Sensitivity Results in Stochastic Analysis

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

Vorgelegt von

Stefan Patrick Koch
aus Karlsruhe

Mannheim, 2019

Dekan: Dr. Bernd Lübcke, Universität Mannheim
Referent: Prof. Dr. Andreas Neuenkirch, Universität Mannheim
Korreferent: Prof. Dr. Samy Tindel, Purdue University

Tag der mündlichen Prüfung: 10.09.2019

Abstract

This thesis consists of two quite distinct topics. In the first and bigger part we show that the Mandelbrot-van Ness representation of fractional Brownian motion is almost surely smooth in the Hurst parameter H. This dependence result is transferred to the solution of a stochastic differential equation driven by fractional Brownian motion if the stochastic differential equation is one-dimensional or $H > \frac{1}{2}$. In the multidimensional case of $H \in (\frac{1}{3}, \frac{1}{2}]$ we use rough path theory to make sense of the differential equations. However, despite it being possible to lift fractional Brownian motion as well as its derivative in H to a rough path via the limit of dyadic approximations, they cannot be lifted jointly in the same way. Nevertheless, we obtain that the solution to a rough stochastic differential equation driven by fractional Brownian motion is locally Lipschitz continuous in H.
In the last part of the thesis we define a directional Malliavin derivative connected to a continuous linear operator. We show that this directional Malliavin derivative being zero is equivalent to some measurability or independence condition on the random variable. Using this result, we obtain that two random variables, whose classical Malliavin derivatives live in orthogonal subspaces, are independent. We also extend the chain rule to directional Malliavin derivatives and a broader class of functions with weaker regularity assumptions.

Zusammenfassung

Diese Arbeit setzt sich aus zwei unterschiedlichen Themenblöcken zusammen. Im ersten und größeren Teil zeigen wir, dass die Mandelbrot-van Ness Darstellung der fraktionalen Brownschen Bewegung glatt im Hurstparameter H ist. Dieses Resultat lässt sich auf die Lösung einer stochastischen Differentialgleichung, welche von der fraktionalen Brownschen Bewegung getrieben wird, übertragen, falls die Differentialgleichung eindimensional ist oder $H > \frac{1}{2}$. Im mehrdimensionalen Fall mit $H \in (\frac{1}{3}, \frac{1}{2}]$ nutzen wir die Rough Path Theorie um den Differentialgleichungen einen Sinn zu geben. Jedoch, obwohl es möglich ist sowohl die fraktionale Brownsche Bewegung als auch ihren Ableitungsprozess mit Hilfe dyadischer Approximationen zu einem Rough Path zu erweitern, lässt sich auf diese Art kein *gemeinsamer* Rough Path konstruieren. Trotzdem können wir zeigen, dass die Lösung einer stochastischen Rough Path Differentialgleichung lokal Lipschitz-stetig in H ist.
Im letzten Teil der Arbeit definieren wir eine Malliavin-Richtungsableitung basierend auf einem beschränkten linearen Operator. Wir zeigen, dass eine Messbarkeits- oder Unabhängigkeitsbedingung an eine Zufallsvariable equivalent dazu ist, dass ihre Malliavin-Richtungsableitung Null ist. Mit Hilfe dieses Resultats lässt sich zeigen, dass zwei Zufallsvariablen, deren klassische Malliavin-Ableitungen in zueinander orthogonalen Unterräumen liegen, unabhängig sind. Des Weiteren verallgemeinern wir die Kettenregel für Malliavin-Richtungsableitungen und schwächere Regularitätsannahmen.

Acknowledgements

First of all, I am deeply grateful to my supervisor Andreas Neuenkirch for offering me this research opportunity in the first place but also for his guidance, advice and help and the many fruitful discussions throughout my research.

Moreover, I want to thank Samy Tindel for co-referring this thesis and the examination committee for kindly agreeing to evaluate my thesis defense.

I am grateful to *Deutsche Forschungsgemeinschaft*, which through the RTG 1953 *Statistical Modeling of Complex Systems* funded my research, enabled the research group to hold many interesting seminars on variety of mathematical topics and allowed me to travel to several conferences to present my results to the scientific community. I would also like to thank all my colleagues in this research group and the institute of mathematics in Mannheim who were always helpful and willing to discuss my mathematical problems.

I want to give a special mention to Philip and Samuel who shared an office with me for most of my time in Mannheim. It has been great to have them around, not only as colleagues but also as friends, to talk about mathematics and to share our troubles on our way to a PhD. Discussion on politics and football or the odd game of *Geoguessr* made my time as a PhD student even more enjoyable.

Last but not least I want to thank my family and close friends, in particular my parents and my partner Sandra, for their unconditional support and love, which helped me through all the ups and downs of writing this thesis.

Contents

1 Introduction

My research as a PhD student started with the aim to analyse parameter sensitivities of a rough volatility model. Like often in research, things work out differently than planned but nevertheless I explain the initial idea of my research in this introduction to motivate why results on such different topics are presented throughout this dissertation. On the one hand we present results on sensitivities of stochastic differential equations (SDEs) driven by fractional Brownian motion (fBm) with respect to the Hurst parameter, on the other hand we work with Malliavin calculus and obtain a characterisation of independence for Malliavin differentiable random variables. Let us first take a look at the original goal of this thesis.

In [13] the authors argue that financial markets should be modelled using a rough volatility model, where the randomness in the volatility stems from a fractional Ornstein-Uhlenbeck process. The fractional Ornstein-Uhlenbeck process differs from a non-fractional one in that the corresponding stochastic differential equation is driven by a fractional Brownian motion instead of a standard Brownian motion, i.e.

$$\mathrm{d}V_t^H = \kappa(\lambda - V_t^H)\,\mathrm{d}t + \theta\,\mathrm{d}B_t^H, \qquad V_0^H = v_0,$$

where $\kappa, \theta > 0,\ \lambda, v_0 \in \mathbb{R}$ and $B^H = (B_t^H)_{t\geq 0}$ is a fractional Brownian motion. This equation has a unique closed form solution that is obtained pathwise (cf. [3]). With the process V from above and inspired by [13] the aim was to look at the parameter sensitivities of the rough volatility model given by

$$\begin{aligned} S_0 &= s_0 \exp(X_t), \\ X_t &= -\frac{1}{2}\int_0^t \sigma^2(V_s)\,\mathrm{d}s + \int_0^t \sigma(V_s)\,\mathrm{d}W_s, \end{aligned} \tag{1.1}$$

where $W = (W_t)_{t\geq 0}$ is a standard Brownian motion and $\sigma : \mathbb{R} \to [0,\infty)$ fulfils certain conditions. We were particularly interested in the Greeks of financial options in this model. Given a terminal time point $T > 0$ the payoff of a European option is given by $f(S_T)$, where f is usually assumed to lie in some function class. The derivatives of the option price with respect to the model parameters are called Greeks, e.g. the option Delta is defined as

$$\Delta = \frac{\partial}{\partial s_0}\mathbb{E}[f(S_T)].$$

Under some smoothness and growth conditions on f the derivative operator can be pulled inside the expectation. This might be unsatisfactory since for many options the payoff function f is not even continuous, let alone differentiable.

But there are results for discontinuous payoff functions in simpler models, e.g. given the model

$$\mathrm{d}S_t = b(S_t)\,\mathrm{d}t + \sigma(S_t)\,\mathrm{d}W_t, \qquad S_0 = s_0$$

Greeks are calculated in [9] using Malliavin calculus. Under certain conditions one obtains, even for some discontinuous f, that

$$\Delta = \frac{\partial}{\partial s_0}\mathbb{E}[f(S_T)] = \mathbb{E}[f(S_T)\Pi],$$

where Π is a so-called Malliavin weight which depends on S_T but not on the function f. The initial goal of this thesis was, motivated by these results, to use Malliavin calculus to obtain Greeks or sensitivities of rough volatility models like (1.1).

As already mentioned this is not the research presented in this thesis but it is strongly linked with what my actual research turned out to be and the initial motivation behind it. My research splits into two, quite distinct parts. We are interested in the sensitivity with respect to the Hurst parameter of fractional Brownian motion and stochastic differential equations driven by it. Since fractional Brownian motion is in general not a semimartingale the usual theory of stochastic differential equations does not apply and other solution concepts are needed. The second part concerns itself with Malliavin calculus, in particular we present a characterisation of independence of two random variables.

An introduction to the different topics can be found in the corresponding chapters, nevertheless a short motivation follows.

1.1 Fractional Brownian motion and SDEs

Fractional Brownian motion is a centred Gaussian process $B^H = (B^H_t)_{t\geq 0}$ with continuous sample paths and covariance

$$\mathbb{E}B^H_s B^H_t = \frac{1}{2}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big), \qquad s,t \geq 0.$$

The parameter $H \in (0,1)$ is called Hurst parameter and for $H = \frac{1}{2}$ the fractional Brownian motion recovers the standard Brownian motion.

In recent years, the analysis of fractional Brownian motion itself and of stochastic differential equations driven by fBm has been an active field of research. However, the dependence of fBm and related SDEs on the Hurst parameter has received only little attention. Continuity of the law with respect to the Hurst parameter has been studied in a series of articles by Jolis and Viles [20, 21, 22, 23] for (iterated) Wiener integrals with respect to fBm, the local time of fBm and for the symmetric Russo-Vallois integral with fBm as an integrator. Moreover, Theorem 43 in [12] implies that the law of SDEs driven by fBm (understood in the rough paths sense) with Hurst parameter $H > 1/4$ depends continuously on the Hurst parameter. A stronger notion of continuous dependence is studied in [46] for scalar SDEs driven by fBm, i.e.

$$\mathrm{d}X^H_t = b(X^H_t)\,\mathrm{d}t + \sigma(X^H_t)\,\mathrm{d}B^H_t, \quad t \in [0,T], \qquad X^H_0 = x_0 \in \mathbb{R}, \tag{1.2}$$

with $b, \sigma : \mathbb{R} \to \mathbb{R}$. Under an ellipticity assumption on σ and otherwise standard smoothness assumptions on the coefficients the authors establish the existence of a constant $C_T > 0$ such that

$$\sup_{t\in[0,T]} \big|\mathbb{E}\varphi(X^H_t) - \mathbb{E}\varphi(X^{\frac{1}{2}}_t)\big| \leq C_T\ (H - \tfrac{1}{2}), \qquad H \in [1/2, 1),$$

for bounded test functions $\varphi \in C^{2+\beta}(\mathbb{R};\mathbb{R})$ with $\beta > 0$. Note that for $H > \frac{1}{2}$ the SDE (1.2) is understood pathwise as a Riemann-Stieltjes integral equation, while for $H = \frac{1}{2}$ it coincides with a Stratonovich SDE. Furthermore, in [46] the authors also establish a similar result for the Laplace-transform of a first passage time of SDE (1.2).

We analyse pathwise dependence on H and show that the Mandelbrot-van Ness representation of fBm (cf. [34]) as well as the solution of an SDE driven by it are differentiable in the Hurst parameter H. In applications, like rough volatility models, H is estimated. Differentiability in H allows to control the error of the SDE solution by the estimation error made in estimating H.

1.2 Malliavin Calculus

Malliavin calculus extends the calculus of variations to the stochastic framework and hence is sometimes called *stochastic calculus of variations*. It was originally developed by Paul Malliavin [33] to find a probabilistic proof to Hörmander's theorem [16]. This theorem gives conditions that guarantee that the law of the solution to an SDE has a smooth density. But nowadays there are more applications for Malliavin calculus. The adjoint of the Malliavin derivative, the divergence operator, enables us to do anticipating stochastic calculus. This extends the Itô integral to non-adapted integrands. The resulting integral is called the Skorokhod integral and allows to consider SDEs, where the solution is not necessarily adapted to the natural filtration of the driving Brownian motion. Furthermore, the integration by parts formula can be applied to obtain parameter sensitivities of SDEs. This way Malliavin calculus is a useful tool in calculating *Greeks* in financial models (cf. [9]). Another application in mathematical finance is the Clark-Ocone theorem, which is helpful in obtaining a replicating portfolio for derivatives.

In this thesis we specifically look at directional Malliavin derivatives. When working with a solution to an SDE driven by a d-dimensional Brownian motion, where $d \geq 2$, it is often easier to consider directional Malliavin derivatives of the solution, i.e. the Malliavin derivative with respect to the i-th Brownian motion, where $i = 1, \dots, d$. The composition of these d directional Malliavin derivatives then coincides with the usual Malliavin derivative (cf. Proposition 5.3.4). Looking at the our definition of a directional Malliavin derivative, we see that this definition already incorporates other existing concepts like the operator D^h, defined by $D^h F = \langle DF, h \rangle_H$ (cf. [38] p.27). But mainly, similar to the operator D^h, it turns out that our directional Malliavin derivative is a great tool for proving results about ordinary Malliavin calculus. In our case, analysing what it means for a random variable to have a directional Malliavin derivative of zero gives insights into the independence of random variables. In particular, we obtain a characterisation of independence for two Malliavin differentiable random variables.

A more detailed introduction is given in Chapter 5

1.3 Outline

Chapter 2, 3 and 4 are based on joint work with my supervisor Andreas Neuenkirch with part of the work already published in *The Mandelbrot-van Ness fractional Brownian motion is infinitely differentiable with respect to its Hurst parameter* [27]. Chapter 5 is based on my paper *Directional Malliavin Derivatives: A Characterisation of Independence and a Generalised Chain Rule* [26].

Chapter 2

We give a short introduction into fractional Brownian motion. Then, to analyse the pathwise dependence of fractional Brownian motion on the Hurst parameter, we need to introduce a coupling for fBms of different Hurst parameters. We study the Mandelbrot-van Ness representation of fBm $B^H = (B_t^H)_{t \geq 0}$ with Hurst parameter $H \in (0, 1)$ and show that for arbitrary fixed $t \geq 0$ the mapping $(0, 1) \ni H \mapsto B_t^H \in \mathbb{R}$ is almost surely infinitely differentiable. Thus, the sample paths of fractional Brownian motion are smooth with respect to H. This allows us to define our own representation, a *Mandelbrot-van Ness type fractional Brownian motion* that is pathwise continuously differentiable in H.

Chapter 3

We analyse the dependence on the Hurst parameter H of solutions to stochastic differential equations driven by fractional Brownian motion. To make sense of this type of SDEs we need to find a suitable solution concept. These concepts will, other than Itô calculus, all work pathwise. In one dimension we use the Doss-Sussmann approach [6, 45]. We show that the solution map, which maps the driving signal of the SDE to its solution, is Frechét differentiable on the space of continuous functions. Then a chain rule argument proves that solutions in the Doss-Sussmann sense are differentiable in the Hurst parameter of the driving fractional Brownian motion.

This result can be replicated in higher dimensions if $H > 1/2$. In this case the SDE can be understood as a pathwise integral equation in the Riemann-Stieltjes sense. Choosing appropriate Banach spaces, it has be shown that the solution map is Frechét differentiable [41] and it is again by a chain rule argument that we can conclude that the solution to the SDE is differentiable in the Hurst parameter.

The multidimensional case of $H \leq 1/2$ is more involved and handled in the next chapter.

Chapter 4

In this chapter we consider the case of multidimensional SDEs, where the driving fractional Brownian motion has a Hurst parameter $H \in (\frac{1}{3}, \frac{1}{2}]$. After a short introduction into rough path theory (cf. [10, 11, 32]), we show that, using the methods developed in [4] and [28], the derivative with respect to the Hurst parameter of a multidimensional fractional Brownian motion can be lifted to a geometric rough path in a natural fashion.

However, the dyadic rough path approximation of fractional Brownian motion together with its derivative in H does not converge in expected p-variation distance. This means that despite it being possible to lift both, fractional Brownian motion as well as its derivative in H, to a rough path via the limit of dyadic approximations, they cannot be lifted jointly in the same way.

In the last part of this chapter we show that a rough SDE driven by a lifted fractional Brownian motion is locally Lipschitz continuous in H. In a very restrictive case, we are able construct a derivative of the solution in a rough paths sense.

Chapter 5

This chapter is about Malliavin calculus. We define a directional Malliavin derivative connected to a continuous linear operator. We show that this directional Malliavin derivative being zero is equivalent to some measurability or independence condition on the random variable. Using this result, we obtain that two random variables, whose classical Malliavin derivatives live in orthogonal subspaces, are independent. We also extend the chain rule to directional Malliavin derivatives and a broader class of functions with weaker regularity assumptions.

2 Fractional Brownian motion

As mentioned in Section 1.1 a fractional Brownian motion is a centred Gaussian process $B^H = (B^H_t)_{t\geq 0}$ with continuous sample paths and covariance

$$\mathbb{E}\big[B^H_s B^H_t\big] = \frac{1}{2}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big), \qquad s,t \geq 0.$$

The parameter $H \in (0,1)$ is called Hurst parameter and for $H = \frac{1}{2}$ fractional Brownian motion coincides with the standard Brownian motion. Before we start working with fractional Brownain motion, we quickly present some features of fBm whose proofs can, for example, be found in [37]. Fractional Brownian motion possesses some characteristics similar to classical Brownian motion like

(i) self-similarity:

$$(a^{-H} B^H_{at})_{t\geq 0} \overset{d}{=} (B^H_t)_{t\geq 0}, \qquad \forall a > 0,$$

(ii) stationary increments:

$$(B^H_{t+s} - B^H_s)_{t\geq 0} \overset{d}{=} (B^H_t)_{t\geq 0}, \qquad \forall s > 0,$$

(iii) time inversion:

$$(t^{2H} B^H_{1/t})_{t\geq 0} \overset{d}{=} (B^H_t)_{t\geq 0},$$

where the d above the equals sign indicates equality in distribution. But there are other properties of Brownian motion that cannot be attained by fractional Brownian motion. In particular, for $H \neq \frac{1}{2}$, the fractional Brownian motion B^H is neither a semimartingale nor a Markov process. Further, the increments of B^H are stationary but no longer independent and are negatively correlated if $H < \frac{1}{2}$. For $H > \frac{1}{2}$ the increments of fractional Brownian motion are positively correlated and exhibit *long-range dependence*, i.e.

$$\sum_{k=1}^{\infty} \rho(k) = \infty,$$

where $\rho(k) = \mathbb{E}[B^H_1(B^H_{k+1} - B^H_k)]$ is the autocovariance of the increments. This long-range dependence was a desired feature in applications in hydrology and mathematical finance and made fractional Brownian motion an interesting object of study (cf. [17]).

2.1 Main result

The aim of this chapter is to analyse the pathwise smoothness with respect to the Hurst parameter. For this we need to choose a specific representation of fBm. Here we choose the so-called Mandelbrot-van Ness representation ([34]). So, let $T > 0$ and $B = (B_t)_{t\in\mathbb{R}}$ be a two-sided Brownian motion on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then

$$B_t^H = C_H \int_{\mathbb{R}} K_H(s,t)\,\mathrm{d}B_s, \quad t \in [0,T],$$

with

$$C_H = \frac{\Big(2H\sin(\pi H)\Gamma(2H)\Big)^{1/2}}{\Gamma(H+1/2)}$$

and

$$K_H(s,t) = \Big(|t-s|^{H-1/2} - |s|^{H-1/2}\Big)\,\mathbf{1}_{(-\infty,0)}(s) + |t-s|^{H-1/2}\,\mathbf{1}_{[0,t)}(s), \qquad (2.1)$$

defines a fBm on $[0,T]$ with Hurst parameter $H \in (0,1)$. Since $x^0 = 1$ we recover in particular that $K_{1/2}(s,t) = \mathbb{1}_{[0,t)}(s)$. Our main result is:

Theorem 2.1.1. *Let $k \in \mathbb{N}$. Then there exists a process $B^{H,k} = (B_t^{H,k})_{t\in[0,T]}$ such that:*

(i) For all $\omega \in \Omega$ the sample paths $(0,1)\times[0,T] \ni (H,t) \mapsto B_t^{H,k}(\omega) \in \mathbb{R}$ are continuous.

(ii) For all $\omega \in \Omega$ and for any fixed $H \in (0,1)$ and $\alpha \in (0,H)$ the sample paths $[0,T] \ni t \mapsto B_t^{H,k}(\omega) \in \mathbb{R}$ are α-Hölder continuous. We even have, for all $0 < a \le b < 1$ and $0 < \gamma < a$, that there exists a constant C depending on $\omega \in \Omega$ such that

$$\sup_{H\in[a,b]} \left|B_t^{H,k}(\omega) - B_s^{H,k}(\omega)\right| \le C(\omega)|t-s|^{\gamma} \qquad s,t \in [0,T].$$

(iii) For all $0 < a \le b < 1$, $t \in [0,T]$ there exists $\Omega_{a,b,k,t} \in \mathcal{A}$ such that $\mathbb{P}(\Omega_{a,b,k,t}) = 1$ and

$$\frac{\partial^k}{\partial H^k} B_t^H(\omega) = B_t^{H,k}(\omega), \qquad H \in [a,b], \quad \omega \in \Omega_{a,b,k,t}.$$

In particular, we have for fixed $t \in [0,T]$ that $B_t^{(\cdot)} \in C^\infty((0,1);\mathbb{R})$ a.s.

2.2 Preliminaries: Stochastic Fubini and continuity

The stochastic Fubini theorem is well known to hold for finite time intervals, see e.g. [43], Theorem 65, p. 211f. From now on let $I \subseteq \mathbb{R}$ be a (possibly infinite) interval. Further, let $\{I_n\}_{n\in\mathbb{N}}$ be an increasing sequence of finite intervals, i.e. $I_n \subseteq I_{n+1}$, such that $I = \bigcup_{n\in\mathbb{N}} I_n$. Let $J \subseteq \mathbb{R}$ be a further interval with $\mu(J) < \infty$, where μ denotes the Lebesgue measure on $\mathbb{R}$. Moreover, we always work on a complete filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in\mathbb{R}}, \mathbb{P})$ and with a two-sided $(\mathcal{F}_t)_{t\in\mathbb{R}}$-Brownian motion $W = (W_t)_{t\in\mathbb{R}}$ on this space.

Lemma 2.2.1. *Let $G^h = (G^h_t)_{t\in I} = (G(t,h))_{t\in I}$ be a measurable and $(\mathcal{F}_t)_{t\in I}$-adapted stochastic process depending on a parameter $h \in J$. We assume*

$$\int_J \int_I \mathbb{E}\big[G^2(s,h)\big]\,\mathrm{d}s\,\mathrm{d}h < \infty. \tag{2.2}$$

Then, we have

$$\int_I \int_J G(s,h)\,\mathrm{d}h\,\mathrm{d}W_s = \int_J \int_I G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h \tag{2.3}$$

almost surely, where both of the above integrals are well defined.

Proof. Using the Jensen inequality yields

$$\begin{aligned}\mathbb{E}\left[\int_I \left(\int_J G(s,h)\,\mathrm{d}h\right)^2 \mathrm{d}s\right] &= \mathbb{E}\left[\mu(J)^2 \int_I \left(\frac{1}{\mu(J)}\int_J G(s,h)\,\mathrm{d}h\right)^2 \mathrm{d}s\right] \\ &\le \mu(J)\mathbb{E}\left[\int_I \int_J G^2(s,h)\,\mathrm{d}h\,\mathrm{d}s\right] < \infty.\end{aligned}$$

Together with assumption (2.2) this shows existence of the integrals in (2.3). Fubini for finite stochastic integrals gives the result for finite I. For infinite I it yields

$$\begin{aligned}\int_I \int_J G(s,h)\,\mathrm{d}h\,\mathrm{d}W_s &= \lim_{n\to\infty} \int_{I_n} \int_J G(s,h)\,\mathrm{d}h\,\mathrm{d}W_s = \lim_{n\to\infty} \int_J \int_{I_n} G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h \\ &= \int_J \int_I G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h \qquad \text{a.s.}\end{aligned}$$

The last equation holds because

$$\begin{aligned}&\mathbb{E}\left[\left(\int_J \int_I G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h - \int_J \int_{I_n} G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h\right)^2\right] \\ &\quad = \mathbb{E}\left[\left(\int_J \int_{I\setminus I_n} G(s,h)\,\mathrm{d}W_s\,\mathrm{d}h\right)^2\right] \le \mu(J)\int_J \mathbb{E}\left[\left(\int_{I\setminus I_n} G(s,h)\,\mathrm{d}W_s\right)^2\right]\mathrm{d}h \\ &\quad = \mu(J)\,\mathbb{E}\left[\int_J \int_{I\setminus I_n} G^2(s,h)\,\mathrm{d}s\,\mathrm{d}h\right] = \mu(J)\int_{I\setminus I_n}\int_J \mathbb{E}\big[G^2(s,h)\big]\,\mathrm{d}h\,\mathrm{d}s \longrightarrow 0\end{aligned}$$

for $n \to \infty$. Here the first inequality is due to the Jensen inequality and the convergence follows from (2.2). Note, this would only prove convergence in $L^2(\Omega)$ but choosing a suitable subsequence implies almost sure convergence. □

The following Theorem is our version of Theorem 2.2 in [18].

Theorem 2.2.2. *Let J be an open interval and $F^H = (F^H_t)_{t\in I} = (F(t,H))_{t\in I}$ be a measurable and $(\mathcal{F}_t)_{t\in I}$-adapted stochastic process depending on $H \in J$. Furthermore, let F be almost surely continuously differentiable in H for all $s \in I$. Assume the following conditions hold:*

(i) We have

$$\mathbb{E}\left[\int_I F^2(s,H)\,\mathrm{d}s\right] < \infty$$

for all $H \in J$.

(ii) We have

$$\mathbb{E}\left[\int_I \left(\frac{\partial}{\partial H} F(s,H)\right)^2 \mathrm{d}s\right] < \infty$$

for all $H \in J$.

(iii) We have

$$\mathbb{E}\left[\int_J \int_I \left(\frac{\partial}{\partial H} F(s,H)\right)^2 \mathrm{d}s\, \mathrm{d}H\right] < \infty.$$

(iv) The functions

$$\begin{aligned} H &\mapsto \int_I F(s,H)\, \mathrm{d}W_s, \\ H &\mapsto \int_I \frac{\partial}{\partial H} F(s,H)\, \mathrm{d}W_s \end{aligned} \tag{2.4}$$

are almost surely continuous.

Then, we have almost surely

$$\frac{\mathrm{d}}{\mathrm{d}H} \int_I F(s,H)\, \mathrm{d}W_s = \int_I \frac{\partial}{\partial H} F(s,H)\, \mathrm{d}W_s, \qquad H \in J.$$

Proof. Let $H, c \in J,\ c \neq H$. By Lemma 2.2.1 it holds almost surely for fixed c and H that

$$\int_I \int_c^H \frac{\partial}{\partial H} F(s,\beta)\, \mathrm{d}\beta\, \mathrm{d}W_s = \int_c^H \int_I \frac{\partial}{\partial H} F(s,\beta)\, \mathrm{d}W_s\, \mathrm{d}\beta.$$

So, the right- and left- hand side of the equation above are modifications of each other (as processes in (c,H)). It follows that there exists $A \in \mathcal{A}$ with $\mathbb{P}(A) = 1$ and

$$\left(\int_I \int_c^H \frac{\partial}{\partial H} F(s,\beta)\, \mathrm{d}\beta\, \mathrm{d}W_s\right)(\omega) = \int_c^H \left(\int_I \frac{\partial}{\partial H} F(s,\beta)\, \mathrm{d}W_s\right)(\omega)\, \mathrm{d}\beta \tag{2.5}$$

for all $\omega \in A$ and for all $c, H \in J \cap \mathbb{Q}$.

We can use the continuity and integrability assumptions to show that these processes are indistinguishable (compare e.g. [24], Problem 1.5, p. 2). Now, let $B \in \mathcal{A}$ with $\mathbb{P}(B) = 1$ be the set on which F is continuously differentiable and the functions in *(iv)* are continuous. Moreover, set $A' = A \cap B$. Then we have $\mathbb{P}(A') = 1$. Consider an arbitrary sequence $\{H_n\}_{n \in \mathbb{N}} \subseteq J \setminus \{H\}$ converging to H. Using (2.5) we have on A' that

$$\begin{aligned} &\frac{1}{H - H_n}\left(\int_I F(s,H)\, \mathrm{d}W_s - \int_I F(s,H_n)\, \mathrm{d}W_s\right) \\ &= \frac{1}{H - H_n} \int_I (F(s,H) - F(s,H_n))\, \mathrm{d}W_s = \frac{1}{H - H_n} \int_I \int_{H_n}^H \frac{\partial}{\partial H} F(s,v)\, \mathrm{d}v\, \mathrm{d}W_s \\ &= \frac{1}{H - H_n} \int_{H_n}^H \int_I \frac{\partial}{\partial H} F(s,v)\, \mathrm{d}W_s\, \mathrm{d}v \\ &\longrightarrow \int_I \frac{\partial}{\partial H} F(s,H)\, \mathrm{d}W_s, \end{aligned}$$

for $n \to \infty$, where the convergence follows from the second assumption in *(iv)*. □

The following is a slightly adapted version of the Kolmogorov-Čentsov theorem and its proof as found in [24] Theorem 2.2.8.

Theorem 2.2.3. *Let A be an interval in $\mathbb{R}$ and $(X_t^a)_{t\in[0,T],a\in A}$ be a parametrised stochastic process on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ that satisfies*

$$\mathbb{E}\Big[\sup_{a\in A}|X_t^a - X_s^a|^\alpha\Big] \leq C|t-s|^{1+\beta}, \qquad s,t\in[0,T].$$

for some $C, \alpha, \beta > 0$ and is almost surely continuous in a. Then, there exists a continuous modification $(\widetilde{X}_t^a)_{t\in[0,T],a\in A}$ of $(X_t^a)_{t\in[0,T],a\in A}$ such that

$$\sup_{a\in A}|\widetilde{X}_t^a(\omega) - \widetilde{X}_s^a(\omega)| \leq c(\omega)|t-s|^\gamma$$

for every $\gamma \in (0, \beta/\alpha)$ and the null set $\{\widetilde{X}_t^a \neq X_t^a\}$ can be chosen independently of a.

Proof. For simplicity of notation take $T = 1$. By Chebyshev's inequality we have

$$\mathbb{P}\Big(\sup_{a\in A}|X_t^a - X_s^a| \geq \varepsilon\Big) \leq \varepsilon^{-\alpha}\mathbb{E}\Big[\sup_{a\in A}|X_t^a - X_s^a|^\alpha\Big] \leq C\varepsilon^{-\alpha}|t-s|^{1+\beta}.$$

Therefore,

$$\mathbb{P}\Big(\sup_{a\in A}\big|X^a_{\frac{k}{2^n}} - X^a_{\frac{k-1}{2^n}}\big| \geq 2^{-n\gamma}\Big) \leq C2^{-n(1+\beta-\alpha\gamma)}$$

and thus

$$\mathbb{P}\Big(\max_{0\leq k\leq 2^n}\sup_{a\in A}\big|X^a_{\frac{k}{2^n}} - X^a_{\frac{k-1}{2^n}}\big| \geq 2^{-n\gamma}\Big) \leq C2^{-n(\beta-\alpha\gamma)}.$$

By the Borel-Cantelli lemma there exists an $\Omega_1 \in \mathcal{A}$ with $\mathbb{P}(\Omega_1) = 1$ and an $n_0(\omega) \in \mathbb{N}$ such that

$$\max_{0\leq k\leq 2^n}\sup_{a\in A}\big|X^a_{\frac{k}{2^n}}(\omega) - X^a_{\frac{k-1}{2^n}}(\omega)\big| < 2^{-n\gamma}, \qquad n \geq n_0(\omega) \tag{2.6}$$

for all $\omega \in \Omega_1$. Let $D_n := \{\frac{k}{2^n} : k = 0, \dots, 2^n\}$ and $D = \bigcup_{n=1}^\infty D_n$. Now fix $\omega \in \Omega_1$ and $n \geq n_0(\omega)$. By induction we show that

$$\sup_{a\in A}|X_t^a(\omega) - X_s^a(\omega)| \leq 2\sum_{j=n+1}^m 2^{-j\gamma} \qquad s,t\in D_m,\ 0<t-s<2^{-n} \tag{2.7}$$

for all $m > n$.

Setting $m = n+1$, we can only have $t = \frac{k}{2^m}$, $s = \frac{k-1}{2^m}$ and (2.7) follows from (2.6). Suppose (2.7) holds for all $m = n+1, \dots, M-1$ and let $s,t \in D_M$ with $0 < t-s < 2^{-n}$. Let $u = \min\{\tau \in D_{M-1} : \tau \geq s\}$ and $v = \max\{\tau \in D_{M-1} : \tau \leq t\}$. So $s \leq u \leq v \leq t$ and $u - s \leq 2^M$, $t - v \leq 2^M$. Therefore, we obtain by (2.6) that

$$\sup_{a\in A}\big|X_u^a(\omega) - X_s^a(\omega)\big| < 2^{-M\gamma},$$
$$\sup_{a\in A}\big|X_t^a(\omega) - X_v^a(\omega)\big| < 2^{-M\gamma}$$

and from (2.7) for $m = M-1$ that

$$\sup_{a\in A} |X_v^a(\omega) - X_u^a(\omega)| \le 2 \sum_{j=n+1}^{M-1} 2^{-j\gamma}.$$

This implies that (2.7) holds for all $m > n$.

For $s,t \in D$ with $0 < t-s < 2^{-n_0(\omega)}$ we chose $n \ge n_0(\omega)$ such that $2^{-(n+1)} \le t-s < 2^{-n}$. Inequality (2.7) yields

$$\sup_{a\in A} |X_t^a(\omega) - X_s^a(\omega)| \le 2 \sum_{j=n+1}^{\infty} 2^{-j\gamma} \le 2\cdot \frac{2^{-\gamma(n+1)}}{1-2^{-\gamma}} \le c|t-s|^{\gamma}, \tag{2.8}$$

where $c = \frac{2}{1-2^{-\gamma}}$. We define $\widetilde{X}_t^a(\omega) \equiv 0$ for $\omega \notin \Omega_1$. Let $\omega \in \Omega_1$ and $t \in [0,1] \cap D^c$. For any sequence $(t_n)_{n\in\mathbb{N}} \subseteq D$ with $t_n \to t$ we have for n,m big enough that

$$\sup_{a\in A} |X_{t_n}^a(\omega) - X_{t_m}^a(\omega)| \le c|t_n - t_m|^{\gamma}.$$

Together with the Cauchy criterion and the fact that D is dense in $[0,1]$, we can extend $X^a(\omega)$ uniquely to a continuous function $\widetilde{X}^a(\omega)$ on $[0,1]$ such that (2.8) holds for $\widetilde{X}$ and all $s,t \in [0,1]$. Thus, we have $X_t = \widetilde{X}_t$ a.s. for all $t \in D$. For $t \in [0,1] \cap D^c$ with $(t_n)_{n\in\mathbb{N}} \subseteq D$ and $t_n \to t$ we have $\widetilde{X}_{t_n} \to \widetilde{X}_t$ a.s. and $X_{t_n} \to X_t$ in probability. This implies $X_t = \widetilde{X}_t$ a.s. and therefore that $\widetilde{X}$ is a modification of X. □

2.3 Smoothness of fBm with respect to the Hurst parameter

The derivatives of K_H with respect to H are given by

$$\frac{\partial^k}{\partial H^k} K_H(s,t) = \Big(|t-s|^{H-1/2} \log^k(|t-s|) - |s|^{H-1/2} \log^k(|s|) \Big) \mathbf{1}_{(-\infty,0)}(s) + |t-s|^{H-1/2} \log^k(|t-s|)\, \mathbf{1}_{[0,t)}(s).$$

The next Lemma implies in particular that these functions belong to $L^2(\mathbb{R}\times[0,T];\mathbb{R})$.

Lemma 2.3.1. *Let $0 < a \le b < 1$ and $k \in \mathbb{N}$. We have*

$$\sup_{H\in[a,b]} \sup_{t\in[0,T]} \int_{\mathbb{R}} \left(\frac{\partial^k}{\partial H^k} K_H(s,t) \right)^2 \mathrm{d}s < \infty.$$

Proof. Let $H \in (a,b)$. We have

$$\int_{\mathbb{R}} \left(\frac{\partial^k}{\partial H^k} K_H(s,t) \right)^2 \mathrm{d}s = \int_0^t (t-s)^{2H-1} \log^{2k}(t-s)\,\mathrm{d}s + \int_{-1}^0 g_{H,k}^2(-s,t)\,\mathrm{d}s + \int_{-\infty}^{-1} g_{H,k}^2(-s,t)\,\mathrm{d}s,$$

where

$$g_{H,k}(s,t) = (t+s)^{H-1/2} \log^k(t+s) - s^{H-1/2} \log^k(s).$$

By substitution we obtain

$$\int_0^t (t-s)^{2H-1}\log^{2k}(t-s)\,\mathrm{d}s \le \int_0^T x^{2H-1}\log^{2k}(x)\,\mathrm{d}x,$$

and so

$$\int_0^t (t-s)^{2H-1}\log^{2k}(t-s)\,\mathrm{d}s \le \int_0^T (x^{2a-1}+x^{2b-1})\log^{2k}(x)\,\mathrm{d}x < \infty. \tag{2.9}$$

Furthermore, we have

$$\begin{aligned}\int_0^1 g_{H,k}^2(s,t)\,\mathrm{d}s &\le 2\int_0^1 (t+s)^{2H-1}\log^{2k}(t+s)\,\mathrm{d}s + 2\int_0^1 s^{2H-1}\log^{2k}(s)\,\mathrm{d}s \\ &= 2\int_t^{1+t} x^{2H-1}\log^{2k}(x)\,\mathrm{d}x + 2\int_0^1 x^{2H-1}\log^{2k}(x)\,\mathrm{d}x \\ &\le 4\int_0^{1+T} x^{2H-1}\log^{2k}(x)\,\mathrm{d}x,\end{aligned}$$

and so

$$\int_0^1 g_{H,k}^2(s,t)\,\mathrm{d} \le 4\int_0^{1+T} (x^{2a-1}+x^{2b-1})\log^{2k}(x)\,\mathrm{d}x. \tag{2.10}$$

Defining $f:(0,\infty)\to\mathbb{R}, x\mapsto x^{H-1/2}\log^k(x)$, we obtain for $-s,t>0$ by Taylor's theorem

$$\begin{aligned}f(-s+t) &= f(-s)+tf'(-s+\xi) \\ &= (-s)^{H-1/2}\log^k(-s) \\ &\quad + \frac{t}{2}(-s+\xi)^{H-3/2}\log^{k-1}(-s+\xi)\left[(2H-1)\log(-s+\xi)+2k\right],\end{aligned} \tag{2.11}$$

for some $\xi\in(0,t)$. This gives

$$\begin{aligned}&\int_1^\infty g_{H,k}^2(s,t)\,\mathrm{d}s \\ &= \int_1^\infty \left(\frac{t}{2}(s+\xi)^{H-3/2}\log^{k-1}(s+\xi)\left[(2H-1)\log(s+\xi)+2k\right]\right)^2\mathrm{d}s \\ &\le \frac{t^2}{4}\int_1^\infty x^{2H-3}\log^{2k-2}(x)\left[(2H-1)\log(x)+2k\right]^2\,\mathrm{d}x \\ &\le \frac{\max\{1,T\}^2}{4}\int_1^\infty x^{2H-3}\log^{2k-2}(x)\left[(2H-1)\log(x)+2k\right]^2\,\mathrm{d}x\end{aligned}$$

and

$$\int_1^\infty g_{H,k}^2(s,t)\,\mathrm{d}s \le \frac{\max\{1,T\}^2}{2}\int_1^\infty x^{2b-3}\log^{2k-2}(x)\left[\log^2(x)+2k^2\right]\,\mathrm{d}x < \infty. \tag{2.12}$$

Putting together (2.9), (2.10) and (2.12), the assertion follows. □

Recall that the Mandelbrot-van Ness fractional Brownian motion B^H is given by

$$B_t^H = C_H\int_{\mathbb{R}} K_H(s,t)\,\mathrm{d}B_s, \quad t\in[0,T],$$

where

$$C_H = \frac{\left(2H\sin(\pi H)\Gamma(2H)\right)^{1/2}}{\Gamma(H+1/2)}.$$

Lemma 2.3.2. *Let $0 < a \le b < 1$ and $k \in \mathbb{N}$. Define a stochastic process $(A_t^{H,k})_{t\in[0,T]}$ by*

$$A_t^{H,k} = \int_{\mathbb{R}} \frac{\partial^k}{\partial H^k} K_H(s,t)\,\mathrm{d}B_s.$$

Then we have:

(i) There exists a modification $\widehat{A}^{H,k} = (\widehat{A}_t^{H,k})_{t\in[0,T]}$ of $A^{H,k} = (A_t^{H,k})_{t\in[0,T]}$ that is jointly continuous in $t \in [0,T]$ and $H \in [a,b]$, and there exists, for every $t \in [0,T]$, a set $\Omega_{a,b,k,t} \in \mathcal{A}$ such that $\mathbb{P}(\Omega_{a,b,k,t}) = 1$ and

$$A_t^{H,k}(\omega) = \widehat{A}_t^{H,k}(\omega), \qquad H \in [a,b], \quad \omega \in \Omega_{a,b,k,t}.$$

(ii) For all $\omega \in \Omega$ the paths $[0,T] \ni t \mapsto \widehat{A}_t^{H,k}(\omega) \in \mathbb{R}$ of any continuous modification of $A^{H,k}$ are α-Hölder continuous for any $\alpha \in (0,H)$. We even have, for all $0 < a \le b < 1$ and $0 < \gamma < a$, that there exists a constant C depending on $\omega \in \Omega$ such that

$$\sup_{H\in[a,b]} \left|\widehat{A}_t^{H,k}(\omega) - \widehat{A}_s^{H,k}(\omega)\right| \le C(\omega)|t-s|^{\gamma} \qquad s,t \in [0,T].$$

Proof. Since $k \in \mathbb{N}$ is fixed we omit k in our notation and write A^H for $A^{H,k}$.
First, let $f \in L^2(I \times \mathbb{R}; \mathbb{R})$ with $\sup_{x\in I}|f(x,\cdot)| \in L^2(\mathbb{R})$ such that for fixed $x \in I$ the mapping $\mathbb{R} \ni y \mapsto f(x,y) \in \mathbb{R}$ is continuous except at a finite number of points. Define $t_i^n = i2^{-n}$ and

$$F^n(x) = \sum_{i=-n2^n}^{n2^n} f(x,t_i^n)(B_{t_{i+1}^n} - B_{t_i^n}).$$

We have

$$\begin{aligned}
\mathbb{E}\left[\sup_{x\in I}|F^n(x)|^2\right] &= \mathbb{E}\left[\sup_{x\in I}\left(\sum_{i=-n2^n}^{n2^n} f(x,t_i^n)(B_{t_{i+1}^n} - B_{t_i^n})\right)^2\right] \\
&= \sum_{i=-n2^n}^{n2^n} \sup_{x\in I} f^2(x,t_i^n)\mathbb{E}\left|B_{t_{i+1}^n} - B_{t_i^n}\right|^2 \\
&= \sum_{i=-n2^n}^{n2^n} \sup_{x\in I} f^2(x,t_i^n)(t_{i+1}^n - t_i^n),
\end{aligned}$$

and since $F^n(x), x \in I$, is a Gaussian process, it follows that

$$\left(\mathbb{E}\left[\sup_{x\in I}|F^n(x)|^{2p}\right]\right)^{1/p} \le C_p \sum_{i=-n2^n}^{n2^n} \sup_{x\in I} f^2(x,t_i^n)(t_{i+1}^n - t_i^n)$$

for some constant $C_p > 0$. Thus, $\sup_{x\in I} |F^n(x)|^{2p}$ is uniformly integrable and taking limits yields

$$\mathbb{E}\Big[\sup_{x\in I}\Big|\int_{\mathbb{R}} f(x,t)\,\mathrm{d}B_t\Big|^{2p}\Big] \le C_p\Big(\int_{\mathbb{R}}\sup_{x\in I}|f(x,t)|^2\,\mathrm{d}t\Big)^p.$$

Therefore, we have

$$\begin{aligned}\mathbb{E}\Big[\sup_{H\in[a,b]}\big|A^H_{t_2}-A^H_{t_1}\big|^{2p}\Big]^{1/p}\\ \le \int_{\mathbb{R}}\sup_{H\in[a,b]}\Big(\frac{\partial^k}{\partial H^k}K_H(s,t_1)-\frac{\partial^k}{\partial H^k}K_H(s,t_2)\Big)^2\,\mathrm{d}s.\end{aligned} \tag{2.13}$$

Without loss of generality let $t_1, t_2 \in [0,T]$ with $t_1 < t_2$ and $t_2 - t_1 < 1$. Using the Taylor expansion in (2.11) yields

$$\begin{aligned}&\int_{-\infty}^{t_1-1}\sup_{H\in[a,b]}\Big(\frac{\partial^k}{\partial H^k}K_H(s,t_1)-\frac{\partial^k}{\partial H^k}K_H(s,t_2)\Big)^2\,\mathrm{d}s\\ &=\int_{-\infty}^{t_1-1}\sup_{H\in[a,b]}\big((t_1-s)^{H-1/2}\log^k(t_1-s)-(t_2-s)^{H-1/2}\log^k(t_2-s)\big)^2\,\mathrm{d}s\\ &=(t_2-t_1)^2\int_{-\infty}^{t_1-1}\sup_{H\in[a,b]}\Big|(-s+t_1+\xi)^{2H-3}\log^{2k-2}(-s+t_1+\xi)\\ &\qquad\cdot\Big[\Big(H-\frac{1}{2}\Big)\log(-s+t_1+\xi)+k\Big]^2\Big|\,\mathrm{d}s\\ &\le (t_2-t_1)^2\int_1^{\infty}\sup_{H\in[a,b]}x^{2H-3}\log^{2k-2}(x)\Big[\frac{1}{2}\log^2(x)+2k^2\Big]\,\mathrm{d}x\\ &\le (t_2-t_1)^2\int_1^{\infty}x^{2b-3}\log^{2k-2}(x)\Big[\frac{1}{2}\log^2(x)+2k^2\Big]\,\mathrm{d}x\end{aligned}$$

where $\xi \in (0, t_2 - t_1)$, and therefore

$$\int_{-\infty}^{t_1-1}\sup_{H\in[a,b]}\Big(\frac{\partial^k}{\partial H^k}K_H(s,t_1)-\frac{\partial^k}{\partial H^k}K_H(s,t_2)\Big)^2\,\mathrm{d}s \le C_1\cdot(t_2-t_1)^2 \tag{2.14}$$

for a constant $C_1 = C_1(a,b,k) > 0$, which depends only on a, b and k.

Recall that we have assumed $t_2 - t_1 < 1$. Using the substitutions $(t_2 - t_1)w = v = s - t_1$, we obtain

$$\begin{aligned}&\int_{t_1}^{t_2}\sup_{H\in[a,b]}\Big(\frac{\partial^k}{\partial H^k}K_H(s,t_1)-\frac{\partial^k}{\partial H^k}K_H(s,t_2)\Big)^2\,\mathrm{d}s\\ &=\int_{t_1}^{t_2}\sup_{H\in[a,b]}(t_2-s)^{2H-1}\log^{2k}(t_2-s)\,\mathrm{d}s=\int_{t_1}^{t_2}(t_2-s)^{2a-1}\log^{2k}(t_2-s)\,\mathrm{d}s\\ &=\int_0^{t_2-t_1}(t_2-t_1-v)^{2a-1}\log^{2k}(t_2-t_1-v)\,\mathrm{d}v\\ &=(t_2-t_1)^{2a}\int_0^1(1-w)^{2a-1}\log^{2k}\big((t_2-t_1)(1-w)\big)\,\mathrm{d}w\end{aligned}$$

$$= (t_2 - t_1)^{2a} \int_0^1 w^{2a-1} \log^{2k}\big((t_2 - t_1)w\big)\,\mathrm{d}w$$
$$\leq 2^{2k-1}(t_2 - t_1)^{2a} \left(\log^{2k}(t_2 - t_1) \int_0^1 w^{2a-1}\,\mathrm{d}w + \int_0^1 w^{2a-1} \log^{2k}(w)\,\mathrm{d}w\right).$$

Thus, there exists a constant $C_2 = C_2(a, b, k) > 0$ such that

$$\int_{t_1}^{t_2} \sup_{H\in[a,b]} \left(\frac{\partial^k}{\partial H^k} K_H(s, t_1) - \frac{\partial^k}{\partial H^k} K_H(s, t_2)\right)^2 \mathrm{d}s \leq C_2 \cdot (t_2 - t_1)^{2a}(1 + \log^{2k}(t_2 - t_1)). \tag{2.15}$$

The substitutions $(t_2 - t_1)w = v = t_1 - s$ provide

$$\int_{t_1-1}^{t_1} \sup_{H\in[a,b]} \left(\frac{\partial^k}{\partial H^k} K_H(s, t_1) - \frac{\partial^k}{\partial H^k} K_H(s, t_2)\right)^2 \mathrm{d}s$$
$$= \int_{t_1-1}^{t_1} \sup_{H\in[a,b]} \big((t_1 - s)^{H-1/2} \log^k(t_1 - s) - (t_2 - s)^{H-1/2} \log^k(t_2 - s)\big)^2\,\mathrm{d}s$$
$$= \int_0^1 \sup_{H\in[a,b]} \big(v^{H-1/2} \log^k(v) - (t_2 - t_1 + v)^{H-1/2} \log^k(t_2 - t_1 + v)\big)^2\,\mathrm{d}v$$
$$= (t_2 - t_1)^{2a} \int_0^{1/(t_2-t_1)} \sup_{H\in[a,b]} \big(w^{H-1/2} \log^k((t_2 - t_1)w) - (1 + w)^{H-1/2} \log^k((t_2 - t_1)(1 + w))\big)^2\,\mathrm{d}w$$
$$\leq (t_2 - t_1)^{2a} \int_0^1 \sup_{H\in[a,b]} \big(w^{H-1/2} \log^k((t_2 - t_1)w) - (1 + w)^{H-1/2} \log^k((t_2 - t_1)(1 + w))\big)^2\,\mathrm{d}w$$
$$+ (t_2 - t_1)^{2a} \int_1^{\infty} \sup_{H\in[a,b]} \big(w^{H-1/2} \log^k((t_2 - t_1)w) - (1 + w)^{H-1/2} \log^k((t_2 - t_1)(1 + w))\big)^2\,\mathrm{d}w$$
$$=: I_1 + I_2.$$

For the first term we obtain

$$I_1 \leq 2(t_2 - t_1)^{2a}\bigg(\int_0^1 w^{2a-1} \log^{2k}((t_2 - t_1)w)\,\mathrm{d}w + \int_0^1 (1 + w)^{2b-1} \log^{2k}((t_2 - t_1)(1 + w))\,\mathrm{d}w\bigg)$$
$$\leq 2^{2k}(t_2 - t_1)^{2a}\bigg(\int_0^1 w^{2a-1} \log^{2k}(w)\,\mathrm{d}w + \int_0^1 (1 + w)^{2b-1} \log^{2k}(1 + w)\,\mathrm{d}w + \int_0^1 w^{2a-1} \log^{2k}(t_2 - t_1)\,\mathrm{d}w + \int_0^1 (1 + w)^{2b-1} \log^{2k}(t_2 - t_1)\,\mathrm{d}w\bigg)$$

and so again the existence of a constant $C_3 = C_3(a, b, k) > 0$ such that

$$I_1 \leq C_3 \cdot (t_2 - t_1)^{2a}(1 + \log^{2k}(t_2 - t_1)). \tag{2.16}$$

Similar to (2.11), we have for $f(x) = x^{H-1/2}\log^k((t_2-t_1)x)$ by Taylor's theorem

$$f(w) - f(1+w) = (w+\xi)^{H-3/2}\log^{k-1}\big((w+\xi)(t_2-t_1)\big) \cdot \left[\left(H-\frac{1}{2}\right)\log\big((w+\xi)(t_2-t_1)\big)+k\right],$$

where $\xi \in (0,1)$. Therefore, we obtain

$$\begin{aligned} I_2 &= (t_2-t_1)^{2a}\int_1^\infty \sup_{H\in[a,b]}\Bigg|(w+\xi)^{2H-3}\log^{2k-2}\big((w+\xi)(t_2-t_1)\big) \\ &\qquad \cdot \left[\left(H-\frac{1}{2}\right)\log\big((w+\xi)(t_2-t_1)\big)+k\right]^2\Bigg|\,\mathrm{d}w \\ &\le (t_2-t_1)^{2a}\int_1^\infty w^{2b-3}\log^{2k-2}\big(w(t_2-t_1)\big) \\ &\qquad \cdot \left[\frac{1}{2}\log^2\big(w(t_2-t_1)\big)+2k^2\right]\,\mathrm{d}w \\ &\le 2^{2k}k^2(t_2-t_1)^{2a}\big(1+\log^{2k}(t_2-t_1)\big)\left(\int_1^\infty w^{2b-3}\big(1+\log^{2k}(w)\big)\,\mathrm{d}w\right) \end{aligned}$$

and thus

$$I_2 \le C_4\cdot(t_2-t_1)^{2a}(1+\log^{2k}(t_2-t_1)). \tag{2.17}$$

for a constant $C_4 = C_4(a,b,k) > 0$.

Putting (2.13) and (2.14) – (2.17) together yields

$$\mathbb{E}\Big[\sup_{H\in[a,b]}\big|A^H_{t_2}-A^H_{t_1}\big|^{2p}\Big] \le K|t_2-t_1|^{2ap}(1+\log^{2k}(t_2-t_1))^p, \tag{2.18}$$

for $t_1, t_2 \in [0,T]$ and some constant $K = K(a,b,k,p,T) > 0$. If we chose $p > (2a)^{-1}$, the assertion follows from Theorem 2.2.3. □

Now Lemma 2.3.1 and Theorem 2.2.2 imply that for every $t \in [0,T]$ and $k \in \mathbb{N}$ there exists a set $\Omega_{a,b,k,t} \in \mathcal{A}$ such that $\mathbb{P}(\Omega_{a,b,k,t}) = 1$ and

$$\frac{\partial}{\partial H}\widehat{A}^{H,k}_t(\omega) = \widehat{A}^{H,k+1}_t(\omega), \qquad H\in[a,b], \quad \omega\in\Omega_{a,b,k,t}.$$

Since $\widehat{A}^{H,0}$ satisfies

$$\widehat{A}^{H,0}_t(\omega) = A^{H,0}_t(\omega) = \left(\int_{\mathbb{R}} K_H(s,t)dB_s\right)(\omega), \qquad H\in[a,b], \quad \omega\in\Omega_{a,b,0,t},$$

the assertions of Theorem 2.1.1 now follow.

This allows us to define a slightly changed representation of a fractional Brownian motion which is differentiable in its Hurst parameter. The jointly continuous process $B^{H,1}$ in Theorem 2.1.1 is given by

$$B^{H,1}_t = \big(\partial_H C_H\big)\int_{\mathbb{R}} K_H(s,t)\,\mathrm{d}B_s + C_H\int_{\mathbb{R}}\partial_H K_H(s,t)\,\mathrm{d}B_s. \tag{2.19}$$

We might now define a second fractional Brownian motion W^H by

$$W_t^H = B_t + \int_{\frac{1}{2}}^{H} B_t^{h,1} \,\mathrm{d}h, \qquad H \in [a,b], \quad t \in [0,1],$$

where $(B_t)_{t\in\mathbb{R}}$ is the same (standard) Brownian motion as in (2.19).

Let $H \in (0,1)$ and chose $a, b \in (0,1)$ such that $a < H < b$ and $\frac{1}{2} \in [a,b]$. Theorem 2.1.1 tells us that for every $t \in [0,1]$, there exists a set $\Omega_{a,b,t} \in \mathcal{A}$ with $\mathbb{P}(\Omega_{a,b,t}) = 1$ and

$$\frac{\partial}{\partial H} B_t^H(\omega) = B_t^{H,1}(\omega), \qquad H \in [a,b], \quad \omega \in \Omega_{a,b,t}.$$

This implies

$$W_t^H(\omega) = B_t(\omega) + \int_{\frac{1}{2}}^{H} B_t^{h,1}(\omega) \,\mathrm{d}h = B_t(\omega) + \int_{\frac{1}{2}}^{H} \frac{\partial}{\partial H} B_t^h(\omega) \,\mathrm{d}h = B_t^H(\omega)$$

for all $\omega \in \Omega_{a,b,t}$ and $H \in [a,b]$. Since W^H and B^H are continuous processes, they are not only modifications of each other but indeed indistinguishable, compare e.g. [24], Problem 1.5, p. 2.

Definition 2.3.3. Let $B = (B_t)_{t\in\mathbb{R}}$ be a two-sided Brownian motion on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For any $H \in (0,1)$ set

$$Z_t^H = \big(\partial_H C_H\big) \int_{\mathbb{R}} K_H(s,t) \,\mathrm{d}B_s + C_H \int_{\mathbb{R}} \partial_H K_H(s,t) \,\mathrm{d}B_s, \qquad t \in [0,1]$$

and

$$B_t^H = B_t + \int_{\frac{1}{2}}^{H} Z_t^h \,\mathrm{d}h = B_t - \int_{H}^{\frac{1}{2}} Z_t^h \,\mathrm{d}h, \qquad t \in [0,1].$$

This defines a fractional Brownian motion, which we call *(Hurst-differentiable) Mandelbrot-van Ness type fractional Brownian motion.* A d-dimensional version is obtained by taking d-independent copies.

Note that for B^H defined in the way of Definition 2.3.3, it holds that

$$\partial_H B_t^H(\omega) = Z_t^H(\omega) = B_t^{H,1}(\omega), \qquad t \in [0,1], \quad \omega \in \Omega,$$

where $B^{H,1}$ is the process from Theorem 2.1.1. From now on we only consider fractional Brownian motions defined in the manner of Definition 2.3.3.

3 Stochastic differential equations

In this chapters we consider stochastic differential equations driven by fractional Brownian motion. As the fractional Brownian motion is in general not a semimartingale the usual Itô theory is not applicable here. Instead we consider several pathwise solution concepts. We then analyse how these solutions depend on the Hurst parameter of the driving fractional Brownian motion.

3.1 Differential equations and solution concepts

In this introduction we discuss how stochastic differential equations driven by fractional Brownian motion can be defined. But before diving into the problem of differential equations, we take a short look at integration theory.

Let $u, v : [0,T] \to \mathbb{R}^d$ be continuous functions. We are interested in a definition of an integral of v with respect to u. For partitions D with $0 = t_0 < t_1 < \cdots < t_n = T$ we might define

$$\int v_t \, \mathrm{d}u_t := \lim_{|D| \to 0} \sum_l v_{t_{l-1}} (u_{t_l} - u_{t_{l-1}}), \tag{3.1}$$

whenever the right-hand side exists, e.g. if u has finite variation. The product in (3.1) is understood component-wise. For example, if $u \in C^1([0,T]; \mathbb{R})$ this integral definition leads to the usual

$$\int v_t \, \mathrm{d}u_t = \int v_t u_t' \, \mathrm{d}t.$$

Recall that we say a continuous function $u : [0,T] \to \mathbb{R}^d$ has finite p-variation for a $p \geq 1$ if

$$\sup_D \sum_l \left| u_{t_l} - u_{t_{l-1}} \right|^p < \infty,$$

where the supremum is taken over all finite partitions D of $[0,T]$. The definition in (3.1) could still make sense for functions u that possess infinite variation but in this case v being continuous is not sufficient for the right-hand side to exist and stronger regularity assumptions on v are need. So, for any $p \geq 1$ let $C^{p\text{-var}}([0,T], \mathbb{R}^d)$ be the set of continuous function with bounded p-variation equipped with the p-variation norm. If $u \in C^{q\text{-var}}([0,T], \mathbb{R}^d)$ and $v \in C^{p\text{-var}}([0,T], \mathbb{R}^d)$ with $p^{-1} + q^{-1} > 1$ the right-hand side of (3.1) exists and the left-hand side is called the Young integral [48]. This integral is continuous in the integrand as well as the integrator.

Having defined a suitable integral we now take a look at differential equations driven by $u : [0,T] \to \mathbb{R}^d$. We consider an integral or differential equation of the following form

$$x_t = x + \int_0^t f(x_s)\,\mathrm{d}u_s, \quad t \in [0,T], \qquad x \in \mathbb{R}^n, \tag{3.2}$$

which is understood to be equivalent to writing

$$\mathrm{d}x_t = f(x_s)\,\mathrm{d}u_s, \quad t \in [0,T], \quad x_0 = x, \qquad x \in \mathbb{R}^n.$$

Under some condition on f this equation is uniquely solvable for $u \in C^{1\text{-var}}([0,T],\mathbb{R}^d)$. Let $\Gamma : C^{1\text{-var}}([0,T],\mathbb{R}^d) \to C^{1\text{-var}}([0,T],\mathbb{R}^n)$, $u \mapsto x$ be the function that maps the driving signal to the solution of the corresponding differential equation, then Γ is continuous. We are interested in extending the domain (and codomain) such that Γ remains a continuous map between appropriate Banach spaces. This allows us to reasonably extend the concept of a solution to (3.2) to the domain of the solution map Γ. Remembering that the Young integral generalises the integral of functions with finite variation, it is reasonable trying to extend the solution map to functions with finite p-variation. So let $p < 2$ and let u have finite p-variation. Under some conditions on f we again obtain that (3.2) has a unique solution, where the integral is understood in the Young sense. Therefore, Γ can indeed be extended to a continuous map from $C^{p\text{-var}}([0,T],\mathbb{R}^d)$ to $C^{p\text{-var}}([0,T],\mathbb{R}^n)$, where $\Gamma(u)$ solves (3.2) (cf. Theorem 1.28 in [32]).

But what happens if the p-variation of u is infinite for all $p < 2$? Let $C_\mathbb{R}([0,T]) = C([0,T];\mathbb{R})$ be the space of continuous functions equipped with the uniform norm. In one dimension it can be shown that there exists a continuous map $\Gamma : C_\mathbb{R}([0,T]) \to C_\mathbb{R}([0,T])$ such that we have, for all $u \in C^1([0,T];\mathbb{R})$, that $\Gamma(u)$ is the solution to (3.2) in the usual sense. Since $C^1([0,T];\mathbb{R})$ is dense in $C_\mathbb{R}([0,T])$, we can extend the solution concept to all driving signals that are merely continuous. This approach follows Doss and Sussman ([6], [45]) and is discussed in Section 3.2. Unfortunately, this approach already breaks down in two dimensions. Consider the differential equation

$$\begin{aligned} \mathrm{d}x_t^1 &= \mathrm{d}u_t^1 \\ \mathrm{d}x_t^2 &= x_t^1\,\mathrm{d}u_t^2, \end{aligned} \tag{3.3}$$

with $x_0 = 0$. This differential equation can be understood in the Young sense for $u \in C^{1\text{-var}}([0,T],\mathbb{R}^d)$ and is then solved by $x_t^1 = u_t^1 - u_0^1$ and $x_t^2 = \int_0^t (u_s^1 - u_0^1)\,\mathrm{d}u_s^2$. However, setting

$$u_t^n = \frac{1}{\sqrt{n\pi}}(\cos(2n\pi t), \sin(2n\pi t)) \in C^{1\text{-var}}([0,T],\mathbb{R}^2)$$

we see that u^n converges uniformly to zero but for all $n \in \mathbb{N}$, we have

$$x_1^{2,n} = 2\int_0^1 \cos^2(2n\pi t)\,\mathrm{d}t = \frac{1}{n\pi}\int_0^{2n\pi} \cos^2(u)\,\mathrm{d}u = \frac{1}{n\pi}\int_0^{2n\pi} \frac{1+\cos(2u)}{2}\,\mathrm{d}u = 1.$$

This implies that a potential solution map can not be continuous in the uniform sense and therefore we cannot extend the concept of a solution to (3.3) to integrators that are merely continuous. Moreover, it can even be shown that such a solution map can in general not be continuous with respect to the 2-variation norm (cf. Proposition 1.30 in [32]). Note that the paths of a Brownian motion have almost surely infinite 2-variation. However, the

Brownian motion has finite quadratic variation around which Itô calculus evolved. But this would leave the pathwise framework we considered so far.

To obtain a continuous solution map for differential equations driven by signals that only have finite p-variation for some $p \geq 2$, we need to enhance the driving path by its iterated integrals to obtain an object called a p-rough path. The enhancing of a path to a rough path is commonly called *lifting*. As the example above has shown, this lift is in general neither unique nor necessarily continuous. However, once we are in the rough path setting, it can be shown that there exists a continuous solution map between appropriate rough path spaces. We will see that we can recover continuity of the solution map if we fix a specific lift and restrict ourselves to the domain of said lift. Rough paths will be considered in more detail in Chapter 4.

Once we constructed a continuous solution map Γ we are interested in finding regularity results on this Γ to be able to transfer the sensitivity results on fractional Brownian motion found in Chapter 2 to the solution $\Gamma(B^H)$ of a pathwise stochastic differential equation driven by a fractional Brownian motion.

3.2 Doss-Sussmann approach

Let B^H be a one-dimensional Mandelbrot-van Ness type fractional Brownian motion as described in Definition 2.3.3. We consider a stochastic differential equation

$$\mathrm{d}X_t^H = b(X_t^H)\,\mathrm{d}t + \sigma(X_t^H)\,\mathrm{d}B_t^H, \quad t \in [0,T], \qquad X_0^H = x_0 \in \mathbb{R}, \tag{3.4}$$

where we assume that

(A1) $b \in C^1(\mathbb{R};\mathbb{R})$ with b' bounded,

(A2) $\sigma \in C^2(\mathbb{R};\mathbb{R})$ with σ' bounded,

and use the so-called Doss-Sussmann solution, see [6, 45]. This is a precursor of the rough paths theory initiated by Lyons in [30, 31], see Remark 3.2.1 for its relation to the rough paths theory.

Let $u \in C([0,T];\mathbb{R})$, $g_1, g_2 : \mathbb{R}^n \to \mathbb{R}^n$ and equip $C([0,T];\mathbb{R}^n)$ with the uniform norm, which we denote by $\|\cdot\|_T$. In [45] a strikingly simple solution concept is introduced for the (formal) ordinary differential equation

$$\mathrm{d}x(t) = g_1(x(t))\,\mathrm{d}t + g_2(x(t))\,\mathrm{d}u_t, \quad t \in [0,T], \qquad x(0) = x_0 \in \mathbb{R}^n. \tag{3.5}$$

Namely, a function $\gamma \in C([0,T];\mathbb{R}^n)$ is called a solution to this equation,

(i) if there exists a continuous map $\Gamma : C([0,T];\mathbb{R}) \to C([0,T];\mathbb{R}^n)$ such that, for every $v \in C^1([0,T];\mathbb{R})$, $\Gamma(v)$ is a classical solution of the ODE

$$x'(t) = g_1(x(t)) + g_2(x(t))v_t', \quad t \in [0,T], \qquad x(0) = x_0,$$

(ii) and $\gamma = \Gamma(u)$.

In particular, if g_1 and g_2 are globally Lipschitz, then (3.5) has a unique solution (in the above sense), see [45].

In the special case $n = 1$, the article [6] even provides a more explicit representation of Γ under slightly stronger assumptions on the coefficients. So consider

$$\mathrm{d}x(t) = b(x(t))\,\mathrm{d}t + \sigma(x(t))\,\mathrm{d}u_t, \quad t \in [0,T], \qquad x(0) = x_0 \in \mathbb{R}, \tag{3.6}$$

and let $b, \sigma : \mathbb{R} \to \mathbb{R}$ be Lipschitz functions with $b \in C^1(\mathbb{R};\mathbb{R})$ and $\sigma \in C^2(\mathbb{R};\mathbb{R})$. Let $T > 0$ and write $C_{\mathbb{R}}([0,T]) = C([0,T];\mathbb{R})$. Further, let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as the solution of

$$\frac{\partial h}{\partial \beta}(\alpha,\beta) = \sigma(h(\alpha,\beta)), \qquad h(\alpha,0) = \alpha, \tag{3.7}$$

and for a given $u \in C_{\mathbb{R}}([0,T])$, let $D \in C^1([0,T];\mathbb{R})$ be the solution of the ODE

$$D'(t) = f(D(t), u_t), \quad t \in [0,T], \qquad D(0) = x_0, \tag{3.8}$$

with $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x,y) = \exp\left(- \int_0^y \sigma'(h(x,s))\,\mathrm{d}s \right) b\big(h(x,y)\big).$$

Then, we have that the unique Doss-Sussmann solution to (3.6) can be written as

$$x(t) = h(D(t), u_t).$$

Moreover, due to Lemma 4 in [6] the Doss-Sussmann map Γ is, in this case, even locally Lipschitz. In Section 3.2.1 this approach is explained in more detail.

Remark 3.2.1. As we have seen before, the Doss-Sussmann theory typically fails if the driving function u is not scalar. This was one of the starting points of the rough paths theory initiated by Lyons in [30, 31]. Roughly speaking, rough paths theory extends and revolutionises the Doss-Sussmann concept by allowing the map Γ to depend on iterated integrals of u and by working in appropriate α-Hölder or p-variation spaces. In particular, if $u \in C^{\beta}([0,T];\mathbb{R})$ for some $\beta > 0$, then due to the local Lipschitzness of Γ, the Doss-Sussmann solution of (3.6) is also a solution in the sense of Definition 10.17 in [11]. The required iterated integrals with respect to $v^0 = \mathrm{id}$ and $v^1 = u$ can be defined as the limit of the iterated (Riemann-Stieltjes) integrals with respect to v^0 and the dyadic piecewise linear interpolation $v^{1,(m)}$ of v^1, i.e.

$$\lim_{m\to\infty} \int_s^t \cdots \int_s^{t_3} \int_s^{t_2} \mathrm{d}v_{t_1}^{i_1,(m)}\,\mathrm{d}v_{t_2}^{i_2,(m)} \cdots \mathrm{d}v_{t_n}^{i_n,(m)}, \qquad 0 \le s \le t \le T,$$

where $n, m \in \mathbb{N}$, $i_k \in \{0,1\}$, $k = 1,\dots,n$, and $v_t^{0,(m)} = t$, respectively, $\Delta_m = T2^{-m}$ and

$$v_t^{1,(m)} = u_{\ell\Delta_m} + \frac{t - \ell\Delta_m}{\Delta_m}\left(u_{(\ell+1)\Delta_m} - u_{\ell\Delta_m}\right), \quad t \in [\ell\Delta_m, (\ell+1)\Delta_m),$$

for $\ell = 0,\dots,2^m - 1$.

Consider now the stochastic integral equation corresponding to SDE (3.4), i.e.

$$X_t^H = x_0 + \int_0^t b(X_s^H)\,\mathrm{d}s + \int_0^t \sigma(X_s^H)\,\mathrm{d}B_s^H, \qquad t \in [0,T].$$

For $H > \frac{1}{2}$ this equation is typically understood as a pathwise Riemann-Stieltjes equation, see e.g. [40], while for $H < \frac{1}{2}$ one can apply the rough paths theory. In all cases the solutions of these equations coincide with the Doss-Sussmann solution, if both exist. This can be seen for $H > \frac{1}{2}$ by an application of the standard change of variable formula for Riemann-Stieltjes integrals, while for $H \leq \frac{1}{2}$ it is a consequence of the Remark above. The Doss-Sussmann solution is also compatible with other integration methods for which the change of variable formula holds, like the symmetric integral (cf. [44]) or the Newton-Côtes functionals ([35], [36]). Note that for $H = \frac{1}{2}$ one recovers the standard Stratonovich solution.

The main result of this chapter is the following theorem.

Theorem 3.2.2. *Under (A1) and (A2) there exists a process $Y^H = (Y_t^H)_{t\in[0,T]}$ with α-Hölder continuous paths for any $\alpha \in (0,H)$ such that*

$$\frac{\partial}{\partial H} X^H = Y^H \ \textit{a.s.}$$

in $C_{\mathbb{R}}([0,T])$ for any $H \in (0,1)$, where X^H is the unique solution of (3.4) in the Doss-Sussmann sense.

Theorem 3.2.2 can be extendend to multi-dimensional SDEs driven by fractional Brownian motion with $H > \frac{1}{2}$, which will be presented in the subsequent chapter. The Fréchet differentiability results given in [41] can be used as a substitute for the Doss-Sussmann representation. This is presented in Section 3.3. The situation is naturally more involved for $\frac{1}{3} < H \leq \frac{1}{2}$ and is treated in Chapter 4.

3.2.1 Doss' results

First note that, because σ is Lipschitz, the differential equation (3.7) has a unique global solution. Integrating both sides of (3.7) with repect to β and differentiating with respect to α quickly leads to

$$\frac{\partial h}{\partial \alpha}(\alpha,\beta) = \exp\left(\int_0^\beta \sigma'\big(h(\alpha,s)\big)\,\mathrm{d}s\right). \tag{3.9}$$

Further, we obtain, for example, for all $\beta_1 \in \mathbb{R}$ that

$$h(\alpha,\beta) = h(h(\alpha,\beta_1),\beta-\beta_1) =: h_{\beta_1}(\alpha,\beta), \qquad \alpha,\beta \in \mathbb{R},$$

since $\frac{\partial}{\partial\beta} h_{\beta_1}(\alpha,\beta) = \sigma(h_{\beta_1}(\alpha,\beta))$ and $h_{\beta_1}(\alpha,\beta_1) = h(\alpha,\beta_1)$. For further properties that can be inferred about the solution h of (3.7) see Lemma 2 in [6].

Lemma 3.2.3. *The differential equation* (3.8) *has a unique, global solution.*

Proof. Let $u \in C_{\mathbb{R}}([0,T])$ be fixed. For convenience we use

$$\widetilde{f} : [0,T] \times \mathbb{R} \to \mathbb{R},\ (t,y) \mapsto f(y,u_t)$$

instead of f for the rest of this proof since (3.8) is for fixed u equivalent to

$$D'(t) = \widetilde{f}(t, D(t)), \qquad D(0) = x_0. \tag{3.10}$$

It is at once clear that $t \to \widetilde{f}(t,y)$ is continuous. Next, we show that $\widetilde{f}$ is locally Lipschitz in y. First, since b is Lipschitz and h is locally Lipschitz, it follows that $b \circ h$ is locally Lipschitz. Moreover, by the Leibniz and chain rule, we have

$$\begin{aligned}
&\left| \frac{\partial}{\partial y} \exp\left(- \int_0^{u_t} \sigma'\big(h(y,s)\big)\,\mathrm{d}s \right) \right| \\
&\quad = \left| \int_0^{u_t} \sigma''\big(h(y,s)\big) \exp\left(- \int_s^{u_t} \sigma'\big(h(y,\tau)\big)\,\mathrm{d}\tau \right) \mathrm{d}s \right| \\
&\quad \le \exp\Big(\|u\|_T \sup_{y \in K; v \in U} \big|\sigma'\big(h(y,v)\big)\big| \Big) \|u\|_T \sup_{y \in K; v \in U} \Big|\sigma''\big(h(y,v)\big)\Big|,
\end{aligned}$$

where $K \subseteq \mathbb{R}$ is a compact set and $U := [-\|u\|_T, \|u\|_T] \subseteq \mathbb{R}$. Thus, the derivative of $y \mapsto \exp\big(\int_0^{u_t} \sigma'\big(h(y,s)\big)\,\mathrm{d}s \big)$ is bounded on compacts and the function thus locally Lipschitz. Together, we obtain that $y \mapsto \widetilde{f}(t,y)$ is locally Lipschitz for all $t \in [0,T]$. This proves that there exists a unique maximal solution to the differential equation (3.10). We show that this solution is uniformly bounded in $t \in [0,T]$ for any fixed $T > 0$. This implies that the solution D of (3.10) is indeed a global solution, i.e. it exists and is finite on any interval $[0,T]$. We derive from the Lipschitz continuity of σ that

$$\sup_{x \in \mathbb{R}} |\sigma'(x)| < \infty.$$

Put

$$\begin{aligned}
C_1 &:= \exp\Big(\|u\|_T \cdot \sup_{x \in \mathbb{R}} |\sigma'(x)| \Big), \\
C_2 &:= \sup_{x \in U} \big| b\big(h(0,x)\big) \big|,
\end{aligned}$$

where again $U := [-\|u\|_T, \|u\|_T]$, and L_b the global Lipschitz constant of b. We have

$$\sup_{x \in \mathbb{R}; t \in [0,T]} \left| \exp\left(- \int_0^{u_t} \sigma'(h(x,s))\,\mathrm{d}s \right) \right| \le \exp\Big(\|u\|_T \sup_{x \in \mathbb{R}; v \in U} \big|\sigma'\big(h(x,v)\big)\big| \Big) \le C_1.$$

Therefore, we obtain

$$\begin{aligned}
|f(t,D(t))| &\le C_1 \big| b\big(h(D(t),u_t)\big) \big| \le C_1 \Big(\big| b\big(h(0,u_t)\big) \big| + \big| b\big(h(D(t),u_t)\big) - b\big(h(0,u_t)\big) \big| \Big) \\
&\le C_1 C_2 + C_1 L_b \big| h(D(t),u_t) - h(0,u_t) \big| \le C_1 C_2 + C_1^2 L_b \big| D(t) \big|.
\end{aligned}$$

Thus, we have

$$|D(t)| \le |x_0| + C_1 C_2 T + \int_0^t C_1^2 L_b |D(s)|\,\mathrm{d}s.$$

Applying Gronwall's inequality yields

$$|D(t)| \le \big(|x_0| + C_1 C_2 T \big) \exp\big(C_1^2 L_b T \big),$$

which concludes the proof. □

Letting D be the unique solution of (3.8), it holds for $u \in C^1([0,T];\mathbb{R})$ that $x(t) = h(D(t), u_t)$ solves (3.6) in the usual ODE sense since

$$\begin{aligned} x(t) - x_0 &= \int_0^t \frac{\partial h}{\partial \alpha}(D(s), u_t) D'(s)\,\mathrm{d}s + \int_0^t \frac{\partial h}{\partial \beta}(D(s), u_s)\,\mathrm{d}u_s \\ &= \int_0^t b(x_s)\,\mathrm{d}s + \int_0^t \sigma(x_s)\,\mathrm{d}u_s, \end{aligned}$$

where we used (3.9). It can also be shown that we can recover the solution D to (3.8) as composition of h and the solution x to (3.6), namely $D(t) = h(x(t), -u_t)$.

The continuity of the solution map, which maps the driving signal $u \in C^1([0,T], \mathbb{R})$ to the solution x of (3.6), with respect to the uniform norm follows from Fréchet differentiability shown in next section. Thus, the solution concept for (3.6) can be extended to all driving signals that are continuous.

3.2.2 Fréchet differentiability of the Doss-Sussmann map

Recall the conditions (A1) and (A2) on b, σ, i.e. $b \in C^1(\mathbb{R};\mathbb{R})$, $\sigma \in C^2(\mathbb{R};\mathbb{R})$ with b', σ' bounded, and let h be given by (3.7). Define

$$\mathcal{D} : C_\mathbb{R}([0,T]) \to C_\mathbb{R}([0,T]), \quad \mathcal{D}(u)(t) = D(t), \quad u \in C_\mathbb{R}([0,T]),\ t \in [0,T],$$

where D is the solution to the ODE (3.8), i.e.

$$D'(t) = f(D(t), u_t), \qquad D(0) = x_0$$

with $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by

$$f(x,y) = \exp\left(- \int_0^y \sigma'(h(x,s))\,\mathrm{d}s \right) b\big(h(x,y)\big).$$

Clearly, f is continuously differentiable under (A1) and (A2).

Due to Lemma 4 in [6] the Doss-Sussmann map

$$\Gamma : C_\mathbb{R}([0,T]) \to C_\mathbb{R}([0,T]), \quad \Gamma(u)(t) = h(\mathcal{D}(u)(t), u_t), \quad u \in C_\mathbb{R}([0,T]), t \in [0,T],$$

is locally Lipschitz. In this section we establish its Fréchet differentiability.

Lemma 3.2.4. *The map $\mathcal{D} : C_\mathbb{R}([0,T]) \to C_\mathbb{R}([0,T])$ is Fréchet differentiable with Fréchet derivative $\mathcal{D}'(u)$ given by*

$$\big[\mathcal{D}'(u)\big]\,(e)(t) = \int_0^t \exp\left(\int_s^t \partial_x f\big(\mathcal{D}(u)(\tau), u_\tau\big)\,\mathrm{d}\tau \right) \partial_y f\big(\mathcal{D}(u)(s), u_s\big) e_s\,\mathrm{d}s$$

for $u, e \in C_\mathbb{R}([0,T]), t \in [0,T]$.

Note that $E(t) = [\mathcal{D}'(u)]\,(e)(t)$ satisfies the linear ordinary differential equation

$$E'(t) = \partial_y f\big(\mathcal{D}(u)(t), u_t\big) e_t + \partial_x f\big(\mathcal{D}(u)(t), u_t\big) E(t), \quad t \in [0,T], \quad E(0) = 0. \tag{3.11}$$

Moreover, since we have

$$\mathcal{D}(u)(t) = h(\Gamma(u)(t), -u_t), \quad u \in C_\mathbb{R}([0,T]), t \in [0,T],$$

see Lemma 2 in [6], the local Lipschitz property of Γ implies that also $\mathcal{D}$ is locally Lipschitz.

Proof. Let $u, e \in C_{\mathbb{R}}([0,T])$, $t \in [0,T]$ and set

$$\Delta^{u,e}(t) = \mathcal{D}(u+e)(t) - \mathcal{D}(u)(t).$$

We have

$$\begin{aligned}
\Delta^{u,e}(t) &= \int_0^t \big(f\big(\mathcal{D}(u+e)(s), u_s + e_s\big) - f\big(\mathcal{D}(u)(s), u_s\big)\big)\ \mathrm{d}s \\
&= \int_0^t \big(f\big(\mathcal{D}(u+e)(s), u_s + e_s\big) - f\big(\mathcal{D}(u+e)(s), u_s\big)\big)\ \mathrm{d}s \\
&\quad + \int_0^t \big(f\big(\mathcal{D}(u+e)(s), u_s\big) - f\big(\mathcal{D}(u)(s), u_s\big)\big)\ \mathrm{d}s \\
&= \int_0^t \left[\int_0^1 \partial_y f\big(\mathcal{D}(u+e)(s), u_s + \lambda e_s\big) d\lambda\right] e_s \,\mathrm{d}s \\
&\quad + \int_0^t \left[\int_0^1 \partial_x f\big(\lambda \mathcal{D}(u+e)(s) + (1-\lambda)\mathcal{D}(u)(s), u_s\big) d\lambda\right] \Delta^{u,e}(s)\,\mathrm{d}s \\
&= \int_0^t \partial_y f\big(\mathcal{D}(u)(s), u_s\big) e_s \,\mathrm{d}s \\
&\quad + \int_0^t \partial_x f\big(\mathcal{D}(u)(s), u_s\big)\Delta^{u,e}(s)\,\mathrm{d}s + R(t,u,e)
\end{aligned}$$

where

$$\begin{aligned}
&R(t,u,e) \\
&= \int_0^t \left[\int_0^1 \big(\partial_y f\big(\mathcal{D}(u+e)(s), u_s + \lambda e_s\big) - \partial_y f\big(\mathcal{D}(u)(s), u_s\big)\big)\ \mathrm{d}\lambda\right] e_s \,\mathrm{d}s \\
&+ \int_0^t \left[\int_0^1 \big(\partial_x f\big(\mathcal{D}(u)(s) + \lambda\Delta^{u,e}(s), u_s\big) - \partial_x f\big(\mathcal{D}(u)(s), u_s\big)\big)\ \mathrm{d}\lambda\right] \Delta^{u,e}(s)\,\mathrm{d}s.
\end{aligned}$$

Using (3.11) we have

$$\begin{aligned}
\Delta^{u,e}(t) - [\mathcal{D}'(u)(e)](t) = \int_0^t &\partial_x f\big(\mathcal{D}(u)(s), u_s\big)\left[\Delta^{u,e}(s) - [\mathcal{D}'(u)(e)](s)\right]\ \mathrm{d}s \\
&+ R(t,u,e)
\end{aligned}$$

and therefore the variation of constants method gives

$$\Delta^{u,e}(t) - [\mathcal{D}'(u)(e)](t) = \int_0^t \exp\left(\int_s^t \partial_x f\big(\mathcal{D}(u)(\tau), u_\tau\big)\,\mathrm{d}\tau\right) R(s,u,e)\,\mathrm{d}s.$$

Thus, we obtain

$$\frac{\|\Delta^{u,e} - [\mathcal{D}'(u)(e)]\|_T}{\|e\|_T} \le T\exp\left(\int_0^T \left|\partial_x f\big(\mathcal{D}(u)(\tau), u_\tau\big)\right|\ \mathrm{d}\tau\right) \cdot \frac{\|R(\cdot,u,e)\|_T}{\|e\|_T}.$$

Since $\mathcal{D}$ is locally Lipschitz, we have that for every $K > 0$ there exists a constant $C_K > 0$ such that

$$\sup_{\|u\|_T \le K}\ \sup_{0<\|e\|_T\le K} \frac{\|\Delta^{u,e}\|_T}{\|e\|_T} \le C_K.$$

Therefore, it follows that for all $u \in C_{\mathbb{R}}([0,T])$ with $\|u\|_T \leq K$ and all $0 \neq e \in C_{\mathbb{R}}([0,T])$ with $\|e\|_T \leq K$ that

$$\frac{\|R(\cdot,u,e)\|_T}{\|e\|_T} \leq T \int_0^1 \sup_{t\in[0,T]} \left|\partial_y f\big(\mathcal{D}(u+e)(t), u_t + \lambda e_t\big) - \partial_y f\big(\mathcal{D}(u)(t), u_t\big)\right| \,\mathrm{d}\lambda$$
$$+ C_K T \int_0^1 \sup_{t\in[0,T]} \left|\partial_x f\big(\mathcal{D}(u)(t) + \lambda \Delta^{u,e}(t), u_t\big) - \partial_x f\big(\mathcal{D}(u)(t), u_t\big)\right| \,\mathrm{d}\lambda.$$

The continuity of f_x, f_y and the local Lipschitzness of $\mathcal{D}$ finally yield

$$\lim_{\|e\|_T \to 0} \frac{\|R(\cdot,u,e)\|_T}{\|e\|_T} = 0$$

and so

$$\lim_{\|e\|_T \to 0} \frac{\|\Delta^{u,e} - [\mathcal{D}'(u)(e)]\|_T}{\|e\|_T} = 0.$$

□

Now the Fréchet differentiability of Γ follows from the representation

$$\Gamma(u)(t) = h(\mathcal{D}(u)(t), u_t), \quad u \in C_{\mathbb{R}}([0,T]),\ t \in [0,T].$$

3.2.3 Smoothness of SDEs with respect to the Hurst parameter

Now let $u : (0,1) \to C_{\mathbb{R}}[0,T]$ be a Fréchet differentiable map and write $u^\lambda = u(\lambda)$, $\lambda \in (0,1)$. The chain rule implies that

$$\frac{\partial}{\partial\lambda}\Gamma(u^\lambda) = \Gamma'(u^\lambda)\frac{\partial}{\partial\lambda}u^\lambda.$$

Lemma 3.2.5. *Let $0 < a \leq b < 1$ and let the fractional Brownian motion B^H be given by Definition 2.3.3. We have that for all $\omega \in \Omega$ the mapping $[a,b] \ni H \mapsto B^H(\omega) \in C_{\mathbb{R}}[0,T]$ is Fréchet differentiable.*

Proof. Fix $\omega \in \Omega$ and let $H, H+\delta \in [a,b]$. From the definition of B^H, we have $\partial_H B^H(\omega) = B^{H,1}(\omega)$, $\omega \in \Omega$, where $B^{H,1}$ is the process from Theorem 2.1.1. Then

$$\begin{aligned}
&\frac{\|B^{H+\delta}(\omega) - B^H(\omega) - \partial_H B^H(\omega)\delta\|_T}{|\delta|} \\
&= \frac{\|B^{H+\delta}(\omega) - B^H(\omega) - B^{H,1}(\omega)\delta\|_T}{|\delta|} \\
&= \sup_{t\in[0,T]} \left| \frac{1}{\delta} \int_H^{H+\delta} \Big(B_t^{h,1}(\omega) - B_t^{H,1}(\omega)\Big) \,\mathrm{d}h \right| \\
&\leq \sup \Big\{ \|B^{h,1}(\omega) - B^{H,1}(\omega)\|_T : \, h \in \, [H - |\delta|, H + |\delta|] \cap [a,b] \Big\}.
\end{aligned}$$

Since $B_t^{h,1}$ is jointly continuous in h and t due to Theorem 2.1.1 the assertion follows. □

Applying this to SDE (3.4) we obtain

$$\begin{aligned}\frac{\partial}{\partial H}X^H &= \frac{\partial}{\partial H}\Gamma(B^H) = \Gamma'(B^H)\partial_H B^H \\ &= \frac{\partial}{\partial \alpha}h\big(\mathcal{D}(B^H), B^H\big)\mathcal{D}'(B^H)\partial_H B^H + \frac{\partial}{\partial \beta}h\big(\mathcal{D}(B^H), B^H\big)\partial_H B^H.\end{aligned}$$

Using equation (3.11), $h \in C^2(\mathbb{R}^2, \mathbb{R})$ as well as the fact that $\mathcal{D}$ is locally Lipschitz and $B^H, B^{H,1}$ are almost surely bounded on $[0,T]$, Theorem 2.1.1 implies that $\frac{\partial}{\partial H}X^H$ is almost surely α-Hölder continuous for any $\alpha \in (0,H)$ and Theorem 3.2.2 follows.

In some cases we are able to obtain an explicit or semi-explicit representation for the derivative $Y^H = \frac{\partial}{\partial H}X^H$. For the linear equation

$$\mathrm{d}X_t^H = \alpha X_t^H\,\mathrm{d}t + \beta X_t^H\,\mathrm{d}B_t^H$$

with $\alpha, \beta \in \mathbb{R}$ we trivially have

$$Y_t^H = X_t^H \cdot \beta \partial_H B_t^H$$

with

$$X_t^H = x_0 \exp\left(\alpha t + \beta B_t^H\right)$$

and the notation $\partial_H B_t^H = \frac{\partial}{\partial H}B_t^H$. In the case of additive noise, e.g. $\sigma(x) = 1$ for all $x \in \mathbb{R}$, the Doss-Sussmann solution simplifies to

$$X_t^H = B_t^H + D(t)$$

and

$$D'(t) = b(B_t^H + D(t)), \qquad D(0) = x_0,$$

since $f(x,y) = b(x+y)$. Thus we have

$$E(t) = \int_0^t \exp\left(\int_s^t b'(X_\tau^H)\,\mathrm{d}\tau\right) b'(X_s^H)e_s\,\mathrm{d}s$$

and therefore

$$\begin{aligned}Y_t^H &= \int_0^t \exp\left(\int_s^t b'(X_\tau^H)\,\mathrm{d}\tau\right) b'(X_s^H)\partial_H B_s^H\,\mathrm{d}s + \partial_H B_t^H \\ &= \int_0^t \exp\left(\int_s^t b'(X_\tau^H)\,\mathrm{d}\tau\right) d\left(\partial_H B_s^H\right)\end{aligned}$$

for $t \in [0,T]$, where the last formula holds due to the integration by parts formula for Riemann-Stieltjes integrals and $\partial_H B_0^H = 0$ a.s.

For non-additive noise, i.e. $\sigma \neq 0$, one expects to obtain

$$Y_t^H = \int_0^t \exp\left(\int_s^t b'(X_u^H)\,\mathrm{d}u + \int_s^t \sigma'(X_u^H)\,\mathrm{d}B_u^H\right)\sigma(X_s^H)\,\mathrm{d}\left(\partial_H B_s^H\right), \qquad t \in [0,T].$$

However, here we are leaving the Doss-Sussmann framework, since e.g. for $\frac{1}{3} < H \leq \frac{1}{2}$ a meaningful interpretation of this object as a rough paths integral would require the construction of a Lévy area for the process $(t, B_t^H, \partial_H B_t^H)_{t\in[0,T]}$.

3.3 Multidimensional SDEs with H bigger 1/2

In this section we show that, for $H > \frac{1}{2}$, the solution of a stochastic differential equation driven by a multidimensional fractional Brownian motion of the type introduced in Definition 2.3.3 is Frechét differentiable in H.

So let B^H be such a m-dimensional Hurst-differentiable Mandelbrot-van Ness type fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$. We consider the integral equation

$$X_t^H = x_0 + \int_0^t b(X_s^H)\,\mathrm{d}s + \int_0^t \sigma(X_s^H)\,\mathrm{d}B_s^H, \qquad t \in [0,T], \tag{3.12}$$

where $b = (b^i)_{1\leq i\leq d} : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma = (\sigma^{ij}) : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $x_0 \in \mathbb{R}^d$ is the initial value of the process X^H. The integrals in (3.12) are understood in the pathwise Riemann-Stieltjes sense. We assume $\sigma^{ij}, b^i \in C_b^3(\mathbb{R}^d)$, which denotes the class of thrice continuously differentiable functions whose partial derivatives up to order 3 are bounded.

We obtain Fréchet differentiability of the solution map by using the solution concept and results in [41]. Let us first introduce the necessary notation.

For any $0 < \lambda < 1$, let $C^\lambda(0,T;\mathbb{R}^d)$ be the space of bounded, Hölder continuous functions $f : [0,T] \to \mathbb{R}^d$ and equip it with the norm

$$\|f\|_\lambda := \|f\|_T + [f]_\lambda,$$

where $\|\cdot\|_T$ denotes the uniform norm on $[0,T]$ and

$$[f]_\lambda := \sup_{0\leq s<t\leq T} \frac{|f(t)-f(s)|}{|t-s|^\lambda}.$$

Let $\alpha \in (0,1/2)$. We denote by $W_1^\alpha(0,T;\mathbb{R}^d)$ the space of measurable functions $f : [0,T] \to \mathbb{R}^d$ such that

$$\|f\|_{\alpha,1} = \sup_{t\in[0,T]} \left(|f(t)| + \int_0^t \frac{|f(t)-f(s)|}{|t-s|^{1+\alpha}}\,\mathrm{d}s \right) < \infty$$

and by $W_2^{1-\alpha}(0,T;\mathbb{R}^d)$ the set of measurable functions $g : [0,T] \to \mathbb{R}^d$ such that

$$\|g\|_{1-\alpha,2} := \sup_{0\leq s<t\leq T} \left(\frac{|g(t)-g(s)|}{|t-s|^{1-\alpha}} + \int_s^t \frac{|g(y)-g(s)|}{|y-s|^{2-\alpha}}\,\mathrm{d}y \right) < \infty.$$

These spaces are naturally closely related to Hölder spaces. Let $f : [0,T] \to \mathbb{R}^d$ be measurable. For $\varepsilon \in (0,\alpha)$, we have

$$\begin{aligned}
\|f\|_{\alpha,1} &\leq \sup_{t\in[0,T]} |f(t)| + \int_0^T |t-s|^{\varepsilon-1} \sup_{t\in[0,T]} \left(\frac{|f(t)-f(s)|}{|t-s|^{\alpha+\varepsilon}} \right) \mathrm{d}s \\
&\leq \|f\|_T + [f]_{\alpha+\varepsilon} \int_0^T x^{\varepsilon-1}\,\mathrm{d}x = \|f\|_T + [f]_{\alpha+\varepsilon} \frac{T^\varepsilon}{\varepsilon} \\
&\leq \max\left\{1, \frac{T^\varepsilon}{\varepsilon}\right\} \|f\|_{\alpha+\varepsilon}.
\end{aligned}$$

It clearly holds that $\|f\|_{1-\alpha} \le \|f\|_{1-\alpha,2}$ and

$$\begin{aligned}\|f\|_{1-\alpha,2} &\le \sup_{0\le s<t\le T} \frac{|f(t)-f(s)|}{|t-s|^{1-\alpha}} + \int_0^T |t-s|^{\varepsilon-1} \sup_{0\le s<t\le T}\left(\frac{|f(t)-f(s)|}{|t-s|^{1-\alpha+\varepsilon}}\right) \mathrm{d}s \\ &\le T^{\varepsilon}[f]_{1-\alpha+\varepsilon} + [f]_{1-\alpha+\varepsilon}\frac{T^{\varepsilon}}{\varepsilon} \\ &\le \frac{T^{\varepsilon}}{\varepsilon}\|f\|_{1-\alpha+\varepsilon}.\end{aligned}$$

Therefore, we obtain

$$C^{\alpha+\varepsilon}(0,T;\mathbb{R}^d) \subseteq W_1^{\alpha}(0,T;\mathbb{R}^d)$$

and

$$C^{1-\alpha+\varepsilon}(0,T;\mathbb{R}^d) \subseteq W_2^{1-\alpha}(0,T;\mathbb{R}^d) \subseteq C^{1-\alpha}(0,T;\mathbb{R}^d).$$

Lemma 3.3.1. *Let $\nu,\lambda \in (\frac{1}{2},1)$ with $\nu<\lambda$ and $f \in C^{\lambda}(0,T;\mathbb{R}^d)$*

(i) There exists a constant $M>0$ that only depends on T,ν,λ such that

$$\|f\|_{\nu,2} \le M\Big([f]_\lambda^{\frac{1+\nu}{1+\lambda}} + [f]_\lambda^{\nu/\lambda}\Big)\Big(\|f\|_T^{1-\nu/\lambda} + \|f\|_T^{\frac{\lambda-\nu}{1+\lambda}}\Big).$$

(ii) Let $(f_n)_{n\in\mathbb{N}} \subseteq C^{\lambda}(0,T;\mathbb{R}^d)$ with $\sup_{n\in\mathbb{N}}[f_n]_\lambda < \infty$ and $\|f_n - f\|_T \to 0$, where $f \in C^{\lambda}(0,T;\mathbb{R}^d)$. Then $\|f_n - f\|_{\nu,2} \to 0$.

Proof. Assertion (ii) is a direct consequence of (i). So let us consider (i). We have

$$\begin{aligned}[f]_\nu &= \sup_{0\le s<t\le T}\frac{|f(t)-f(s)|}{|t-s|^{\nu}} = \sup_{0\le s<t\le T}\left(\frac{|f(t)-f(s)|}{|t-s|^{\lambda}}\right)^{\nu/\lambda}|f(t)-f(s)|^{1-\nu/\lambda} \\ &\le 2^{1-\nu/\lambda}[f]_\lambda^{\nu/\lambda}\|f\|_T^{1-\nu/\lambda},\end{aligned}$$

and using the same technique

$$\begin{aligned}\int_s^t \frac{|f(y)-f(s)|}{(y-s)^{1+\nu}}\,\mathrm{d}y &\le \big(2\|f\|_T\big)^{\frac{\lambda-\nu}{1+\lambda}} \int_s^t \left(\frac{|f(y)-f(s)|}{|y-s|^{1+\lambda}}\right)^{\frac{1+\nu}{1+\lambda}} \mathrm{d}y \\ &\le 2^{\frac{\lambda-\nu}{1+\lambda}}\|f\|_T^{\frac{\lambda-\nu}{1+\lambda}}[f]_\lambda^{\frac{1+\nu}{1+\lambda}} \int_s^t (y-s)^{-\frac{1+\nu}{1+\lambda}}\,\mathrm{d}y \\ &\le \left(\frac{1+\lambda}{\lambda-\nu}\right)(2T)^{\frac{\lambda-\nu}{1+\lambda}}\|f\|_T^{\frac{\lambda-\nu}{1+\lambda}}[f]_\lambda^{\frac{1+\nu}{1+\lambda}}.\end{aligned}$$

Therefore, there exists a constant $M>0$ depending only on T,ν,λ such that

$$\begin{aligned}\|f\|_\nu &\le [f]_\nu + \sup_{0\le s<t\le T}\int_s^t \frac{|f(y)-f(s)|}{(y-s)^{1+\nu}}\,\mathrm{d}y \\ &\le 2^{1-\nu/\lambda}[f]_\lambda^{\nu/\lambda}\|f\|_T^{1-\nu/\lambda} + \left(\frac{1+\lambda}{\lambda-\nu}\right)(2T)^{\frac{\lambda-\nu}{1+\lambda}}\|f\|_T^{\frac{\lambda-\nu}{1+\lambda}}[f]_\lambda^{\frac{1+\nu}{1+\lambda}} \\ &\le M\Big([f]_\lambda^{\frac{1+\nu}{1+\lambda}} + [f]_\lambda^{\nu/\lambda}\Big)\Big(\|f\|_T^{1-\nu/\lambda} + \|f\|_T^{\frac{\lambda-\nu}{1+\lambda}}\Big).\end{aligned}$$

□

Fix $\omega \in \Omega$, let $\alpha < \frac{1}{2}$ and let $a, b \in (0,1)$ such that $1-\alpha < a < H < b < 1$. We now use the result above to show that

$$B^{(\cdot)}(\omega) : (a,b) \to W_2^{1-\alpha}(0,T;\mathbb{R}^m); \; H \mapsto B^H(\omega)$$

is a Frechét differentiable map. Recall that for fractional Brownian motion as given by Definition 2.3.3, it holds that $\partial_H B_t^H(\omega) = B_t^{H,1}(\omega), \; t \in [0,1]$.

Let $(\delta_n)_{n\in\mathbb{N}}$ be an arbitrary sequence such that $\delta_n \to 0$ and $H + \delta_n \in [a,b]$ for all $n \in \mathbb{N}$. We define $f_n : [0,T] \to \mathbb{R}^d$ by

$$f_n(t) = \frac{B_t^{H+\delta_n}(\omega) - B_t^H(\omega) - B_t^{H,1}(\omega)\delta_n}{\delta_n} = \frac{1}{\delta_n}\int_H^{H+\delta_n} B_t^{h,1}(\omega) - B_t^{H,1}(\omega)\,\mathrm{d}h.$$

Let $\beta \in (1-\alpha, a)$. Using Theorem 2.1.1, there exists, for fixed $\omega \in \Omega$, a positive constant C such that

$$\begin{aligned} |f_n(t) - f_n(s)| &= \left| \frac{1}{\delta_n}\int_H^{H+\delta_n} B_t^{h,1}(\omega) - B_s^{h,1}(\omega) - B_t^{H,1}(\omega) + B_s^{H,1}(\omega)\,\mathrm{d}h \right| \\ &\le 2 \sup_{h\in[a,b]} |B_t^{h,1}(\omega) - B_s^{h,1}(\omega)| \\ &\le C|t-s|^\beta. \end{aligned}$$

Thus, we obtain

$$\sup_{n\in\mathbb{N}} [f_n]_\beta \le \sup_{0\le s<t\le T} \frac{C|t-s|^\beta}{|t-s|^\beta} = C < \infty.$$

In the proof of Lemma 3.2.5, we have further shown that $\|f_n\|_T \to 0$ for $n \to \infty$. By Lemma 3.3.1, this yields

$$\lim_{|\delta|\to 0} \frac{\|B_t^{H+\delta}(\omega) - B_t^H(\omega) - B_t^{H,1}(\omega)\delta\|_{1-\alpha,2}}{|\delta|} = 0.$$

Therefore, $B^{(\cdot)}(\omega) : (a,b) \to W_2^{1-\alpha}(0,T;\mathbb{R}^m); \; H \mapsto B^H(\omega)$ is Fréchet differentiable. We can now combine this result with Proposition 4 in [41], which states the following.

Proposition 3.3.2. *Let $\alpha \in (0, \frac{1}{2})$ and $g \in W_2^{1-\alpha}(0,T;\mathbb{R}^m)$. Denote by $x \in W_1^\alpha(0,T;\mathbb{R}^d)$ the solution of*

$$x_t = x_0 + \int_0^t b(x_s)\,\mathrm{d}s + \int_0^t \sigma(x_s)\,\mathrm{d}g_s, \qquad t \in [0,T].$$

The mapping

$$\Gamma : W_2^{1-\alpha}(0,T;\mathbb{R}^m) \to W_1^\alpha(0,T;\mathbb{R}^d); \; g \mapsto x(g)$$

is Fréchet differentiable. For $h \in W_2^{1-\alpha}(0,T;\mathbb{R}^m)$ its derivative is given by

$$\big(\Gamma'(g)h\big)(t) = \int_0^t \Phi_t(s)\,\mathrm{d}h_s,$$

where $\Phi_t(s) \in \mathbb{R}^{d\times m}$ is defined as follows. Letting ∂_k denote the derivative with respect to the k-th variable, $s \mapsto \Phi_t(s)$ satisfies

$$\Phi_t^{ij}(s) = \sigma^{ij}(x_s) + \sum_{k=1}^d \int_s^t \partial_k b^i(x_u)\Phi_u^{k,j}(s)\,\mathrm{d}u + \sum_{k=1}^d\sum_{l=1}^m \int_s^t \partial_k \sigma^{il}(x_u)\Phi_u^{k,j}(s)\,\mathrm{d}g_u^l$$

for $0 \le s \le t \le T$ and $\Phi_t^{ij}(s) = 0$ for $s > t$, where $i = 1,\dots,d, \; j = 1,\dots,m$.

Applying this proposition to our situation, we obtain, by the chain rule, that

$$\frac{\partial}{\partial H} X^H(\omega) = \frac{\partial}{\partial H}\Gamma(B^H(\omega)) = \Gamma'(B^H(\omega))B^{H,1}(\omega) = \int_0^t \Phi_t(s)\,\mathrm{d}B_s^{H,1}(\omega),$$

where $X^H = \Gamma(B^H)$ is the pathwise solution to equation (3.12), and where $\Phi_t(s)$ depends on ω and H and is given by

$$\begin{aligned}\Phi_t^{ij}(s) = \sigma^{ij}\big(X_s^H(\omega)\big) &+ \sum_{k=1}^d \int_s^t \partial_k b^i\big(X_u^H(\omega)\big)\Phi_u^{k,j}(s)\,\mathrm{d}u \\ &+ \sum_{k=1}^d \sum_{l=1}^m \int_s^t \partial_k \sigma^{il}\big(X_u^H(\omega)\big)\Phi_u^{k,j}(s)\,\mathrm{d}(B_u^H(\omega))^l\end{aligned}$$

for $0 \leq s \leq t \leq T$ and $\Phi_t^{ij}(s) = 0$ for $s > t$, where $i = 1, \dots, d,\ j = 1, \dots, m$. In the equation above $(B_u^H(\omega))^l$ denotes the l-th element of the m-dimensional vector $B_u^H(\omega)$.

4 Rough paths

In this chapter we consider stochastic differential equations driven by a multidimensional fractional Brownian motion with Hurst parameter $H \leq \frac{1}{2}$. However, in this case the paths of fBm have almost surely infinite 2-variation and we saw in Section 3.1 that the usual integration and SDE theory fails. Thus, a more involved theory is needed and we hopefully motivated why rough path theory is a suitable tool to tackle SDEs of this type.

The theory of rough path was initially developed by Terry Lyons [31] and allows to consider differential equations, where the driving signal is rough. The idea is that enhancing rough processes with their iterated integrals restores continuity properties in the (rough) integration theory, e.g. one of the first application of the theory states that the solution to a Stratonovich stochastic differential equation is a continuous map of the tuple consisting of the driving Brownian motion and its Lévy area. Since then the topic of rough paths has been an active field of research, see e.g. [12, 14, 15, 28, 29, 42].

The aim of this chapter is to analyse rough stochastic differential equations of the type

$$\mathrm{d}Y_t^H = f(Y_t^H)\,\mathrm{d}B_t^H, \qquad Y_0^H = y_0, \tag{4.1}$$

where f is a suitable function and the driving multidimensional fractional Brownian motion has Hurst parameter $H \leq \frac{1}{2}$. Therefore, a rough path over the driving fBm needs to be constructed.

So, in Section 4.1, we first give a short introduction to some aspects of rough path theory and recall a few definitions and results. We then present the partly adapted methods in [4, 28], which allow fractional Brownian motion to be lifted to a rough path. The idea is to bound the p-variation of a rough path by its values at dyadic points. This enables us to control the first and second level paths and we obtain, by taking the limit of its dyadic approximations, a geometric rough path over fractional Brownian motion. We show that the same construction can be used to lift the derivative process of fBm with respect to the Hurst parameter to a geometric rough path. However, this method fails when trying to jointly lift fBm *and* its derivative process because the dyadic second level approximations diverge in expected p-variation distance.

Nevertheless, we can show that the solution Y^H to (4.1) is locally Lipschitz in H in p-variation distance and in a very restrictive case a derivative of Y^H in a rough path sense can be constructed.

For a general introduction to rough paths, there exists a growing literature of monographs on the topic, e.g. [10, 11, 32].

4.1 Introduction to rough paths

Since the rough path theory is not needed in its full generality, some of our definitions might be more restrictive than in the classical theory. For simplicity we also restrict ourselves to the time horizon of $[0, 1]$ but with some possible adjustment to constants, one could choose any finite time frame. This introduction is mainly based on the the lecture notes in [32].

4.1.1 Preliminaries

Let V be a Banach space and $V^{\otimes n}$ its n-th tensor power. We endow the tensor spaces $V^{\otimes n},\ n \in \mathbb{N}$ with an admissible norm that means that for all $n, m \in \mathbb{N}$

$$\begin{aligned} \|\sigma v\| &= \|v\|, \quad v \in V^{\otimes n},\ \sigma \in \mathfrak{S}_n, \\ \|v \otimes w\| &\leq \|v\|\|w\|, \quad v \in V^{\otimes n},\ w \in V^{\otimes m}, \end{aligned}$$

where $\mathfrak{S}_n$ denotes the symmetric group. Note that here and throughout this work we might use the same notation for norms on different spaces, yet it should remain clear from context which norm and space is meant.

Remark 4.1.1. In case V is a Hilbert space, we will always use the tensor product of Hilbert spaces which defines a norm on the tensor powers that satisfies

$$\|v \otimes w\| = \|v\|\|w\|, \quad v \in V^{\otimes n},\ w \in V^{\otimes m},$$

for all $n, m \in \mathbb{N}$.

Let us take a look at the common case of $V = \mathbb{R}^d$ for some $d \in \mathbb{N}$. With $\{e_i : i \in \{1, \dots, d\}\}$ denoting the standard orthonormal basis in $\mathbb{R}^d$, we can write any element $v \in V^{\otimes n}$ as

$$v = \sum_{i_1,\dots,i_n=1}^{d} \alpha_{i_1,\dots,i_n}(e_{i_1} \otimes \cdots \otimes e_{i_n}), \quad \alpha_{i_1,\dots,i_n} \in \mathbb{R},$$

and define its norm $\|\cdot\|$ as

$$\|v\| := \left(\sum_{i_1,\dots,i_n=1}^{n} |\alpha_{i_1,\dots,i_n}|^2 \right)^{1/2}. \tag{4.2}$$

Let $T^{(n)}(V)$ denote the truncated tensor algebra

$$T^{(n)}(V) = \bigoplus_{k=0}^{n} V^{\otimes k}, \quad n \in \mathbb{N},$$

where $V^0 := \mathbb{R}$. The space $T^{(n)}(V)$ is equipped with addition and multiplication, which are, for $\mathbf{v} = (\mathbf{v}^0, \dots, \mathbf{v}^n), \mathbf{w} = (\mathbf{w}^0, \dots, \mathbf{w}^n) \in T^{(n)}(V)$, given by

$$\begin{aligned} \mathbf{v} + \mathbf{w} &= (\mathbf{v}^0 + \mathbf{w}^0, \dots, \mathbf{v}^n + \mathbf{w}^n), \\ \mathbf{v} \otimes \mathbf{w} &= (\mathbf{z}^0, \dots, \mathbf{z}^n), \end{aligned}$$

where $\mathbf{z}^k = \sum_{i=0}^{k} \mathbf{v}^i \otimes \mathbf{w}^{k-i}$.

A natural norm on $T^{(n)}(V)$ is given by

$$\|\mathbf{v}\| = |\mathbf{v}^0| + \sum_{k=1}^{n} \|\mathbf{v}^k\|, \qquad \mathbf{v} = (\mathbf{v}^0, \dots, \mathbf{v}^n) \in T^{(n)}(V). \tag{4.3}$$

Since $V = \mathbb{R}^d$ is a Hilbert space there is a second natural choice of norm on $T^{(n)}(\mathbb{R}^d)$. Defining the inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^0 \mathbf{w}^0 + \sum_{k=1}^{n} \langle \mathbf{v}^k, \mathbf{w}^k \rangle, \qquad \mathbf{v} = (\mathbf{v}^0, \dots, \mathbf{v}^n),\ \mathbf{w} = (\mathbf{w}^0, \dots \mathbf{w}^n) \in T^{(n)}(\mathbb{R}^d)$$

on $T^{(n)}(\mathbb{R}^d)$ gives rise to the norm

$$\|\mathbf{v}\|_2 = \sqrt{|\mathbf{v}^0|^2 + \sum_{k=1}^{n} \|\mathbf{v}^k\|^2}, \qquad \mathbf{v} = (\mathbf{v}^0, \dots, \mathbf{v}^n) \in T^{(n)}(\mathbb{R}^d).$$

But this norm is equivalent to the norm presented in (4.3) and in what follow it makes no difference which norm we choose. It is, however, important to note that convergence in $T^{(n)}(V)$ is equivalent to the convergence of all its elements in $V^{\otimes k}, k = 0, \dots, n$. This reasoning is not restricted to the case of $V = \mathbb{R}^d$ but it is sufficient for our needs and hopefully more accessible.

From now on let $\Delta := \{(s,t) : 0 \le s \le t \le 1\}$. Next we define a multiplicative functional in $T^{(n)}(V)$ (cf. Def 3.1 in [32]).

Definition 4.1.2. Let $n \in \mathbb{N}$ and $\mathbf{w} : \Delta \to T^{(n)}(V)$ be a continuous map. We write $\mathbf{w}_{s,t}$ for the value of $\mathbf{w}$ evaluated at $(s,t) \in \Delta$ and

$$\mathbf{w}_{s,t} = (\mathbf{w}^0_{s,t}, \mathbf{w}^1_{s,t}, \dots, \mathbf{w}^n_{s,t}) \in \mathbb{R} \oplus V \oplus \cdots \oplus V^{\oplus n}.$$

We call $\mathbf{w}$ an *multiplicative functional (of degree* n or *in* $T^{(n)}(V)$*)* if $\mathbf{w}^0 \equiv 1$ and

$$\mathbf{w}_{s,u} \otimes \mathbf{w}_{u,t} = \mathbf{w}_{s,t}, \qquad 0 \le s \le u \le t \le 1. \tag{4.4}$$

Equation (4.4) is also called *Chen's relation.* As we later use the space $T^{(2)}(V)$, let us note that for a multiplicative functional $\mathbf{w} = (1, \mathbf{w}^1, \mathbf{w}^2)$ in $T^{(2)}(V)$ equation (4.4) translates to

$$\begin{aligned} \mathbf{w}^1_{s,t} &= \mathbf{w}^1_{s,u} + \mathbf{w}^1_{u,t}, \\ \mathbf{w}^2_{s,t} &= \mathbf{w}^2_{s,u} + \mathbf{w}^2_{u,t} + \mathbf{w}^1_{s,u} \otimes \mathbf{w}^1_{u,t}, \end{aligned}$$

for all $0 \le s \le u \le t \le 1$.

4.1.2 Geometric rough paths

For the rest of this work let $V = \mathbb{R}^d$ and the norms $|\cdot| : V^{\otimes n} \to \mathbb{R}$ on its tensor powers given by (4.2).

Definition 4.1.3. Let $\mathbf{w} = (1, \mathbf{w}^1, \dots, \mathbf{w}^n)$ be a multiplicative functional in $T^{(n)}(\mathbb{R}^d)$ and $p \geq 1$. We say $\mathbf{w}$ has finite p-variation if

$$\max_{1 \leq i \leq n} \sup_D \sum_l |\mathbf{w}^i_{t_{l-1},t_l}|^{p/i} < \infty,$$

where the supremum runs over all finite subdivisions $D = \{t_l\}$ of $[0, 1]$.

Definition 4.1.4. Let $p \geq 1$ and $\mathbf{w}$ be a multiplicative functional in $T^{(n)}(V)$, $n \geq \lfloor p \rfloor$ with finite p-variation. Then we call $\mathbf{w}$ a *p-rough path* in V. The space of p-rough paths in V is commonly denoted by $\Omega_p(V)$.

The Extension Theorem (see e.g. [32]) states that, for any $n \geq \lfloor p \rfloor$, a multiplicative functional $\mathbf{w}$ in $T^{(\lfloor p \rfloor)}(V)$ with finite p-variation can be uniquely extend to a multiplicative functional in $T^{(n)}(V)$ which retains finite p-variation. Therefore, if we speak of a p-rough path we can consider it to be a multiplicative functional in $T^{(n)}(V)$ for any $n \geq \lfloor p \rfloor$. With this in mind, we have $\Omega_q(V) \subseteq \Omega_p(V)$ for $1 \leq q \leq p$.

To get an idea of how this unique extension looks, let $x : [0, 1] \to \mathbb{R}^d$ be a continuous function with finite variation. For any $n \in \mathbb{N}$ we can construct a 1-rough path $\mathbf{x}$ in $T^{(n)}(\mathbb{R}^d)$ by setting

$$\mathbf{x}_{s,t} = (1, \mathbf{x}^1_{s,t}, \dots, \mathbf{x}^n_{s,t}), \qquad (s,t) \in \Delta,$$

where $\mathbf{x}^i_{s,t}$ denotes the i-th iterated integral of x over the interval $[s, t]$ with $(s, t) \in \Delta$, i.e.

$$\begin{aligned} \mathbf{x}^i_{s,t} &= \int_{s<u_1<\cdots<u_i<t} \mathrm{d}x_{u_1} \otimes \cdots \otimes \mathrm{d}x_{u_i} \\ &= \int_{s<u_{k+1}<\cdots<u_i<t} \mathbf{x}^k_{s,u_{k+1}} \otimes \mathrm{d}x_{u_{k+1}} \otimes \cdots \otimes \mathrm{d}x_{u_i}, \end{aligned} \tag{4.5}$$

for $1 \leq k < i$. Ignoring the term in the middle, equation (4.5) can also be used to extend multiplicative functionals with finite q-variation by setting $k = \lfloor q \rfloor$. Note that $\mathbf{x}^k$ has finite $\frac{q}{k}$-variation, and since $\frac{k}{q} + q^{-1} > 1$, the integrals on the right-hand side of (4.5) are well defined.

Definition 4.1.5. Let $p \geq 1$. For all multiplicative functionals $\mathbf{w}, \mathbf{v}$ on $T^{(\lfloor p \rfloor)}(V)$ with finite p-variation we define the *p-variation distance* as

$$d_p(\mathbf{w}, \mathbf{v}) = \left(\max_{1 \leq i \leq \lfloor p \rfloor} \sup_D \sum_l \left| \mathbf{w}^i_{t_{l-1},t_l} - \mathbf{v}^i_{t_{l-1},t_l} \right|^{p/i} \right)^{i/p},$$

where $D = \{t_l\}$ runs over all finite subdivisions of $[0, 1]$.

The function d_p is usually just a pseudo-metric but since rough paths do not have a starting value (or an identical starting value is chosen) d_p is indeed a metric as

$$|\mathbf{v}^i_{s,t} - \mathbf{w}^i_{s,t}| \leq |\mathbf{v}^i_{0,0} - \mathbf{w}^i_{0,0}| + \left(\sup_D \sum_l \left| \mathbf{w}^i_{t_{l-1},t_l} - \mathbf{v}^i_{t_{l-1},t_l} \right|^{p/i} \right)^{i/p}.$$

Definition 4.1.6. Let $p \geq 1$. If a p-rough path $\mathbf{w}$ is the limit in p-variation distance of a sequence of 1-rough paths, we call $\mathbf{w}$ a *geometric rough path.* The set of all geometric rough paths in V is denoted by $G\Omega_p(V)$.

4.2 Rough paths and their dyadic approximations

In this section, we show that the p-variation of a multiplicative functional in $T^{(2)}(\mathbb{R}^d)$ can be bounded by values of the multiplicative functional at its dyadic points. In the second part we infer some properties of a dyadic rough path approximation. For the original results, see [4] and [28]. Throughout this section we denote the dyadic points in $[0,1]$ by $t_k^n := k/2^n,\ k = 0, \dots, 2^n, n \in \mathbb{N}$.

4.2.1 Controlling p-variation by dyadic points

The following lemma and its proof is given as Lemma 2 in [28].

Lemma 4.2.1. *Let* $\mathbf{w}$ *be a multiplicative functional in* $T^{(2)}(\mathbb{R}^d)$*. Then, for* $i = 1, 2$*,* p *satisfying* $p/i > 1$ *and any* $\gamma > p/i - 1$*, there exists a constant* C_i*, depending only on* p, γ, i *such that for all* $(s,t) \in \Delta$*,*

$$\sup_D \sum_l \left|\mathbf{w}^i_{t_{l-1},t_l}\right|^{p/i} \le C_i \sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s \le t^n_{k-1} < t^n_k \le t}}^{2^n} \sum_{j=1}^{i} \left|\mathbf{w}^j_{t^n_{k-1},t^n_k}\right|^{p/j},$$

where the supremum runs over all finite subdivisions D *of* $[s,t]$*.*

Proof. As the specific boundary points of the closed interval are not important to the proof, we can assume, without loss of generality, that $[s,t] = [0,1]$. By Hölder's inequality we have for positive $(a_n)_{n\in\mathbb{N}}$ and any $\gamma > q - 1 > 0$ that

$$\left(\sum_{n\in\mathbb{N}} a_n\right)^q = \left(\sum_{n\in\mathbb{N}} a_n \frac{n^{\gamma/q}}{n^{\gamma/q}}\right)^q \le \left(\sum_{n\in\mathbb{N}} \frac{1}{n^{\gamma/(q-1)}}\right)^{q-1} \sum_{n\in\mathbb{N}} n^{\gamma} a_n^q = C(q,\gamma) \sum_{n\in\mathbb{N}} n^{\gamma} a_n^q. \quad (4.6)$$

Now fix an subinterval $[a,b] \subseteq [0,1]$. Let $n_0 \in \mathbb{N}$ be the smallest number such that $[a,b]$ contains a dyadic interval $[t^{n_0}_{k_0-1}, t^{n_0}_{k_0}]$, with $1 \le k_0 \le 2^{n_0}$. If $[t^{n_0}_{k_0-1}, t^{n_0}_{k_0}] = [a,b]$, we stop. If $t^{n_0}_{k_0} < b$, we choose the smallest $n_1 > n_0$ with $1 \le k_1 \le 2^{n_1}$ implicitly defined by $t^{n_0}_{k_0} = t^{n_1}_{k_1-1}$ such that $[t^{n_1}_{k_1-1}, t^{n_1}_{k_1}] \subseteq [t^{n_0}_{k_0}, b]$. Carrying on, we obtain an increasing sequence (n_j) with corresponding (k_j) such that

$$t^{n_0}_{k_0-1} < t^{n_0}_{k_0} = t^{n_1}_{k_1-1} < t^{n_1}_{k_1} = t^{n_2}_{k_2-1} < t^{n_2}_{k_2} = \dots < t^{n_j}_{k_j} \le b,$$

where the sequence is finite if there exists a $j \in \mathbb{N}$ such that $t^{n_j}_{k_j} = b$ and infinite with $t^{n_j}_{k_j} \to b,\ j \to \infty$ otherwise. The same procedure can be applied to the left end point of $[t^{n_0}_{k_0-1}, t^{n_0}_{k_0}]$ which yields another increasing sequence $(\overline{n}_j)$ with corresponding $(\overline{k}_j)$ such that

$$t^{n_0}_{k_0} > t^{n_0}_{k_0-1} = t^{\overline{n}_0}_{k_0-1} = t^{\overline{n}_1}_{\overline{k}_1} > t^{\overline{n}_1}_{\overline{k}_1-1} = t^{\overline{n}_2}_{\overline{k}_2} > \dots > t^{\overline{n}_j}_{\overline{k}_j-1} \ge a$$

where the sequence is finite if there exists a $j \in \mathbb{N}$ such that $t^{\overline{n}_j}_{\overline{k}_j-1} = a$ and infinite with $t^{\overline{n}_j}_{\overline{k}_j-1} \to a,\ j \to \infty$ otherwise. The summations and unions over j in the remainder of this proof depend on the construction above, e.g. whether the sequences above are finite or not. For the benefit of a simpler notation we do not write down these limits but they should be clear from the construction of the sequences $(n_j), (\overline{n}_j)$. Set $u_0 := t^{n_0}_{k_0}$, $u_j := t^{n_j}_{k_j}$ and $u_{-j} := t^{\overline{n}_{j-1}}_{\overline{k}_{j-1}-1}$ for $j > 0$. Our construction then yields

$$[a, b] = \bigcup_j [u_{j-1}, u_j],$$

where the intervals are dyadic and disjunct expect for common boundary points. Since $\mathbf{w}$ is multiplicative functional, and in particular continuous, we have

$$\mathbf{w}^1_{a,b} = \mathbf{w}^1_{a,u_{-N-1}} + \sum_{j=-N}^{N} \mathbf{w}^1_{u_{j-1},u_j} + \mathbf{w}^1_{u_N,b} = \sum_j \mathbf{w}^1_{u_{j-1},u_j}.$$

Using the triangle inequality and (4.6) yields

$$\begin{aligned}
|\mathbf{w}^1_{a,b}|^p &\le C\left(|\mathbf{w}^1_{t^{n_0}_{k_0-1},t^{n_0}_{k_0}}|^p + \Big(\sum_j |\mathbf{w}^1_{t^{n_j}_{k_j-1},t^{n_j}_{k_j}}|\Big)^p + \Big(\sum_j |\mathbf{w}^1_{t^{\overline{n}_j}_{\overline{k}_j-1},t^{\overline{n}_j}_{\overline{k}_j}}|\Big)^p\right) \\
&\le C\left(|\mathbf{w}^1_{t^{n_0}_{k_0-1},t^{n_0}_{k_0}}|^p + \sum_j j^\gamma |\mathbf{w}^1_{t^{n_j}_{k_j-1},t^{n_j}_{k_j}}|^p + \sum_j j^\gamma |\mathbf{w}^1_{t^{\overline{n}_j}_{\overline{k}_j-1},t^{\overline{n}_j}_{\overline{k}_j}}|^p\right) \qquad (4.7) \\
&\le C_1\left(|\mathbf{w}^1_{t^{n_0}_{k_0-1},t^{n_0}_{k_0}}|^p + \sum_j n_j^\gamma |\mathbf{w}^1_{t^{n_j}_{k_j-1},t^{n_j}_{k_j}}|^p + \sum_j \overline{n}_j^\gamma |\mathbf{w}^1_{t^{\overline{n}_j}_{\overline{k}_j-1},t^{\overline{n}_j}_{\overline{k}_j}}|^p\right),
\end{aligned}$$

where $C_1 > 0$ is a constant only depending on p and γ. Now let $D := \{0 = s = t_0 < \cdots < t_m = t = 1\}$ be a finite partition. Apply the procedure above to every subinterval $[t_{l-1}, t_l]$ and we obtain partitions P^l of $[t_{l-1}, t_l]$ that only contain, possibly infinitely many, dyadic intervals, such that

$$\sum_l |\mathbf{w}^1_{t_{l-1},t_l}|^p \le C_1 \sum_l \sum_{[t^{(l)}_{j-1},t^{(l)}_j]\in P^l} \widetilde{n}^\gamma_{j,l} |\mathbf{w}^1_{t^{(l)}_{j-1},t^{(l)}_j}|^p,$$

where $\widetilde{n}_{j,l}$ is given by $2^{-\widetilde{n}_{j,l}} = t^{(l)}_j - t^{(l)}_{j-1}$. Since any dyadic interval occurs at most once on the right-hand side, summing over all dyadic intervals can only increase the right-hand side and we obtain, after rearranging the sum, that

$$\sum_l |\mathbf{w}^1_{t_{l-1},t_l}|^p \le C_1 \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |\mathbf{w}^1_{t^n_{k-1},t^n_k}|^p.$$

Since the bound on the right-hand side is independent of the partition used on the left-hand side the assertion follows. In the same way as for the first level path the multiplicative nature of $\mathbf{w}$ yields

$$\mathbf{w}^2_{a,b} = \sum_j \mathbf{w}^2_{u_{j-1},u_j} + \sum_{k>j} \mathbf{w}^1_{u_{j-1},u_j} \otimes \mathbf{w}^1_{u_{k-1},u_k} \qquad (4.8)$$

and therefore

$$\left|\mathbf{w}^2_{a,b}\right|^{p/2} \le 2^{p/2-1}\left|\sum_j \mathbf{w}^2_{u_{j-1},u_j}\right|^{p/2} + 2^{p/2-1}\left|\sum_j\sum_{k>j} \mathbf{w}^1_{u_{j-1},u_j} \otimes \mathbf{w}^1_{u_{k-1},u_k}\right|^{p/2}$$

$$\leq 2^{p/2-1}\Bigg(\bigg(\sum_j |\mathbf{w}^2_{u_{j-1},u_j}|\bigg)^{p/2} + \bigg(\sum_j |\mathbf{w}^1_{u_{j-1},u_j}|\bigg)^{p}\Bigg).$$

In the same way as above, this leads to

$$\sum_l |\mathbf{w}^2_{t_{l-1},t_l}|^{p/2} \leq C_2\Bigg(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |\mathbf{w}^2_{t^n_{k-1},t^n_k}|^{p/2} + \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} |\mathbf{w}^1_{t^n_{k-1},t^n_k}|^{p}\Bigg),$$

which concludes the proof. □

We want to use Lemma 4.2.1 to control continuity in p-variation of multiplicative functionals. The result for first level paths is a direct consequence of Lemma 4.2.1 and, as we do not use any properties of higher level paths in the proof, we can extend this result to the first level paths of multiplicative functionals in $T^{(n)}(\mathbb{R}^d)$, $n \in \mathbb{N}$.

Corollary 4.2.2. *Let $\mathbf{w}, \mathbf{v}$ be two multiplicative functionals in $T^{(n)}(\mathbb{R}^d)$, $p > 1$ and $\gamma > 0$. Then we have for all $(s,t) \in \Delta$ that*

$$\sup_D \sum_l |\mathbf{w}^1_{t_{l-1},t_l} - \mathbf{v}^1_{t_{l-1},t_l}|^p \leq C_1 \sum_{n=1}^{\infty} n^\gamma \sum_{\substack{k=1 \\ s\leq t^n_{k-1}<t^n_k\leq t}}^{2^n} |\mathbf{w}^1_{t^n_{k-1},t^n_k} - \mathbf{v}^1_{t^n_{k-1},t^n_k}|^p, \tag{4.9}$$

where the supremum runs over all finite subdivisions $D = \{t_l\}$ of $[s,t]$.

The result for second level paths is a less direct consequence and presented in the following lemma which can be found as Lemma 3 in [28].

Lemma 4.2.3. *Let $\mathbf{w}, \mathbf{v}$ be two multiplicative functionals in $T^{(2)}(\mathbb{R}^d)$. For any $p > 2$ and $\gamma > p/2 - 1$, there exists a constant C_2, only depending on γ and p, such that*

$$\begin{aligned}
&\sup_D \sum_l \big|\mathbf{w}^2_{t_{l-1},t_l} - \mathbf{v}^2_{t_{l-1},t_l}\big|^{p/2} \\
&\quad\leq C_2 \sum_{n=1}^{\infty} n^\gamma \sum_{\substack{k=1 \\ s\leq t^n_{k-1}<t^n_k\leq t}}^{2^n} \big|\mathbf{w}^2_{t^n_{k-1},t^n_k} - \mathbf{v}^2_{t^n_{k-1},t^n_k}\big|^{p/2} \\
&\quad + C_2\Bigg(\sum_{n=1}^{\infty} n^\gamma \sum_{\substack{k=1 \\ s\leq t^n_{k-1}<t^n_k\leq t}}^{2^n} \big|\mathbf{w}^1_{t^n_{k-1},t^n_k} - \mathbf{v}^1_{t^n_{k-1},t^n_k}\big|^{p}\Bigg)^{1/2} \\
&\qquad\qquad \times \Bigg(\sum_{n=1}^{\infty} n^\gamma \sum_{\substack{k=1 \\ s\leq t^n_{k-1}<t^n_k\leq t}}^{2^n} \big|\mathbf{w}^1_{t^n_{k-1},t^n_k}\big|^{p} + \big|\mathbf{v}^1_{t^n_{k-1},t^n_k}\big|^{p}\Bigg)^{1/2},
\end{aligned} \tag{4.10}$$

where the supremum runs over all finite subdivisions $D = \{t_l\}$ of $[s,t]$.

Proof. This proof builds on the work in proof of Lemma 4.2.1 and uses the same notation. We can again assume, without loss of generality, that $[s,t] = [0,1]$. Let $D := \{0 = s = t_0 < \cdots < t_m = t = 1\}$ be a finite partition. Using relation (4.8) for $\mathbf{w}^2_{t_{l-1},t_l}$ and $\mathbf{v}^2_{t_{l-1},t_l}$, we have

$$\begin{aligned}\mathbf{w}^2_{t_{l-1},t_l} - \mathbf{v}^2_{t_{l-1},t_l} = \sum_j \Big[&(\mathbf{w}^2_{u_{j-1},u_j} - \mathbf{v}^2_{u_{j-1},u_j}) \\ &+ \sum_{k>j} (\mathbf{w}^1_{u_{j-1},u_j} - \mathbf{v}^1_{u_{j-1},u_j}) \otimes \mathbf{w}^1_{u_{k-1},u_k} \\ &+ \sum_{k>j} \mathbf{v}^1_{u_{j-1},u_j} \otimes (\mathbf{w}^1_{u_{k-1},u_k} - \mathbf{v}^1_{u_{k-1},u_k}) \Big],\end{aligned} \tag{4.11}$$

where the (u_j) depend on l. Since $|x+y+z|^q \leq 3^{q-1}(|x|^q+|y|^q+|z|^q)$ for $x, y, z \in \mathbb{R},\ q \geq 1$, we split the right-hand side above into three parts and look at them separately. The first term can be handled the same way as in the proof of Lemma 4.2.1 and we obtain

$$\sum_l \Big| \sum_j (\mathbf{w}^2_{u_{j-1},u_j} - \mathbf{v}^2_{u_{j-1},u_j}) \Big|^{p/2} \leq C \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big| \mathbf{w}^2_{t^n_{k-1},t^n_k} - \mathbf{v}^2_{t^n_{k-1},t^n_k} \big|^{p/2}.$$

Moreover, by (4.6), we have for the second term of (4.11) that

$$\begin{aligned}&\Big| \sum_j \sum_{k>j} (\mathbf{w}^1_{u_{j-1},u_j} - \mathbf{v}^1_{u_{j-1},u_j}) \otimes \mathbf{w}^1_{u_{k-1},u_k} \Big|^{p/2} \\ &\qquad \leq \Big(\sum_j \big| \mathbf{w}^1_{u_{j-1},u_j} - \mathbf{v}^1_{u_{j-1},u_j} \big| \Big)^{p/2} \Big(\sum_j \big| \mathbf{w}^1_{u_{j-1},u_j} \big| \Big)^{p/2} \\ &\qquad \leq \Big(\sum_j j^\gamma \big| \mathbf{w}^1_{u_{j-1},u_j} - \mathbf{v}^1_{u_{j-1},u_j} \big|^p \Big)^{1/2} \Big(\sum_j j^\gamma \big| \mathbf{w}^1_{u_{j-1},u_j} \big|^p \Big)^{1/2}.\end{aligned}$$

Again applying the methods of the proof of Lemma 4.2.1 yields

$$\begin{aligned}&\sum_l \Big| \sum_j \sum_{k>j} (\mathbf{w}^1_{u_{j-1},u_j} - \mathbf{v}^1_{u_{j-1},u_j}) \otimes \mathbf{w}^1_{u_{k-1},u_k} \Big|^{p/2} \\ &\qquad \leq C \Big(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big| \mathbf{w}^1_{t^n_{k-1},t^n_k} - \mathbf{v}^1_{t^n_{k-1},t^n_k} \big|^p \Big)^{1/2} \Big(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big| \mathbf{w}^1_{t^n_{k-1},t^n_k} \big|^p \Big)^{1/2}.\end{aligned}$$

The third term on the right-hand side of (4.11) can be treated equivalently to the second and the assertion follows. □

Remark 4.2.4. Let $\mathbf{w}, \mathbf{v}$ be two multiplicative functionals in $T^{(2)}(\mathbb{R}^d)$. Looking at the second level components $\mathbf{w}^2 = (\mathbf{w}^2(i,j))_{i,j=1,\dots d}$ Lemma 4.2.3 can be rewritten in the following way. For any $p > 2$ and $\gamma > \frac{p}{2} - 1$, there exists a constant C_2, only depending on γ and p, such that

$$\begin{aligned}
&\sup_D \sum_l \left|\mathbf{w}^2(i,j)_{t_{l-1},t_l} - \mathbf{v}^2(i,j)_{t_{l-1},t_l}\right|^{p/2} \\
&\quad \le C_2 \sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^2(i,j)_{t_{k-1}^n,t_k^n} - \mathbf{v}^2(i,j)_{t_{k-1}^n,t_k^n}\right|^{p/2} \\
&\quad + C_2 \Bigg(\sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^1(i)_{t_{k-1}^n,t_k^n} - \mathbf{v}^1(i)_{t_{k-1}^n,t_k^n}\right|^{p} \\
&\qquad\qquad + \left|\mathbf{w}^1(j)_{t_{k-1}^n,t_k^n} - \mathbf{v}^1(j)_{t_{k-1}^n,t_k^n}\right|^{p} \Bigg)^{1/2} \\
&\qquad\qquad \times \left(\sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^1(j)_{t_{k-1}^n,t_k^n}\right|^{p} + \left|\mathbf{v}^1(i)_{t_{k-1}^n,t_k^n}\right|^{p} \right)^{1/2} \\
&\quad \le C_2 \sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^2(i,j)_{t_{k-1}^n,t_k^n} - \mathbf{v}^2(i,j)_{t_{k-1}^n,t_k^n}\right|^{p/2} \\
&\quad + C_2 \left(\sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^1_{t_{k-1}^n,t_k^n} - \mathbf{v}^1_{t_{k-1}^n,t_k^n}\right|^{p} \right)^{1/2} \\
&\qquad\qquad \times \left(\sum_{n=1}^{\infty} n^{\gamma} \sum_{\substack{k=1 \\ s\le t_{k-1}^n < t_k^n \le t}}^{2^n} \left|\mathbf{w}^1_{t_{k-1}^n,t_k^n}\right|^{p} + \left|\mathbf{v}^1_{t_{k-1}^n,t_k^n}\right|^{p} \right)^{1/2},
\end{aligned}$$

where the supremum runs over all finite subdivisions $D = \{t_l\}$ of $[s,t]$.

4.2.2 Dyadic approximation

To lift a function to a rough paths we first consider rough paths over its dyadic approximations and then check whether these converge to a limiting rough path. In this section we present some results on such dyadic approximations. The results were originally presented in Section 3.2 of [4].

Let $w : [0,1] \to \mathbb{R}^d$ be a function and $m \in \mathbb{N}$. We then denote by $w(m)$ the linear interpolation of w through the dyadic points $t_k^m := k/2^m,\ k = 0,\dots,2^m$, i.e.

$$w(m)_t = w_{t_{k-1}^m} + 2^m(t - t_{k-1}^m)\Delta_k^m w, \qquad \text{for} \quad t_{k-1}^m \le t < t_k^m,$$

where $\Delta_k^m w := w(m)_{t_{k-1}^m,t_k^m} := w_{t_k^m} - w_{t_{k-1}^m}$. Since $w(m)$ has finite variation, we are able to define the smooth rough path

$$\mathbf{w}(m)_{s,t} = (1, \mathbf{w}(m)^1_{s,t}, \mathbf{w}(m)^2_{s,t}),$$

where $\mathbf{w}(m)^i_{s,t}$ is the i-th iterated integral of $w(m)$ over the interval $[s,t]$.

We now prove some formulas for the rough path $\mathbf{w}(m)$. The following statements hold for any $k = 1,\dots,2^n$.

(i) If $m \leq n$, we have

$$\mathbf{w}(m)^1_{t^n_{k-1},t^n_k} = 2^{m-n}\Delta^m_l w, \tag{4.12}$$

where $l \in \mathbb{N}$ is the unique number that satisfies

$$\frac{l-1}{2^m} \leq \frac{k-1}{2^n} < \frac{k}{2^n} \leq \frac{l}{2^m}. \tag{4.13}$$

For the second level paths and l as defined in (4.13), we obtain

$$\mathbf{w}(m)^2_{t^n_{k-1},t^n_k} = \frac{1}{2}(2^{m-n})^2(\Delta^m_l w)^{\otimes 2} \tag{4.14}$$

since

$$\begin{aligned}
\mathbf{w}(m)^2_{t^n_{k-1},t^n_k} &= \int_{t^n_{k-1}}^{t^n_k} w(m)_{t^n_{k-1},t} \otimes \mathrm{d}w(m)_t \\
&= 2^{2n}\int_{t^n_{k-1}}^{t^n_k} (t - t^n_{k-1}) w(m)_{t^n_{k-1},t^n_k} \otimes w(m)_{t^n_{k-1},t^n_k}\,\mathrm{d}t \\
&= 2^{2n}(w(m)_{t^n_{k-1},t^n_k})^{\otimes 2}\int_0^{2^{-n}} t\,\mathrm{d}t = \frac{1}{2}(\mathbf{w}(m)^1_{t^n_{k-1},t^n_k})^{\otimes 2} \\
&= \frac{1}{2}(2^{m-n})^2(\Delta^m_l w)^{\otimes 2},
\end{aligned}$$

where we make use of (4.12).

(ii) If $m \geq n$, we have

$$\mathbf{w}(m)^1_{t^n_{k-1},t^n_k} = \sum_{j=2^{m-n}(k-1)+1}^{2^{m-n}k} \Delta^m_j w \tag{4.15}$$

by Chen's relation, which, for first level paths, is just stating the property of a telescopic sum.

For the second level paths, we also make use of Chen's relation and obtain

$$\begin{aligned}
\mathbf{w}(m)^2_{t^n_{k-1},t^n_k} &= \sum_{j=2^{m-n}(k-1)+1}^{2^{m-n}k} \mathbf{w}(m)^2_{t^m_{j-1},t^m_j} \\
&\quad + \sum_{j=2^{m-n}(k-1)+2}^{2^{m-n}k} \mathbf{w}(m)^1_{t^m_{2^{m-n}(k-1)},t^m_{j-1}} \otimes \mathbf{w}(m)^1_{t^m_{j-1},t^m_j} \\
&= \sum_{j=2^{m-n}(k-1)+1}^{2^{m-n}k} \frac{(\Delta^m_j w)^{\otimes 2}}{2} \\
&\quad + \sum_{j=2^{m-n}(k-1)+2}^{2^{m-n}k} \sum_{i=2^{m-n}(k-1)+1}^{j-1} \Delta^m_i w \otimes \Delta^m_j w.
\end{aligned} \tag{4.16}$$

Lemma 4.2.5. *Let $n \in \mathbb{N}_0$, $k \in \{1, \dots, 2^n\}$ and $m \geq n$. Then we have*

$$\mathbf{w}(m+1)^2_{t^n_{k-1}, t^n_k} - \mathbf{w}(m)^2_{t^n_{k-1}, t^n_k} = \frac{1}{2} \sum_{j=2^{m-n}(k-1)+1}^{2^{m-n}k} \Delta^{m+1}_{2j-1} w \otimes \Delta^{m+1}_{2j} w - \Delta^{m+1}_{2j} w \otimes \Delta^{m+1}_{2j-1} w.$$

Proof. Let us first note that

$$\Delta^m_j w = w_{\frac{j}{2^m}} - w_{\frac{j-1}{2^m}} = w_{\frac{2j}{2^{m+1}}} - w_{\frac{2j-1}{2^{m+1}}} + w_{\frac{2j-1}{2^{m+1}}} - w_{\frac{2j-2}{2^{m+1}}} = \Delta^{m+1}_{2j} w + \Delta^{m+1}_{2j-1} w.$$

To shorten notation, we write

$$\overline{m} = 2^{m-n} k \qquad \underline{m} = 2^{m-n}(k-1).$$

Obviously, this implies $\overline{m+1} = 2\overline{m}$ and $\underline{m+1} = 2\underline{m}$. We prove the Lemma by using equation (4.16), which trivially holds for $m = n$. We split the sums appearing in (4.16) into two parts

$$\mathbf{w}(m)^2_{t^n_{k-1}, t^n_k} = \frac{a(m)}{2} + b(m),$$

where

$$a(m) = \sum_{j=\underline{m}+1}^{\overline{m}} (\Delta^m_j w)^{\otimes 2},$$

$$b(m) = \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=\underline{m}+1}^{j-1} \Delta^m_i w \otimes \Delta^m_j w.$$

We have

$$a(m+1) = \sum_{j=2\underline{m}+1}^{2\overline{m}} (\Delta^{m+1}_j w)^{\otimes 2} = \sum_{j=\underline{m}+1}^{\overline{m}} (\Delta^{m+1}_{2j} w)^{\otimes 2} + (\Delta^{m+1}_{2j-1} w)^{\otimes 2}$$

and

$$\begin{aligned} a(m) = \sum_{j=\underline{m}+1}^{\overline{m}} (\Delta^m_j w)^{\otimes 2} &= \sum_{j=\underline{m}+1}^{\overline{m}} (\Delta^{m+1}_{2j} w + \Delta^{m+1}_{2j-1} w)^{\otimes 2} \\ &= \sum_{j=\underline{m}+1}^{\overline{m}} (\Delta^{m+1}_{2j} w)^{\otimes 2} + (\Delta^{m+1}_{2j-1} w)^{\otimes 2} \\ &+ \sum_{j=\underline{m}+1}^{\overline{m}} \Delta^{m+1}_{2j} w \otimes \Delta^{m+1}_{2j-1} w + \Delta^{m+1}_{2j-1} w \otimes \Delta^{m+1}_{2j} w. \end{aligned}$$

So we obtain

$$a(m+1) - a(m) = -\sum_{j=\underline{m}+1}^{\overline{m}} \Delta^{m+1}_{2j} w \otimes \Delta^{m+1}_{2j-1} w + \Delta^{m+1}_{2j-1} w \otimes \Delta^{m+1}_{2j} w$$

for the difference of the two. Moreover, we have

$$\begin{aligned} b(m+1) &= \sum_{j=2\underline{m}+2}^{2\overline{m}} \sum_{i=2\underline{m}+1}^{j-1} \Delta_i^{m+1} w \otimes \Delta_j^{m+1} w \\ &= \sum_{j=\underline{m}+1}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-1} \Delta_i^{m+1} w \otimes \Delta_{2j}^{m+1} w + \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-2} \Delta_i^{m+1} w \otimes \Delta_{2j-1}^{m+1} w \end{aligned}$$

and

$$\begin{aligned} b(m) &= \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=\underline{m}+1}^{j-1} (\Delta_{2i}^{m+1} w + \Delta_{2i-1}^{m+1} w) \otimes \Delta_j^m w \\ &= \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-2} \Delta_i^{m+1} w \otimes (\Delta_{2j}^{m+1} w + \Delta_{2j-1}^{m+1} w) \\ &= \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-2} \Delta_i^{m+1} w \otimes \Delta_{2j}^{m+1} w + \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-2} \Delta_i^{m+1} w \otimes \Delta_{2j-1}^{m+1} w. \end{aligned}$$

Subtracting the two yields

$$\begin{aligned} b(m+1) - b(m) &= \Delta_{2\underline{m}+1}^{m+1} w \otimes \Delta_{2\underline{m}+2}^{m+1} w + \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-1} \Delta_i^{m+1} w \otimes \Delta_{2j}^{m+1} w \\ &\qquad - \sum_{j=\underline{m}+2}^{\overline{m}} \sum_{i=2\underline{m}+1}^{2j-2} \Delta_i^{m+1} w \otimes \Delta_{2j}^{m+1} w \\ &= \Delta_{2\underline{m}+1}^{m+1} w \otimes \Delta_{2\underline{m}+2}^{m+1} w + \sum_{j=\underline{m}+2}^{\overline{m}} \Delta_{2j-1}^{m+1} w \otimes \Delta_{2j}^{m+1} w \\ &= \sum_{j=\underline{m}+1}^{\overline{m}} \Delta_{2j-1}^{m+1} w \otimes \Delta_{2j}^{m+1} w \end{aligned}$$

and the assertion follows. □

4.3 Paths of fBm and its derivative as rough paths

We remind the reader, that, as in the chapters above, we consider a fractional Brownian B^H of the type presented in Definition 2.3.3.

Let $n \in \mathbb{N}$ and $k \in \{1, \dots, 2^n\}$. Then, with the notation introduced above, we have that

$$\mathbb{E}\big[\Delta_k^n B^H \Delta_k^n B^H\big] = \mathbb{E}\big[(\Delta_k^n B^H)^2\big] = \frac{1}{2^{2Hn}}.$$

The following lemma, stated in [4], gives an estimate of the right-hand side in the case non-overlapping increments.

Lemma 4.3.1. *Let $n \in \mathbb{N}$ and $k, l \in \{1, \dots, 2^n\}$ such that $|k-l| \geq 1$. Then there exists a positive constant C depending only on H such that*

$$\left|\mathbb{E}\big[(\Delta_l^n B^H)(\Delta_k^n B^H)\big]\right| \leq C\frac{|k-l|^{2H-2}}{2^{2Hn}}.$$

For $H = \frac{1}{2}$ the constant on the right-hand side can be chosen to be zero.

Proof. By the properties of fBm, we have

$$\mathbb{E}\big[(B_t^H - B_s^H)^2\big] = |t-s|^{2H}, \qquad s, t \geq 0.$$

Let $s, t, u, v \geq 0$. We obtain, by expanding the product, that

$$\begin{aligned}
&\mathbb{E}\big[(B_t^H - B_s^H)(B_u^H - B_v^H)\big] \\
&\quad = \frac{1}{2}\mathbb{E}\big[(B_t^H - B_v^H)^2 + (B_s^H - B_u^H)^2 - (B_t^H - B_u^H)^2 - (B_s^H - B_v^H)^2\big] \\
&\quad = \frac{1}{2}\big(|t-v|^{2H} + |s-u|^{2H} - |t-u|^{2H} - |s-v|^{2H}\big).
\end{aligned}$$

So, for all $n \in \mathbb{N}$ and $k, l \in \{1, \dots, 2^n\}$, it follows that

$$\mathbb{E}\big[\Delta_k^n B^H \Delta_l^n B^H\big] = \frac{1}{2(2^n)^{2H}}\big(|k-l-1|^{2H} + |k-l+1|^{2H} - 2|k-l|^{2H}\big).$$

Letting $|k-l| \geq 1$, this can be written as

$$\mathbb{E}\big[\Delta_k^n B^H \Delta_l^n B^H\big] = \frac{|k-l|^{2H}}{2^{2Hn+1}}\big((1-|k-l|^{-1})^{2H} + (1+|k-l|^{-1})^{2H} - 2\big)$$

and for $|k-l| = 1$ we obtain

$$\mathbb{E}\big[\Delta_k^n B^H \Delta_l^n B^H\big] = \frac{2^{2H} - 2}{2^{2Hn+1}}.$$

Let $|k-l| \geq 2$ and put $f : [0, \infty) \to \mathbb{R};\ x \mapsto x^{2H}$. It holds by Taylor's Theorem that

$$\begin{aligned}
(1-|k-l|^{-1})^{2H} &= f(1-|k-l|^{-1}) = 1 - f'(1)|k-l|^{-1} + f''(\xi_1)|k-l|^{-2}, \\
(1+|k-l|^{-1})^{2H} &= f(1+|k-l|^{-1}) = 1 + f'(1)|k-l|^{-1} + f''(\xi_2)|k-l|^{-2},
\end{aligned}$$

where $\xi_1 \in (1-|k-l|^{-1}, 1) \subseteq [\frac{1}{2}, 1]$ and $\xi_1 \in (1, 1+|k-l|^{-1},) \subseteq [1, \frac{3}{2}]$. Therefore, we obtain

$$\left|\mathbb{E}\big[\Delta_k^n B^H \Delta_l^n B^H\big]\right| \leq \frac{|k-l|^{2H-2}}{2^{2Hn}} 2H(2H-1) \sup_{x \in [1/2, 3/2]} \left|x^{2H-2}\right|.$$

Putting the two results together yields

$$\left|\mathbb{E}\big[\Delta_k^n B^H \Delta_l^n B^H\big]\right| \leq C\frac{|k-l|^{2H-2}}{2^{2Hn}}, \qquad |k-l| \geq 1.$$

□

Writing ∂_H for $\frac{\partial}{\partial H}$, our aim is the construction of an extension of $(B^H, \partial_H B^H)$ to a rough path by using the results in Chapter 2. As this extension is trivial in the case of $H > \frac{1}{2}$ we restrict ourself to $H \le \frac{1}{2}$ until stated otherwise.

Let $T > 0$ and $0 \le s < t \le s + \tau < t + \tau \le T$. Note that by substituting in their integral representations one quickly sees that $(B^H, \partial_H B^H)$ is (shift) stationary. Therefore, we have for $\lambda_1, \lambda_2 \in \{0, 1\}$ that

$$\begin{aligned} &\mathbb{E}\big[(\partial_H^{\lambda_1} B^H_{t+\tau} - \partial_H^{\lambda_1} B^H_{s+\tau})(\partial_H^{\lambda_2} B^H_t - \partial_H^{\lambda_2} B^H_s)\big] \\ &\qquad = \mathbb{E}\big[(\partial_H^{\lambda_1} B^H_{(t-s)+\tau} - \partial_H^{\lambda_1} B^H_\tau)(\partial_H^{\lambda_2} B^H_{(t-s)} - \partial_H^{\lambda_2} B^H_0)\big]. \end{aligned} \tag{4.17}$$

Let $\varepsilon_1 > 0$ such that $\varepsilon_1 \neq \frac{1}{2} - H$ and $\varepsilon \in (0, H)$, which ensures $\varepsilon \neq H - \frac{1}{2}$. Writing $H_\varepsilon = H - \varepsilon$, we obtain by Lemma 4.5.1 and (2.19), (4.17) that there exists a constant C depending only on $\varepsilon, \varepsilon_1$ and H such that

$$\begin{aligned} &\mathbb{E}\big[(\partial_H^{\lambda_1} B^H_{t+\tau} - \partial_H^{\lambda_1} B^H_{s+\tau})(\partial_H^{\lambda_2} B^H_t - \partial_H^{\lambda_2} B^H_s)\big] \\ &\le C\bigg(\int_{-\infty}^{-1} \big|(\nu + \tau - u)^{H-1/2+\varepsilon_1} - (\tau - u)^{H-1/2+\varepsilon_1}\big| \\ &\qquad\qquad \times \big|(\nu - u)^{H-1/2+\varepsilon_1} - (-u)^{H-1/2+\varepsilon_1}\big| \,\mathrm{d}u \\ &+ \int_{-1}^{0} \big|(\nu + \tau - u)^{H_\varepsilon-1/2} - (\tau - u)^{H_\varepsilon-1/2}\big| \\ &\qquad\qquad \times \big|(\nu - u)^{H_\varepsilon-1/2} - (-u)^{H_\varepsilon-1/2}\big| \,\mathrm{d}u \\ &+ \int_0^{\nu} \big|(\nu + \tau - u)^{H_\varepsilon-1/2} - (\tau - u)^{H_\varepsilon-1/2}\big|(t - u)^{H_\varepsilon-1/2} \,\mathrm{d}u\bigg) \\ &=: C(I_1 + I_2 + I_3), \end{aligned} \tag{4.18}$$

where $\nu = t - s$.

The next Lemma is essential in ensuring that we can apply the procedure used in [4] to not only lift a multidimensional fractional Brownian motion to a rough path as the authors did but to also lift its derivative process.

Lemma 4.3.2. *Let $n \in \mathbb{N}$ and $k, l \in \{1, \ldots, 2^n\}$ such that $|k-l| \ge 1$. For every $\varepsilon \in (0, H)$ there exists a positive constant C depending only on H (restricted to $H \le \frac{1}{2}$) and ε such that for any $\lambda_1, \lambda_2 \in \{0, 1\}$ we have*

$$\Big|\mathbb{E}\big[\Delta_k^n\big(\partial_H^{\lambda_1} B^H\big)\Delta_l^n\big(\partial_H^{\lambda_2} B^H\big)\big]\Big| \le C\frac{|k-l|^{H_\varepsilon-3/2}}{2^{2H_\varepsilon n}}, \tag{4.19}$$

where $H_\varepsilon = H - \varepsilon \in (0, \frac{1}{2})$.

Proof. We set $\nu := t-s$ and let $\tau > 0$. If we show that there exists a constant C depending only on H and ε such that for any $\lambda_1, \lambda_2 \in \{0, 1\}$ we have

$$\Big|\mathbb{E}\big[(\partial_H^{\lambda_1} B^H_{t+\tau} - \partial_H^{\lambda_1} B^H_{s+\tau})(\partial_H^{\lambda_2} B^H_t - \partial_H^{\lambda_2} B^H_s)\big]\Big| \le C\tau^{2H_\varepsilon}\eta^{H_\varepsilon+3/2}$$

for all $0 \le s < t \le 1$ with $\eta := \nu/\tau \in \{1\} \cup [0, \frac{1}{2}]$, then the assertion follows by setting $\nu = \frac{1}{2n}$, $\tau = \frac{|k-l|}{2^n}$, which implies $\eta = \frac{1}{|k-l|}$. Thanks to (4.18) we only need to bound the integrals I_1, I_2, I_3.

Set $\varepsilon_1 := \frac{1-H}{2} \neq \frac{1-2H}{2} = \frac{1}{2} - H$. We have $2(H+\varepsilon_1) - 3 = H - 2$. Therefore, using Taylor's theorem yields

$$\begin{aligned}\int_{-\infty}^{-1} &\left|(\nu-u)^{H-1/2+\varepsilon_1} - (-u)^{H-1/2+\varepsilon_1}\right|^2 \mathrm{d}u \\ &= (H-1/2+\varepsilon_1)^2\nu^2 \int_{-\infty}^{-1} (-u+\xi_u)^{2(H+\varepsilon_1)-3} \,\mathrm{d}u \\ &\leq C\nu^2 \int_1^\infty x^{2(H+\varepsilon_1)-3} \,\mathrm{d}x \leq C\nu^2,\end{aligned} \tag{4.20}$$

where $\xi = \xi_u \in (0,\nu)$ and C depends only on H. In the same way we obtain

$$\begin{aligned}\int_{-\infty}^{-1} \left|(\nu+\tau-u)^{H-1/2+\varepsilon_1} - (\tau-u)^{H-1/2+\varepsilon_1}\right|^2 \mathrm{d}u &\leq C\nu^2 \int_{1+\tau}^\infty x^{2(H+\varepsilon_1)-3} \,\mathrm{d}x \\ &\leq C\nu^2 \int_1^\infty x^{2(H+\varepsilon_1)-3,} \,\mathrm{d}x \\ &\leq C\nu^2.\end{aligned} \tag{4.21}$$

Thus, the Cauchy-Schwarz inequality together with (4.20), (4.21) yields $I_1 \leq C\nu^2$, where C depends only on H. As $\tau \leq 1$ we have $\tau^{2H_\varepsilon - 2} \geq 1$ and obtain

$$I_1 \leq C\tau^{2H_\varepsilon}\left(\frac{\nu}{\tau}\right)^2. \tag{4.22}$$

Since $H \leq \frac{1}{2}$, we have that for $u \in (-\infty, 0)$

$$\big((\nu+\tau-u)^{H_\varepsilon-1/2} - (\tau-u)^{H_\varepsilon-1/2}\big)\big((\nu-u)^{H_\varepsilon-1/2} - (-u)^{H_\varepsilon-1/2}\big) \geq 0$$

and for $u \in (0,\nu)$

$$\big((\nu+\tau-u)^{H_\varepsilon-1/2} - (\tau-u)^{H_\varepsilon-1/2}\big)(\nu-u)^{H_\varepsilon-1/2} \leq 0.$$

Further, it holds that

$$\begin{aligned}0 &\geq C_H^{-2} \cdot \mathbb{E}\big[(B^{H_\varepsilon}_{t+\tau} - B^{H_\varepsilon}_{s+\tau})(B^{H_\varepsilon}_t - B^{H_\varepsilon}_s)\big] \\ &= \int_{-\infty}^0 \big((\nu+\tau-u)^{H_\varepsilon-1/2} - (\tau-u)^{H_\varepsilon-1/2}\big)\big((\nu-u)^{H_\varepsilon-1/2} - (-u)^{H_\varepsilon-1/2}\big) \,\mathrm{d}u \\ &+ \int_0^\nu \big((\nu+\tau-u)^{H_\varepsilon-1/2} - (\tau-u)^{H_\varepsilon-1/2}\big)(\nu-u)^{H_\varepsilon-1/2} \,\mathrm{d}u,\end{aligned}$$

where C_H is the normalising constant of the Mandelbrot-van Ness representation. Therefore, we obtain

$$I_2 + I_3 \leq 2I_3. \tag{4.23}$$

Substituting $u/\tau = v = \eta - w$ yields

$$\begin{aligned}I_3 &= \int_0^\nu (\nu-u)^{H_\varepsilon-1/2}\left|(\nu+\tau-u)^{H_\varepsilon-1/2} - (\tau-u)^{H_\varepsilon-1/2}\right| \mathrm{d}u \\ &= \tau^{2H_\varepsilon} \int_0^\eta (\eta-v)^{H_\varepsilon-1/2}\left|(\eta+1-v)^{H_\varepsilon-1/2} - (1-v)^{H_\varepsilon-1/2}\right| \mathrm{d}v \\ &= \tau^{2H_\varepsilon} \int_0^\eta w^{H_\varepsilon-1/2}\left|(1+w)^{H_\varepsilon-1/2} - (1-\eta+w)^{H_\varepsilon-1/2}\right| \mathrm{d}w.\end{aligned} \tag{4.24}$$

Thus, we obtain for $\eta = 1$ that

$$\begin{aligned} I_3 &\le \tau^{2H_\varepsilon} \int_0^1 w^{H_\varepsilon - 1/2} \big| (1+w)^{H_\varepsilon - 1/2} - w^{H_\varepsilon - 1/2} \big| \, \mathrm{d}w \\ &\le \tau^{2H_\varepsilon} \left(\int_0^1 w^{H_\varepsilon - 1/2} \, \mathrm{d}w + \int_0^1 w^{2H_\varepsilon - 1} \, \mathrm{d}w \right) \\ &\le C\tau^{2H_\varepsilon} \le C\tau^{2H_\varepsilon}\eta^2. \end{aligned} \tag{4.25}$$

Now consider the case of $\eta \in (0, \frac{1}{2}]$. By Taylor's theorem we have for any $w \ge 0$ that

$$\begin{aligned} &\big| (1+w)^{H_\varepsilon - 1/2} - (1+w-\eta)^{H_\varepsilon - 1/2} \big| \\ &\qquad = \eta \Big| H_\varepsilon - \frac{1}{2} \Big| (1+w-\xi)^{H_\varepsilon - 3/2} \\ &\qquad \le \eta \Big| H_\varepsilon - \frac{1}{2} \Big| (1-\eta)^{H_\varepsilon - 3/2} \le \eta \Big| H_\varepsilon - \frac{1}{2} \Big| \left(\frac{1}{2} \right)^{H_\varepsilon - 3/2} \\ &\qquad \le C\eta, \end{aligned}$$

where $\xi \in (0, \eta)$. Therefore, plugging this into (4.24), we obtain

$$I_3 \le C\tau^{2H_\varepsilon}\eta \int_0^\eta w^{H_\varepsilon - 1/2} \, \mathrm{d}w = C\tau^{2H_\varepsilon}\eta^{H_\varepsilon + 3/2}. \tag{4.26}$$

Putting together (4.22), (4.23), (4.25), (4.26) and using the fact that $\eta^2 \le \eta^{H_\varepsilon + 3/2}$ yields that there exists a constant C depending only on H and ε, such that

$$I_1 + I_2 + I_3 \le C\tau^{2H_\varepsilon}\eta^{H_\varepsilon + 3/2},$$

for all $0 \le s < t \le 1$ and $\tau > 0$ such that $\eta = \nu/\tau \in \{1\} \cup [0, \frac{1}{2}]$. The proof can now be concluded by the argument given at the beginning. $\square$

Remark 4.3.3. Let $\lambda_1, \lambda_2 \in \{0, 1\}$ and $H_\varepsilon = H - \varepsilon$, where $\varepsilon \in (0, H)$. It follows from (4.18) and the arguments used in the proof of Lemma 2.3.2 that there exists a C depending only on H and ε such that

$$\Big| \mathbb{E}\big[(\partial_H^{\lambda_1} B_t^H - \partial_H^{\lambda_1} B_s^H)(\partial_H^{\lambda_2} B_t^H - \partial_H^{\lambda_2} B_s^H) \big] \le C(t-s)^{2H_\varepsilon}$$

for all $0 \le s < t \le 1$ and therefore we have

$$\Big| \mathbb{E}\big[\Delta_k^n (\partial_H^{\lambda_1} B^H) \Delta_k^n (\partial_H^{\lambda_2} B^H) \big] \Big| \le C 2^{-2H_\varepsilon n}$$

for all $n \in \mathbb{N}$ and $k \in \{1, \dots, 2^n\}$.

From now on we consider a d-dimension Mandelbrot-van Ness type fractional Brownian motion $B^H = (B^{H,(1)}, \dots, B^{H,(d)})$, as given in Definition 2.3.3. To simplify the notation, we sometimes omit the H and write

$$B = B^H = (B^{H,(1)}, \dots, B^{H,(d)}) = (B^1, \dots, B^d).$$

Further let D^j denote the derivative process $D^j = \partial_H B^j$ and $D = (D^1, \dots, D^d)$. We are interested in constructing a rough path that extends

$$X = X^H = (X^1, \dots, X^{2d}) = (B^1, \dots, B^d, D^1, \dots, D^d) = (B, D).$$

But we will see in Section 4.3.2 that this not possible with the tools given here. However, we can construct a rough path over D.

We derive the following lemma from Lemma 4.3.2, which serve the purpose of Lemma 12 in [4] but is more involved due to the interdependence of components.

Lemma 4.3.4. *For $i, j \in \{1, \dots, 2d\}$ we have*

$$\Big|\mathbb{E}\Big[\big(\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j - \Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big)^2\Big]\Big| \leq C2^{-4H_\varepsilon m}.$$

Let $k > l$ and $i, j \in \{1, \dots, 2d\}$ with $i \bmod d \neq j \bmod d$. Then

$$\begin{aligned}\Big|\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j - \Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big)\big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\big)\big]\Big|\\ \leq C\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}},\end{aligned}$$

where C depends only on H and ε.

Proof. In the case $i = j$ the terms inside the expectations are zero and the assertion is trivial. From now on we only consider $i \neq j$.

Letting $k = l$, we have

$$\begin{aligned}&\Big|\mathbb{E}\Big[\big(\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j - \Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big)^2\Big]\Big|\\ &\qquad\leq \mathbb{E}\Big[\big(\Delta^{m+1}_{2k-1}X^i\big)^2\big(\Delta^{m+1}_{2k}X^j\big)^2\Big] + \mathbb{E}\Big[\big(\Delta^{m+1}_{2k}X^i\big)^2\big(\Delta^{m+1}_{2k-1}X^j\big)^2\Big]\\ &\qquad+ 2\Big|\mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j\big]\Big|.\end{aligned} \tag{4.27}$$

Using Remark 4.3.3 and the fact that the increments are Gaussian we have

$$\mathbb{E}\Big[\big(\Delta^{m+1}_{2k-1}X^i\big)^2\big(\Delta^{m+1}_{2k}X^j\big)^2\Big] \leq \sqrt{\mathbb{E}\Big[\big(\Delta^{m+1}_{2k-1}X^i\big)^4\Big]}\sqrt{\mathbb{E}\Big[\big(\Delta^{m+1}_{2k}X^j\big)^4\Big]} \leq C2^{-4mH_\varepsilon}.$$

The same can be done for the second term in (4.27). For the third term we use Isserlis' theorem to obtain

$$\begin{aligned}&\Big|\mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j\big]\Big|\\ &\quad= \Big|\mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big]\mathbb{E}\big[\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j\big]\\ &\quad+ \mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^i\big]\mathbb{E}\big[\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2k}X^j\big]\\ &\quad+ \mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k}X^j\big]\mathbb{E}\big[\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2k-1}X^i\big]\Big|\\ &\quad\leq \frac{3C}{2^{4mH_\varepsilon}},\end{aligned}$$

where the individual summands can be bound using Lemma 4.3.2 and Remark 4.3.3.

Now let $k > l$ and $i \bmod d \neq j \bmod d$. This implies that X^i and X^j are independent.

Using stationarity and independence we obtain

$$\begin{aligned}
&\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j - \Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big)\\
&\qquad\qquad\qquad\cdot\big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\big)\big]\\
&\qquad = \mathbb{E}\big[\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j\big]\\
&\qquad - \mathbb{E}\big[\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j\Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\big]\\
&\qquad - \mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j\big] \qquad (4.28)\\
&\qquad + \mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\big]\\
&\qquad = 2\mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2l}X^i\big]\mathbb{E}\big[\Delta^{m+1}_{2k}X^j\Delta^{m+1}_{2l}X^j\big]\\
&\qquad - \mathbb{E}\big[\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2l}X^i\big]\mathbb{E}\big[\Delta^{m+1}_{2k}X^j\Delta^{m+1}_{2l-1}X^j\big]\\
&\qquad - \mathbb{E}\big[\Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2l-1}X^i\big]\mathbb{E}\big[\Delta^{m+1}_{2k-1}X^j\Delta^{m+1}_{2l}X^j\big].
\end{aligned}$$

Thus, applying (4.19) from Lemma 4.3.2 yields

$$\begin{aligned}
&\Big|\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}X^i\Delta^{m+1}_{2k}X^j - \Delta^{m+1}_{2k}X^i\Delta^{m+1}_{2k-1}X^j\big)\big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\big)\big]\Big|\\
&\qquad \le \frac{C}{2^{4H_\varepsilon m}}\Big(|2k-2l|^{2H_\varepsilon-3} + |2k-2l-1|^{H_\varepsilon-3/2}|2k-2l+1|^{H_\varepsilon-3/2}\Big)\\
&\qquad \le C\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}.
\end{aligned}$$

This concludes the proof.

□

4.3.1 Control of first level paths

We continue to denote by $X = X^H$ the process consisting of a Mandelbrot-van Ness type fractional Brownian motion and its derivative. As above we also restrict ourselves to the case of $H \le \frac{1}{2}$. In the same way as at the beginning of Section 4.2.2 we denote by $X(m)$ the dyadic approximation of X, i.e.

$$X(m)_t = X_{t^m_{k-1}} + 2^m(t - t^m_{k-1})\Delta^m_k X, \qquad \text{for} \qquad t^m_{k-1} \le t < t^m_k.$$

We define the smooth rough path

$$\mathbf{X}(m)_{s,t} = (1, \mathbf{X}(m)^1_{s,t}, \mathbf{X}(m)^2_{s,t}),$$

where $\mathbf{X}(m)^i_{s,t}$ is the i-th iterated (pathwise) integral of $X(m)$ over the interval $[s,t]$.

In this section we show, further following the approach of [4], that the first level paths $\mathbf{X}(m)^1$ converge in a rough path sense to $\mathbf{X}^1 = (X_t - X_s)_{(s,t)\in\Delta}$. The idea is to bound the expected p-variation distance between $\mathbf{X}^1(m)$ and $\mathbf{X}^1 = (X_t - X_s)_{(s,t)\in\Delta}$ by their distance at dyadic points via Corollary 4.2.2. If that bound decreases fast enough in m we can apply a Borel-Cantelli argument to obtain almost sure convergence in p-variation.

The following Lemma is given as Proposition 1 in [28], but we need to adjust for the different (co-)variance of X compared to a Wiener process.

Lemma 4.3.5. *Let $p > \frac{1}{H}$. For any $\gamma > 0$, it holds that*

$$\sup_m \mathbb{E}\Big[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\big|^p\Big] < \infty.$$

Proof. If $n \le m$, we have that $\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} = \Delta^n_k X$ and thus

$$\sum_{k=1}^{2^n} \big|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\big|^p = \sum_{k=1}^{2^n} \big|\Delta^n_k X\big|^p.$$

Let $n > m$ and note that for fixed $l \in \{1,\dots,2^m\}$ we have

$$\#\big\{k \in \{1,\dots,2^n\} : [t^n_{k-1},t^n_k] \subseteq [t^m_{l-1},t^m_l]\big\} = 2^{n-m}.$$

Equation (4.12) yields

$$\sum_{k=1}^{2^n} \big|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\big|^p = \sum_{k=1}^{2^n}\sum_{l=1}^{2^m} \sum_{t^m_{l-1}\le t^n_{k-1}<t^n_k\le t^m_l} \big|2^{m-n}\Delta^m_l X\big|^p = (2^{m-n})^{p-1}\sum_{l=1}^{2^m}\big|\Delta^m_l X\big|^p.$$

So by splitting up the sum, we obtain

$$\begin{aligned}
&\mathbb{E}\Bigg[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\big|^p\Bigg] \\
&= \mathbb{E}\Bigg[\sum_{n=1}^{m} n^\gamma \sum_{k=1}^{2^n} \big|\Delta^n_k X\big|^p\Bigg] + \mathbb{E}\Bigg[\sum_{n=m+1}^{\infty} n^\gamma (2^{m-n})^{p-1}\sum_{l=1}^{2^m}\big|\Delta^m_l X\big|^p\Bigg] \\
&= \sum_{n=1}^{m} n^\gamma \sum_{k=1}^{2^n} \mathbb{E}\Big[\big|\Delta^n_k X\big|^p\Big] + \sum_{n=m+1}^{\infty} n^\gamma (2^{m-n})^{p-1}\sum_{l=1}^{2^m}\mathbb{E}\Big[\big|\Delta^m_l X\big|^p\Big] \\
&= C_p\Bigg(\sum_{n=1}^{m} n^\gamma 2^n \mathbb{E}\Big[\big|\Delta^n_1 X\big|^2\Big]^{p/2} + \sum_{n=m+1}^{\infty} n^\gamma 2^{mp-n(p-1)}\mathbb{E}\Big[\big|\Delta^m_1 X\big|^2\Big]^{p/2}\Bigg),
\end{aligned}$$

where we use that the increments are stationary and Gaussian. Let $\varepsilon \in (0, H - \frac{1}{p})$, so $H_\varepsilon = H - \varepsilon \in (\frac{1}{p}, H)$. Following Remark 4.3.3, it holds that

$$\mathbb{E}\Big[\big|\Delta^n_1 X\big|^2\Big] \le C 2^{-2H_\varepsilon n},$$

where C depends only on H and ε. Therefore, we have

$$\sum_{n=1}^{m} n^\gamma 2^n \mathbb{E}\Big[\big|\Delta^n_1 X\big|^2\Big]^{p/2} \le C\sum_{n=1}^{m} n^\gamma 2^{-n(pH_\varepsilon - 1)}$$

and

$$\begin{aligned}
\sum_{n=m+1}^{\infty} n^\gamma 2^{mp-n(p-1)}\mathbb{E}\Big[\big|\Delta^m_1 X\big|^2\Big]^{p/2} &\le C\sum_{n=m+1}^{\infty} n^\gamma 2^{-mp(H_\varepsilon-1)-n(p-1)} \\
&\le C\sum_{n=m+1}^{\infty} n^\gamma 2^{-n(pH_\varepsilon-1)}.
\end{aligned}$$

Together this yields

$$\mathbb{E}\Bigg[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \big|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\big|^p\Bigg] \le C\sum_{n=1}^{\infty} n^\gamma 2^{-n(pH_\varepsilon-1)}, \tag{4.29}$$

which is finite since $p > 1/H_\varepsilon$. As the right-hand side does not depend on m, the assertion follows. □

In the same way as Corollary 4 in [28] this results leads to the following corollary.

Corollary 4.3.6. *For any* $p > \frac{1}{H}$

$$\mathbb{E}\left[\sup_m \sup_D \sum_l \left|\mathbf{X}(m)^1_{t_{l-1},t_l}\right|^p\right] < \infty$$

and therefore

$$\sup_m \sup_D \sum_l \left|\mathbf{X}(m)^1_{t_{l-1},t_l}\right|^p < \infty \ a.s.$$

Proof. In the proof of Lemma 4.3.5 we saw for $m < n$ that

$$\sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p = \left(2^{m-n}\right)^{p-1} \sum_{l=1}^{2^m} \left|\Delta^m_l X\right|^p.$$

Further,

$$\begin{aligned}\left(2^{m-n}\right)^{p-1} \sum_{l=1}^{2^m} \left|\Delta^m_l X\right|^p &\le \left(2^{m-n}\right)^{p-1} \sum_{l=1}^{2^m} \left|\Delta^{m+1}_{2l} X + \Delta^{m+1}_{2l+1} X\right|^p \\ &\le \left(2^{m+1-n}\right)^{p-1} \sum_{l=1}^{2^{m+1}} \left|\Delta^{m+1}_l X\right|^p\end{aligned}$$

for $m \le n$ and

$$\sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p = \sum_{k=1}^{2^n} \left|\Delta^n_k X\right|^p$$

for any $m \ge n$. Thus, we have that

$$m \mapsto \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p$$

is increasing in m for any fixed $n \in \mathbb{N}$. Applying Lemma 4.2.1 and the monotone convergence theorem yields

$$\begin{aligned}\mathbb{E}\left[\sup_m \sup_D \sum_l \left|\mathbf{X}(m)^1_{t_{l-1},t_l}\right|^p\right] &\le C_1 \mathbb{E}\left[\sup_m \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\right] \\ &= C_1 \mathbb{E}\left[\lim_{m\to\infty} \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\right] \\ &= C_1 \lim_{m\to\infty} \mathbb{E}\left[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\right],\end{aligned}$$

with some $\gamma > p - 1$. The right-hand side of the equation above is finite by Lemma 4.3.5. □

Let us define $\mathbf{X}^1_{s,t} := X_{s,t} = X_t - X_s,\ (s,t) \in \Delta$, then $(1, \mathbf{X}^1)$ is a multiplicative functional in $T^{(1)}(\mathbb{R}^{2d})$. We can adapt Theorem 2 in [28] to our situation and obtain the following.

Proposition 4.3.7. *For any $p > \frac{1}{H}$, we have*

$$\lim_{m\to\infty} \sup_D \sum_l \left|\mathbf{X}(m)^1_{t_{l-1},t_l} - \mathbf{X}^1_{t_{l-1},t_l}\right|^p = 0$$

almost surely, where $\mathbf{X}^1_{t_{l-1},t_l} = X_{t_{l-1},t_l} = X_{t_l} - X_{t_{l-1}}$.

Proof. Let $\gamma > p-1$. As $\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} = \Delta^n_k X = X_{t^n_{k-1},t^n_k}$ for $n \le m$, we obtain

$$\begin{aligned}\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X_{t^n_{k-1},t^n_k}\right|^p &= \sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X_{t^n_{k-1},t^n_k}\right|^p \\ &\le 2^{p-1} \sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p + 2^{p-1} \sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\Delta^n_k X\right|^p.\end{aligned}$$

In the same way as in the proof of Lemma 4.3.5, we have

$$\mathbb{E}\Big[\sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\Big] \le C \sum_{n=m+1}^{\infty} n^\gamma 2^{-n(pH_\varepsilon - 1)}$$

and

$$\mathbb{E}\Big[\sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\Delta^n_k X\right|^p\Big] \le C \sum_{n=m+1}^{\infty} n^\gamma 2^{-n(pH_\varepsilon - 1)},$$

where H_ε, as a reminder, is given by $H_\varepsilon = H - \varepsilon$ for some $\varepsilon \in (0, H - \frac{1}{p})$. Setting $\alpha := (pH_\varepsilon - 1)$, there exists a C depending only on γ such that $n^\gamma \le C 2^{n\alpha/2},\ n \in \mathbb{N}$. Since

$$\sum_{n=m+1}^{\infty} 2^{-n\alpha/2} = \frac{2^{-m\alpha/2}}{2^{\alpha/2} - 1},$$

we obtain

$$\mathbb{E}\Big[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X_{t^n_{k-1},t^n_k}\right|^p\Big] \le C 2^{-m(pH_\varepsilon - 1)/2}, \tag{4.30}$$

where C depends only on p, γ, H and ε. So it follows, e.g. from Markov's inequality and the Borel-Cantelli lemma, that the term within the expectation converges almost surely to 0 as $m \to \infty$. By Corollary 4.2.2 we have

$$\sup_D \sum_l \left|\mathbf{X}(m)^1_{t_{l-1},t_l} - X_{t_{l-1},t_l}\right|^p \le C_1 \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X_{t^n_{k-1},t^n_k}\right|^p \xrightarrow{m\to\infty} 0,\ a.s.$$

□

4.3.2 Divergence of the second level paths

Let $H \in (\frac{1}{4}, \frac{1}{2}]$. In this section we consider a one-dimensional fractional Brownian motion $B = B^H$ and let $\partial_H B = \partial_H B^H$ denote its pathwise derivative in H. We will often omit writing the superscript H for better readability.

We define $Z = (B^H, \partial_H B^H)$ and denote by $Z(m)$ its dyadic approximation. A smooth rough path is then given by

$$\mathbf{Z}(m)_{s,t} = (1, \mathbf{Z}(m)^1_{s,t}, \mathbf{Z}(m)^2_{s,t}),$$

where $\mathbf{Z}(m)^k_{s,t}$ is the k-th iterated (pathwise) integral of $Z(m)$ over the interval $[s,t]$. We use the following notation for the elements of the second level paths

$$\mathbf{Z}(m)^2_{s,t} = \{\mathbf{Z}^{i,j}(m)^2_{s,t}\}_{i,j=1,2}, \qquad (s,t) \in \Delta.$$

To show that $\mathbf{Z}(m)$ does not converge in expected p-variation distance, we first need to prove two helping lemmata.

Lemma 4.3.8. *Let $m \in \mathbb{N}$, then*

$$\Big(\mathbb{E}\big[\Delta^m_1 B \Delta^m_2(\partial_H B)\big] - \mathbb{E}\big[\Delta^m_2 B \Delta^m_1(\partial_H B)\big]\Big)^2 \geq C_H^4 2^{-4mH},$$

where the C_H is the constant from the Mandelbrot-van Ness representation.

Proof. In what follows we write t_k for $t^m_k = k2^{-m}$. By Lemma 4.5.2, we have

$$\begin{aligned}&\Big((t_2-u)^{H-1/2}\log(t_2-u) - (t_1-u)^{H-1/2}\log(t_1-u)\Big)\Big((t_1-u)^{H-1/2} - (-u)^{H-1/2}\Big)\\ &\quad\geq \Big((t_2-u)^{H-1/2} - (t_1-u)^{H-1/2}\Big)\Big((t_1-u)^{H-1/2}\log(t_1-u) - (-u)^{H-1/2}\log(-u)\Big)\end{aligned}$$

for $u < 0$. Therefore, using the Mandelbrot-van Ness integral representation, we obtain

$$\begin{aligned}
&C_H^{-2}\Big(\mathbb{E}\big[\Delta^m_1 B \Delta^m_2(\partial_H B)\big] - \mathbb{E}\big[\Delta^m_2 B \Delta^m_1(\partial_H B)\big]\Big)\\
&= \int_{-\infty}^0 \Big((t_2-u)^{H-1/2}\log(t_2-u) - (t_1-u)^{H-1/2}\log(t_1-u)\Big)\Big((t_1-u)^{H-1/2} - (-u)^{H-1/2}\Big)\mathrm{d}u\\
&- \int_{-\infty}^0 \Big((t_2-u)^{H-1/2} - (t_1-u)^{H-1/2}\Big)\Big((t_1-u)^{H-1/2}\log(t_1-u) - (-u)^{H-1/2}\log(-u)\Big)\mathrm{d}u\\
&+ \int_0^{t_1} \Big((t_2-u)^{H-1/2}\log(t_2-u) - (t_1-u)^{H-1/2}\log(t_1-u)\Big)(t_1-u)^{H-1/2}\mathrm{d}u\\
&- \int_0^{t_1} \Big((t_2-u)^{H-1/2} - (t_1-u)^{H-1/2}\Big)(t_1-u)^{H-1/2}\log(t_1-u)\mathrm{d}u\\
&\geq \int_0^{t_1} \Big((t_2-u)^{H-1/2}\log(t_2-u) - (t_1-u)^{H-1/2}\log(t_1-u)\Big)(t_1-u)^{H-1/2}\mathrm{d}u\\
&- \int_0^{t_1} \Big((t_2-u)^{H-1/2} - (t_1-u)^{H-1/2}\Big)(t_1-u)^{H-1/2}\log(t_1-u)\mathrm{d}u\\
&= \int_0^{2^{-m}} \Big((2^{-m}+v)^{H-1/2}\log(2^{-m}+v) - v^{H-1/2}\log(v)\Big)v^{H-1/2}\\
&\qquad\qquad - \Big((2^{-m}+v)^{H-1/2} - v^{H-1/2}\Big)v^{H-1/2}\log(v)\mathrm{d}v
\end{aligned}$$

$$= \int_0^{2^{-m}} (2^{-m}+v)^{H-1/2} v^{H-1/2} \big(\log(2^{-m}+v) - \log(v)\big) \, \mathrm{d}v.$$

For $v \in (0, 2^{-m})$, we have

$$(v + 2^{-m})^{H-1/2} \geq (2^{-m} + 2^{-m})^{H-1/2} = 2^{H-1/2+m/2-mH}$$

and $\log(2^{-m} + v) \geq \log(2^{-m})$. Therefore, as the integrand is positive, we have

$$\begin{aligned}
& C_H^{-2} \Big| \mathbb{E}\big[\Delta_2^m B \Delta_1^{m+1}(\partial_H B)\big] - \mathbb{E}\big[\Delta_1^m B \Delta_2^m(\partial_H B)\big] \Big| \\
& \geq \int_0^{2^{-m}} (2^{-m} + v)^{H-1/2} v^{H-1/2} \big(\log(2^{-m} + v) - \log(v) \big) \, \mathrm{d}v \\
& \geq 2^{H-1/2+m/2-mH} \int_0^{2^{-m}} v^{H-1/2} \big(\log(2^{-m}) - \log(v) \big) \, \mathrm{d}v \\
& = 2^{H-1/2+m/2-mH} \bigg(-m \log(2) \int_0^{2^{-m}} v^{H-1/2} \, \mathrm{d}v - \int_0^{2^{-m}} v^{H-1/2} \log(v) \, \mathrm{d}v \bigg).
\end{aligned}$$

It holds that

$$\int_0^{2^{-m}} v^{H-1/2} \, \mathrm{d}v = \frac{2^{-m(H+1/2)}}{H + \frac{1}{2}}$$

and integration by parts yields

$$\int_0^{2^{-m}} v^{H-1/2} \log(v) \, \mathrm{d}v = -m \log(2) \frac{2^{-m(H+1/2)}}{H + \frac{1}{2}} - \frac{2^{-m(H+1/2)}}{\big(H + \frac{1}{2}\big)^2}.$$

Thus, we obtain

$$\Big(\mathbb{E}\big[\Delta_2^m B \Delta_1^{m+1}(\partial_H B)\big] - \mathbb{E}\big[\Delta_1^m B \Delta_2^m(\partial_H B)\big] \Big)^2 \geq C_H^4 \frac{2^{2H-1}}{\big(H + \frac{1}{2}\big)^4} \, 2^{-4mH}.$$

For

$$0 < H \leq \frac{1}{2} < \frac{4 - \log(2)}{\log(4)}$$

it holds that

$$\frac{\partial}{\partial H} \left(\frac{2^{2H-1}}{\big(H + \frac{1}{2}\big)^4} \right) = \frac{4^{H+2}}{(2H+1)^5} \big((2H+1) \log(2) - 4 \big) \leq 0,$$

and therefore

$$\frac{2^{2H-1}}{\big(H + \frac{1}{2}\big)^4} \geq \frac{2^{\frac{2}{2}-1}}{\big(\frac{1}{2} + \frac{1}{2}\big)^4} = 1.$$

□

These results can now be used to show the next lemma.

Lemma 4.3.9. *Let $k \neq l$ and $\varepsilon \in (0, \frac{1}{4})$. Then*

$$\begin{aligned}\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B) - \Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\big)&\\ \cdot\big(\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B) - \Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big)\big]&\\ \geq C_H^4 2^{-4(m+1)H} + O\bigg(\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}\bigg).&\end{aligned}$$

Proof. Let $k \neq l$. We have

$$\begin{aligned}\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B) - \Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\big)&\\ \cdot\big(\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B) - \Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big)\big]&\\ = \mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ - \mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big]&\\ - \mathbb{E}\big[\Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ + \mathbb{E}\big[\Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big].&\end{aligned} \tag{4.31}$$

We use Isserlis' theorem, also called Wick formula for Gaussian random vectors, to obtain

$$\begin{aligned}\mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ = \mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\big]\mathbb{E}\big[\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ + \mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2l-1}B\big]\mathbb{E}\big[\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ + \mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]\mathbb{E}\big[\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l-1}B\big].&\end{aligned}$$

Applying Lemma 4.3.2 and using stationarity yields

$$\begin{aligned}\mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]&\\ = \mathbb{E}\big[\Delta^{m+1}_{1}B\Delta^{m+1}_{2}(\partial_H B)\big]^2 + O\bigg(\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}\bigg).&\end{aligned}$$

In the same way, we obtain

$$\begin{aligned}&\mathbb{E}\big[\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B)\Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big]\\ &\quad= \mathbb{E}\big[\Delta^{m+1}_{1}B\Delta^{m+1}_{2}(\partial_H B)\big]\mathbb{E}\big[\Delta^{m+1}_{2}B\Delta^{m+1}_{1}(\partial_H B)\big] + O\bigg(\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}\bigg),\\ &\mathbb{E}\big[\Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B)\big]\\ &\quad= \mathbb{E}\big[\Delta^{m+1}_{2}B\Delta^{m+1}_{1}(\partial_H B)\big]\mathbb{E}\big[\Delta^{m+1}_{1}B\Delta^{m+1}_{2}(\partial_H B)\big] + O\bigg(\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}\bigg),\\ &\mathbb{E}\big[\Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big]\\ &\quad= \mathbb{E}\big[\Delta^{m+1}_{2}B\Delta^{m+1}_{1}(\partial_H B)\big]^2 + O\bigg(\frac{(k-l)^{2H_\varepsilon-3}}{2^{4H_\varepsilon m}}\bigg).\end{aligned}$$

Plugging everything back into (4.31) leads to

$$\begin{aligned}\mathbb{E}\big[\big(\Delta^{m+1}_{2k-1}B\Delta^{m+1}_{2k}(\partial_H B) - \Delta^{m+1}_{2k}B\Delta^{m+1}_{2k-1}(\partial_H B)\big)&\\ \cdot\big(\Delta^{m+1}_{2l-1}B\Delta^{m+1}_{2l}(\partial_H B) - \Delta^{m+1}_{2l}B\Delta^{m+1}_{2l-1}(\partial_H B)\big)\big]&\end{aligned}$$

$$= \mathbb{E}\big[\Delta_1^{m+1} B \Delta_2^{m+1}(\partial_H B)\big]^2 - 2\mathbb{E}\big[\Delta_2^{m+1} B \Delta_1^{m+1}(\partial_H B)\big]\mathbb{E}\big[\Delta_1^{m+1} B \Delta_2^{m+1}(\partial_H B)\big]$$
$$+ \mathbb{E}\big[\Delta_2^{m+1} B \Delta_1^{m+1}(\partial_H B)\big]^2 + O\Big(\frac{(k-l)^{2H_\varepsilon - 3}}{2^{4H_\varepsilon m}}\Big)$$
$$= \Big(\mathbb{E}\big[\Delta_1^{m+1} B \Delta_2^{m+1}(\partial_H B)\big] - \mathbb{E}\big[\Delta_2^{m+1} B \Delta_1^{m+1}(\partial_H B)\big]\Big)^2 + O\Big(\frac{(k-l)^{2H_\varepsilon - 3}}{2^{4H_\varepsilon m}}\Big).$$

The assertion now follows from Lemma 4.3.8. □

This finally enables us to prove the main theorem of this section.

Theorem 4.3.10. *Let $H \in (\frac{1}{4}, \frac{1}{2}]$ and $p \leq 4$. There is* ***no*** *stochastic rough paths $\mathbf{Z} \in G\Omega_p(\mathbb{R}^d)$ such that*

$$\mathbb{E}\big[d_p(\mathbf{Z}(m), \mathbf{Z})\big] \xrightarrow{m \to \infty} 0.$$

Proof. The iterated integral of fractional Brownian motion with respect to its derivative in H is found on the off-diagonal component of $\mathbf{Z}(m)^2$. We therefore only consider the component $(\mathbf{Z}^{1,2}(m)^2_{s,t})$. By Lemma 4.2.5 we have that

$$\mathbf{Z}^{1,2}(m+1)^2_{0,1} - \mathbf{Z}^{1,2}(m)^2_{0,1} = \frac{1}{2}\sum_{k=1}^{2^m}\big(\Delta_{2k-1}^{m+1} B \Delta_{2k}^{m+1}(\partial_H B) - \Delta_{2k}^{m+1} B \Delta_{2k-1}^{m+1}(\partial_H B)\big).$$

Let $\varepsilon \in (0, H - \frac{1}{4})$. We have

$$\Big|\sum_{k>l} C\frac{(k-l)^{2H_\varepsilon - 3}}{2^{4H_\varepsilon m}}\Big| \leq \frac{C}{2^{4H_\varepsilon m}}\sum_{k=2}^{2^m}\sum_{l=1}^{k-1}|k-l|^{2H_\varepsilon - 3} \leq C 2^{m - 4H_\varepsilon m}\sum_{l=1}^{\infty} l^{2H_\varepsilon - 3} = o(1). \quad (4.32)$$

This yields, using Lemma 4.3.9, that

$$\mathbb{E}\Big[\big(\mathbf{Z}^{1,2}(m+1)^2_{0,1} - \mathbf{Z}^{1,2}(m)^2_{0,1}\big)^2\Big]$$
$$= \frac{1}{4}\sum_{k=1}^{2^m}\sum_{l=1}^{2^m}\mathbb{E}\Big[\big(\Delta_{2k-1}^{m+1} B \Delta_{2k}^{m+1}(\partial_H B) - \Delta_{2k}^{m+1} B \Delta_{2k-1}^{m+1}(\partial_H B)\big)$$
$$\cdot \big(\Delta_{2l-1}^{m+1} B \Delta_{2l}^{m+1}(\partial_H B) - \Delta_{2l}^{m+1} B \Delta_{2l-1}^{m+1}(\partial_H B)\big)\Big]$$
$$\geq \frac{1}{2}\sum_{k>l}\mathbb{E}\Big[\big(\Delta_{2k-1}^{m+1} B \Delta_{2k}^{m+1}(\partial_H B) - \Delta_{2k}^{m+1} B \Delta_{2k-1}^{m+1}(\partial_H B)\big)$$
$$\cdot \big(\Delta_{2l-1}^{m+1} B \Delta_{2l}^{m+1}(\partial_H B) - \Delta_{2l}^{m+1} B \Delta_{2l-1}^{m+1}(\partial_H B)\big)\Big]$$
$$\geq \frac{C_H^4}{2} \cdot 2^{-4(m+1)H}\sum_{k=2}^{2^m}\sum_{l=1}^{k-1} 1 + o(1) \geq \frac{C_H^4}{8} \cdot 2^{-4mH}\sum_{k=1}^{2^m - 1} k + o(1)$$
$$= C_H^4 2^{-4mH}\frac{(2^m - 1)2^m}{16} + o(1) \geq \frac{C_H^4}{16}\big(1 - 2^{-m(4H-1)}\big) + o(1)$$
$$\xrightarrow{m \to \infty} \frac{C_H^4}{16} \neq 0.$$

It follows that there exists **no** stochastic rough paths $\mathbf{Z} \in G\Omega_p(\mathbb{R}^2)$ such that

$$\mathbb{E}\big[d_4(\mathbf{Z}(m), \mathbf{Z})\big] \xrightarrow{m \to \infty} 0$$

and therefore $\mathbf{Z}(m)$ does also not converge in expected p-variation distance for any $p \leq 4$. □

We like to give some intuitive understanding of this result in the case $H = \frac{1}{2}$. For the approximated iterated integral of the derivative in H with respect to the (fractional) Brownian motion itself, we have

$$\int_0^1 \big(\partial_H B_u(m) - \partial_H B_0(m)\big)\,\mathrm{d}B_u(m) = \frac{1}{2}\sum_{k=1}^{2^m} \big(\partial_H B_{t_k^m} + \partial_H B_{t_{k-1}^m}\big)\big(B_{t_k^m} - B_{t_{k-1}^m}\big), \quad (4.33)$$

where $B(m)$ is the m-th dyadic approximation of the (fractional) Brownian motion $B = B^{\frac{1}{2}}$. If $\partial_H B$ were an semi-martingale adapted to the natural filtration of $(B_t)_{t\in[0,1]}$, the right-hand side of (4.33) would converge in $L^2(\Omega)$ to the Stratonovich integral

$$\int_0^1 \partial_H B_u \circ \mathrm{d}B_u.$$

However, it is easily seen from the integral representation of $\partial_H B$ that $\partial_H B_u$ not only depends on $(B_t)_{t\in[0,u]}$ but also on $(B_t)_{t<0}$. Furthermore, the proof of Theorem 4.3.10 shows that the dyadic approximations of the second level paths do not converge in $L^2(\Omega)$ at all. Therefore, it seems that we fail to find a sensible definition for the symmetric integral that would appear as the limit of the terms in equation (4.33) and would extend the Stratonovich integral to this specific non-adapted integrand.

4.3.3 Control of some second level paths

The convergence of the first order paths was handled in Section 4.3.1 and we have seen in Section 4.3.2 that $\mathbf{X}(m)$ does not converge in expected p-variation distance. However, the complications of Section 4.3.2 only occur for elements $\mathbf{X}^{i,j}(m)^2$ with $i \bmod d = j \bmod d$, where, with the same notation as above, we denote by $\mathbf{X}^{i,j}(m)^2$ the elements of second level paths of $\mathbf{X}(m)^2$, i.e.

$$\mathbf{X}(m)^2_{s,t} = \{\mathbf{X}^{i,j}(m)^2_{s,t}\}_{i,j=1,\dots,2d}, \qquad (s,t)\in\Delta.$$

So, our aim is to show that, for fixed $i, j \in \{1,\dots,2d\}$ with $i \bmod d \neq j \bmod d$, there exists a unique function $\mathbb{X}^{i,j} : \Delta \to \mathbb{R}$, which is given as the limit of $\mathbf{X}^{i,j}(m)^2$ in $\frac{p}{2}$-variation.

The following is our adapted version of Proposition 17 in [4].

Proposition 4.3.11. *Let $H > \frac{1}{3}$, $\min\{2, \frac{1}{H}\} < p \le 4$ and $\varepsilon \in (0, H - \frac{1}{p})$. For $i, j \in \{1,\dots,2d\}$ with $i \bmod d \neq j \bmod d$, there exists a constant C depending only on d, p, H and ε such that*

(i) for $m < n$

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] \le C 2^{-np} 2^{mp(1-H_\varepsilon)},$$

(ii) for $m \ge n$

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] \le C (2^{m-n})^{p/4} 2^{-mpH_\varepsilon},$$

where, as above $H_\varepsilon = H - \varepsilon$. These inequalities also hold for $i = j$, where, for $m \ge n$, we even have

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,i}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,i}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] = 0.$$

Proof. For $m < n$ we have by (4.14)

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] \le C2^{p(m-n)}\Big(\mathbb{E}\Big[\big|\Delta^{m+1}_{l(m+1)}X\big|^p\Big] + \mathbb{E}\Big[\big|\Delta^m_{l(m)}X\big|^p\Big]\Big) \le C2^{-np}2^{mp(1-H_\varepsilon)},$$

where $l(m)$ and $l(m+1)$ are given by (4.13). Now let us consider the case $m \ge n$. By Lemma 4.2.5 we have to control

$$\Big|\sum_l \Delta^{m+1}_{2l-1}X^i \otimes \Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i \otimes \Delta^{m+1}_{2l-1}X^j\Big|^2 = \sum_l \Big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\Big)^2 + 2\sum_{r<l}\Big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\Big) \times \Big(\Delta^{m+1}_{2r-1}X^i\Delta^{m+1}_{2r}X^j - \Delta^{m+1}_{2r}X^i\Delta^{m+1}_{2r-1}X^j\Big),$$

where r, l run from $2^{m-n}(k-1)+1$ to $2^{m-n}k$. Using Lemma 4.3.4, we have

$$A^{r,l}_{i,j} := \mathbb{E}\Big[\Big(\Delta^{m+1}_{2l-1}X^i\Delta^{m+1}_{2l}X^j - \Delta^{m+1}_{2l}X^i\Delta^{m+1}_{2l-1}X^j\Big) \times \Big(\Delta^{m+1}_{2r-1}X^i\Delta^{m+1}_{2r}X^j - \Delta^{m+1}_{2r}X^i\Delta^{m+1}_{2r-1}X^j\Big)\Big] \le C2^{-4mH_\varepsilon}(l-r)^{2H_\varepsilon-3}.$$

Also using the first part of Lemma 4.3.4 leads to

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)_{t^n_{k-1},t^n_k}\big|^2\Big] \le C2^{m-n}2^{-4mH_\varepsilon} + C2^{-4mH_\varepsilon}\sum_{l=2}^{2^{m-n}}\sum_{r=1}^{l-1}\frac{1}{(l-r)^{3-2H_\varepsilon}} \le C2^{m-n}2^{-4mH_\varepsilon}$$

because

$$\sum_{l=2}^{2^{m-n}}\sum_{r=1}^{l-1}\frac{1}{(l-r)^{3-2H_\varepsilon}} \le 2^{m-n}\sum_{r=1}^{2^{m-n}}\frac{1}{r^{3-2H_\varepsilon}} \le 2^{m-n}\sum_{r=1}^{\infty}\frac{1}{r^{3-2H_\varepsilon}} \le C2^{m-n}.$$

By Hölder's inequality this yields

$$\mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] \le \mathbb{E}\Big[\big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^2\Big]^{p/4} \le C(2^{m-n})^{p/4}2^{-mpH_\varepsilon}$$

for any $p \in (\min\{2, \frac{1}{H}\}, 4]$.

□

This proposition allows us to finally prove convergence of the aforementioned components of the second level path $\mathbf{X}(m)^2$.

Theorem 4.3.12. *Let $H > \frac{1}{3}$ and $2 < p \leq 4$ with $p > \frac{1}{H}$. Then, for $i, j \in \{1, \ldots, 2d\}$ with $i \bmod d \neq j \bmod d$, there exists a unique function $\mathbb{X}^{i,j}$ on Δ such that*

$$\lim_{m\to\infty} \sup_D \sum_l \left|\mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l} - \mathbb{X}^{i,j}_{t_{l-1},t_l}\right|^{p/2} = 0.$$

The result also holds for $i = j$.

Proof. Fix $i, j \in \{1, \ldots, 2d\}$ with $i \bmod d \neq j \bmod d$ (or $i = j$). Note that the placeholder constant C might only depend on d, H, p and on the variables ε, ν, γ introduced later in the proof. Let $\gamma > \frac{p}{2} - 1$. By Remark 4.2.4 we have

$$\begin{aligned}
&\sup_D \sum_l \left|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\right|^{p/2} \\
&\quad\leq C \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\right|^{p/2} \\
&\quad+ C\Bigg(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m+1)^1_{t^n_{k-1},t^n_k} - \mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\Bigg)^{1/2} \\
&\qquad\qquad \times \Bigg(\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m+1)^1_{t^n_{k-1},t^n_k}\right|^p + \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\Bigg)^{1/2}.
\end{aligned}$$

Fix an $\varepsilon \in (0, H - \frac{1}{p})$ and set $H_\varepsilon := H - \varepsilon \in (\frac{1}{p}, H)$. Following Lemma 4.3.5 we obtain

$$\mathbb{E}\Bigg[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p\Bigg] < C$$

and

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m+1)^1_{t^n_{k-1},t^n_k} - \mathbf{X}(m)^1_{t^n_{k-1},t^n_k}\right|^p \\
&\quad\leq \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m+1)^1_{t^n_{k-1},t^n_k} - X^1_{t^n_{k-1},t^n_k}\right|^p + \sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X^1_{t^n_{k-1},t^n_k}\right|^p,
\end{aligned}$$

where by (4.30) we have that

$$\mathbb{E}\Bigg[\sum_{n=1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}(m)^1_{t^n_{k-1},t^n_k} - X^1_{t^n_{k-1},t^n_k}\right|^p\Bigg] \leq C 2^{-m(pH_\varepsilon - 1)/2}.$$

It remains to consider the first summand which we will split into two parts. Fix an $\nu \in (1, pH_\varepsilon)$, then there exists a constant C such that $n^\gamma \leq C 2^{n(\nu-1)}$. By using (i) in Proposition 4.3.11, we have

$$\mathbb{E}\Bigg[\sum_{n=m+1}^{\infty} n^\gamma \sum_{k=1}^{2^n} \left|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\right|^{p/2}\Bigg]$$

$$\leq \sum_{n=m+1}^{\infty} n^{\gamma} \sum_{k=1}^{2^n} C2^{-np}2^{mp(1-H_\varepsilon)} = C2^{mp(1-H_\varepsilon)} \sum_{n=m+1}^{\infty} n^{\gamma}2^{-n(p-1)}$$
$$\leq C2^{mp(1-H_\varepsilon)} \sum_{n=m+1}^{\infty} 2^{-n(p-\nu)} \leq C2^{mp(1-H_\varepsilon)}2^{-m(p-\nu)}$$
$$\leq C2^{-m(pH_\varepsilon-\nu)}.$$

Further, applying (ii) in Proposition 4.3.11, yields

$$\begin{aligned}\mathbb{E}\Big[\sum_{n=1}^{m} n^{\gamma} \sum_{k=1}^{2^n} \big|\mathbf{X}^{i,j}(m+1)^2_{t^n_{k-1},t^n_k} - \mathbf{X}^{i,j}(m)^2_{t^n_{k-1},t^n_k}\big|^{p/2}\Big] \\ \leq \sum_{n=1}^{m} n^{\gamma} \sum_{k=1}^{2^n} (2^{m-n})^{p/4}2^{-mpH_\varepsilon} \leq C2^{-mp(H_\varepsilon-1/4)} \sum_{n=1}^{m} n^{\gamma}2^{n(1-p/4)} \\ \leq C2^{-mp(H_\varepsilon-1/4)} \sum_{n=1}^{m} 2^{n(\nu-p/4)} \leq C2^{-mp(H_\varepsilon-1/4)}2^{m(\nu-p/4)} \\ \leq C2^{-m(pH_\varepsilon-\nu)}.\end{aligned}$$

Putting everything together yields

$$\begin{aligned}\mathbb{E}\Big[\sup_D \sum_l \big|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\big|^{p/2}\Big] &\leq C\Big(2^{-m(pH_\varepsilon-\nu)} + 2^{-m(pH_\varepsilon-1)/4}\Big) \\ &\leq C2^{-m(pH_\varepsilon-\nu)/4},\end{aligned}$$

and in particular we obtain

$$\mathbb{E}\Big[\sum_{m=1}^{\infty} \Big(\sup_D \sum_l \big|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\big|^{p/2}\Big)^{2/p}\Big] < \infty.$$

Therefore,

$$\sum_{m=1}^{\infty} \Big(\sup_D \sum_l \big|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\big|^{p/2}\Big)^{2/p} < \infty$$

almost surely. This yields, for $k > r$ that

$$\begin{aligned}\|\mathbf{X}^{i,j}(k)^2 - \mathbf{X}^{i,j}(r)^2\|_{p/2} &\leq \sum_{m=r}^{k-1} \Big(\sup_D \sum_l \big|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\big|^{p/2}\Big)^{2/p} \\ &\leq \sum_{m=r}^{\infty} \Big(\sup_D \sum_l \big|\mathbf{X}^{i,j}(m+1)^2_{t_{l-1},t_l} - \mathbf{X}^{i,j}(m)^2_{t_{l-1},t_l}\big|^{p/2}\Big)^{2/p} \\ &\xrightarrow{k,r\to\infty} 0,\end{aligned}$$

where

$$\|\mathbb{Y}\|_{p/2} = \Big(\sup_D \sum_l |\mathbb{Y}_{t_{l-1},t_l}|^{p/2}\Big)^{2/p}$$

denotes the $\frac{p}{2}$-variation norm. So, $(\mathbf{X}^{i,j}(m)^2)_{m\in\mathbb{N}}$ is almost surely a Cauchy sequence in the $\frac{p}{2}$-variation norm. Since

$$\{x : \Delta \to \mathbb{R} : x \text{ continuous}, \|x\|_{p/2} < \infty \text{ and } x_{0,0} = 0\}$$

endowed with the $\frac{p}{2}$-variation norm is a Banach space, we have that $(\mathbf{X}^{i,j}(m)^2)_{m\in\mathbb{N}}$ is convergent and denote its limit by $\mathbb{X}^{i,j}$. □

Let the Hurst parameter $H \in (\frac{1}{3}, \frac{1}{2}]$ and $\frac{1}{H} < p \leq 4$. Proposition 4.3.7 and Theorem 4.3.12 imply that there exist multiplicative functionals $\mathbf{B} = (1, \mathbf{B}^1, \mathbf{B}^2)$ and $\mathbf{D} = (1, \mathbf{D}^1, \mathbf{D}^2)$ in $T^{(2)}(\mathbb{R}^d)$ such that

$$\lim_{m\to\infty} d_p(\mathbf{B}(m), \mathbf{B}) = 0, \qquad \lim_{m\to\infty} d_p(\mathbf{D}(m), \mathbf{D}) = 0, \qquad p > \frac{1}{H},$$

where $\mathbf{B}(m)$ and $\mathbf{D}(m)$ are the dyadic rough path approximations of B^H and $D = \partial_H B^H$, respectively. Thus, we have $\mathbf{B}, \mathbf{D} \in G\Omega_p(\mathbb{R}^d)$ for $p \in (\frac{1}{H}, 3)$. However, Section 4.3.2 has shown that fractional Brownian motion together with its derivative in H can not be lifted as a joint rough paths in the same way. Nevertheless, the process might still be lifted to a (potentially non-geometric) rough path using other methods, like for example the ones used in [42].

We can conclude that the most natural way of lifting a multidimensional stochastic process to a rough path might fail due to some non-trivial interdependence of its components.

4.4 Dependence of rough SDEs on the Hurst parameter

Lifting fractional Brownian motion to a rough path enables us to analyse the dependence of the solution to a rough SDE driven by fBm on the Hurst parameter. The following definition can e.g. be in found [11].

Definition 4.4.1. Let A, B be two Banach spaces and $\gamma > 0$. A map $F : U \to V$ is γ-Lipschitz if F is $\lfloor\gamma\rfloor$-times continuously (Fréchet-)differentiable and there exists a constant $M > 0$ such that the supremum norm of its k-th derivative, $k = 0, \dots, \lfloor\gamma\rfloor$ is bounded by M as well as the $(\gamma - \lfloor\gamma\rfloor)$-Hölder norm of the $\lfloor\gamma\rfloor$-th derivative. The smallest such constant is denoted by $|F|_{\mathrm{Lip}^\gamma}$. The space of all γ-Lipschitz functions $F : U \to V$ is denoted by $\mathrm{Lip}^\gamma(U, V)$ or simply $\mathrm{Lip}^\gamma(U)$ if $U = V$.

Let $H \in (\frac{1}{3}, \frac{1}{2}]$, $\frac{1}{H} < p < \gamma < 3$ and $f \in \mathrm{Lip}^\gamma(\mathbb{R}^m, \mathrm{L}(\mathbb{R}^d, \mathbb{R}^m))$. We consider the rough SDE

$$\mathrm{d}Y_t^H = f(Y_t^H)\,\mathrm{d}B_t^H, \qquad Y_0^H = y_0, \tag{4.34}$$

where $y_0 \in \mathbb{R}^m$. This equation has a unique global solution (see e.g. Theorem 5.3 in [32]) but we will restrict ourselves to $t \in [0, 1]$. Note that we sometimes use the same notation for the process and the lifted rough path however the meaning should always be clear from context.

4.4.1 Local Lipschitz continuity

Let $p \in (2,3)$, $H \in (\frac{1}{p}, \frac{1}{2}]$ and $\mathbf{B}^H$ be the fractional Brownian motion from Definition 2.3.3 lifted to a geometric rough path.

Lemma 4.4.2. *Let $p \in (2,3)$, $\alpha \in (\frac{1}{p}, \frac{1}{2}]$ and $\mathbf{B}^H$ be the lifted geometric rough path of a d-dimensional fractional Brownian motion $(B^H)_{t\in[0,1]}$. There exists a positive constant $C = C(\omega)$, $\omega \in \Omega$ such that*

$$d_p(\mathbf{B}^H, \mathbf{B}^h) \leq C|H-h| \qquad H, h \in (\alpha, 1/2],$$

almost surley i.e.

$$(1/p, 1/2] \to G\Omega_p(\mathbb{R}^d) : H \mapsto \mathbf{B}^H$$

is almost surely locally Lipschitz continuous p-variation distance.

Proof. The following calculations hold almost surely and finite value of the positive constant $C = C(\omega)$ might change from line to line. Let $H, h \in (\alpha, 1/2]$, where, without loss of generality, $h \leq H$. By Theorem 2.1.1, we have

$$\sup_{\delta\in[h,H]} \|\partial_H B^\delta\|_{p\text{-var}} \leq C.$$

It further holds

$$\|B^H - B^h\|_{p\text{-var}} \leq |H-h| \sup_{\delta\in[h,H]} \|\partial_H B^\delta\|_{p\text{-var}} \leq C|H-h|.$$

Let $\mathbb{B}^H$ denote the second level path of $\mathbf{B}^H$ and $\mathbb{B}^{H,n}$ its piecewise linear dyadic approximation (equivalently for h). Applying the Extension Theorem (cf. [32], Theorem 3.7), we obtain

$$\begin{aligned}\|\mathbb{B}^H - \mathbb{B}^h\|_{\frac{p}{2}-\text{var}} &= \lim_{n\to\infty} \|\mathbb{B}^{H,n} - \mathbb{B}^{h,n}\|_{\frac{p}{2}-\text{var}} \leq C \lim_{n\to\infty} \|B^{H,n} - B^{h,n}\|_{p\text{-var}} \\ &\leq |H-h| \sup_{\delta\in[h,H]} \|\partial_H B^\delta\|_{p\text{-var}} \leq C|H-h|\end{aligned}$$

and the assertion follows. □

Let I_f be the solution or Ito-Lyons map that maps the driving signal to the solution of the rough differential equation, i.e.

$$I_f : G\Omega_p(\mathbb{R}^d) \to G\Omega_p(\mathbb{R}^m) : x \to z,$$

where z is the RDE solution to

$$\mathrm{d}z_t = f(z_t)\,\mathrm{d}x_t, \qquad z_0 = y_0.$$

Since the Ito-Lyons map is locally Lipschitz continuous (compare e.g. [11], Corollary 10.39) we obtain that

$$(1/p, 1/2] \to G\Omega_p(\mathbb{R}^m) : H \mapsto Y^H = I_f(\mathbf{B}^H),$$

where Y^H is the solution to (4.34) and $G\Omega_p(\mathbb{R}^m)$ is endowed with the p-variation distance, is almost surely locally Lipschitz continuous as the composition of two locally Lipschitz continuous functions.

As $\mathbf{B}^H$ is a stochastic process, one might also consider local Lipschitzness in a more probabilistic sense. Therefore, we present a second approach by which we obtain local Lipschitzness of the fractional Brownian rough path in *expected* p-variation. Before we present this result, we introduce some notation that is needed to prove it.

Let $(X_t)_{t\in[0,1]}$ be a one-dimensional stochastic process and $s,t,u,v \in [0,1]$, where $s<t$ and $u<v$. With the notation above, i.e. $X_{s,t} = X_t - X_s$, we define

$$R_X\begin{pmatrix} s,t \\ u,v \end{pmatrix} := \mathbb{E}[X_{s,t}X_{u,v}]$$

and its two-dimensional q-variation on $[s,t]^2 \subseteq [0,1]^2$ as

$$\|R_X\|_{q\text{-var},[s,t]^2} := \left(\sup_{D_1,D_2}\sum_{k,l}\left|R_X\begin{pmatrix} t_{k-1},t_k \\ t_{l-1},t_l \end{pmatrix}\right|^q\right)^{1/q}, \qquad q\geq 1,$$

where the supremum runs over all subdivisions $D_1=\{t_k\}, D_2=\{t_l\}$ of $[s,t]$.

Proposition 4.4.3. *Let $p\in(2,3)$, $\alpha\in(\frac{1}{p},\frac{1}{2}]$ and $\mathbf{B}^H$ be the lifted geometric rough path of a d-dimensional fractional Brownian motion $(B^H)_{t\in[0,1]}$. We have that*

$$\mathbb{E}\big[d_p(\mathbf{B}^H,\mathbf{B}^h)\big] \leq C|H-h| \qquad H,h\in(\alpha,1/2],$$

i.e.

$$(1/p,1/2] \to G\Omega_p(\mathbb{R}^d) : H \mapsto \mathbf{B}^H$$

is locally Lipschitz continuous in expected p-variation distance.

Proof. Let $H,h>\alpha>\frac{1}{p}$ and without loss of generality we can assume $H>h$. Furthermore, let $s,t,u,v\in[0,1]$, where $s<t$ and $u<v$. We denote by $B^{H,(i)}$ the i-th component of the d-dimensional fBm. We have

$$\begin{aligned} R_{B^{H,(i)}}\begin{pmatrix} s,t \\ u,v \end{pmatrix} &= \mathbb{E}\big[B^{H,(i)}_{s,t}B^{H,(i)}_{u,v}\big] \leq \sqrt{\mathbb{E}\Big[\big(B^{H,(i)}_{s,t}\big)^2\Big]\mathbb{E}\Big[\big(B^{H,(i)}_{u,v}\big)^2\Big]} = (t-s)^H(v-u)^H \\ &\leq (t-s)^\alpha(v-u)^\alpha. \end{aligned}$$

Using the calculations in the proof of Lemma 2.3.2, in particular (2.18), we have that there exists a constant C depending only on α such that

$$\begin{aligned} R_{B^{H,(i)}-B^{h,(i)}}\begin{pmatrix} s,t \\ u,v \end{pmatrix} &= \mathbb{E}\big[(B^{H,(i)}_{s,t}-B^{h,(i)}_{s,t})(B^{H,(i)}_{u,v}-B^{h,(i)}_{u,v})\big] \\ &\leq (H-h)^2\sqrt{\mathbb{E}\Big[\sup_{\theta\in[h,H]}\big(\partial_H B^{\theta,(i)}_{s,t}\big)^2\Big]\mathbb{E}\Big[\sup_{\theta\in[h,H]}\big(\partial_H B^{\theta,(i)}_{u,v}\big)^2\Big]} \\ &\leq C(H-h)^2(t-s)^\alpha(v-u)^\alpha. \end{aligned}$$

Further, we have

$$\|R_{B^{H,(i)}}\|_{\frac{1}{2\alpha}\text{-var},[s,t]^2} \le \|R_{B^{H,(i)}}\|_{\frac{1}{\alpha}\text{-var},[s,t]^2} \le \left(\sup_{D_1,D_2}\sum_{k,l}|t_k-t_{k-1}|\,|t_l-t_{l-1}|\right)^{\alpha}$$
$$\le \left(\sup_{D_1}\sum_{k}|t_k-t_{k-1}|\right)^{2\alpha} \le |t-s|^{2\alpha},$$

where the supremum runs over all subdivisions $D_1=\{t_k\},\ D_2=\{t_l\}$ of $[s,t]$. In the same way we obtain

$$\|R_{B^{h,(i)}}\|_{\frac{1}{2\alpha}\text{-var},[s,t]^2} \le |t-s|^{2\alpha},$$
$$\|R_{B^{H,(i)}-B^{h,(i)}}\|_{\frac{1}{2\alpha}\text{-var},[s,t]^2} \le C(H-h)^2|t-s|^{2\alpha}.$$

As the p-variation on $[0,1]$ is bounded by the $\frac{1}{p}$-Hölder semi-norm, Theorem 10.5 in [10] yields

$$\mathbb{E}\big[d_p(\mathbf{B}^H,\mathbf{B}^h)\big] \le C|H-h|.$$

□

To infer local Lipschitz continuity of the RDE solution in expected p-variation, we would need to prove an integrability condition on the constant, which would depend on $\omega\in\Omega$, appearing in the local Lipschitz condition of the Ito-Lyons map.

4.4.2 A candidate for a derivative

Consider the rough SDE

$$\mathrm{d}Y_t^H = f(Y_t^H)\,\mathrm{d}B_t^H,\quad Y_0^H = y_0,$$

with a suitable function f. We have seen that Y^H is locally Lipschitz in H. If Y^H is differentiable in H, we would expect its derivative $Z^H=\partial_H Y^H$ to satisfy a rough SDE of the form

$$\mathrm{d}Z_t^H = f'(Y_t^H)Z_t^H\,\mathrm{d}B_t^H + f(Y_t^H)\,\mathrm{d}\big(\partial_H B_t^H\big),\qquad Z_0^H=0. \tag{4.35}$$

However, this rough SDE relies on the existence of a *joint* rough path over B^H and its derivative in H. Section 4.3.2 has shown that, with the method used here, we were unable to construct such a rough path. Nevertheless, we are able to find a derivative in the rough path sense in one very specific setting.

Let $(\xi_t,\eta_t)\in G\Omega_p(\mathbb{R}^{2d})$. Then, $(y_t)_{t\in[0,1]}$ is well-defined by setting

$$y_t=\int_0^t \xi_s\,\mathrm{d}\eta_s,$$

where the integral is understood in the rough path sense. Writing

$$\mathrm{d}\begin{pmatrix}x_t\\ y_t\end{pmatrix} = \begin{pmatrix}\mathrm{d}\xi_t\\ \xi_t\,\mathrm{d}\eta_t\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}\mathrm{d}\xi_t + \begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}x_t\\ y_t\end{pmatrix}\mathrm{d}\eta_t,$$

this is a direct consequence of Theorem 10.52 in [11].

Therefore, letting $H > \frac{1}{3}$ and $p \in (\frac{1}{H}, 3)$, the rough integral equation

$$\begin{aligned} X_t^H &= W_t^H, \\ Y_t^H &= \int_0^t W_s^H \, \mathrm{d}B_s^H, \end{aligned} \tag{4.36}$$

where B^H and W^H are two independent fractional Brownian motions of the type defined in Definition 2.3.3, has a unique global solution. Furthermore, we have that

$$Z_t^H = \int_0^t \partial_H W_s^H \, \mathrm{d}B_s^H + \int_0^t W_s^H \, \mathrm{d}(\partial_H B_s^H) \tag{4.37}$$

exists globally. Let $B^{H,n}, W^{H,n}$ be the sequences of dyadic or smooth approximations, i.e. 1-rough paths converging to (B^H, W^H) in d_p distance. We have that Z^H is the p-variation limit of the (pathwise) ODE solution

$$Z_t^{H,n} = \int_0^t \partial_H W_s^{H,n} \, \mathrm{d}B_s^{H,n} + \int_0^t W_s^{H,n} \, \mathrm{d}(\partial_H B_s^{H,n}) = \partial_H Y_t^{H,n},$$

despite $(B^{H,n}, W^{H,n}, \partial_H B^{H,n}, \partial_H W^{H,n})$ not converging to a joint rough path in expected d_p distance. The relationship between Z^H and Y^H as well as their approximations is displayed in the figure below.

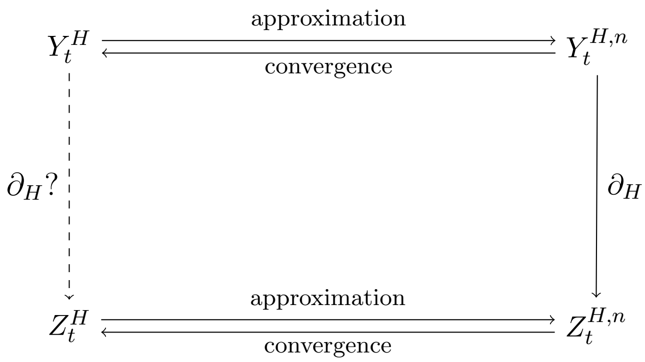

Figure 4.1: This diagram displays the relations between Y^H defined in (4.36), Z^H defined in (4.37) and their smooth approximations for a fixed time point $t \in [0, 1]$.

To sum it up, in section 4.4.1 we obtained that Y^H is locally Lipschitz in H. Moreover, we have that $Z^{H,n}$ is the derivative in H of $Y^{H,n}$ for all $n \in \mathbb{N}$. Further, $Z^{H,n}$ converges in the rough path sense to a limiting process Z^H. Therefore, this limit might be called a derivative in H of Y^H in an approximating rough path sense. Note that this does not imply that the first level of the rough path Y_t^H is differentiable in H for $t \in [0, 1]$ with derivative Z_t^H.

4.5 Auxiliary results

In this section we present the proofs to two technical lemmas that were used in deriving the results above.

Lemma 4.5.1. *Let $T > 0$, $t \in [0, T]$ and $\varepsilon > 0$. There exists a constant $C > 0$ depending only on T, ε and H such that the following points hold.*

(i) Suppose $\varepsilon \neq H - \frac{1}{2}$ and let $x \in (0, T]$, then we have

$$\begin{aligned} |x^{H-1/2}| &\leq C x^{H-1/2-\varepsilon}, \\ |x^{H-1/2}\log(x)| &\leq C x^{H-1/2-\varepsilon}, \\ |(t+x)^{H-1/2} - x^{H-1/2}| &\leq C\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big|, \\ |(t+x)^{H-1/2}\log(t+x) - x^{H-1/2}\log(x)| &\leq C\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big|. \end{aligned}$$

(ii) Suppose $\varepsilon \neq \frac{1}{2} - H$ and let $x \in [1, \infty)$, then we have

$$\begin{aligned} |x^{H-1/2}| &\leq x^{H-1/2+\varepsilon}, \\ |x^{H-1/2}\log(x)| &\leq C x^{H-1/2+\varepsilon}, \\ |(t+x)^{H-1/2} - x^{H-1/2}| &\leq C\big|(t+x)^{H-1/2+\varepsilon} - x^{H-1/2+\varepsilon}\big|, \\ |(t+x)^{H-1/2}\log(t+x) - x^{H-1/2}\log(x)| &\leq C\big|(t+x)^{H-1/2+\varepsilon} - x^{H-1/2+\varepsilon}\big|. \end{aligned}$$

Proof.

(i) We first note that for $y \in (0, T]$ there exists a positive constant C depending only on ε and T such that

$$1 \leq C y^{-\varepsilon}, \qquad |\log(y)| \leq C y^{-\varepsilon}.$$

Thus, the first two inequalities follow. Further, we have

$$\begin{aligned} |(t+x)^{H-1/2} - x^{H-1/2}| &= \left|H - \frac{1}{2}\right| \int_x^{t+x} y^{H-3/2}\,\mathrm{d}y \leq \left|H - \frac{1}{2}\right| C \int_x^{t+x} y^{H-3/2-\varepsilon}\,\mathrm{d}y \\ &= \frac{|H-1/2|}{|H-1/2-\varepsilon|} C\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big| \\ &\leq C\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big| \end{aligned}$$

and

$$\begin{aligned} \big|(t+x)^{H-1/2}\log(x+t) - x^{H-1/2}\log(x)\big| &\leq \int_x^{t+x} y^{H-3/2}\big[|H-1/2||\log(y)| + 1\big]\,\mathrm{d}y \\ &\leq C\left(|H-1/2| \int_x^{t+x} y^{H-3/2-\varepsilon}\,\mathrm{d}y + \int_x^{x+t} y^{H-3/2-\varepsilon}\,\mathrm{d}y\right) \\ &\leq C \int_x^{t+x} y^{H-3/2-\varepsilon}\,\mathrm{d}y = \frac{C}{|H-1/2-\varepsilon|}\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big| \\ &\leq C\big|(t+x)^{H-1/2-\varepsilon} - x^{H-1/2-\varepsilon}\big|. \end{aligned}$$

(ii) We note that for $y \in [1, \infty)$ there exists a constant $C > 0$ depending only on ε such that

$$1 \leq y^{\varepsilon}, \qquad |\log(y)| \leq C y^{\varepsilon}.$$

Therefore, the first two inequalities follow. Moreover, we have

$$\begin{aligned}|(t+x)^{H-1/2}-x^{H-1/2}| &= \left|H-\frac{1}{2}\right|\int_x^{t+x} y^{H-3/2}\,\mathrm{d}y \leq \left|H-\frac{1}{2}\right|\int_x^{t+x} y^{H-3/2+\varepsilon}\,\mathrm{d}y\\ &\leq \frac{|H-1/2|}{|H-1/2+\varepsilon|}\left|(t+x)^{H-1/2+\varepsilon}-x^{H-1/2+\varepsilon}\right|\end{aligned}$$

and

$$\begin{aligned}\left|(t+x)^{H-1/2}\log(x+t)-x^{H-1/2}\log(x)\right| &\leq \int_x^{t+x} y^{H-3/2}\big[|H-1/2||\log(y)|+1\big]\,\mathrm{d}y\\ &\leq C\int_x^{x+t} y^{H-3/2+\varepsilon}\,\mathrm{d}y \leq \frac{C}{|H-1/2+\varepsilon|}\left|(t+x)^{H-1/2+\varepsilon}-x^{H-1/2+\varepsilon}\right|\\ &\leq C\left|(t+x)^{H-1/2+\varepsilon}-x^{H-1/2+\varepsilon}\right|.\end{aligned}$$

□

Lemma 4.5.2. *Let $b>a>0,\ u>0$ and $\alpha\in\mathbb{R}$. Then we have*

$$\begin{aligned}&\Big((b+u)^\alpha-(a+u)^\alpha\Big)\Big((a+u)^\alpha\log(a+u)-u^\alpha\log(u)\Big)\\ &\qquad\leq \Big((b+u)^\alpha\log(b+u)-(a+u)^\alpha\log(a+u)\Big)\Big((a+u)^\alpha-u^\alpha\Big).\end{aligned}$$

Proof. We first make a preliminary observation. Let $c_1<c_2<c_3$ be numbers in $\mathbb{R}$, f,g functions mapping from $I\supseteq[c_1,c_3]$ to $\mathbb{R}$. Further, we assume $f(x)\geq 0,\ x\in[c_1,c_3]$ and g is monotonically increasing on $[c_1,c_3]$. Then, we have

$$\begin{aligned}\Big(\int_{c_1}^{c_2} f(x)g(x)\,\mathrm{d}x\Big)\Big(\int_{c_2}^{c_3} f(x)\,\mathrm{d}x\Big) &\leq g(c_2)\Big(\int_{c_1}^{c_2} f(x)\,\mathrm{d}x\Big)\Big(\int_{c_2}^{c_3} f(x)\,\mathrm{d}x\Big)\\ &\leq \Big(\int_{c_1}^{c_2} f(x)\,\mathrm{d}x\Big)\Big(\int_{c_2}^{c_3} f(x)g(x)\,\mathrm{d}x\Big).\end{aligned}$$

Applying this result yields

$$\begin{aligned}&\Big((b+u)^\alpha-(a+u)^\alpha\Big)\Big((a+u)^\alpha\log(a+u)-u^\alpha\log(u)\Big)\\ &= \Big(\alpha\int_{a+u}^{b+u} x^{\alpha-1}\,\mathrm{d}x\Big)\Big(\int_u^{a+u} x^{\alpha-1}\big(\alpha\log(x)+1\big)\,\mathrm{d}x\Big)\\ &= \alpha^2\Big(\int_{a+u}^{b+u} x^{\alpha-1}\,\mathrm{d}x\Big)\Big(\int_u^{a+u} x^{\alpha-1}\log(x)\,\mathrm{d}x\Big)+\alpha\Big(\int_{a+u}^{b+u} x^{\alpha-1}\,\mathrm{d}x\Big)\Big(\int_u^{a+u} x^{\alpha-1}\,\mathrm{d}x\Big)\\ &\leq \alpha^2\Big(\int_{a+u}^{b+u} x^{\alpha-1}\log(x)\,\mathrm{d}x\Big)\Big(\int_u^{a+u} x^{\alpha-1}\,\mathrm{d}x\Big)+\alpha\Big(\int_{a+u}^{b+u} x^{\alpha-1}\,\mathrm{d}x\Big)\Big(\int_u^{a+u} x^{\alpha-1}\,\mathrm{d}x\Big)\\ &= \Big(\int_{a+u}^{b+u} x^{\alpha-1}\big(\alpha\log(x)+1\big)\,\mathrm{d}x\Big)\Big(\alpha\int_u^{a+u} x^{\alpha-1}\,\mathrm{d}x\Big)\\ &= \Big((b+u)^\alpha\log(b+u)-(a+u)^\alpha\log(a+u)\Big)\Big((a+u)^\alpha-u^\alpha\Big).\end{aligned}$$

□

5 Directional Malliavin Calculus

This chapter is separated into three main parts. The first part covers the definition and study of the directional Malliavin derivative, the second gives a characterisation of independence, which is the main result of this chapter, and in the third part we extend the chain rule of Malliavin calculus to the directional Malliavin derivative and a broader class of functions.

5.1 Introduction

We consider an isonormal Gaussian process $W = \{W(h), h \in H\}$ associated with a separable Hilbert space H and defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Two types of directional Malliavin derivatives are widely used in the literature and both are covered by the definition of directional Malliavin derivative that we will introduce later on. The first one is given by

$$D^h F = \langle DF, h \rangle_H, \qquad h \in H,$$

for smooth random variables F, where $\langle \cdot, \cdot \rangle_H$ denotes the inner product on H, and which appears, among others, in [5, 19, 38]. Further, letting $B = (B_t)_{t \geq 0}$ be a d-dimensional Brownian motion, $H = L^2([0,T], \mathbb{R}^d)$ and $W(h) = \int_0^T h(t) \, \mathrm{d}B_t$, $h \in H$, we have that $D^{(j)}$, the Malliavin derivative with respect to the j-th Brownian motion, is a directional Malliavin derivative used e.g. in [39].

It is well-known that $DF = 0$ is equivalent to F being almost surely constant. This raises the question whether the directional Malliavin derivative being zero also corresponds to a different property of the random variable F. To give an intuition, we take a look at the result in the context of the example $H = L^2([0,T], \mathbb{R}^d)$, using $d = 2$. It is clear that if F is measurable with respect to $\sigma^{(1)} = \sigma(B_t^{(1)} : t \in [0,T])$, then $D^{(2)} F = 0$. It turns out that the converse also holds. This is done, in this example, by first proving that $D^{(2)} F = 0$ implies that F is independent of $\sigma^{(2)} = \sigma(B_t^{(2)} : t \in [0,T])$. In a second step we show that independence of $\sigma^{(2)}$ is close enough to measurability with respect to $\sigma^{(1)}$ to allow for the reverse statement. This result can be used to shed some new light on the characterisation of independence of random variables. In [47] the authors have shown that $\langle DF, DG \rangle = 0$ a.s. is not sufficient to ensure independence of $F, G \in \mathbb{D}^{1,2}$ and conjectured that the conditions that imply independence have to be more complicated. We show that only slightly stricter conditions suffice, namely, if there exists a closed subspace $\mathcal{H}$ of H such that almost surely $DF \in \mathcal{H}$ and $DG \in \mathcal{H}^\perp$, it follows that $F, G \in \mathbb{D}^{1,1}$ are independent. These results are presented in Section 5.4.

In Section 5.5 we derive a chain rule for our directional Malliavin derivative that also extends the existing chain rule in standard Malliavin calculus. Letting $p, d \in \mathbb{N}$ and

$F = (F^1, \dots, F^d)$ be a d-dimensional random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ where $F^i \in \mathbb{D}^{1,p},\ i \in \{1, \dots, d\}$, the chain rule for Malliavin calculus states that, for a continuously differentiable Lipschitz function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we have $\varphi(F) \in \mathbb{D}^{1,p}$ and

$$D\varphi(F) = \sum_{i=1}^{d} \partial_i \varphi(F) DF^i. \tag{5.1}$$

Let $L : H \to \mathcal{H}$ be a bounded linear operator. The directional Malliavin derivative D^L, which we will define later on, extends the standard Malliavin derivative in the sense that $D^L F = LDF,\ F \in \mathbb{D}^{1,2}$. We obtain a chain rule for this directional derivative and a less restrictive class of functions stating that, under certain conditions on φ and for $F^i \in \mathbb{D}^{1,p,L},\ i \in \{1, \dots, d\}$, we have

$$D^L \varphi(F) = \sum_{i \in J} \partial_i \varphi(F) D^L F^i,$$

where

$$J = \{1, \dots, d\} \setminus \{ i \mid F^i \text{ independent of } \sigma(W(h) : h \in \ker(L)^{\perp}) \}.$$

This helps e.g. to check Malliavin differentiability in the Heston model (see [2]) as the square root is not globally Lipschitz but nevertheless an admissible function in our theorem.

Some more elementary lemmata that we used can be found in Section 5.6. But first we introduce the notation and state some preliminary results in Section 5.2 before defining our directional Malliavin derivative in Section 5.3.

5.2 Preliminaries

Let H be a separable Hilbert space. A stochastic process $W = \{W(h), h \in H\}$ that is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called an *isonormal Gaussian process* (associated with or on H) if, for every $n \in \mathbb{N}$ and all $h_1, \dots, h_n \in H$, we have that $(W(h_1), \dots, W(h_n))$ is a centred normal random vector with covariance given by

$$\mathbb{E}[W(h_k)W(h_l)] = \langle h_k, h_l \rangle_H, \qquad k, l = 1, \dots, n.$$

From now on we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, where the σ-algebra $\mathcal{F}$ is generated by the isonormal Gaussian process W.

The following definitions and conventions are in line with [38]. Denote by $C_p^{\infty}(\mathbb{R}^d)$ all functions $f : \mathbb{R}^d \to \mathbb{R}$ that are infinitely often differentiable, and f and all its partial derivatives have polynomial growth. We define $\mathcal{S}$ to be the set of all random variables of the form

$$F = f(W(h_1), \dots, W(h_n)),$$

where $n \in \mathbb{N}$, $f \in C_p^{\infty}(\mathbb{R}^n)$ and $h_1, \dots, h_n \in H$. This set is called the set of smooth random variables. Similarly we define $\mathcal{S}_b$ to be the set of all smooth random variables where

$$f \in C_b^{\infty}(\mathbb{R}^n) := \{ g \in C^{\infty}(\mathbb{R}^n) : g \text{ and all its partial derivatives are bounded} \}.$$

It holds that $\mathcal{S}_b \subseteq \mathcal{S}$ and both are dense in $L^p(\Omega)$. On $\mathcal{S}$ the Malliavin derivative is defined as

$$DF = \sum_{i=1}^{n} \partial_i f(W(h_1), \dots, W(h_n))h_i$$

and $\mathbb{D}^{1,p}$ denotes the closure of $\mathcal{S}$ with respect to the norm

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

The same definition can be extended to Hilbert space-valued random variables. Let $\mathcal{H}$ be a Hilbert space and $\mathcal{S}_{\mathcal{H}}$ a family of $\mathcal{H}$-valued random variables of the form

$$F = \sum_{i=1}^{n} F_j h_j,$$

where $F_j \in \mathcal{S},\ h_j \in \mathcal{H}$ for all $j \in \{1, \dots, n\}$. Define $DF = \sum_{j=1}^{n} DF_j \otimes h_j$. We denote by $\mathbb{D}^{1,p}(\mathcal{H})$ the closure of $\mathcal{S}_{\mathcal{H}}$ with respect to the norm

$$\|F\|_{1,p,\mathcal{H}} = \big(\mathbb{E}[\|F\|_{\mathcal{H}}^p] + \mathbb{E}[\|DF\|_{H\otimes\mathcal{H}}^p]\big)^{\frac{1}{p}}.$$

Note that $\mathcal{S}_{\mathcal{H}}$ is dense in $L^2(\Omega; \mathcal{H})$. This way it is possible to define higher order Malliavin derivatives D^k and their respective domains $\mathbb{D}^{k,p}$.

Our first auxiliary result is the following small lemma.

Lemma 5.2.1. *Let $\mathscr{B} = \{e_j, j \in I\}$ be an orthonormal basis of H, where $I = \{1, \dots, N\}$ or $I = \mathbb{N}$, depending on the dimension of H. Define*

$$\mathscr{S} := \{F \in \mathcal{S}_b : F = f(W(e_1), \dots, W(e_n)),\ n \in I,\ f \in C_b^\infty(\mathbb{R}^n)\}.$$

Then $\mathscr{S}$ is dense in $\mathcal{S}_b$ and therefore in $L^p(\Omega)$.

Proof. We prove the result for infinite dimensional H. The proof for finite dimensional H follows trivially.
Let $F = f(W(h_1), \dots, W(h_m)) \in \mathcal{S}_b$, i.e. $f \in C_b^\infty(\mathbb{R}^m)$ and $h_1, \dots, h_m \in H$. We have that

$$h_i := \sum_{j=1}^{\infty} \underbrace{\langle h_i, e_j\rangle_H}_{=:\lambda_{ij}} e_j.$$

Because of the linearity of W, there exists some $g_n \in C_b^\infty(\mathbb{R}^n)$ such that

$$F_n := f\bigg(W\Big(\sum_{j=1}^{n} \lambda_{1j} e_j\Big), \dots, W\Big(\sum_{j=1}^{n} \lambda_{mj} e_j\Big)\bigg) = g_n(W(e_1), \dots, W(e_n)).$$

So, $F_n \in \mathscr{S}$ for all $n \in \mathbb{N}$. Since all $W(h)$, $h \in H$ are normally distributed with mean zero and variance $\|h\|_H^2$, there exists a constant $c_p > 0$ only depending on p such that

$$\bigg\|W(h_i) - W\Big(\sum_{j=1}^{n} \lambda_{ij} e_j\Big)\bigg\|_{L^p(\Omega)}^p = \mathbb{E}\bigg[\Big|W\Big(\sum_{j=n+1}^{\infty} \lambda_{ij} e_j\Big)\Big|^p\bigg] \le c_p\Big(\sum_{j=n+1}^{\infty} \lambda_{ij}^2\Big)^{p/2}.$$

Because the right-hand side converges to zero as $n \to \infty$ and f is Lipschitz continuous, we obtain $F_n \xrightarrow{L^p(\Omega)} F$. □

5.3 Directional Malliavin derivative

In this section we generalise the idea of Malliavin derivatives to the concept of directional Malliavin derivatives. This section follows the work in [1].

Let $\mathcal{H}$ be a Hilbert space and $L : H \to \mathcal{H}$ a bounded linear operator. On the set $\mathcal{S}$ of smooth random variables, we define the directional Malliavin derivative D^L as $L \circ D$, i.e.

$$D^L F = \sum_{i=1}^{m} \partial_i f(W(h_1), \ldots, W(h_m)) L h_i,$$

where $F = f(W(h_1), \ldots, W(h_m))$, $f \in C_p^\infty(\mathbb{R}^d)$, $h_1, \ldots, h_m \in H$. This implies that $D^L F = LDF$ for all $F \in \mathcal{S}$.

Lemma 1.2.1 and 1.2.2 in [38] state the following: Let $F, G \in \mathcal{S}$ and $h \in H$. Then

$$\mathbb{E}[\langle DF, h \rangle_H] = \mathbb{E}[FW(h)] \tag{5.2}$$

and

$$\mathbb{E}[G\langle DF, h \rangle_H] = \mathbb{E}[-F\langle DG, h \rangle_H + FGW(h)].$$

With the help of these result we can prove the corresponding statements for D^L.

Lemma 5.3.1. *Let $F, G \in \mathcal{S}$ and $h \in \mathcal{H}$. We denote the adjoint of L by $L^\star$. We have*

$$\mathbb{E}[\langle D^L F, h \rangle_{\mathcal{H}}] = \mathbb{E}[FW(L^\star h)] \tag{5.3}$$

and

$$\mathbb{E}[G\langle D^L F, h \rangle_{\mathcal{H}}] = \mathbb{E}[-F\langle D^L G, h \rangle_{\mathcal{H}} + FGW(L^\star h)]. \tag{5.4}$$

Proof. Using (5.2) yields

$$\mathbb{E}[\langle D^L F, h \rangle_{\mathcal{H}}] = \mathbb{E}[\langle DF, L^\star h \rangle_H] = \mathbb{E}[FW(L^\star h)].$$

To prove (5.4) first note that by linearity of L we have

$$D^L(FG) = L(D(FG)) = L(FDG + GDF) = FD^L G + GD^L F.$$

Using this result and (5.3) we obtain

$$\mathbb{E}[FGW(L^\star h)] = \mathbb{E}[\langle D^L(FG), h \rangle_{\mathcal{H}}] = \mathbb{E}[\langle FD^L G, h \rangle_{\mathcal{H}} + \langle GD^L F, h \rangle_{\mathcal{H}}].$$

□

The next Proposition can be found in [1].

Proposition 5.3.2. *The operator D^L is closable from $L^p(\Omega)$ to $L^p(\Omega; \mathcal{H})$.*

Proof. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{S}$ converging to zero in $L^p(\Omega)$ such that $D^L F_n$ converges to η in $L^p(\Omega;\mathcal{H})$. By equation (5.4) we have for any $h \in \mathcal{H}$ and any

$$G \in \{F \in \mathcal{S}_b : FW(L^\star h) \text{ is bounded}\} =: \beta(h)$$

that

$$\lim_{n\to\infty} \mathbb{E}[\langle D^L F_n, h\rangle_{\mathcal{H}} G] = \lim_{n\to\infty} \mathbb{E}[-F_n \langle D^L G, h\rangle_{\mathcal{H}} + F_n G W(L^\star h)] = 0,$$

since $\langle D^L G, h\rangle_{\mathcal{H}}$ and $GW(L^\star h)$ are bounded. It remains to be shown that $\beta(h)$ is dense in $\mathcal{S}_b$ which is itself dense in $L^p(\Omega)$. Then, $\eta = 0$ in $L^p(\Omega;\mathcal{H})$ and the assertion follows. So, let $G \in \mathcal{S}_b$ and set $G_n := G\exp\left(-\frac{W(L^\star h)^2}{n}\right)$ for $n \in \mathbb{N}$. Then we have that $(G_n)_{n\in\mathbb{N}}$ is a sequence in $\beta(h)$ with $G_n \xrightarrow{L^p(\Omega)} G$. □

This proposition allows us to define $\mathbb{D}^{1,p,L}$ as the domain of D^L in $L^p(\Omega)$, i.e. $\mathbb{D}^{1,p,L}$ is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p,L} = \left(\mathbb{E}[|F|^p] + \mathbb{E}[\|D^L F\|_{\mathcal{H}}^p]\right)^{\frac{1}{p}}.$$

For $p = 2$, the space $\mathbb{D}^{1,2,L}$ is a Hilbert space with the inner product

$$\langle F, G\rangle_{1,2,L} = \mathbb{E}[FG] + \mathbb{E}[\langle D^L F, D^L G\rangle_{\mathcal{H}}].$$

We remark that a different approach would be to define

$$\widetilde{D}^L : \mathbb{D}^{1,p} \to L^p(\Omega;\mathcal{H});\ F \mapsto L(DF).$$

In fact we have $\mathbb{D}^{1,p} \subseteq \mathbb{D}^{1,p,L}$ and $\widetilde{D}^L F = D^L F$ for $F \in \mathbb{D}^{1,p}$ but in general $\mathbb{D}^{1,p} \neq \mathbb{D}^{1,p,L}$.

Remark 5.3.3. Similar to the divergence operator δ in standard Malliavin calculus it is possible to define δ^L as the adjoint of D^L and many properties of δ carry over to δ^L, for example the following properties.

(i) Let $G \in L^2(\Omega)$, $u \in L^2(\Omega;\mathcal{H})$. If it holds for all $F \in \mathcal{S}_b$ that

$$\mathbb{E}[\langle D^L F, u\rangle_{\mathcal{H}}] = \mathbb{E}[FG],$$

then $u \in \operatorname{Dom}\delta^L$ and $\delta^L(u) = G$.

(ii) We have $\operatorname{Dom}\delta^L = (L^\star)^{-1}(\operatorname{Dom}\delta)$ and $\delta^L = \delta \circ L^\star$. This implies that δ^L is a closed operator.

(iii) Let $F \in \mathbb{D}^{1,2,L}$ and $u \in \operatorname{Dom}\delta^L$ such that $Fu \in L^2(\Omega;\mathcal{H})$. Then $Fu \in \operatorname{Dom}\delta^L$ and

$$\delta^L(Fu) = F\delta^L(u) - \langle D^L F, u\rangle_{\mathcal{H}}.$$

The next proposition shows that in some cases, which include the ones usually considered, directional Malliavin differentiability implies Malliavin differentiability. This is to be expected as the Malliavin derivative is a kind of weak derivative. In some set-ups this might make it easier to check for Malliavin differentiability.

Proposition 5.3.4. *Let $d \in \mathbb{N}$ and H_j, $j \in \{1, \dots, d\}$ orthogonal subspaces of H, such that $H = \bigoplus_{j=1}^d H_j$. We denote by $L_j : H \to H_j$ the projections of H onto H_j. If $F \in \bigcap_{j=1}^d \mathbb{D}^{1,p,L_j}$, then $F \in \mathbb{D}^{1,p}$ and*

$$DF = \sum_{j=1}^d D^{L_j} F.$$

Proof. It is evident that there exists a sequence $(F_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}_b$ such that $F_n \xrightarrow{L^p(\Omega)} F$. We have, for some $m = m(n) \in \mathbb{N}$, that

$$D^{L_j} F_n = \sum_{i=1}^m \partial_i f_n(W(h_1), \dots, W(h_m)) L_j h_i,$$

where $f_n \in C_b^\infty(\mathbb{R}^m)$. Since $\sum_{j=1}^d L_j$ is the identity on H, it follows that

$$\sum_{j=1}^d D^{L_j} F_n = \sum_{i=1}^m \left[\partial_i f_n(W(h_1), \dots, W(h_m)) \Big(\sum_{j=1}^d L_j \Big) h_i \right] = DF_n.$$

Since the left hand side of the equation converges in $L^p(\Omega; H)$ to $\sum_{j=1}^d D^{L_j} F$ and the operator D is closed, we obtain $F \in \mathbb{D}^{1,p}$ and

$$DF = \sum_{j=1}^d D^{L_j} F.$$

□

The following is a common example of a directional Malliavin derivative. Let $T > 0$ and consider $H = L^2([0,T]; \mathbb{R}^d)$ and the isonormal Gaussian process $W = \{W(h) : h \in H\}$ that is defined by a Wiener integral over a d-dimensional Brownian motion $(B_t)_{t\in[0,T]} = ((B_t^{(1)}, \dots, B_t^{(d)})^\top)_{t\in[0,T]}$. Putting $\mathcal{H} = L^2([0,T]; \mathbb{R})$ and defining

$$L_j : H \to \mathcal{H};\ L_j h = h_j, \qquad \text{where } h = (h_1, \dots, h_d)^\top \in H$$

for $j \in \{1, \dots, d\}$, we can understand $D^{L_j} := D^{(j)}$ as the (directional) Malliavin derivative with respect to the j^{th} Brownian motion. If $F \in \mathbb{D}^{1,1}$, then

$$DF = \begin{pmatrix} (DF)_1 \\ \vdots \\ (DF)_d \end{pmatrix} = \begin{pmatrix} D^{(1)} F \\ \vdots \\ D^{(d)} F \end{pmatrix}.$$

5.4 Characterisation of independence

In this section we present what could be inferred about $F \in \mathbb{D}^{1,p,L}$ if $D^L F = 0$. This result allows us to formulate a condition on the Malliavin derivatives that implies independence of the random variables.

The following lemma is a direct consequence of Lemma 1.2.4 in [38].

Lemma 5.4.1. *Let $\sigma^{\ker^\perp}$ denote the σ-algebra generated by $\{W(h) : h \in \ker(L)^\perp\}$. Then we have that $H_L^\perp := \ker(L)^\perp$ with the inner product of H is a Hilbert space and the set*

$$T^\perp = \{1, W(h)G - \langle DG, h\rangle_H : G \in \mathcal{S}_b^\perp,\ h \in H_L^\perp\},$$

where

$$\mathcal{S}_b^\perp := \{F = f(W(h_1), \dots, W(h_m)) : f \in C_b^\infty(\mathbb{R}^m),\ h_1, \dots, h_m \in H_L^\perp\}$$

is a total set in $L^2(\Omega, \sigma^{\ker^\perp}, \mathbb{P})$.

Proposition 5.4.2. *Let $F \in \mathbb{D}^{1,1,L}$. If F is measurable with respect to the σ-algebra $\sigma^{\ker} := \sigma\big(W(h) : h \in \ker(L)\big)$, then $D^L F = 0$. On the other hand, $D^L F = 0$ implies that F is independent of $\sigma^{\ker^\perp}$. Note that equality is meant in the $L^1(\Omega; \mathcal{H})$ sense.*

Proof. First we assume that F is $\sigma^{\ker}$-measurable. Then, there exists a sequence $(F_n)_{n\in\mathbb{N}}$, where

$$F_n = f(W(h_1), \dots, W(h_m)), \quad f \in C_b^\infty(\mathbb{R}^m),\ h_1, \dots, h_m \in \ker(L)$$

for all $n \in \mathbb{N}$ and $F_n \xrightarrow{n\to\infty} F$ in $L^1(\Omega)$. We have $D^L F_n = 0$ for all $n \in \mathbb{N}$ and thus $D^L F = 0$.

Now we suppose that $D^L F = 0$. It holds that $L : H_L^\perp \to \operatorname{im}(L)$ is an isomorphism and consequently so is $L^\star : \operatorname{im}(L) \to H_L^\perp$. Let $G \in \mathcal{S}_b^\perp$ be arbitrary and bounded by $c > 0$ and fix an $h \in H_L^\perp$. There exists a $g \in \operatorname{im}(L) \subseteq \mathcal{H}$ such that $h = L^\star g$ and we have

$$\mathbb{E}[\langle DG, h\rangle_H] = \mathbb{E}[\langle DG, L^\star g\rangle_H] = \mathbb{E}[\langle D^L G, g\rangle_{\mathcal{H}}].$$

Let $(F_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}_b$ such that $F_n \xrightarrow{L^1(\Omega)} F$ and $\mathbb{E}[\|D^L F_n\|_{\mathcal{H}}] \to 0$ as $n \to \infty$.

In addition, let $\psi : \mathbb{R} \to \mathbb{R}$ be a bounded, measurable function. As the law of F, denoted by $\mathbb{P}^F$, is a Radon measure on the Borel sets of $\mathbb{R}$, Lusin's Theorem (see e.g. [8], Theorem 7.10) states that ψ can be approximated in $L^2(\mathbb{R}, \mathbb{P}^F)$ by continuous, compactly supported functions. The approximations can be chosen to be uniformly bounded by $\|\psi\|_\infty$. A mollifying argument yields that there exists a sequence $(\psi_N)_{N\in\mathbb{N}} \subseteq C_b^\infty(\mathbb{R})$ such that $\psi_N \to \psi$ in $L^2(\mathbb{R}, \mathbb{P}^F)$, or, in other words,

$$\mathbb{E}[(\psi_N(F) - \psi(F))^2] \xrightarrow{N\to\infty} 0.$$

For the moment let $N \in \mathbb{N}$ be fixed. So, we have $\psi_N \in C_b^\infty(\mathbb{R})$ and, for all $n \in \mathbb{N}$, $F_n, G \in \mathcal{S}$, which implies $\psi_N(F_n)G \in \mathcal{S}$. It follows by equation (5.3) that

$$\mathbb{E}[\langle D^L(\psi_N(F_n)G), g\rangle_{\mathcal{H}}] = \mathbb{E}[\psi_N(F_n) W(L^\star g) G]. \tag{5.5}$$

Note that, for $X = x(W(h_1), \dots, W(h_n)), Y = y(W(h_1), \dots, W(h_n)) \in \mathcal{S}$, we have

$$\begin{aligned} D^L(XY) = \sum_{i=1}^n \Big[& x(W(h_1), \dots, W(h_n)) \partial_i y(W(h_1), \dots, W(h_n)) \\ & + \partial_i x(W(h_1), \dots, W(h_n)) y(W(h_1), \dots, W(h_n)) \Big] L h_i \\ = X D^L Y & + Y D^L X, \end{aligned} \tag{5.6}$$

and

$$D^L(\psi_N(X)) = \psi_N'(X) \sum_{j=1}^{n} \partial_j x(W(h_1), \ldots, W(h_n)) Lh_j = \psi_N'(X) D^L X. \qquad (5.7)$$

Using (5.5)-(5.7) and $h = L^\star g$, we obtain

$$\begin{aligned} &\mathbb{E}[\psi_N(F_n)(W(h)G - \langle DG, h\rangle_H)] \\ &\qquad = \mathbb{E}[\psi_N(F_n)W(L^\star g)G - \langle D^L(\psi_N(F_n)G), g\rangle_{\mathcal{H}})] \\ &\qquad\quad + \mathbb{E}[G\langle D^L\psi_N(F_n), g\rangle_{\mathcal{H}}] \\ &\qquad = \mathbb{E}[G\psi_N'(F_n)\langle D^L F_n, g\rangle_{\mathcal{H}}] \\ &\qquad \le c\gamma_N \mathbb{E}[\|D^L F_n\|_{\mathcal{H}}]\|g\|_{\mathcal{H}} \xrightarrow{n\to\infty} 0, \end{aligned}$$

where $\gamma_N = \sup_{x\in\mathbb{R}} |\psi_N'(x)|$ is the Lipschitz constant of ψ_N. In particular, using dominated convergence and the continuity of ψ_N, we obtain $\mathbb{E}[\psi_N(F)(W(h)G - \langle DG, h\rangle_H)] = 0$ for all $N \in \mathbb{N}$, and thus

$$\mathbb{E}[\psi(F)(W(h)G - \langle DG, h\rangle_H)] = 0. \qquad (5.8)$$

Let X be a bounded $\sigma^{\ker^\perp}$-measurable random variable. Then $X \in L^2(\Omega)$ and by Lemma 5.4.1 there exist $Y_i \in T^\perp$ and $a_i \in \mathbb{R}$, $i \in \mathbb{N}$ such that

$$X_n := \sum_{i=1}^{n} a_i Y_i \xrightarrow{L^2(\Omega)} X, \text{ as } n \to \infty.$$

The linear functional $\phi : L^2(\Omega) \to \mathbb{R}$, $X \mapsto \mathbb{E}[\psi(F)(X - \mathbb{E}X)]$ is continuous and, by (5.8), we have $\mathbb{E}[\psi(F)(X_n - \mathbb{E}X_n)] = 0$ for all $n \in \mathbb{N}$. Thus, $\mathbb{E}[\psi(F)X] = \mathbb{E}[\psi(F)]\mathbb{E}[X]$. The choices of the bounded, measurable function ψ and the bounded $\sigma^{\ker^\perp}$-measurable random variable X were arbitrary. Consequently, F is independent of $\sigma^{\ker^\perp}$. □

The following proposition provides a useful characterisation of independence of random variables. This result, being of rather basic nature, was surely proven before but unfortunately we were unable to find it or references to it in the literature.

Proposition 5.4.3. *Let $(\Omega, \mathscr{A}, \mathcal{P})$ be a probability space and $\mathscr{A} = \sigma(\sigma_1 \cup \sigma_2)$, where σ_1, σ_2 are two independent σ-algebras. A random variable $X \in L^1(\Omega, \mathscr{A}, \mathcal{P})$ is independent of σ_2 if and only if there exists a σ_1-measurable random variable $\widetilde{X} \in L^1(\Omega, \sigma_1, \mathcal{P}) \subseteq L^1(\Omega, \mathscr{A}, \mathcal{P})$ such that $X = \widetilde{X}$ almost surely.*

Proof. First, let $\widetilde{X} \in L^1(\Omega, \mathscr{A}, \mathbb{P})$ be a σ_1-measurable random variable and $X = \widetilde{X}$ almost surely. For any bounded σ_2-measurable random variable G and any bounded measurable function $h : \mathbb{R} \to \mathbb{R}$ we have

$$\mathbb{E}[h(X)G] = \mathbb{E}[h(\widetilde{X})G] = \mathbb{E}[h(\widetilde{X})\mathbb{E}[G|\sigma_1]] = \mathbb{E}[h(\widetilde{X})]\mathbb{E}[G] = \mathbb{E}[h(X)]\mathbb{E}[G].$$

This implies that X is independent of σ_2.

It remains to show the reverse implication. Assume X is independent of σ_2 and define $\widetilde{X} := \mathbb{E}[X|\sigma_1]$. The properties of the conditional expectation give us $\widetilde{X} \in L^1(\Omega, \mathscr{A}, \mathcal{P})$ and $\widetilde{X}$ is σ_1-measurable. We have that $\Pi := \{A \cap B : A \in \sigma_1,\ B \in \sigma_2\}$ is a π-system with

$\sigma(\Pi) = \mathscr{A}$. To see this, we note that any $A \in \sigma_1$ or $B \in \sigma_2$ is clearly also an element of Π and therefore $\sigma_1 \cup \sigma_2 \subseteq \Pi$, which implies $\mathscr{A} = \sigma(\sigma_1 \cup \sigma_2) \subseteq \sigma(\Pi)$. As finite intersection of elements in $\mathscr{A}$ are also in $\mathscr{A}$, we have $\Pi \subseteq \mathscr{A}$, which implies $\sigma(\Pi) \subseteq \mathscr{A}$. We put $C := A \cap B \in \Pi$, where $A \in \sigma_1$ and $B \in \sigma_2$. Because X and $\widetilde{X}$ are both independent of σ_2, we obtain

$$\mathbb{E}[\mathbb{1}_C(X - \widetilde{X})] = \mathbb{E}[\mathbb{1}_A \mathbb{1}_B (X - \widetilde{X})] = \mathbb{E}[\mathbb{1}_B]\mathbb{E}\big[\mathbb{1}_A\big(X - \mathbb{E}[X|\sigma_1]\big)\big] = 0$$

because $\mathbb{E}[\mathbb{1}_A(X - \mathbb{E}[X|\sigma_1])] = 0$ by the definition of conditional expectation. Applying Lemma 5.6.1 yields $X = \mathbb{E}[X|\sigma_1] = \widetilde{X}$ almost surely. □

Proposition 5.4.3 allows us to reformulate and improve Proposition 5.4.2 into Theorem 5.4.4 below. The Theorems 5.4.4 and 5.4.5 constitute one of the main results of this chapter.

Theorem 5.4.4. *Let $F \in \mathbb{D}^{1,1,L}$. The following statements are equivalent.*

(i) $D^L F = 0$ *in* $L^1(\Omega; \mathcal{H})$.

(ii) F *is independent of* $\sigma^{\ker^\perp}$.

(iii) There exists a random variable $G \in L^1(\Omega)$ *such that* $F = G$ *a.s. and* G *is* $\sigma^{\ker}$*-measurable.*

Proof.

(i) ⇒ *(ii)* Let $D^L F = 0$. By Proposition 5.4.2 we have that F is independent of $\sigma^{\ker^\perp}$.

(ii) ⇒ *(iii)* Let F be independent of $\sigma^{\ker^\perp}$. It follows from Proposition 5.4.3 that there exists a $\sigma^{\ker}$-measurable random variable G such that $F = G$ almost surely.

(iii) ⇒ *(i)* Let $X, Y \in L^1(\Omega)$ with $X = Y$ *a.s.*, then $X = Y$ in $L^1(\Omega)$. By the definition of the operator D^L we have $X \in \mathbb{D}^{1,1,L}$ if and only if $Y \in \mathbb{D}^{1,1,L}$ and in case $X \in \mathbb{D}^{1,1,L}$ it holds that $D^L X = D^L Y$ in $L^1(\Omega; \mathcal{H})$. Therefore, $G \in \mathbb{D}^{1,1,L}$ and $D^L F = D^L G = 0$ by Proposition 5.4.2.

□

From this theorem we can derive a condition on the standard Malliavin derivatives of two random variables that implies independence of said random variables.

Theorem 5.4.5. *Let $F, G \in \mathbb{D}^{1,1}$. If there exists a closed subspace $\mathcal{H}$ of H such that*

$$DF \in \mathcal{H} \text{ a.s.} \qquad \text{and} \qquad DG \in \mathcal{H}^\perp \text{ a.s.},$$

then F and G are independent.

Proof. Let L be the projection of H onto $\mathcal{H}$. Then $D^L G = 0$. Theorem 5.4.4 yields that G is independent of $\sigma^{\ker^\perp}$ and there exits a random variable $\widetilde{G} \in L^1(\Omega)$ such that $\widetilde{G} = G$ *a.s.* and $\widetilde{G}$ is $\sigma^{\ker}$-measurable. In the same way we obtain F is independent of $\sigma^{\ker}$ and it follows that F and G are independent. □

Using a result in [47], the reverse implication can be proven in the case of $H = L^2([0,T])$ and under slightly stricter conditions.

Proposition 5.4.6. *Let* $H = L^2([0,T])$ *and* $W(h) = \int_0^T h(t)\,\mathrm{d}W_t$*. Suppose* $F, G \in \mathbb{D}^{1,2}$*. Then the following are equivalent:*

(i) There exists a closed subspace $\mathcal{H}$ *of* H *such that*

$$DF \in \mathcal{H} \text{ a.s.} \qquad \text{and} \qquad DG \in \mathcal{H}^\perp \text{ a.s.}$$

(ii) The random variables F *and* G *are independent.*

Proof. Theorem 5.4.5 proves $(i) \Rightarrow (ii)$. Now let F, G be independent. The random variables can be expanded into a series of multiple stochastic Wiener integrals

$$F = \sum_{n=0}^{\infty} I_n(f_n), \qquad G = \sum_{n=0}^{\infty} I_n(g_n),$$

where $f_n, g_n \in L^2([0,T]^n)$ are symmetric functions. For $n \in \mathbb{N}_0$, denote by J_n the projection onto the n-th Wiener chaos. For $n, m \in \mathbb{N}_0$, we have

$$\begin{aligned}\mathbb{P}(J_nF \in A, J_mG \in B) &= \mathbb{P}(F \in J_n^{-1}(A), G \in J_m^{-1}(B)) \\ &= \mathbb{P}(F \in J_n^{-1}(A))\mathbb{P}(G \in J_m^{-1}(B)) \\ &= \mathbb{P}(J_nF \in A)\mathbb{P}(J_mG \in B)\end{aligned}$$

for all $A, B \in \mathcal{B}(\mathbb{R})$. Thus, $J_nF = I_n(f_n)$ and $J_mG = I_m(g_m)$ are independent for all $n, m \in \mathbb{N}$. Define

$$\mathcal{H} := \left\{ \varphi \in L^2([0,T]) \ : \ \left\| \int_0^T g_m(t,\cdot)\varphi(t)\,\mathrm{d}t \right\|_{L^2([0,T]^{m-1})} = 0, \ \forall m \in \mathbb{N} \right\},$$

which is a closed subspace of H.

In what follows let $\cdot$ and $\bullet$ be placeholders for different variables. In iterated integrals we always integrate over the variables represented by $\cdot$ and never over those represented by $\bullet$. The justification of the stochastic Fubini results used in this proof is given in Lemma 5.6.2.

Let $m \in \mathbb{N}$ and $\varphi \in \mathcal{H}$. Applying stochastic Fubini, we have almost surely

$$\begin{aligned}\langle DI_m(g_m), \varphi\rangle_{L^2([0,T])} &= m \int_0^T I_{m-1}(g_m(t,\cdot))\varphi(t)\,\mathrm{d}t \\ &= mI_{m-1}\Big(\int_0^T g_m(t,\cdot)\varphi(t)\,\mathrm{d}t\Big) = 0,\end{aligned}$$

and it follows

$$DI_m(g_m) = mI_{m-1}(g_m(t,\cdot)) \in \mathcal{H}^\perp \text{ a.s.}$$

for all $m \in \mathbb{N}$. Theorem 6 in [47] states that

$$\|f_n \otimes_1 g_m\|_{L^2([0,T]^{m+n-2})} = 0$$

for any choice of $n, m \in \mathbb{N}$, where

$$f_n \otimes_1 g_m = \int_0^T f_n(t,\cdot)g_m(t,\bullet)\,\mathrm{d}t.$$

Again applying stochastic Fubini, we obtain for any $n, m \in \mathbb{N}$ that

$$\begin{aligned}\int_0^T D_t I_n(f_n) g_m(t,\bullet)\,\mathrm{d}t &= n\int_0^T I_{n-1}(f_n(t,\cdot))g_m(t,\bullet)\,\mathrm{d}t \\ &= nI_{n-1}\Big(\int_0^T f_n(t,\cdot)g_m(t,\bullet)\,\mathrm{d}t\Big) = 0 \ a.s,\end{aligned}$$

where the last zero denotes the zero function in $L^2([0,T]^{m-1})$. Thus,

$$DI_n(f_n) = nI_{n-1}(f_n(t,\cdot)) \in \mathcal{H} \ a.s.$$

for all $n \in \mathbb{N}$. Since $\mathcal{H}$ and $\mathcal{H}^\perp$ are closed subspaces it follows that

$$\begin{aligned} t \mapsto D_t F &= \sum_{n=1}^\infty nI_{n-1}(f_n(t,\cdot)) \in \mathcal{H} \ a.s., \\ t \mapsto D_t G &= \sum_{m=1}^\infty mI_{m-1}(g_m(t,\cdot)) \in \mathcal{H}^\perp \ a.s. \end{aligned}$$

□

It might be conjectured that the statement above holds for general $F, G \in \mathbb{D}^{1,1}$ and the additional assumptions in Proposition 5.4.6 are only an artefact of the proof.

The following example shows that, for $F, G \in \mathbb{D}^{1,2}$, the condition $\langle DF, DG\rangle = 0$ a.s. is not sufficient to imply independence of F and G.

Example 5.4.7. Let $W(h) = \int_0^1 h(t)\,\mathrm{d}B_t$, $h \in H = L^2([0,1],\mathbb{R})$ and $B = (B_t)_{t\geq 0}$ a standard Brownian motion. Put

$$\begin{aligned} F &= \alpha(B_1 + 1) \\ G &= \alpha(B_1 - 1), \end{aligned}$$

where $\alpha \in C^\infty(\mathbb{R})$ is nonnegative function with support on the unit interval and $\int_{\mathbb{R}} \alpha(x)\,\mathrm{d}x = 1$. Then F, G are not independent as

$$\mathbb{E}[F]\mathbb{E}[G] > 0 = \mathbb{E}[FG].$$

But using the chain rule, which is also presented in the next section, we obtain

$$\begin{aligned} D_t F &= \alpha'(B_1 + 1)\mathbb{1}_{[0,1]}(t) \\ D_t G &= \alpha'(B_1 - 1)\mathbb{1}_{[0,1]}(t), \end{aligned}$$

and therefore $\langle DF, DG\rangle_{L^2} = 0$.

5.5 Chain rule in Malliavin calculus

In this section let $p, d \in \mathbb{N}$, $F = (F^1, \dots, F^d)$ be a d-dimensional random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\|\cdot\|$ denote the Euclidean norm on $\mathbb{R}^d$. We want to quickly restate the standard chain rule in Malliavin calculus that can, e.g., be found in [38], Proposition 1.2.3.

Proposition 5.5.1. *Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function with bounded derivative and $F^i \in \mathbb{D}^{1,p}$, $i \in \{1, \dots, d\}$, then $\varphi(F) \in \mathbb{D}^{1,p}$ and (5.1) holds, i.e.*

$$D\varphi(F) = \sum_{i=1}^{d} \partial_i \varphi(F) DF^i.$$

Our aim is to transfer this result to the directional Malliavin derivative and find a larger class of function such that (5.1) still holds. This chapter is based on [1], where similar results are presented for the Malliavin derivative.

Let $f : \mathbb{R}^d \to \mathbb{R}$, $I \subseteq \{1, \dots, d\}$ and $e_i \in \mathbb{R}^d$ the vector that has a one in the i-th position and zeros otherwise. We make the following definitions

(i) We say that f is Lipschitz continuous in direction I if there exists a constant $\gamma > 0$ such that for all $x \in \mathbb{R}^d$ and $h \in \mathbb{R}$ we have

$$|f(x + he_i) - f(x)| \leq \gamma \|h\|, \qquad i \in I.$$

(ii) We say that f is locally Lipschitz in direction I if for every $x \in \mathbb{R}^d$ there exist positive constants $\varepsilon(x)$ and $\gamma(x)$ such that for all $\|h\| \leq \varepsilon(x)$ we have

$$|f(x + he_i) - f(x)| \leq \gamma(x) \|h\|, \qquad i \in I.$$

(iii) For $p \in \mathbb{N}$, we say $f \in C_I^p(\mathbb{R}^d)$ if, for all $k \leq p$ and $i_1, \dots, i_k \in I$, we have that the partial derivative $\partial_{i_1, \dots i_k} f$ exists and is continuous on $\mathbb{R}^d$. Further, define

$$C_I^\infty(\mathbb{R}^d) = \bigcap_{p \in \mathbb{N}} C_I^p(\mathbb{R}^d).$$

Let $\alpha \in C^\infty(\mathbb{R}^d)$ be a nonnegative function with support on the unit ball and $\int_{\mathbb{R}^d} \alpha(x) \, \mathrm{d}x = 1$. Then, for $n \in \mathbb{N}$, we define

$$\alpha_n : \mathbb{R}^d \to \mathbb{R}, \, x \mapsto n^d \alpha(nx).$$

This so-called mollifier function is needed in the proofs that follow. To simplify notation for the rest of Section 5.5, we make the following definition. If $g : \mathbb{R}^d \to \mathbb{R}$ is not partially differentiable at $x \in \mathbb{R}^d$ in the i-th component, we set $\partial_i g(x) := 0$.

The proof of the following lemma is transferred to the end of this chapter and can be found in Section 5.6.

Lemma 5.5.2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function and set $f_n = f * \alpha_n$, $n \in \mathbb{N}$ with α_n as defined above. The following properties hold:*

(i) For all $n \in \mathbb{N}$ we have $\displaystyle\int_{\mathbb{R}^d} \|x\| \alpha_n(x) \, \mathrm{d}x \leq \frac{1}{n}$.

(ii) Let f be continuous at $x_0 \in \mathbb{R}^d$. Then $f_n(x_0) \to f(x_0)$ for $n \to \infty$.

(iii) Let f be continuous on $\mathbb{R}^d$. Then $f_n \in C^\infty(\mathbb{R}^d)$.

(iv) *In addition to the continuity assumption in (iii), let f be Lipschitz continuous in direction $I \subseteq \{1,\dots,d\}$ with Lipschitz constant γ. Then, $\|\partial_i f_n\|_\infty \leq \gamma$ for all $i \in I$. Moreover, for higher partial derivatives of f_n we have that for every $k \in \mathbb{N}$ there exists $c_k > 0$ such that*

$$\sup_{x\in\mathbb{R}^d} \left|\partial_{i_1,\dots i_k} f_n(x)\right| \leq c_k$$

for all $i_1,\dots,i_k \in I$.

(v) *Assume that f is locally Lipschitz continuous in direction I. Then*

$$\partial_i(f * \alpha_n) = \partial_i f * \alpha_n$$

almost everywhere for all $i \in I$.

As the following assumption will be needed in all the chain rule results that follow, we state it here once and only refer to it henceforth.

Assumption 5.5.3. Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ and $F = (F^1,\dots,F^d)$ be a d-dimensional random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with $F^i \in \mathbb{D}^{1,p,L}$, $i \in \{1,\dots,d\}$, and

$$J := \{1,\dots,d\} \setminus \{i \mid F^i \text{ independent of } \sigma^{\ker^\perp}\},$$

where $\sigma^{\ker^\perp} = \sigma(W(h), h \in \ker(L)^\perp)$ is the same as in Lemma 5.4.1 above.

Note that it follows from Assumption 5.5.3 and Theorem 5.4.4 that $DF^i = 0$ for all $i \notin J$. We now have the necessary notation to extend Proposition 5.5.1 to the directional derivative. The result is generalised step-by-step by making the conditions on φ less restrictive, e.g. while the first proposition assumes φ to be bounded, the final result (Theorem 5.5.7) does not require boundedness.

Proposition 5.5.4. *Under Assumption 5.5.3, let φ be bounded, continuous and $\varphi \in C^1_J(\mathbb{R}^d)$ with bounded partial derivatives $\partial_i\varphi$, $i \in J$. Then $\varphi(F) \in \mathbb{D}^{1,p,L}$ and*

$$D^L\varphi(F) = \sum_{i\in J} \partial_i\varphi(F) D^L F^i. \tag{5.9}$$

Proof. Because $F^i \in \mathbb{D}^{1,p,L}$, there exists a sequence $(F_k)_{k\in\mathbb{N}} = ((F^1_k,\dots,F^d_k)^\top)_{k\in\mathbb{N}}$ with $(F^i_k)_{k\in\mathbb{N}} \subseteq \mathcal{S}_b$, $i \in \{1,\dots,d\}$ and F_k converging component-wise in $\mathbb{D}^{1,p,L}$ to F. We can write

$$F_k = f_k(W(h_1),\dots,W(h_m)) = \begin{pmatrix} f^1_k(W(h_1),\dots,W(h_m)) \\ \vdots \\ f^d_k(W(h_1),\dots,W(h_m)) \end{pmatrix} = \begin{pmatrix} F^1_k \\ \vdots \\ F^d_k \end{pmatrix},$$

where $h_1,\dots,h_m \in H$ and $f_k = (f^1_k,\dots,f^d_k)^\top \in C^\infty_p(\mathbb{R}^m)$. We define $\varphi_n := \varphi * \alpha_n$, where α_n is the mollifier function from above. We have $\varphi_n \circ f_k \in C^\infty_p(\mathbb{R}^m)$ and obtain by definition that

$$D^L\varphi_n(F_k) = \sum_{j=1}^m \partial_j(\varphi_n \circ f_k)(W(h_1),\dots,W(h_m))Lh_j$$

$$= \sum_{i=1}^{d} \sum_{j=1}^{m} \partial_i \varphi_n\big(f_k(W(h_1), \dots, W(h_m))\big) \partial_j f_k^i(W(h_1), \dots, W(h_m)) L h_j$$
$$= \sum_{i=1}^{d} \partial_i \varphi_n(F_k) D^L F_k^i.$$

By Theorem 5.4.4 the sequence $(F_k)_{k\in\mathbb{N}}$ can be chosen such that $F_k^i \subseteq \mathcal{S}_b^\perp,\ i \notin J,\ k \in \mathbb{N}$, where $\mathcal{S}_b^\perp$ is defined in Lemma 5.4.1. This yields $D^L F_k^i = 0,\ i \notin J,\ k \in \mathbb{N}$ and thus

$$D^L \varphi_n(F_k) = \sum_{i\in J} \partial_i \varphi_n(F_k) D^L F_k^i.$$

Since $F_k^i \xrightarrow{L^p(\Omega)} F^i$ as $k \to \infty$, there exists a subsequence $(F_{k_l})_{l\in\mathbb{N}}$ such that this subsequence converges almost surely to F. We choose such a subsequence as our initial sequence $(F_k)_{k\in\mathbb{N}}$, i.e. we can assume w.l.o.g. that $F_k \xrightarrow{k\to\infty} F$ almost surely. It remains to show that

$$\lim_{n\to\infty} \lim_{k\to\infty} \|\varphi_n(F_k) - \varphi(F)\|_{1,p,L} = 0.$$

So, the limits in this proof, if not state otherwise, are obtained by first letting $k \to \infty$ and then $n \to \infty$. Using the triangle inequality we obtain

$$\|\varphi_n(F_k) - \varphi(F)\|_{L^p(\Omega)} \leq \|\varphi_n(F_k) - \varphi_n(F)\|_{L^p(\Omega)} + \|\varphi_n(F) - \varphi(F)\|_{L^p(\Omega)}.$$

Because φ_n is continuous and bounded by $\|\varphi\|_\infty$, we have that $|\varphi_n(F_k) - \varphi_n(F)|$ converges almost surely to zero as $k \to \infty$ and applying dominated convergence yields that the first summand converges to zero. By Lemma 5.5.2(ii), we have that $\varphi_n(F)$ converges pointwise to $\varphi(F)$ as $n \to \infty$. Using again dominated convergence, we see that the second summand converges to zero. Moreover, for $i \in J$, the triangle inequality yields

$$\begin{aligned}\|\partial_i \varphi_n(F_k) D^L F_k^i - \partial_i \varphi(F) D^L F^i\|_{L^p(\Omega;\mathcal{H})} &\leq \|\partial_i \varphi_n(F_k)(D^L F_k^i - D^L F^i)\|_{L^p(\Omega;\mathcal{H})} \\ &\quad + \|(\partial_i \varphi_n(F_k) - \partial_i \varphi_n(F)) D^L F^i\|_{L^p(\Omega;\mathcal{H})} \\ &\quad + \|(\partial_i \varphi_n(F) - \partial_i \varphi(F)) D^L F^i\|_{L^p(\Omega;\mathcal{H})}.\end{aligned}$$

Note that $|\partial_i \varphi_n|$ and $|\partial_i \varphi|$ are bounded by some constant C. So the first summand is bounded by

$$C \|D^L F_k^i - D^L F^i\|_{L^p(\Omega;\mathcal{H})},$$

which converges to zero as $k \to \infty$. The absolute value of the term inside the last norm is bounded by $2C|D^L F^i| \in L^p(\Omega;\mathcal{H})$ and by Lemma 5.5.2(ii)

$$\partial_i \varphi_n(F(\omega)) \xrightarrow{n\to\infty} \partial_i \varphi(F(\omega))$$

for all $\omega \in \Omega$. So the third summand converges to zero as $n \to \infty$ by the dominated convergence theorem. The absolute value of the term inside the norm of the second summand is also bounded by $2C|D^L F^i| \in L^p(\Omega;\mathcal{H})$ and since $F_k \to F$ *a.s.*, we have by the continuous mapping theorem and dominated convergence that the second summand converges to zero as $k \to \infty$. Thus, we have shown that

$$\lim_{n\to\infty} \lim_{k\to\infty} \|\varphi_n(F_k) - \varphi(F)\|_{1,p,L} = 0$$

and the proof is complete. □

Lemma 5.5.5. *Under Assumption 5.5.3, let φ be Lipschitz continuous in direction J with Lipschitz constant γ. Further, suppose that there exists a set $N \in \mathcal{B}(\mathbb{R}^d)$ with $\mathbb{P}(F \in N) = 0$ such that φ is bounded, continuous, and continuously differentiable in direction J on $\mathbb{R}^d \setminus N$. Then, $\varphi(F) \in \mathbb{D}^{1,p,L}$ and (5.9) holds.*

Proof. We set $\varphi_n := \varphi * \alpha_n$. By property (ii) in Lemma 5.5.2 we have $\varphi_n(F) \to \varphi(F)$ *a.s.* and it follows by dominated convergence that

$$\varphi_n(F) \xrightarrow{L^p(\Omega)} \varphi(F).$$

By property (iv) of Lemma 5.5.2 we have that φ_n is continuously differentiable in direction J and its first order partial derivatives are bounded by γ. Now let $\omega \in \Omega_0 := \{\omega \in \Omega : F(\omega) \notin N\}$ be fixed and $i \in J$. Property (v) in Lemma 5.5.2 implies

$$\partial_i \varphi_n(F(\omega)) = (\partial_i \varphi * \alpha_n)(F(\omega)).$$

Since $\partial_i \varphi$ is continuous at $F(\omega)$, property (ii) in Lemma 5.5.2 yields

$$\partial_i \varphi_n(F(\omega)) \xrightarrow{n \to \infty} \partial_i \varphi(F(\omega)).$$

Thus, we have $\partial_i \varphi_n(F) D^L F^i \to \partial_i \varphi(F) D^L F^i$ almost surely. Because $|\partial_i \varphi_n(F)| \leq \gamma$ and $D^L F^i \in L^p(\Omega, \mathcal{H})$, the dominated convergence theorem yields

$$\partial_i \varphi_n(F) D^L F^i \xrightarrow{L^p(\Omega;\mathcal{H})} \partial_i \varphi(F) D^L F^i.$$

□

Corollary 5.5.6. *Under Assumption 5.5.3, let $B \in \mathcal{B}(\mathbb{R}^d)$ with $\mathbb{P}(F \in B) = 1$. We assume that on B the function φ is bounded and continuous as well as continuously partially differentiable in direction J. Further, suppose $\varphi_{|_B}$ is Lipschitz in direction J. Then $\varphi(F) \in \mathbb{D}^{1,p,L}$ and relation (5.9) holds.*

Proof. By Kirszbraun's Theorem, see e.g. Theorem 2.10.43 in [7], there exists an extension $\widetilde{\varphi}$ of $\varphi_{|_B}$ on $\mathbb{R}^d$ such that $\widetilde{\varphi}$ is globally Lipschitz continuous in direction J with the same Lipschitz constant as $\varphi_{|_B}$. Since $\Omega \setminus B$ is a $\mathbb{P}^F$-null set Proposition 5.5.5 yields that (5.9) holds for $\widetilde{\varphi}$. The result now follows from the fact that $\varphi(F) = \widetilde{\varphi}(F)$ in $L^p(\Omega)$. □

Theorem 5.5.7. *Under Assumption 5.5.3, let φ be locally Lipschitz in direction J on a closed set $B \in \mathcal{B}(\mathbb{R}^d)$, where $\mathbb{P}(F \in B) = 1$. Further, suppose that φ is continuous as well as continuously differentiable in direction J on $B \setminus N$, where $\mathbb{P}(F \in N) = 0$. In addition, we assume $\varphi(F) \in L^p(\Omega)$ and $\partial_i \varphi(F) D^L F^i \in L^p(\Omega; \mathcal{H})$ for all $i \in J$. Then the chain rule (5.9) holds.*

Proof. The proof is divided into two steps. We first suppose that φ is also bounded and show that (5.9) holds and then extend this result to the more general setting stated in the theorem.

Step 1: So, let φ be bounded and let $(a_n)_{n \in \mathbb{N}}$ be a sequence in $(0, \infty)$ such that

$$\mathbb{P}(F^i \neq a_n, \forall i \in \{1, \ldots, d\}) = 1, \ \forall n \in \mathbb{N}$$

and $a_n \to \infty$ as $n \to \infty$. Set $\varphi_n(x) = \varphi(-a_n \vee x \wedge a_n)$, where the minimum and maximum are understood component-wise, i.e.

$$-a_n \vee x \wedge a_n := \begin{pmatrix} h_n(x_1) \\ \vdots \\ h_n(x_d) \end{pmatrix}, \qquad h_n : \mathbb{R} \to \mathbb{R};\ y \mapsto \begin{cases} -a_n, & y < -a_n \\ y, & -a_n \leq y \leq a_n \\ a_n, & y > a_n \end{cases}.$$

Define

$$A_n := \{y \in B \setminus N \,|\, \forall i \in \{1, \dots, d\} : y_i \neq a_n\}.$$

We have $\mathbb{P}(F \in A_n) = 1$, φ_n is, on A_n, continuous and continuously differentiable in direction J, and $\varphi_n|_{A_n}$ is globally Lipschitz in direction J. Thus, $\varphi_n(F) \in \mathbb{D}^{1,p,L}$ and (5.9) holds for all φ_n by Corollary 5.5.6. We have $\varphi_n \to \varphi$ pointwise and $\|\varphi_n\|_\infty \leq \|\varphi\|_\infty$. Therefore, by dominated convergence,

$$\varphi_n(F) \xrightarrow{L^p(\Omega)} \varphi(F).$$

Moreover, we have $|\partial_i \varphi_n(x)| \leq |\partial_i \varphi(x)|,\ x \in \mathbb{R}^d$ and $|D^L F^i| \in L^p(\Omega, \mathcal{H})$ for $i \in J$, and it follows

$$\partial_i \varphi_n(F) D^L F^i \xrightarrow{L^p(\Omega,\mathcal{H})} \partial_i \varphi(F) D^L F^i$$

for all $i \in J$.

Step 2: We now drop the assumption of φ being bounded and let $(b_n)_{n \in \mathbb{N}}$ be a sequence in $(0, \infty)$ such that $\mathbb{P}(|\varphi(F)| = b_n) = 0$ for all $n \in \mathbb{N}$ and $b_n \to \infty$ as $n \to \infty$. With a similar notation to above we set $\varphi_n(x) := -b_n \vee \varphi(x) \wedge b_n$. It follows that φ_n is bounded, locally Lipschitz in direction J on B, and partially continuously differentiable in direction J for all $x \in B \setminus \big(N \cup \{x : |\varphi(x)| = b_n\}\big)$. By *step 1*, the chain rule holds for all φ_n. Using the dominated convergence theorem we obtain

$$\varphi_n(F) \xrightarrow{L^p(\Omega)} \varphi(F) \qquad \text{and} \qquad \partial_i \varphi_n(F) D^L F^i \xrightarrow{L^p(\Omega;\mathcal{H})} \partial_i \varphi(F) D^L F^i.$$

□

Note that choosing L as the identity operator, Theorem 5.5.7 also gives a more general chain rule result for the standard Malliavin derivative.

In the context of an absolute continuous random variable F on $\mathbb{R}$, the function φ, in general, cannot be discontinuous for a chain rule to hold. Consider, e.g., $\varphi : \mathbb{R} \to \mathbb{R},\ x \mapsto \mathbb{1}_{(-\infty,0]}(x)$ and $F = B_1 = W(\mathbb{1}_{[0,1]})$ in the setup of Example 5.4.7. As for $A \in \mathcal{F}$, $\mathbb{1}_A$ is Malliavin differentiable if and only if $\mathbb{P}(A) \in \{0, 1\}$ (cf. e.g. Proposition 1.2.6 in [38]), we have that $\varphi(F) = \mathbb{1}_{(-\infty,0]}(B_1) = \mathbb{1}_{\{B_1 \leq 0\}}$ is not Malliavin differentiable.

5.6 Auxiliary results

This section serves as an appendix for Chapter 5.

Lemma 5.6.1. *Let $(\Omega, \mathscr{A}, \mathcal{P})$ be a probability space and $Y \in L^1(\Omega, \mathscr{A}, \mathcal{P})$. Further, we assume that Π is a π-system, i.e. a non-empty family of subsets of Ω that is closed under finite intersection, with $\sigma(\pi) = \mathscr{A}$. If $\mathbb{E}[\mathbb{1}_A Y] = 0$ for all $A \in \Pi$, then $Y = 0$ almost surely.*

Proof. Let $Y_+ = Y\mathbb{1}_{\{Y\geq 0\}}$ and $Y_- = -Y\mathbb{1}_{\{Y<0\}}$. Then $Y = Y_+ - Y_-$ and we define measures ν_1, ν_2 on Π as

$$\nu_1(A) = \mathbb{E}[\mathbb{1}_A Y_+] \quad \text{and} \quad \nu_2(A) = \mathbb{E}[\mathbb{1}_A Y_-], \qquad A \in \Pi.$$

For any $A \in \Pi$ we have

$$0 = \mathbb{E}[\mathbb{1}_A Y] = \nu_1(A) - \nu_2(A)$$

and therefore ν_1 and ν_2 coincide on a π-system that generates the σ-algebra $\mathscr{A}$. It follows that $\nu_1 = \nu_2$ on $\mathscr{A}$ (see e.g. Lemma 1.42 in [25]). Thus, we have

$$\mathbb{E}[\mathbb{1}_B Y] = \nu_1(B) - \nu_2(B) = 0, \qquad \forall B \in \mathscr{A}.$$

Plugging in $B = \{Y \geq 0\} \in \mathscr{A}$ and $B = \{Y < 0\} \in \mathscr{A}$ gives us the assertion. $\square$

As above we denote by $I_p(g)$ the multiple stochastic Wiener integral over $g \in L^2([0,T]^p)$. In what follows let $\cdot$ and $\bullet$ be placeholders for different variables. In iterated integrals we always integrate over the variables represented by $\cdot$ and never over those represented by $\bullet$. To simplify notation in the following lemma we set $L^2([0,T]^0) := \mathbb{R}$ and I_0 the identity function on $\mathbb{R}$.

Lemma 5.6.2. *Let $p, q \in \mathbb{N}$ and $g \in L^2([0,T]^p), f \in L^2([0,T]^q)$. Then we have*

$$\Big\| \int_0^T I_{p-1}(g(t,\cdot)) f(t,\bullet)\,\mathrm{d}t - I_{p-1}\Big(\int_0^T g(t,\cdot) f(t,\bullet)\,\mathrm{d}t\Big)\Big\|_{L^2([0,T]^{q-1})} = 0 \quad a.s.$$

Proof. We write L^2_m for $L^2([0,T]^m)$. Let $(g^n)_{n\in\mathbb{N}}$ $((f^n)_{n\in\mathbb{N}})$ be a sequence of bounded, continuous functions approximating g (f) in L^2_p (L^2_q) with $|g^n(x)| \leq |g(x)|$ for all $x \in [0,T]^p$ $(|f^n(x)| \leq |f(x)|$ for all $x \in [0,T]^q)$ and all $n \in \mathbb{N}$. Stochastic Fubini (e.g. Theorem 64 in [43], p.210) yields that, for fixed $t \in [0,T]$ and $n \in \mathbb{N}$,

$$\int_0^T I_{p-1}(g^n(t,\cdot)) f^n(t,\bullet)\,\mathrm{d}t = I_{p-1}\Big(\int_0^T g^n(t,\cdot) f^n(t,\bullet)\,\mathrm{d}t\Big) \tag{5.10}$$

almost surely. The continuity of f^n, g^n together with a density argument yields that the null set for which (5.10) does not hold can be chosen simultaneously for all $t \in [0,T]$. It follows that

$$\Big\| \int_0^T I_{p-1}(g^n(t,\cdot)) f^n(t,\bullet)\,\mathrm{d}t - I_{p-1}\Big(\int_0^T g^n(t,\cdot) f^n(t,\bullet)\,\mathrm{d}t\Big)\Big\|_{L^2_{q-1}} = 0$$

almost surely for all $n \in \mathbb{N}$. As $L^2(\Omega)$ convergence implies almost sure convergence along a suitable subsequence, we consider such subsequences whenever we look at limits in the remainder of this proof. By dominated convergence, we obtain

$$\Big\| \int_0^T g^n(t,\cdot) f^n(t,\bullet)\,\mathrm{d}t - \int_0^T g(t,\cdot) f(t,\bullet)\,\mathrm{d}t \Big\|_{L^2_{p+q-2}} \xrightarrow{n\to 0} 0. \tag{5.11}$$

This implies with the help of the standard Fubini theorem that

$$\mathbb{E}\Big[\Big\| I_{p-1}\Big(\int_0^T g^n(t,\cdot) f^n(t,\bullet)\,\mathrm{d}t\Big) - I_{p-1}\Big(\int_0^T g(t,\cdot) f(t,\bullet)\,\mathrm{d}t\Big)\Big\|^2_{L^2_{q-1}}\Big]$$

$$= \frac{1}{(p-1)!}\Big\| \int_0^T g^n(t,\cdot)f^n(t,\bullet)\,\mathrm{d}t - \int_0^T g(t,\cdot)f(t,\bullet)\,\mathrm{d}t\Big\|^2_{L^2_{p+q-2}} \xrightarrow{n\to 0} 0,$$

and thus

$$\Big\| I_{p-1}\Big(\int_0^T g^n(t,\cdot)f^n(t,\bullet)\,\mathrm{d}t\Big) - I_{p-1}\Big(\int_0^T g(t,\cdot)f(t,\bullet)\,\mathrm{d}t\Big)\Big\|_{L^2_{q-1}} \xrightarrow{n\to 0} 0 \tag{5.12}$$

almost surely. It is easy to see that

$$I_{p-1}(g^n(t,\cdot)) \xrightarrow{L^2_1} I_{p-1}(g(t,\cdot)) \quad a.s.$$

as $n \to \infty$ and therefore

$$\int_0^T I_{p-1}(g^n(t,\cdot))f^n(t,\bullet)\,\mathrm{d}t \xrightarrow{L^2_{q-1}} \int_0^T I_{p-1}(g(t,\cdot))f(t,\bullet)\,\mathrm{d}t \quad a.s. \tag{5.13}$$

as $n \to \infty$. Putting (5.11) – (5.13) together yields

$$\begin{aligned}
&\Big\| \int_0^T I_{p-1}(g(t,\cdot))f(t,\bullet)\,\mathrm{d}t - I_{p-1}\Big(\int_0^T g(t,\cdot)f(t,\bullet)\,\mathrm{d}t\Big)\Big\|_{L^2([0,T]^{q-1})} \\
&= \lim_{n\to\infty} \Big\| \int_0^T I_{p-1}(g^n(t,\cdot))f^n(t,\bullet)\,\mathrm{d}t - I_{p-1}\Big(\int_0^T g^n(t,\cdot)f^n(t,\bullet)\,\mathrm{d}t\Big)\Big\|_{L^2_{q-1}} \\
&= 0 \quad a.s.
\end{aligned}$$

□

For the rest of this section we denote by $B_r(z)$ the ball around $z \in \mathbb{R}^d$ with radius $r > 0$ and by $\overline{B_r(z)}$ its closure, i.e.

$$\begin{aligned}
B_r(z) &:= \{y \in \mathbb{R}^d : \|z-y\| < r\}, \\
\overline{B_r(z)} &:= \{y \in \mathbb{R}^d : \|z-y\| \le r\}.
\end{aligned}$$

The proof of the next lemma can be found in standard text books on analysis. As the reader might not be familiar with mollifiers, we nevertheless give its proof.

Lemma 5.6.3. *Let $\beta \in C_0^\infty(\mathbb{R}^d)$, i.e. an infinitely differentiable and compactly supported function, and $f : \mathbb{R}^d \to \mathbb{R}$ be continuous. Then $f * \beta$ is continuous.*

Proof. Fix $q \in \mathbb{R}$ such that $\operatorname{supp}\beta \subseteq B_q(0)$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^d$ converging to $x \in \mathbb{R}^d$. W.l.o.g. $\|x - x_n\| \le q$. For $y \in \mathbb{R} \setminus \overline{B_{2q}(x)}$ we have

$$\|y - x_n\| = \|y - x - x_n + x\| \ge \big|\|y-x\| - \|x_n - x\|\big| \ge q.$$

Thus, $\beta(x_n - y) = 0$ for $y \notin \overline{B_{2q}(x)}$ and

$$\big|f(y)\beta(x_n - y)\big| \le \|\beta\|_\infty |f(y)| \mathbb{1}_{\overline{B_{2q}(x)}}(y),$$

where the right hand side is integrable. By dominated convergence, we have

$$\begin{aligned}
\lim_{n\to\infty}(f*\beta)(x_n) &= \lim_{n\to\infty}\int_{\mathbb{R}^d} f(y)\beta(x_n-y)\,\mathrm{d}y = \int_{\mathbb{R}^d} f(y)\lim_{n\to\infty}\beta(x_n-y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d} f(y)\beta(x-y)\,\mathrm{d}y = (f*\beta)(x).
\end{aligned}$$

□

Recall, letting $\alpha \in C^\infty(\mathbb{R}^d)$ be a nonnegative function with support on the unit ball and $\int_{\mathbb{R}^d} \alpha(x)\,\mathrm{d}x = 1$, we define

$$\alpha_n : \mathbb{R}^d \to \mathbb{R},\ x \mapsto n^d \alpha(nx), \qquad n \in \mathbb{N}.$$

The following lemma was given in the text.

Lemma 5.5.2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function and set $f_n = f * \alpha_n$, $n \in \mathbb{N}$ with α_n as defined above. The following properties hold:*

(i) For all $n \in \mathbb{N}$ we have $\displaystyle\int_{\mathbb{R}^d} \|x\|\alpha_n(x)\,\mathrm{d}x \le \frac{1}{n}$.

(ii) Let f be continuous at $x_0 \in \mathbb{R}^d$. Then $f_n(x_0) \to f(x_0)$ for $n \to \infty$.

(iii) Let f be continuous on $\mathbb{R}^d$. Then $f_n \in C^\infty(\mathbb{R}^d)$.

(iv) In addition to the continuity assumption in (iii), let f be Lipschitz continuous in direction $I \subseteq \{1,\dots,d\}$ with Lipschitz constant γ. Then, $\|\partial_i f_n\|_\infty \le \gamma$ for all $i \in I$. Moreover, for higher partial derivatives of f_n we have that for every $k \in \mathbb{N}$ there exists $c_k > 0$ such that

$$\sup_{x\in\mathbb{R}^d} \left|\partial_{i_1,\dots i_k} f_n(x)\right| \le c_k$$

for all $i_1,\dots,i_k \in I$.

(v) Assume that f is locally Lipschitz continuous in direction I. Then

$$\partial_i(f * \alpha_n) = \partial_i f * \alpha_n$$

almost everywhere for all $i \in I$.

Proof.

(i) We have

$$\int_{\mathbb{R}^d} \|x\|\alpha_n(x)\,\mathrm{d}x = \int_{\{x:\|x\|\le 1/n\}} \|x\|\alpha_n(x)\,\mathrm{d}x \le \frac{1}{n}.$$

(ii) Let $\varepsilon \ge 0$. Since f is continuous at x_0 there exists an $N \in \mathbb{N}$ such that

$$|f(x) - f(x_0)| \le \varepsilon, \quad x \in \left\{x \in \mathbb{R}^d : \|x - x_0\| \le \frac{1}{N}\right\}.$$

Thus, we have for $n \ge N$ that

$$\begin{aligned} |f_n(x_0) - f(x_0)| &= |(f * \alpha_n)(x_0) - f(x_0)| \\ &= \left|\int_{\mathbb{R}^d} f(x_0 - y)\alpha_n(y)\,\mathrm{d}y - f(x_0)\int_{\mathbb{R}^d} \alpha_n(y)\,\mathrm{d}y\right| \\ &\le \int_{\mathbb{R}^d} |f(x_0 - y) - f(x_0)|\alpha_n(y)\,\mathrm{d}y \le \varepsilon. \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, the assertion follows.

(iii) Let $i \in \{1, \dots, d\}$ and $x, y \in \mathbb{R}^d$. We have for $h < 1$ that

$$\left| \frac{\alpha_n(x - y + he_i) - \alpha_n(x - y)}{h} \right| \leq \|\partial_i \alpha_n\|_\infty \mathbb{1}_{B_2(x)}(y),$$

where $e_i \in \mathbb{R}^d$ denotes the vector that has a one in the i-th position and zeros otherwise. Therefore, we obtain by dominated convergence that

$$\begin{aligned} \partial_i(f * \alpha_n)(x) &= \lim_{h \to 0} \int_{\mathbb{R}^d} f(y) \frac{1}{h} \big(\alpha_n(x - y + he_i) - \alpha_n(x - y)\big) \, \mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(y) \lim_{h \to 0} \frac{1}{h} \big(\alpha_n(x - y + he_i) - \alpha_n(x - y)\big) \, \mathrm{d}y \\ &= (f * \partial_i \alpha_n)(x). \end{aligned}$$

So, by Lemma 5.6.3 and the calculations above $f * \alpha_n$ is partially differentiable in direction $i \in \{1, \dots, d\}$ with continuous partial derivatives $f * \partial_i \alpha_n$. For $k \in \mathbb{N}$ and $j = (j_1, \dots, j_k) \in \{1, \dots, d\}^k$ we define the operator $\Delta_j := \frac{\partial^k}{\partial_{j_1} \dots \partial_{j_k}}$. Iterating the calulation above then yields $\Delta_j(f * \alpha_n) = f * (\Delta_j \alpha_n)$.

(iv) To show the boundedness consider

$$\begin{aligned} |\partial_i(f * \alpha_n)(x)| &= \lim_{h \to 0} \left| \frac{1}{h} \big((f * \alpha_n)(x + he_i) - (f * \alpha_n)(x)\big) \right| \\ &= \lim_{h \to 0} \left| \frac{1}{h} \left(\int_{\mathbb{R}^d} f(x + he_i - y) \alpha_n(y) \, \mathrm{d}y - \int_{\mathbb{R}^d} f(x - y) \alpha_n(y) \, \mathrm{d}y \right) \right| \\ &\leq \lim_{h \to 0} \int_{\mathbb{R}^d} \left| \frac{1}{h} \big(f(x + he_i - y) - f(x - y)\big) \right| \alpha_n(y) \, \mathrm{d}y \\ &\leq \gamma \int_{\mathbb{R}^d} \alpha_n(y) \, \mathrm{d}y = \gamma, \end{aligned}$$

for all $i \in I$. Replacing α_n by $\Delta_j \alpha_n$, where $j = (j_1, \dots, j_k) \in I^k$ in the calculation above yields

$$|\partial_i(\Delta_j(f * \alpha_n))(x)| \leq \gamma \int_{\mathbb{R}^d} |\Delta_j \alpha_n(y)| \, \mathrm{d}y < \infty,$$

for all $i \in I$.

(v) First note that a function that is locally Lipschitz continuous in direction I is Lipschitz continuous in direction I on every compact set. Let $x \in \mathbb{R}^d$ be arbitrary but fixed and $i \in I$. In the same way as in (iv) we obtain

$$\partial_i(f * \alpha_n)(x) = \lim_{h \to 0} \int_{\mathbb{R}^d} \frac{1}{h} \big(f(x + he_i - y) - f(x - y)\big) \alpha_n(y) \, \mathrm{d}y$$

For $\|y\| > 1/n$ the integrand is zero and for $\|y\| \leq 1/n$ (and assuming $h < 1$) we have that $x + he_i - y, x - y \in B_2(x)$. Since f is locally Lipschitz in direction I, f is Lipschitz continuous in direction I on $\overline{B_2(x)}$ with some Lipschitz constant $\gamma(x) \geq 0$. It follows

$$\left| \frac{1}{h} \big(f(x + he_i - y) - f(x - y)\big) \right| \alpha_n(y) \leq \gamma(x) \alpha_n(y),$$

where the right-hand side is integrable with respect to y and independent of h. By Stepanov's Theorem (a consequence of Rademacher's Theorem, compare [7] Theorem 3.1.9) $\partial_i f$ exists almost everywhere and we obtain by dominated convergence

$$\begin{aligned}
&\lim_{h\to 0}\int_{\mathbb{R}^d}\frac{1}{h}\big(f(x+he_i-y)-f(x-y)\big)\alpha_n(y)\,\mathrm{d}y \\
&\quad=\int_{\mathbb{R}^d}\lim_{h\to 0}\frac{1}{h}\big(f(x+he_i-y)-f(x-y)\big)\alpha_n(y)\,\mathrm{d}y \\
&\quad=\int_{\mathbb{R}^d}\partial_i f(x-y)\alpha_n(y)\,\mathrm{d}y=(\partial_i f*\alpha_n)(x).
\end{aligned}$$

□

Bibliography

[1] M. Altmayer. *Quadrature of discontinuous SDE functionals using Malliavin integration by parts.* PhD thesis, University of Mannheim, 2015.

[2] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo quadrature of discontinuous payoffs in the generalized Heston model using Malliavin integration by parts. *SIAM Journal of Finacial Mathematics*, 6(1):22–52, 2015.

[3] P. Cheridito, H. Kawaguchi, and M. Maejima. Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability*, 8, 2003.

[4] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional brownian motions. *Probability Theory and Related Fields*, 122:108–140, 2002.

[5] A. Deya and S. Tindel. Malliavin Calculus for Fractional Heat Equation. In *Malliavin Calculus and Stochastic Analysis*, pages 361–384. Springer US, 2013.

[6] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Annals de l'Institut Henri Pointcaré*, 13(2):99–125, 1977.

[7] H. Federer. *Geometric measure theory.* Springer, 1969.

[8] G. B. Folland. *Real analysis: modern techniques and their applications.* Wiley, 2nd edition, 1999.

[9] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi. Applications of Malliavin calaculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3 (4):391–412, 1999.

[10] P. Friz and M. Hairer. *A Course on Rough Paths.* Springer, 2014.

[11] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths.* Cambridge University Press, 2010.

[12] P. Friz and N. Victoir. Differential equations driven by Gaussian signals. *Ann. Inst. H. Poincaré Probab. Statist.*, 46(2):369–413, 2010.

[13] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[14] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216:86–140, 2004.

[15] M. Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4): 693–721, 2010.

[16] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, pages 147–171, 1967.

[17] H. E. Hurst. Long-term storage capacity in reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 1951.

[18] J. E. Hutton and P. I. Nelson. Interchanging the order of differentiation and stochastic integration. *Stochastic Processes and their Applications*, 18(2):371–377, 1984.

[19] P. Imkeller, G. dos Reis, W. Salkeld, and G. Smith. Differentiability of SDEs with drifts of super-linear growth. *eprint arXiv:1803.06947*, 2018.

[20] M. Jolis and N. Viles. Continuity with respect to the Hurst parameter of the laws of the multiple fractional integrals. *Stochastic Processes and their Applications*, 117(9): 1189–1207, 2007.

[21] M. Jolis and N. Viles. Continuity in law with respect to the Hurst parameter of the local time of the fractional Brownian motion. *Journal of Theoretical Probability*, 20 (2):133–152, 2007.

[22] M. Jolis and N. Viles. Continuity in the Hurst parameter of the law of the Wiener integral with respect to the fractional Brownian motion. *Statistics & Probability Letters*, 80(7-8):566–572, 2010.

[23] M. Jolis and N. Viles. Continuity in the Hurst parameter of the law of the symmetric integral with respect to the fractional Brownian motion. *Stochastic Processes and their Applications*, 120(9):1651–1679, 2010.

[24] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, Second edition, 1998.

[25] A. Klenke. *Probability Theory*. Springer, 2nd edition, 2014.

[26] S. Koch. Directional malliavin derivatives: A characterisation of independence and a generalised chain rule. *Communications on Stochastic Analysis*, 12(2), 2018.

[27] S. Koch and A. Neuenkirch. The Mandelbrot-van Ness fractional Brownian motion is infinitely differentiable with respect to its Hurst parameter. *Discrete & Continuous Dynamical Systems-B*, 2019.

[28] M. Ledoux, T. Lyons, and Z. Qian. Lévy area of Wiener processes in Banach spaces. *The Annals of Probability*, 30(2):546–578, 2002.

[29] M. Ledoux, Z. Qian, and T. Zhang. Large deviations and support theorem for diffusion processes via rough paths. *Stochastic Processes and their Applications*, 102(2):265–283, 2002.

[30] T. Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L.C. Young. *Mathematical Research Letters*, 1:451–464, 1994.

[31] T. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[32] T. J. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths : Ecole d'Eté de Probabilités de Saint-Flour XXXIV - 2004*. Springer, 2007.

[33] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pages 195–263, 1978.

[34] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, 10(4):422–437, 10 1968.

[35] I. Nourdin. *Calcul stochastique généralisé et applications au mouvement brownien fractionaire; Estimation non-paramétrique de la volatilité et test d'adéquation*. PhD thesis, Université Henri Poincaré - Nancy I, 2004.

[36] I. Nourdin. A simple theory for the study of SDEs driven by a fractional Brownian motion, in dimension one. In *Séminaire de Probabilités XLI*, pages 181–197. Springer, 2008.

[37] I. Nourdin. *Selected Aspects of Fractional Brownian Motion*. Springer, 2012.

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

[39] D. Nualart and E. Pardoux. Stochastic Calculus with Anticipating Integrands. *Probability Theory and Related Fields*, 78:535–581, 1988.

[40] D. Nualart and A. Răşcanu. Differential equations driven by fractional Brownian motion. *Collect. Math.*, 53(1):55–81, 2002.

[41] D. Nualart and B. Saussereau. Malliavin calculus for stochastic differential equations driven by a fractional brownian motion. *Stochastic Processes and their Applications*, 119:391–409, 2009.

[42] D. Nualart and S. Tindel. A construction of the rough path above fractional Brownian motion using Volterra's representation. *The Annals of Probability*, 39(3):1061–1096, 2011.

[43] P. E. Protter. *Stochastic integration and differential equations : Version 2.1*. Stochastic modelling and applied probability ; 21. Springer, 2. ed., corr. 3. print edition, 2005.

[44] F. Russo and P. Vallois. Forward, backward and symmetric stochastic integration. *Probability Theory and Related Fields*, 97:403–421, 1993.

[45] H. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Annals of Probability*, 6(1):19–41, 1978.

[46] D. Talay and A. Richard. Noise sensitivity of functionals of fractional Brownian motion driven stochastic differential equations: Results and perspectives. In *Springer Proc. Math. Stat.*, volume 208, pages 219–235, 2017.

[47] A. S. Üstünel and M. Zakai. On independence and conditioning on Wiener space. *The Annals of Probability*, 17(4), 1989.

[48] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica*, 67:251–282, 1936.