of* l *and* Ψ_n *in place of* Ψ. *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 (2) in the Markovian case, which will be used to construct the solution to the BSDE, namely

$$\begin{cases} \partial_t u(t) + \dfrac{1}{2}\Delta u(t) = -\nabla u^*(t)b(t) - f\big(t, \cdot, u(t), \nabla u(t)\big), \\ u(T) = \Phi. \end{cases} \qquad (20)$$

We note that the term Δu (as in PDE (18) 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 (20) is a function u that satisfies

$$u(t) = P(T - t)\Phi - \int_t^T P(r - t)\big(\nabla u^*(r)b(r)\big)\,\mathrm{d}r$$

$$- \int_t^T P(r - t)f\big(r, \cdot, u(r), \nabla u(r)\big)\,\mathrm{d}r \qquad (21)$$

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 (20) was first studied in [18] on a bounded domain $D \subset \mathbb{R}^d$ and with $f \equiv 0$. It was then solved in $\mathbb{R}^d$ in [13] with $f = 0$, and in [20] with f nonzero. Related non-linear PDEs with rough coefficients have been studied with similar techniques in [17,19,21]. We remark in particular that in [19] the author applies analytical results on a quadratic rough PDE to study a quadratic rough BSDE. Ideas used there are similar to what has been done in [20], where 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 (2). 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 2.7 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 2.7 (we will abuse the notation and write f for both). Then $\tilde{f}$ satisfies

the required conditions, as explained in [20], Remark 2.5, in particular $\tilde{f}$ is Lipschitz continuous in the Sobolev spaces

$$\|\tilde{f}(t, u, v) - \tilde{f}(t, u', v')\|_{H_p^0} \le c\big(\|u - u'\|_{H_p^{1+\delta}} + \|v - v'\|_{H_p^\delta}\big). \tag{22}$$

Theorem 5, and Lemmata 5 and 8 in [20] give the following existence, uniqueness and regularity result.

Theorem 4.5 (Issoglio, Jing). *Under Assumption* 2.6 *and Assumption* 2.7 *there exists a unique mild solution* u *to* (20) *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 [20] 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 (2) when $b \in C([0, T]; H_q^{-\beta})$ in the *Markovian setting*. The Markovian setting means here 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 generalised functions 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 Section 5.1 below. In Section 5.2, we will show existence (and uniqueness) of a solution to BSDE (2) according to Definition 3.3 when b is a rough driver.

5.1. Properties for the occupation time operator $A^{W,W}$

In this section, we show how to extend the operator $A^{W,Y}$ to generalised functions. 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 the *occupation time operator* $A^{W,W}$, where we recall that

$$A^{W,W}: \quad C_c\big([0, T] \times \mathbb{R}^d; \mathbb{R}^d\big) \to \mathcal{C} \tag{23}$$

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

$$A_t^{W,W}(l) = \left(\int_0^t l^*(r, W_r) I_d \, \mathrm{d}r\right)^* = \int_0^t l(r, W_r) \, \mathrm{d}r, \tag{24}$$

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 2.3

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:

$$
\begin{aligned}
A_t^{W,Y}(l) &= \left(\int_0^t l^*(r, W_r)\,\mathrm{d}[W, Y]_r \right)^* \\
&= \left(\int_0^t l^*(r, W_r)\nabla\gamma(r, W_r)\,\mathrm{d}r \right)^* \\
&= \int_0^t \nabla\gamma^*(r, W_r)l(r, W_r)\,\mathrm{d}r \\
&= A_t^{W,W}(\nabla\gamma^*l).
\end{aligned}
\tag{25}
$$

By Theorem 4.5 $u \in C^{0,1}$ and so equation (25) holds true also in the case where γ is replaced by the solution u of PDE (20).

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* $-\frac{1}{2} < s \le 0$ *and* $2 \le r < \infty$, *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 the occupation time operator $A^{W,W}$ for smooth l, and this representation will still hold in the rough case.

Proposition 5.3 (Chain rule – smooth case).

(i) *Let Assumption 2.6 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 (19). Then for the integral operator* $A^{W,W}$ *given in (23) we have the representation*

$$
A_t^{W,W}(l) = \phi(t, W_t) - \phi(0, W_0) - \int_0^t \nabla\phi^*(r, W_r)\,\mathrm{d}W_r,
\tag{26}
$$

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

(ii) *The map on the right-hand side of (26) is continuous with respect to* $\phi \in C^{0,1}$.

We note that the structure of the representation (26) does not change when Ψ changes (although obviously the actual function ϕ changes when Ψ changes).

Proof. We prove *part (ii)* first. By linearity, it is enough to prove it for $\phi = 0$. Let $\phi_n \in C^{0,1}$ such that $\phi_n \to 0$ in the same space. Clearly, $\phi_n(\cdot, W)$ converges uniformly to 0 a.s., and in particular

uniformly in probability. Setting $f_n = \nabla \phi_n^*$ it remains to show that

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

According to [22], Proposition 2.26, it is enough to show that

$$\int_0^T |f_n(r, W_r)|^2 \, dr \to 0 \tag{27}$$

in probability. Now $f_n \to 0$ uniformly on each compact by assumption, which implies that (27) holds a.s.

Next, we prove *part (i)*. 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 2.6. 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 (19), 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 (24), the PDE (18) and Itô's formula we get

$$\begin{aligned}
A_t^{W,W}(l_n) &= \int_0^t l_n(r, W_r) \, dr \\
&= \int_0^t \left(\partial_t \phi_n(r, W_r) + \frac{1}{2} \Delta \phi_n(r, W_r) \right) dr \\
&= \phi_n(t, W_t) - \phi_n(0, W_0) - \int_0^t \nabla \phi_n^*(r, W_r) \, dW_r,
\end{aligned} \tag{28}$$

for $0 \leq t \leq T$. By Corollary 4.4, we have that $\phi_n \to \phi$ in $C^{0,1}$, thus applying part (ii) we conclude that

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

and the proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The following proposition will be used to extend the occupation time operator $A^{W,W}$ to a suitable space of generalised functions, see Remark 5.5, part 1.

Proposition 5.4. *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 (19) 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 part (i)) and taking the u.c.p.-limit in $\mathcal{C}$ as $n \to \infty$, we get by Proposition 5.3 part (ii)

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

The continuity of the occupation time operator $A^{W,W}$ at 0 implies the continuity everywhere by linearity. $\qquad\square$

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 occupation time 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 occupation time operator $A^{W,W}$ to the space $E = C([0, T]; H_p^{-\beta})$, as explained below.

Remark 5.5.

1. By Lemma 5.1 and Proposition 5.4, 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 2.6. So $A^{W,W}$ is well-defined according to Definition 3.2.

2. Clearly the extended operator $A^{W,W}$ defined in Remark 5.5 part 1. is continuous, that is, 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 (18) when $l \in C([0, T]; H_p^{-\beta})$.

Proposition 5.6 (Chain rule – rough case). *Let Assumption 2.6 hold, $l \in C([0, T]; H_p^{-\beta})$ and ϕ be given by (19) for a terminal condition $\Psi \in H_p^{1+\delta}$. Then for all $t \in [0, T]$ we have the representation*

$$A_t^{W,W}(l) = \phi(t, W_t) - \phi(0, W_0) - \int_0^t \nabla \phi^*(r, W_r) \, dW_r. \tag{29}$$

Moreover, $A^{W,W}(l)$ is a martingale-orthogonal 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 5.5 part 2. and the chain rule for the smooth case (Proposition 5.3 part (i)), we get

$$A_\cdot^{W,W}(l) = \lim_{n \to \infty} A_\cdot^{W,W}(l_n)$$

$$= \lim_{n \to \infty} \left(\phi_n(\cdot, W) - \phi_n(0, W_0) - \int_0^\cdot \nabla \phi_n^*(r, W_r) \, dW_r \right).$$

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 Proposition 5.3 part (ii), we can take the u.c.p. limit in $\mathcal{C}$ when $n \to \infty$ and we get

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

To show that $A^{W,W}(l)$ is a martingale-orthogonal process we use the representation (29) 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 2.3

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

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 (29) gives

$$\left[-\int_0^\cdot \nabla \phi^*(r, W_r) \, dW_r, N \right]_t = -\int_0^t \nabla \phi^*(r, W_r) \, d[W, N]_r,$$

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

Remark 5.7.

1. The terminology *occupation time operator* for $A^{W,W}$ comes from the extension of the density occupation formula

$$\int_0^t g(W_s) \, ds = \int_{\mathbb{R}} g(a) L_t^W(a) \, da,$$

   where L^W is the Brownian local time. If g is not a function but $g = h'$, where h is a bounded Borel function, the extension of the right-hand side is possible by Bouleau–Yor formula, see [3]. If X is a semimartingale, there the authors introduce an integral $\int_{\mathbb{R}} h(a) L_t^X(da)$. Clearly when $X = W$ the integral is well-defined because L^W is itself a semimartingale.

2. In the literature, one can find various Itô type formulae involving stochastic processes, formally of the type $\int_0^t g(X_s) \, d[X]_s$ (like in [3] above), where X is a semimartingale and g is

a Schwartz distribution. For example in [31] where X is a (multidimensional) semimartingale and $g = \text{Hess } f$, the integral is formally expressed as the covariation $[\nabla f(X), X]$. In the special case when $X = W$ is a Brownian motion (so $[W]_t \equiv t$) those papers expanded $f(W_t)$ for some $f \in C^1(\mathbb{R})$ (resp. $f \in C^1(\mathbb{R}^d)$) and g is the distribution Δf. In particular those formulae focused on the pointwise composition $f(X)$.

3. Using a different approach, [38] expanded abstractly $T \star \delta_{W_t}$, where T is a Schwarz distribution and W is a standard Brownian motion. Taking T associated with a C^1 function f, this would imply the expansion of the function $x_0 \mapsto f(W_t + x_0)$. By an easy adaptation of Itô's formula shown in [38] one gets dx_0-a.e.

$$f(W_t + x_0) = f(W_0 + x_0) + \int_0^t \nabla f(W_s + x_0) \, dX_s + \mathcal{A}_t(\Delta f)(x_0),$$

where $x_0 \mapsto \mathcal{A}_t(\Delta f)(x_0)$ is (for each t) a random field a.s. associated with the random distribution

$$\varphi \mapsto \int_{\mathbb{R}^d} \mathcal{A}_t(\Delta f)(x_0)\varphi(x_0) \, dx_0 = \int_0^t (\Delta f \star \varphi)(W_s + x_0) \, ds. \tag{30}$$

4. This can be linked to the occupation time operator, indeed the right-hand side of (30) can be seen as $A_t^{W,W}(\Delta f \star \varphi(x_0 + \cdot))$. By continuity with respect to x_0 it is possible to extend Itô's formula to every x_0. At this point, if we formally take $\varphi = \delta_{x_0}$, then we recover the chain rule stated in Proposition 5.6 in the special case where f is time-independent. The rigorous proof however, would need mollifications of δ_{x_0} and a limiting procedure, which in essence is the same idea we used (translated in our context) when we defined the extended operator $A^{W,W}$.

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 2.4 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. $\qquad\square$

Proposition 5.9. *Let Assumption* 2.6 *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*

$$A^{W,Y}(l) = A^{W,W}(\nabla \gamma^* l), \tag{31}$$

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 5.5 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 (31). Taking into account (25) 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. $\square$

Remark 5.10. We observe that in [20] the authors deal with the singular integral term $\int_0^t Z_s b(s, W_s) \, ds$ by replacing it with known terms. In particular, they *define it using the chain rule* (29) with $l = \nabla u^* b$ but without proving it. 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 $g \in C([0, T]; H_r^s) \cap L_{loc}^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ with $-\frac{1}{2} < s \le 0$ and $2 \le r < \infty$, 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 $\int_0^{\cdot} g(s, W_s) \, ds$. The proof of the theorem below is postponed to the Appendix for ease of reading.

Theorem 5.11. *Let* $g \in C([0, T]; H_r^s) \cap L_{loc}^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ *with* $-\frac{1}{2} < s \le 0$ *and* $2 \le r < \infty$, *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*

$$A_{\cdot}^{W,W}(g) = \int_0^{\cdot} g(s, W_s) \, ds. \tag{32}$$

Note that the operator $A^{W,W}$ is well-defined for example, if $s = -\beta$ and $r = p$ see Remark 5.5.

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

$$\int_0^t g(s, W_s) \, ds = \phi(t, W_t) - \phi(0, W_0) - \int_0^t \nabla \phi^*(s, W_s) \, dW_r,$$

where ϕ *is the solution of* (18) *with* $\Psi \in H_p^{1+\delta}$, *given by* (19).

Proof. This follows by Theorem 5.11 and Proposition 5.6 with $l = g$. $\qquad\square$

5.2. Existence for the BSDE and Feynman–Kac representation

Here we show that the solution of PDE (20) can be used to construct a solution to BSDE (2) when $b \in C([0, T]; H_q^{-\beta})$. In particular, in Theorem 5.13 we construct a solution to BSDE (2) with $\xi = \Phi(W_T)$ using the solution to the associated PDE. As a corollary of Theorem 5.13, we obtain a Feynman–Kac representation in Corollay 5.14.

For ease of reading, we rewrite the formal meaning of the BSDE (2) under consideration:

$$Y_t = \Phi(W_T) + \int_t^T Z_r b(r, W_r) \, dr + \int_t^T f(r, W_r, Y_r, Z_r) \, dr - \int_t^T Z_r \, dW_r.$$

Theorem 5.13. *Let Assumption 2.6 and Assumption 2.7 hold and let $b \in C([0, T]; H_q^{-\beta})$. We denote by u be the unique mild solution to (20). Then $Y_t = u(t, W_t)$ is a solution of (2) 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

$$A_t^{W,Y}(b) = A_t^{W,W}(\nabla u^* b). \tag{33}$$

This is a martingale-orthogonal process by Proposition 5.6 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

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

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

$$\int_0^t f\left(r, W_r, Y_r, \frac{d[Y, W]_r}{dr}\right) dr = \int_0^t f(r, W_r, u(r, W_r), \nabla u(r, W_r)) \, dr$$

$$= \int_0^t \tilde{f}(r, W_r) \, dr, \tag{34}$$

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 $\tilde{f} \in L_{\text{loc}}^\infty([0, T] \times \mathbb{R}^d)$. 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 2.7 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

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

Moreover by (33) and the linearity of $A^{W,W}$ one gets

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

Now we apply the chain rule to $A_t^{W,W}(-\nabla u^* b - \tilde{f})$, namely Proposition 5.6 with $l = -\nabla u^* b - \tilde{f}$ on the RHS of (18). Note that in this case (29) holds for $\phi = u$ because the function u verifies (19) with $l = -\nabla u^* b - \tilde{f}$, see indeed (21). Thus, we get

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

so that

$$M_t = \int_0^t \nabla u^*(r, W_r) \, dW_r,$$

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

Corollary 5.14. *Under the hypothesis of Theorem 5.13 we have the Feynman–Kac (implicit) representation for the solution u of PDE (20) given by*

$$u(s, x_0) = \mathbb{E}\Bigg[\Phi(x_0 + W_{T-s})$$
$$+ \int_s^T f\big(r, W_r + x_0, u(r, W_r + x_0), \nabla u(r, W_r + x_0)\big) \, dr$$
$$+ A_T^{W,W}\big((\nabla u^* b)(x_0 + \cdot)\big) - A_s^{W,W}\big((\nabla u^* b)(x_0 + \cdot)\big) \Bigg]$$

for all $s \in [0, T]$ and $x_0 \in \mathbb{R}^d$.

Proof. For ease of proof, we show the result for $s = 0$. We set $\hat{f}(t, x, y, z) := f(t, x + x_0, y, z)$, $\hat{\Phi}(x) := \Phi(x + x_0)$ and (formally) $\hat{b}(t, x) := b(t, x + x_0)$. Let $\hat{u}$ be the solution of PDE (20)

where the coefficients b, f, Φ are replaced by $\hat{b}$, $\hat{f}$ and $\hat{\Phi}$. It is easy to see that $\hat{u}(t, x) = u(t, x + x_0)$, where u is the solution to the original PDE.

If we now consider BSDE (2) where the coefficients b, f, Φ are replaced by $\hat{b}$, $\hat{f}$ and $\hat{\Phi}$, then by Theorem 5.13 we know that $Y_t = \hat{u}(t, W_t) = u(t, x_0 + W_t)$ is a solution according to Definition 3.3. In particular, we have for all $t \in [0, T]$

$$Y_t = \hat{\Phi}(W_T) + \int_t^T \hat{f}\left(r, W_r, Y_r, \frac{[Y, W]_r}{\mathrm{d}r}\right) \mathrm{d}r$$
$$+ A_T^{W,Y}(\hat{b}) - A_t^{W,Y}(\hat{b}) - (M_T - M_t),$$

where M_t is an $\mathcal{F}_t$-martingale. We now use the explicit expression of Y in terms of u and Corollary 2.3 to replace the bracket, and taking the expectation we get for $t = 0$

$$u(0, x_0) = \mathbb{E}\left[\Phi(W_T + x_0)\right.$$
$$+ \int_0^T f\left(r, W_r + x_0, u(r, W_r + x_0), \nabla u(r, W_r + x_0)\right) \mathrm{d}r$$
$$\left. + A_T^{W,W}\left((\nabla u^* b)(x_0 + \cdot)\right)\right],$$

having used Proposition 5.9 in the last step to replace $A^{Y,W}$ with $A^{W,W}$. $\qquad\square$

We conclude with a result about the uniqueness of the solution Y in the class $Y_t = \gamma(t, W_t)$ for certain γs.

Proposition 5.15. *Let Assumption 2.6 and Assumption 2.7 hold and let $b \in C([0, T]; H_q^{-\beta})$. If the solution of (2) 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 (2) according to Definition 3.3 and let us denote by

$$M_t^i := Y_t^i - Y_0^i + A_t^{W,Y^i}(b) + \int_0^t f\left(r, W_r, Y_r^i, \frac{\mathrm{d}[Y^i, W]_r}{\mathrm{d}r}\right) \mathrm{d}r, \tag{35}$$

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

$$(\nabla\gamma^i)^* b \in C([0, T]; H_p^{-\beta}) \tag{36}$$

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

$$A_t^{Y^i, W}(b) = A_t^{W,W}\left((\nabla\gamma^i)^* b\right). \tag{37}$$

Furthermore by Corollary 2.3, we have

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

$$
= \int_0^t f\left(r, W_r, \gamma^i(r, W_r), \nabla\gamma^i(r, W_r)\right) \mathrm{d}r
$$

$$
= \int_0^t \tilde{f}^i(r, W_r)\, \mathrm{d}r, \tag{38}
$$

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

$$
\tilde{f}^i \in L^\infty_{\mathrm{loc}} \cap C\left([0, T]; L^p\right), \tag{39}
$$

which can be proven similarly to the considerations below (34) in the proof of the previous existence theorem. Thus we can apply Theorem 5.11, so $(38) = A_t^{W,W}(\tilde{f}^i)$. By (37) and the additivity of $A^{W,W}$ we have

$$
M_t^i = Y_t^i - Y_0^i + A_t^{W,W}\left((\nabla\gamma^i)^* b + \tilde{f}^i\right). \tag{40}
$$

Let us consider the PDE

$$
\begin{cases}
\partial_t h^i(t) + \dfrac{1}{2}\Delta h^i(t) = \left(\nabla\gamma^i\right)^*(t)b(t) + f\left(t, \cdot, \gamma^i, \nabla\gamma^i\right), \\
h^i(T) = 0,
\end{cases} \tag{41}
$$

which is PDE (18) 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 (39) and (36)) on the right-hand side in place of l. We denote by h^i, $i = 1, 2$ the corresponding (mild solution) expression (19), which belongs to $C([0, T]; C^{1+\alpha})$ by Lemma 4.3. Then $(\nabla h^i)^*$ is bounded. By the chain rule (Proposition 5.6), we get

$$
A_t^{W,W}\left((\nabla\gamma^i)^* b + \tilde{f}^i\right) = h^i(t, W_t) - h^i(0, W_0) - \int_0^t \left(\nabla h^i\right)^*(r, W_r)\, \mathrm{d}W_r. \tag{42}
$$

Plugging (42) into (40), we get

$$
M_t^i = \gamma^i(t, W_t) - \gamma^i(0, W_0) + h^i(t, W_t) - h^i(0, W_0)
$$

$$
- \int_0^t \left(\nabla h^i\right)^*(r, W_r)\, \mathrm{d}W_r.
$$

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

$$
\gamma^i(t, W_t) + h^i(t, W_t) = -\left(M_T^i - M_t^i\right) - \int_t^T \left(\nabla h^i\right)^*(r, W_r)\, \mathrm{d}W_r
$$

$$
+ \gamma^i(T, W_T) + h^i(T, W_T)
$$

$$
= \Phi(W_T) - \left(\tilde{M}_T^i - \tilde{M}_t^i\right), \tag{43}
$$

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 (41) 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 (43) can be represented by

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

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

$$\gamma^1(t, x) + h^1(t, x) = \gamma^2(t, x) + h^2(t, x), \tag{44}$$

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

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

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 (44). 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

$$\hat{h}^2(t) - \hat{h}^1(t) = \int_0^t P(t - r)\big((\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r))^*\hat{b}(r)\big)\,\mathrm{d}r$$

$$+ \int_0^t P(t - r)(\hat{f}(r, \hat{\gamma}^2(r), \nabla\hat{\gamma}^2(r) - \hat{f}(r, \hat{\gamma}^1(r), \nabla\hat{\gamma}^1(r)))\,\mathrm{d}r.$$

Taking the ρ-equivalent norm (see (12)) of the difference above, we have

$$\|\hat{h}^2 - \hat{h}^1\|_{C([0,T]; H_p^{1+\delta})}^{(\rho)}$$

$$= \sup_{0 \le t \le T} e^{-\rho t}\|\hat{h}^2(t) - \hat{h}^1(t)\|_{H_p^{1+\delta}}$$

$$\le \sup_{0 \le t \le T} e^{-\rho t}\|\int_0^t P(t - r)\big((\nabla\hat{\gamma}^2(r) - \nabla\hat{\gamma}^1(r))^*\hat{b}(r)\big)\,\mathrm{d}r\|_{H_p^{1+\delta}}$$

$$+ \sup_{0 \le t \le T} e^{-\rho t}\|\int_0^t P(t - r)(\hat{f}(r, \hat{\gamma}^2(r), \nabla\hat{\gamma}^2(r)$$

$$- \hat{f}\big(r, \hat{\gamma}^1(r), \nabla \hat{\gamma}^1(r)\big)\big) \, dr \|_{H_p^{1+\delta}}$$

$$=: (A) + (B).$$

To bound the first term we use the pointwise product estimate for fixed time $r \in [0, T]$ (Lemma 2.4), the mapping property (8) of the semigroup, and the definition of the ρ-equivalent norm (12). We get

$$(A) \le c \sup_{0 \le t \le T} \int_0^t e^{-\rho t} (t-r)^{-\frac{1+\delta+\beta}{2}} \|\hat{b}(r)\|_{H_q^{-\beta}} \|\nabla \hat{\gamma}^2(r) - \nabla \hat{\gamma}^1(r)\|_{H_p^\delta} \, dr$$

$$\le c \|\hat{b}\|_{C([0,T]; H_q^{-\beta})} \sup_{0 \le t \le T} \int_0^t e^{-\rho(t-r)} (t-r)^{-\frac{1+\delta+\beta}{2}}$$

$$\cdot e^{-\rho r} \|\hat{\gamma}^2(r) - \hat{\gamma}^1(r)\|_{H_p^{1+\delta}} \, dr$$

$$\le c \|\hat{\gamma}^2 - \hat{\gamma}^1\|_{C([0,T]; H_p^{1+\delta})}^{(\rho)} \sup_{0 \le t \le T} \int_0^t e^{-\rho(t-r)} (t-r)^{-\frac{1+\delta+\beta}{2}} \, dr$$

$$\le c \rho^{\frac{\delta+\beta-1}{2}} \|\hat{\gamma}^2 - \hat{\gamma}^1\|_{C([0,T]; H_p^{1+\delta})}^{(\rho)}, \tag{45}$$

having used the Gamma function and the bound

$$\int_0^t e^{-\rho r} r^\alpha \, dr \le \Gamma(\alpha+1) \rho^{-(\alpha+1)}$$

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 (22) of $\hat{f}$ so we get

$$(B) \le c \sup_{0 \le t \le T} \int_0^t e^{-\rho(t-r)} (t-r)^{-\frac{1+\delta}{2}}$$

$$\cdot e^{-\rho r} \big(c \|\hat{\gamma}^2(r) - \hat{\gamma}^1(r)\|_{H_p^{1+\delta}} + \|\nabla \hat{\gamma}^2(r) - \nabla \hat{\gamma}^1(r)\|_{H_p^\delta}\big) \, dr$$

$$\le c \rho^{\frac{\delta-1}{2}} \|\hat{\gamma}^2 - \hat{\gamma}^1\|_{C([0,T]; H_p^{1+\delta})}^{(\rho)}. \tag{46}$$

Collecting the estimates (45) and (46), we get

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

so

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

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 that $Y^1 = Y^2$. $\qquad\square$

Note that, if one wanted to generalize this framework to the non-Markovian case, one possibility would be to use functional Itô calculus. This approach however, would need an analytical study of a path-dependent PDE like (6), and the current analytical tools seem insufficient at the moment.

Appendix: A technical lemma and proofs of Lemma 5.1 and Theorem 5.11

We first state and prove a technical lemma that is used in the proofs below.

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 $K \subseteq \bigcup_{i=1}^m B(a_i, \delta)$. For a given $a \in H$ there exists $j \in \{1, \ldots, m\}$ such that $a \in B(a_j, \delta)$. Then we write

$$\|P_N a - a\| \leq \|P_N(a - a_j)\| + \|P_N a_j - a_j\| + \|a_j - a\|$$

$$\leq (1+c)\|a - a_j\| + \max_{i=1,\ldots,m} \|P_N a_i - a_i\|$$

$$\leq (1+c)\delta + \max_{i=1,\ldots,m} \|P_N a_i - a_i\|,$$

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

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

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

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

as wanted. $\qquad\square$

Before proving Lemma 5.1, we introduce the Haar wavelet functions and illustrate their use within the context of fractional Sobolev spaces H_r^s. For simplicity of notation, we recall only the

case of the Haar wavelets on $\mathbb{R}$ (see [37], 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

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

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

and

$$h_{j,m}(x) := h_M(2^j x - m), \quad j \in \mathbb{N}_0, m \in \mathbb{Z}.$$

Then the family

$$\{h_{j,m}, j \in \mathbb{N}_0 \cup \{-1\}, m \in \mathbb{Z}\} \tag{47}$$

is an unconditional basis of $H_r^s(\mathbb{R})$ for $2 \le r \le \infty$ and $-\frac{1}{2} < s < \frac{1}{r}$ by [37], 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 $d \ge 1$ is given in Theorem 2.21, (ii). Moreover for any $h \in H_r^s(\mathbb{R})$, 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

$$\mu_{j,m} := 2^{j(s-\frac{1}{r}+1)} \int_{\mathbb{R}} h(x) h_{j,m}(x) \, dx, \tag{48}$$

and the integral has to be interpreted as a dual pairing as mentioned in [37], Theorem 2.9, see also [36], Remark 1.14. Rewriting the same series with a different notation $\bar{\mu}_{j,m} := 2^j \int_{\mathbb{R}} h(x) h_{j,m}(x) \, dx$ we get another equivalent representation for h given by

$$h = \sum_{j=-1}^{\infty} \sum_{m \in \mathbb{Z}} \bar{\mu}_{j,m} h_{j,m}. \tag{49}$$

Defining the projector P_N as

$$P_N h := \sum_{j=-1}^{N} \sum_{m=-N}^{N} \bar{\mu}_{j,m} h_{j,m}, \tag{50}$$

for h of the form (49), then clearly $P_N h \in H_r^s(\mathbb{R})$ and

$$\|h - P_N h\|_{H_r^s(\mathbb{R}^d)} \to 0, \tag{51}$$

as $N \to \infty$.

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

$$\|P_N h\|_{H_r^s(\mathbb{R})} \le \|h\|_{H_r^s(\mathbb{R})}.$$

This can be seen as follows. We denote by $\mu(h)$ the collection of $\mu_{j,m}$ given by (48) for some h. Then for $2 \le r \le \infty$ the map $h \mapsto \mu(h)$ is an isomorphism between H_r^s 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 [37] 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 $\|\mu_{j,m}(P_N h)\|_{f_{r2}^-} \le \|\mu_{j,m}(h)\|_{f_{r2}^-}$ and this together with the isomorphism implies $\|P_N h\|_{H_r^s(\mathbb{R})} \le \|h\|_{H_r^s(\mathbb{R})}$ as stated.

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

Step 1. Density of $C_c^\infty(\mathbb{R}^d)$ in $H_r^s(\mathbb{R}^d)$. It is a known result that $C_c^\infty(\mathbb{R}^d)$ is dense in $H_r^s(\mathbb{R}^d)$ for all $1 < r < \infty$ and $-\infty < s < \infty$. For a proof see, for example, [35], Theorem in Section 2.3.2, part (b).

Step 2. Non-smooth approximating sequence for $l \in C([0,T]; H_r^s(\mathbb{R}^d))$. We consider $d = 1$ in the proof for simplicity of notation and explanation, but the same methodology extends to the case $d \ge 1$, see, for example, [37], 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 (47). Now let $l \in C([0,T]; H_r^s(\mathbb{R}))$ and let $t \in [0,T]$. We recall that $(P_N)_N$ defined by (50) is a family of linear operators acting on $H_r^s(\mathbb{R})$. The coefficients $\bar{\mu}$ of $P_N l(t)$ are now parametrized by time, namely $\bar{\mu}_{j,m}(t) = 2^j \int_{\mathbb{R}} l(t,x) h_{j,m}(x) \, dx$. By (51), we have that $P_N l(t) \to l(t)$ in $H_r^s(\mathbb{R})$ 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 $C([0,T]; H_r^s(\mathbb{R}))$ hence $P_N l \in C([0,T]; H_r^s(\mathbb{R}))$. We will now show that $P_N l \to l$ in $C([0,T]; H_r^s(\mathbb{R}))$, namely that

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \|l(t) - P_N l(t)\|_{H_r^s(\mathbb{R})} = 0. \tag{52}$$

To prove this, we want to use Lemma A.1 with the compact $K := \{l(t) : t \in [0,T]\} \subset H_r^s(\mathbb{R})$ and the projection P_N defined by (50). The family of functions $t \mapsto P_N l(t)$ is bounded in N in the space $C([0,T]; H_r^s(\mathbb{R}))$ by Remark A.2. Since $\{l(t), t \in [0,T]\}$ is a compact set in H_r^s then we can apply Lemma A.1 and we get (52).

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 $C_c^\infty([0,T] \times \mathbb{R})$ and such that $\sup_{t \in [0,T]} \|l_N(t) - \tilde{l}_N(t)\|_{H_r^s(\mathbb{R})} < \varepsilon$. 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 $j \in \{-1, 0, \ldots, N\}$ and $m \in \{-N, \ldots, 0, \ldots, N\}$. 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

$$\|\tilde{l}_N(t) - l_N(t)\|_{H_r^s(\mathbb{R})} = \|\sum_{j,m} \mu_{j,m}(t)(h_{j,m} - \tilde{h}_{j,m})\|_{H_r^s(\mathbb{R})}$$

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

$$< \varepsilon. \qquad \square$$

Below we give the proof of Theorem 5.11.

Proof of Theorem 5.11. The proof is split in two steps. In Step 1, we show that (32) holds for $g \in C([0, T]; H_r^s) \cap L_{loc}^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^d)$ 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 $h \in H_r^s(\mathbb{R}^d)$ by

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

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

$$P_N(A^{-s/2}h) = A^{-s/2}(P_N h). \tag{53}$$

Indeed by the definition of the norm in the $H_r^s(\mathbb{R}^d)$-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

$$\mathcal{F}(A^{-s/2}(P_N h))(\xi) = \left(1 + \frac{\xi^2}{2}\right)^{-s/2} \mathcal{F}(P_N h)(\xi)$$

$$= \left(1 + \frac{\xi^2}{2}\right)^{-s/2} \mathcal{F}(h)(\xi)\mathcal{F}(\phi_N)(\xi)$$

$$= \mathcal{F}(A^{-s/2}h)(\xi)\mathcal{F}(\phi_N)(\xi).$$

Taking the inverse Fourier transform on both sides we obtain the commutation property as stated in (53). Now it easily follows that

$$P_N h \to h \quad \text{in } H_r^s(\mathbb{R}^d), \text{ as } N \to \infty \tag{54}$$

for every $h \in H_r^s(\mathbb{R}^d)$, 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 (53). Moreover, P_N is a contraction in the same spaces, namely

$$\|P_N h\|_{H_r^s(\mathbb{R}^d)} \leq \|h\|_{H_r^s(\mathbb{R}^d)}. \tag{55}$$

This can be seen by observing that

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

where we have used (53), and the latter is bounded by $\|A^{-s/2}h\|_{H_r^0(\mathbb{R}^d)} = \|h\|_{H_r^s(\mathbb{R}^d)}$ because P_N is a contraction operator in $H_r^0(\mathbb{R}^d) = L^r(\mathbb{R}^d)$. Property (55) 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 $H_r^s(\mathbb{R}^d)$. Indeed for any sequence $t_k \to t$ we have

$$\|P_N g(t_k, \cdot) - P_N g(t, \cdot)\|_{H_r^s(\mathbb{R}^d)} = \|P_N(g(t_k, \cdot) - g(t, \cdot))\|_{H_r^s(\mathbb{R}^d)}$$
$$\leq \|g(t_k, \cdot) - g(t, \cdot)\|_{H_r^s(\mathbb{R}^d)},$$

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

$$P_N g \to g \tag{56}$$

in $C([0, T]; H_r^s(\mathbb{R}^d))$, we use Lemma A.1 with $H = H_r^s(\mathbb{R}^d)$. We can do so since the family of operators $(P_N)_N$ is linear and equibounded in $H_r^s(\mathbb{R}^d)$ by (55), and it fulfils (54). Thus, defining the compact K in $H_r^s(\mathbb{R}^d)$ by $K := \{g(t) : t \in [0, T]\}$ we have

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

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 (32) holds for g replaced by $P_N g$ thanks to (24). Moreover by (56), we can apply Remark 5.5, part 2. and get

$$\lim_{N \to \infty} A^{W,W}(P_N g) = A^{W,W}(g) \quad \text{in } \mathcal{C}. \tag{57}$$

Finally, we can see that

$$\int_0^{\cdot} P_N g(s, W_s)\, ds \to \int_0^{\cdot} g(s, W_s)\, ds \tag{58}$$

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

$$E\left[\sup_{0 \leq t \leq T} \left| \int_0^t (P_N g - g)(s, W_s)\, ds \right| \right]$$

$$\leq E\left[\sup_{0 \leq t \leq T} \int_0^t |P_N g - g|(s, W_s)\, ds \right]$$

$$\leq \int_0^T \int_{\mathbb{R}^d} |P_N g - g|(s, y)\, p_s(y)\, dy\, ds, \tag{59}$$

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 $|P_N g(s, x)| \leq \|g\|_{L^\infty([0,T] \times \mathbb{R}^d)} \leq C$, because g is bounded by assumption. This, together with the fact that

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

implies that (59) 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 (59) converges to 0. This implies (58) and with (57) we conclude.

Step 2. General case $g \in C([0, T]; H_r^s(\mathbb{R}^d)) \cap L^\infty_{loc}([0, T] \times \mathbb{R}^d)$.

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

$$\int_0^{t \wedge \tau_M} g(s, W_s)\, ds = \int_0^{t \wedge \tau_M} g_M(s, W_s)\, ds. \tag{60}$$

On the other hand we know that g_M is bounded and has compact support by definition, and that $g_M \in C([0, T]; H_r^s(\mathbb{R}^d))$ because g is in the same space and χ_M is smooth (using the pointwise multipliers property, see [36], Section 2.2.2). So Step 1 applies to g_M

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

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

$$A_{t \wedge \tau_M}^{W,W}(g_M) = \int_0^{t \wedge \tau_M} g_M(s, W_s)\, ds. \tag{61}$$

Now we want to show that

$$A^{W,W}_{\cdot \wedge \tau_M}(g_M) = A^{W,W}_{\cdot \wedge \tau_M}(g).$$ (62)

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 $g_M^n := g^n \chi_M$ for each n, and this is an approximating sequence for g_M in $C([0,T]; H_r^s(\mathbb{R}^d))$. Indeed the linear map $\phi \mapsto \phi \chi_M$ is continuous in $C([0,T]; H_r^s(\mathbb{R}^d))$ by [36], equation (2.50), namely there exists a constant $c(M)$ only dependent on χ_M such that

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

Then

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

and since g^n converges to g in $C([0,T]; H_r^s(\mathbb{R}^d))$ then so does g_M^n to g_M.

For each n we have

$$A^{W,W}_{\cdot \wedge \tau_M}(g_M^n) = A^{W,W}_{\cdot \wedge \tau_M}(g^n)$$ (63)

because both sides are defined explicitly and the two functions coincide before τ_M. We note that $A^{W,W}_{\cdot}(g_M^n)$ (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 (63) also converge u.c.p., hence we get (62). This, together with (60) and (61) gives

$$\int_0^{\cdot \wedge \tau_M} g(s, W_s)\, ds = A^{W,W}_{\cdot \wedge \tau_M}(g).$$ (64)

For almost all ω there exists $n_0(\omega)$ such that for all $M > n_0(\omega)$ we have $\tau_M(\omega) \geq T$, then taking the limit as $M \to \infty$ of (64) we conclude. □

Acknowledgments

References

[1] Barrasso, A. and Russo, F. (2017). Backward stochastic differential equations with no driving martingale, Markov processes and associated pseudo partial differential equations. Part II: Decoupled mild solutions and examples. Preprint. Available at hal-01505974.

[2] Bismut, J.-M. (1973). Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* **44** 384–404. MR0329726 https://doi.org/10.1016/0022-247X(73)90066-8

[3] Bouleau, N. and Yor, M. (1981). Sur la variation quadratique des temps locaux de certaines semimartingales. *C. R. Acad. Sci., Sér. 1 Math.* **292** 491–494. MR0612544

[4] Cheridito, P. (2003). Arbitrage in fractional Brownian motion models. *Finance Stoch.* **7** 533–553. MR2014249 https://doi.org/10.1007/s007800300101

[5] Cheridito, P. and Nam, K. (2017). BSE's, BSDE's and fixed-point problems. *Ann. Probab.* **45** 3795–3828. MR3729615 https://doi.org/10.1214/16-AOP1149

[6] Coviello, R., di Girolami, C. and Russo, F. (2011). On stochastic calculus related to financial assets without semimartingales. *Bull. Sci. Math.* **135** 733–774. MR2838099 https://doi.org/10.1016/j.bulsci.2011.06.008

[7] Davies, E.B. (1989). *Heat Kernels and Spectral Theory. Cambridge Tracts in Mathematics* **92**. Cambridge: Cambridge Univ. Press. MR0990239 https://doi.org/10.1017/CBO9780511566158

[8] Delarue, F. and Diel, R. (2016). Rough paths and 1d SDE with a time dependent distributional drift: Application to polymers. *Probab. Theory Related Fields* **165** 1–63. MR3500267 https://doi.org/10.1007/s00440-015-0626-8

[9] Diehl, J. and Zhang, J. (2017). Backward stochastic differential equations with Young drift. *Probab. Uncertain. Quant. Risk* **2** Paper No. 5. MR3663305 https://doi.org/10.1186/s41546-017-0016-5

[10] El Karoui, N., Peng, S. and Quenez, M.C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71. MR1434407 https://doi.org/10.1111/1467-9965.00022

[11] Engelbert, H.-J. and Wolf, J. (1998). Strong Markov local Dirichlet processes and stochastic differential equations. *Teor. Veroyatn. Primen.* **43** 331–348. MR1679006 https://doi.org/10.1137/S0040585X97976829

[12] Erraoui, M., Ouknine, Y. and Sbi, A. (1998). Reflected solutions of backward stochastic differential equations with distribution as terminal condition. *Random Oper. Stoch. Equ.* **6** 1–16. MR1610189 https://doi.org/10.1515/rose.1998.6.1.1

[13] Flandoli, F., Issoglio, E. and Russo, F. (2017). Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.* **369** 1665–1688. MR3581216 https://doi.org/10.1090/tran/6729

[14] Flandoli, F., Russo, F. and Wolf, J. (2003). Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.* **40** 493–542. MR1988703

[15] Flandoli, F., Russo, F. and Wolf, J. (2004). Some SDEs with distributional drift. II. Lyons–Zheng structure, Itô's formula and semimartingale characterization. *Random Oper. Stoch. Equ.* **12** 145–184. MR2065168 https://doi.org/10.1163/156939704323074700

[16] Gozzi, F. and Russo, F. (2006). Weak Dirichlet processes with a stochastic control perspective. *Stochastic Process. Appl.* **116** 1563–1583. MR2269216 https://doi.org/10.1016/j.spa.2006.04.009

[17] Hinz, M., Issoglio, E. and Zähle, M. (2014). Elementary pathwise methods for nonlinear parabolic and transport type stochastic partial differential equations with fractal noise. In *Modern Stochastics and Applications. Springer Optim. Appl.* **90** 123–141. Cham: Springer. MR3236072 https://doi.org/10.1007/978-3-319-03512-3_8

[18] Issoglio, E. (2013). Transport equations with fractal noise – existence, uniqueness and regularity of the solution. *Z. Anal. Anwend.* **32** 37–53. MR3061360 https://doi.org/10.4171/ZAA/1473

[19] Issoglio, E. (2019). A non-linear parabolic PDE with a distributional coefficient and its applications to stochastic analysis. *J. Differential Equations* **267** 5976–6003. MR3996791 https://doi.org/10.1016/j.jde.2019.06.014

[20] Issoglio, E. and Jing, S. (2018). Forward–backward SDEs with distributional coefficients. *Stochastic Process*. *Appl*., in press. https://doi.org/10.1016/j.spa.2019.01.001

[21] Issoglio, E. and Zähle, M. (2015). Regularity of the solutions to SPDEs in metric measure spaces. *Stoch*. *Partial Differ. Equ. Anal. Comput*. **3** 272–289. MR3350453 https://doi.org/10.1007/s40072-015-0048-8

[22] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

[23] Lunardi, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. *Progress in Nonlinear Differential Equations and Their Applications* **16**. Basel: Birkhäuser. MR1329547 https://doi.org/10.1007/978-3-0348-9234-6

[24] Mathieu, P. (1994). Zero white noise limit through Dirichlet forms, with application to diffusions in a random medium. *Probab. Theory Related Fields* **99** 549–580. MR1288070 https://doi.org/10.1007/BF01206232

[25] Pardoux, É. and Peng, S.G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett*. **14** 55–61. MR1037747 https://doi.org/10.1016/0167-6911(90)90082-6

[26] Pardoux, E. and Răşcanu, A. (2014). *Stochastic Differential Equations*, *Backward SDEs*, *Partial Differential Equations*. *Stochastic Modelling and Applied Probability* **69**. Cham: Springer. MR3308895 https://doi.org/10.1007/978-3-319-05714-9

[27] Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. *Applied Mathematical Sciences* **44**. New York: Springer. MR0710486 https://doi.org/10.1007/978-1-4612-5561-1

[28] Pham, H. (2009). *Continuous-Time Stochastic Control and Optimization with Financial Applications*. *Stochastic Modelling and Applied Probability* **61**. Berlin: Springer. MR2533355 https://doi.org/10.1007/978-3-540-89500-8

[29] Runst, T. and Sickel, W. (1996). *Sobolev Spaces of Fractional Order*, *Nemytskij Operators*, *and Nonlinear Partial Differential Equations*. *De Gruyter Series in Nonlinear Analysis and Applications* **3**. Berlin: de Gruyter. MR1419319 https://doi.org/10.1515/9783110812411

[30] Russo, F. and Trutnau, G. (2007). Some parabolic PDEs whose drift is an irregular random noise in space. *Ann*. *Probab*. **35** 2213–2262. MR2353387 https://doi.org/10.1214/009117906000001178

[31] Russo, F. and Vallois, P. (1995). The generalized covariation process and Itô formula. *Stochastic Process*. *Appl*. **59** 81–104. MR1350257 https://doi.org/10.1016/0304-4149(95)93237-A

[32] Russo, F. and Vallois, P. (2007). Elements of stochastic calculus via regularization. In *Séminaire de Probabilités XL*. *Lecture Notes in Math*. **1899** 147–185. Berlin: Springer. MR2409004 https://doi.org/10.1007/978-3-540-71189-6_7

[33] Russo, F. and Wurzer, L. (2017). Elliptic PDEs with distributional drift and backward SDEs driven by a càdlàg martingale with random terminal time. *Stoch*. *Dyn*. **17** 1750030. MR3646613 https://doi.org/10.1142/S0219493717500307

[34] Seignourel, P. (2000). Discrete schemes for processes in random media. *Probab. Theory Related Fields* **118** 293–322. MR1800534 https://doi.org/10.1007/PL00008743

[35] Triebel, H. (1978). *Interpolation Theory, Function Spaces, Differential Operators*. *North-Holland Mathematical Library* **18**. Amsterdam: North-Holland. MR0503903

[36] Triebel, H. (2008). *Function Spaces and Wavelets on Domains*. *EMS Tracts in Mathematics* **7**. Zürich: Eur. Math. Soc. MR2455724 https://doi.org/10.4171/019

[37] Triebel, H. (2010). *Bases in Function Spaces*, *Sampling*, *Discrepancy*, *Numerical Integration*. *EMS Tracts in Mathematics* **11**. Zürich: Eur. Math. Soc. MR2667814 https://doi.org/10.4171/085

[38] Üstünel, A.S. (1982). A generalization of Itô's formula. *J. Funct. Anal*. **47** 143–152. MR0664333 https://doi.org/10.1016/0022-1236(82)90102-1