Fractional Stochastic Calculus via Stochastic Sewing

zur Erlangung des Grades eines Doktors der Naturwissenschaften

am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

vorgelegt von

Toyomu Matsuda

Berlin, 2023

Erstgutachter:	Prof. Dr. Nicolas Perkowski
Zweitgutachter:	Prof. Dr. Ivan Nourdin

Tag der Disputation: 12.09.2023

Abstract

The thesis explores stochastic calculus for fractional Brownian motion. Our approach builds upon a novel technique called stochastic sewing, originally introduced by Lê [*Electron. J. Probab.* 25:1-55, 2020]. The stochastic sewing has been effectively used to obtain sharp estimates on stochastic Riemann sums.

The main result of the thesis is an extension of Lê's stochastic sewing, which we refer to as the shifted stochastic sewing. This extension takes advantage of asymptotic decorrelation in stochastic Riemann sums and can be seen as a combination of Lê's stochastic sewing and the asymptotic independence formulated by Picard [*Ann. Probab.* 36(6): 2235-2279, 2008].

As applications of the shifted stochastic sewing, we address two important problems in fractional stochastic calculus. Firstly, we characterize the local time of the fractional Brownian motion via level crossings, extending the classical work of Lévy to the fractional setting. Secondly, we establish the pathwise uniqueness of Young and rough differential equations driven by fractional Brownian motion. This result optimizes the regularity of the noise coefficient, which is consistent with the Brownian setting.

Additionally, we demonstrate strong regularization by fractional noise for differential equations with integrable drifts. This result can be viewed as a fractional analogue of the celebrated work by Krylov and Röckner [*Probab. Theory Relat. Fields* 131: 154–196, 2005].

Zusammenfassung

In dieser Arbeit wird der stochastische Kalkül für die gebrochene Brownsche Bewegung untersucht. Unser Ansatz basiert auf einer neuen Technik namens "stochastic sewing", die ursprünglich von Lê [*Electron. J. Probab.* 25:1-55, 2020] eingeführt wurde. Das "stochastic sewing" wird effektiv eingesetzt, um optimale Abschätzungen für stochastische Riemann-Summen zu erhalten.

Das Hauptergebnis dieser Arbeit ist eine Erweiterung von Lês "stochastic sewing", die wir als verschobenes "stochastic sewing" bezeichnen. Diese Erweiterung macht sich die asymptotische Dekorrelation in stochastischen Riemann-Summen zunutze und kann als eine Kombination von Lês "stochastic sewing" und der von Picard formulierten asymptotischen Unabhängigkeit gesehen werden [*Ann. Probab.* 36(6): 2235-2279, 2008].

Als Anwendungen des verschobenen "stochastic sewing" behandeln wir zwei wichtige Probleme des gebrochenen stochastischen Kalküls. Zum einen charakterisieren wir die Lokalzeit der gebrochenen Brownschen Bewegung durch Überquerungen von Niveaulinien, und erweitern damit die klassische Arbeit von Lévy auf den gebrochenen Fall. Zum anderen etablieren wir die pfadweise Eindeutigkeit von Young- und irregulären Differentialgleichungen, die durch gebrochene Brownsche Bewegung angetrieben werden. Dieses Ergebnis optimiert die Regularitätsannahmen des Diffusionskoeffizienten, in Einklang mit dem Brownschen Fall.

Zusätzlich zeigen wir eine starke Regularisierung durch gebrochenes Rauschen für Differentialgleichungen mit integrierbarem Drift. Dieses Ergebnis kann als ein gebrochenes Analogon der berühmten Arbeit von Krylov und Röckner [*Probab. Theory Relat. Fields* 131: 154-196, 2005] angesehen werden.

Selbstständigkeitserklärung

Name: Matsuda
Vorname: Matsuda

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht. Diese Dissertation wurde in gleicher oder ähnlicher Form noch in keinem früheren Promotionsverfahren eingereicht.

Acknowledgement

First and foremost, I would like to express my deepest gratitude to Nicolas Perkowski. All the results presented in this thesis originate from his profound mathematical insights. His unwavering support, even for seemingly mundane results, has been invaluable. It is nearly impossible to imagine completing my PhD without his supervision.

I am also deeply grateful to Ivan Nourdin for agreeing to referee this thesis.

In addition to Nicolas, I would like to extend my warm thanks to my other coauthors — Oleg Butkovsky, Purba Das, Khoa Lê, Rafał Łochowski, Avi Mayorcas, and Willem van Zuijlen. I feel truly blessed to have had the opportunity to collaborate with such talented mathematicians. Special thanks go to Khoa and Willem. Obviously, this thesis would not exist without Khoa's profound works. Besides, he has essentially acted as a second supervisor for this thesis. Willem has also acted as a second supervisor for the project on Anderson Hamiltonians, although it is not covered in this thesis.

Nicolas has built an exceptional research group, and I have had the privilege of sharing wonderful moments with its members — Ana, Helena, Henri, Huanyu, Immanuel, Lukas, Tomasso, Wei, Willem, Xiaohao. We have had countless enjoyable discussions during our lunches, coffee breaks, and social outings. Many of these discussions have served as the starting point for some results presented in this thesis. I would also like to express my gratitude to Silvia, the group's secretary, for her tremendous support and engaging conversations.

My PhD work has been financially supported by the IRTG 2544 "Stochastic Analysis in Interaction," which has provided an exceptional working environment. I am especially grateful to Peter Bank, the spokesperson of the program, for his tireless efforts in maintaining such a stimulating environment. I would also like to acknowledge my colleagues in the program — Adrian, Alexander, Alexandra, Dave, Emanuel, Florian, Francesca, Gideon, Hannes, Konstantins, Laura, Likai, Philipp, Robert, Sascha, Sebastian, Simon, Thomas, Weile, and Yizheng — for making my time in Germany so enriching. I would also like to thank Peter Friz for organizing seminars and facilitating discussions related to the topics covered in this thesis. Additionally, I am grateful to Jean-Dominique Deuschel for serving as a BMS mentor. The program has also provided me with the opportunity to spend valuable time in Oxford, where significant progress was made. I extend my thanks to Rama Cont for hosting me, and to my fellow PhD researchers, Ruhong and Sarah, for the enjoyable time we had together.

I would like to acknowledge Kajino Naotaka and Yuzuro Inahama, my supervisors during my Bachelor's and Master's studies in Japan. Without their guidance and support, I would not have embarked on a PhD journey in stochastic analysis. I would also like to express my gratitude to Xue-Mei Li, whose interest in my work has motivated me to continue as a postdoc under her.

Finally, I would like to thank my parents, who have *always* trusted in me.

松田 響

Contents

Introduction **1**
Bibliography 15

1 Shifted stochastic sewing lemma **17**
1.1 Introduction and the main theorem 18
1.2 Proof of the main theorem 22
1.3 Integration along fractional Brownian motions 28
1.4 Proofs of technical results 38
Bibliography 47

2 Level crossings of fractional Brownian motions **49**
2.1 Introduction 50
2.1.1 Main results 52
2.2 Variations along Lebesgue partitions 56
2.2.1 Elementary results 56
2.2.2 Convergence of the variations 61
2.3 Local time via Lebesgue partitions 66
2.3.1 Heuristics 66
2.3.2 Convergence to the local time 68
2.3.3 Lemiuex type result 83
2.4 A fractional excursion measure 86
Bibliography 93

3 Young and rough differential equations driven by fractional Brownian motion **95**
3.1 Introduction 96
3.2 Preliminaries 101
3.2.1 Function spaces and heat semigroup estimates 102
3.2.2 SDE driven by fractional Brownian motion 104

3.2.3 Stochastic estimates . . . 107
3.3 Young case . . . 109
3.3.1 Young integral with irregular integrand . . . 110
3.3.2 Pathwise uniqueness . . . 116
3.4 Rough case . . . 121
3.4.1 Rough integral with irregular integrand . . . 123
3.4.2 Iterated integrals . . . 130
3.4.3 Pathwise uniqueness . . . 134
Bibliography . . . 139

4 Strong regularization of differential equations with integrable drifts by fractional noise . . . 143
4.1 Introduction . . . 144
4.1.1 Literatures and discussions . . . 148
4.1.2 Strategies of proofs . . . 149
4.2 Weak well-posedness . . . 152
4.2.1 Preliminaries . . . 153
4.2.2 Weak solutions and stochastic estimates . . . 156
4.3 Strong well-posedness . . . 165
4.3.1 Some technical estimates . . . 166
4.3.2 Pathwise uniqueness . . . 168
4.4 Stability . . . 174
4.4.1 VMO estimates . . . 177
4.4.2 Lifting paths . . . 181
4.4.3 Gronwall's argument in the probabilistic setting . . . 182
Bibliography . . . 187

Introduction

During the 19th century, Robert Brown made a significant observation regarding the irregular movement of particles within a medium. This motion, now widely recognized as Brownian motion, or the Wiener process in honor of the mathematician who laid its mathematical groundwork, has had a profound impact on various fields, including modern mathematics. Brownian motion, denoted by W, is a centered Gaussian process characterized by the following property (in one dimension):

$$\mathbb{E}[(W_t - W_s)^2] = t - s, \quad s < t.$$

The process exhibits independent increments and possesses martingale and Markovian properties, which paved the way for the development of a comprehensive theory on Brownian motion. In the 1940s, Itô initiated the field of stochastic calculus, which involves the calculus with respect to Brownian motion. This field has evolved into one of the most fruitful areas in mathematics, as demonstrated in the monograph [RY99].

However, in practical applications, Brownian motion is often considered too ideal. To address this, the *fractional Brownian motion* B^H, indexed by $H \in (0, 1)$, was introduced. It is a centered Gaussian process characterized by the following property:

$$\mathbb{E}[(B_t^H - B_s^H)^2] = (t - s)^{2H}, \quad s < t.$$

The parameter H represents the roughness of the process, as depicted in Figure 1. When $H = 1/2$, the process reduces to the standard Brownian motion. In other cases, the process exhibits correlated increments and is neither a martingale nor Markovian. Kolmogorov [Kol40] first introduced this process, and it was later popularized by Mandelbrot [MV68; Man82]. Naturally, the field of *fractional stochastic calculus* emerged to handle calculus involving fractional Brownian motion.

Since fractional Brownian motion is neither a martingale nor Markovian, many of the arguments used in Itô's stochastic calculus cannot be directly applied to fractional stochastic calculus. Consequently, researchers have developed two main tools in fractional stochastic calculus. The first tool involves pathwise arguments, such as Young's integration theory

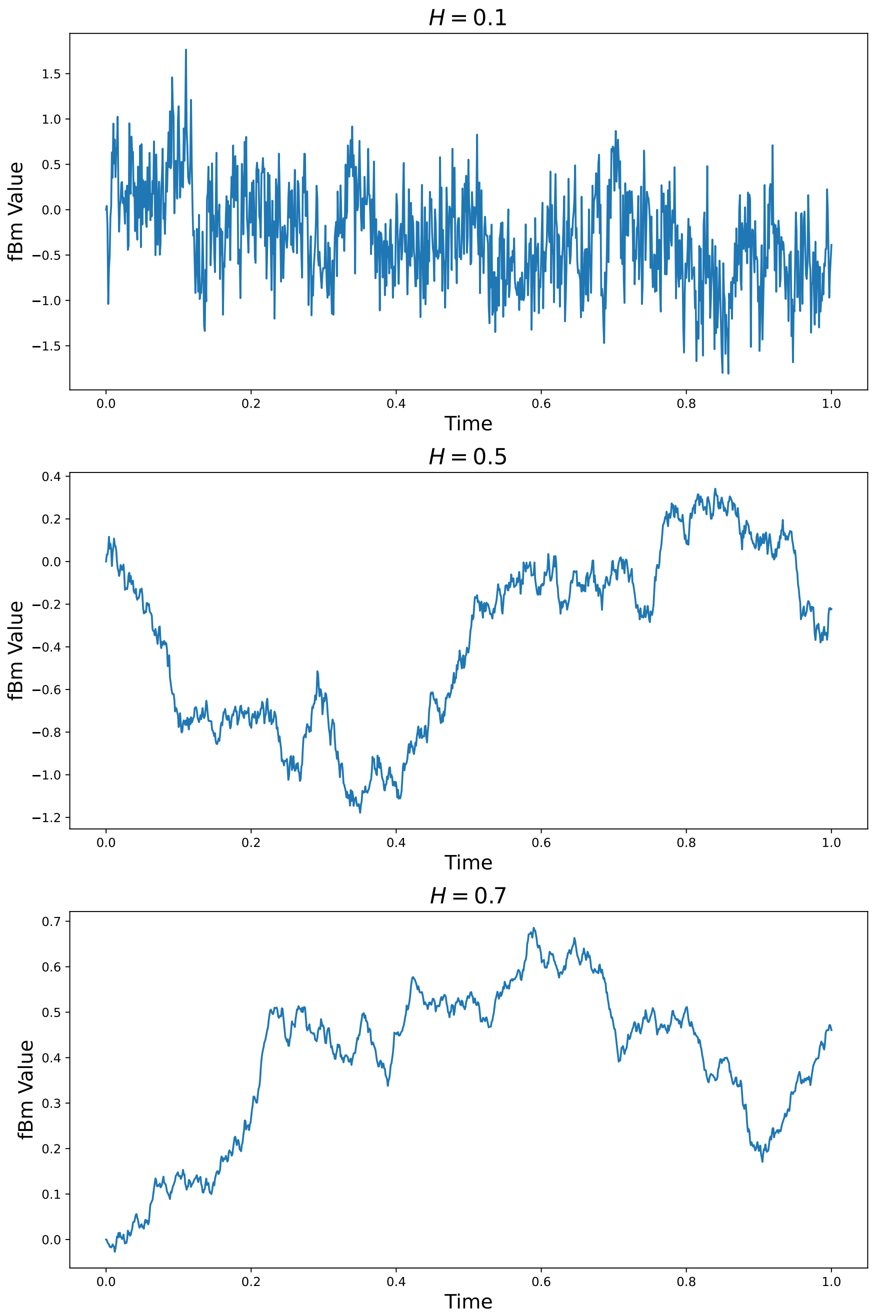


Figure 1: Simulating fractional Brownian motions for $H = 0.1, 0.5, 0.7$.

[You36], Lyons' rough path theory [Lyo98], and Zähle's fractional calculus [Zäh98]. These approaches fix the realization of the fractional Brownian motion, and perform pathwise analysis. The second tool is Malliavin calculus. This calculus was first invented by Malliavin [Mal78] to obtain a probabilistic proof of Hörmander's theorem [Hör67], but it turns out to be also useful to study probabilistic aspects of the fractional Brownian motion [Nua06; Nou12].

However, it is important to note that these tools often yield less precise results compared to the classical Brownian setting. For example, consider the stochastic integral

$$\int_0^T f(B_r^H)\,\mathrm{d}B_r^H \tag{1}$$

for $H \in (1/4, 1)$, and for $H < 1/2$, the integral is understood as a rough integral. Typically, the integral (in multi dimensions) is defined for functions f with Hölder regularity $(1-H)/H$. However, it is natural to suspect that this definition is not optimal, as the Itô integral (1) is well-defined for any bounded measurable f when $H = 1/2$.

Recently, Lê [Lê20] combined the martingale inequality (Burkholder–Davis–Gundy inequality) and Gubinelli's sewing lemma [Gub04] to obtain the *stochastic sewing lemma*. This lemma provides sharp stochastic estimates on stochastic Riemann sums, including the stochastic integral

$$\int_0^T f(B_r^H)\mathrm{d}r,$$

where f can be an irregular function or even a distribution. The stochastic sewing lemma quickly gained recognition for its innovation and has become a central force in the recent development of regularization by noise.

This thesis aims to provide a new perspective on fractional stochastic calculus through the stochastic sewing lemma. Our results are on par with their Brownian counterparts. For instance, we establish the well-definedness of the integral (1) for f of Hölder regularity $(1/(2H) - 1 + \varepsilon)$, for any positive ε. Our main contribution is a novel version of the stochastic sewing lemma, which we call the *shifted* stochastic sewing (Chapter 1). This new version offers the advantage of capturing the asymptotic decorrelation in the stochastic Riemann sums. It can be viewed as a combination of Lê's stochastic sewing and the asymptotic independence introduced by Picard [Pic08]. As applications of the shifted stochastic sewing, we investigate partitions defined by level crossings of fractional Brownian motions (Chapter 2) and study Young and rough differential equations driven by fractional Brownian motions (Chapter 3). Additionally, we derive precise results on regularization by fractional noise for integrable drifts (Chapter 4), which significantly improve upon previous works and align with the results obtained by Krylov and Röckner for the Brownian case [KR05].

In the following sections, we provide more detailed descriptions of each chapter.

Chapter 1: Shifted stochastic sewing

Chapter 1 is the most important part of the thesis, which serves as the foundation for Chapters 2 and 3. The content of this chapter is based on joint work with Nicolas Perkowski.

In the fields of analysis and probability theory, the convergence and the estimate of Riemann sums plays a crucial role. These sums are expressed as

$$\sum_{[s,t]\in\pi} A_{s,t}, \tag{2}$$

where π represents a partition of the interval $[0,T]$. The focus lies on the limit as the mesh size

$$|\pi| := \max_{[s,t]\in\pi} |t-s|$$

tends to 0. The term $A_{s,t}$ is called a germ. For example, when $A_{s,t} := f(s)(t-s)$, we consider a Riemann sum approximation of $\int_0^T f(s)\,ds$. Similarly, when $A_{s,t} := X_s(W_t - W_s)$, where W is a Brownian motion and X is an adapted process, we study the Itô approximation of the stochastic integral $\int_0^T X_r\,dW_r$.

Gubinelli [Gub04], inspired by Lyons' results on almost multiplicative functionals in the theory of rough paths [Lyo98], established the remarkable *sewing lemma*. This lemma states that if the quantity

$$\delta A_{s,u,t} := A_{s,t} - A_{s,u} - A_{u,t}, \quad 0 \le s < u < t \le T,$$

satisfies $|\delta A_{s,u,t}| \lesssim |t-s|^{1+\varepsilon}$ for some $\varepsilon > 0$, then the sums (2) converge. The sewing lemma has proven to be immensely powerful, leading to numerous applications and extensions in the field. Notably, it has been utilized for defining rough integrals, as described in the monographs [Gub04; FH20].

When $(A_{s,t})_{s\le t}$ is random and we aim to prove the convergence of the sums (2), Gubinelli's sewing lemma is often insufficient. For instance, if $A_{s,t} := (W_t - W_s)^2$, the sums converge in $L^m(\mathbb{P})$, $m < \infty$, to the quadratic variation of the Brownian motion. However, we only expect the bound

$$\|\delta A_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim_m |t-s|,$$

and hence we cannot apply the sewing lemma.

In his seminal work, Lê [Lê20] obtained a stochastic version of Gubinelli's sewing lemma. Just as Gubinelli's sewing lemma plays an important role in *pathwise* stochastic calculus, Lê's stochastic sewing lemma does so in *probabilistic* stochastic calculus. In particular, the discovery of the stochastic sewing has significantly advanced the field of regularization by noise.

A concrete statement of the stochastic sewing lemma is as follows. If $(A_{s,t})_{s<t}$ is a stochastic germ adapted to a filtration $(\mathcal{F}_t)$ and if

$$\|\delta A_{s,u,t}\|^2_{L^m(\mathbb{P})} + \|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]\|_{L^m(\mathbb{P})} \leq \Gamma(t-s)^{1+\varepsilon},$$

for $s < u < t$, $m \in [2, \infty)$, $\Gamma \in (0, \infty)$ and $\varepsilon > 0$, then the Riemann sums $\sum_{[s,t]\in\pi} A_{s,t}$, where π is a partition of some fixed interval $[0, T]$, converge in $L^m(\mathbb{P})$ as the mesh size of π tends to 0. The strength of the stochastic sewing lemma lies in the fact that we only need to assume $\left(\frac{1}{2} + \varepsilon\right)$-regularity for $\|\delta A_{s,u,t}\|_{L^m(\mathbb{P})}$, although we also need to consider the regularization effect encoded in the estimate $\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]\|_{L^m(\mathbb{P})} \lesssim (t-s)^{1+\varepsilon}$. Furthermore, if we denote by $\mathcal{A}_T$ the limit of the Riemann sums in the interval $[0, T]$, we have a quantitative bound

$$\|\mathcal{A}_{s,t}\|_{L^m(\mathbb{P})} \lesssim_{m,\varepsilon} \Gamma(t-s)^{\frac{1+\varepsilon}{2}}.$$

That is, we can transfer the estimate of $A_{s,t}$ to that of $\mathcal{A}_{s,t}$.

Sometimes, it is difficult to observe the regularization effect through $\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]\|_{L^m(\mathbb{P})}$. The easiest example is $A_{s,t} = |B^H_t - B^H_s|^{1/H}$, the $1/H$-variation of the fractional Brownian motion B^H. For this example, it is not possible to estimate $\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]$, although the convergence of the Riemann sums (along equipartitions) is well known.

Chapter 1 of the thesis presents an extension of Lê's stochastic sewing (Theorem 1.1.1), relaxing the estimate of the conditional expectation $\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]$. We replaced it with

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v]\|_{L^m(\mathbb{P})} \lesssim (s-v)^{-\alpha}(t-s)^{1+\varepsilon}, \quad v < s < u < t, \quad \alpha < \frac{1}{2} + \varepsilon. \tag{3}$$

Because of this new condition, where the conditioning is shifted from $\mathcal{F}_s$ to $\mathcal{F}_v$, we call this extension the *shifted* stochastic sewing lemma. The case where $\alpha = 0$ and $v = s$ corresponds to Lê's stochastic sewing. The version of the mild shifting, namely the case where $\alpha = 0$ and $v = s - M(t-s)$ for a fixed positive constant M, is obtained by Gerencsér [Ger22]. Our extension allows us to take advantage of the *asymptotic* effect of regularization, inspired by [Pic08].

For the example $A_{s,t} = |B_t - B_s|^{1/H}$, we can prove estimates of the form (3). Additionally, as a more interesting application, we demonstrate the convergence of Itô approximations and Stratonovich approximations under low regularity assumptions, which can be viewed as a simplification and improvement of [Nou12, Theorem 3.5]. More precisely, in Section 1.3 we prove

$$\exists \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} f(B_s^H)(B_t^H - B_s^H) \quad \text{in } L^m(\mathbb{P}) \text{ for } H \in (1/2, 1) \text{ and } f \in L^\infty(\mathbb{R}^d)$$

and with $H \in (1/4, 1/2)$ and $\gamma > \frac{1}{2H} - 1$ we prove

$$\exists \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \frac{f(B_s^H) + f(B_t^H)}{2}(B_t^H - B_s^H) \quad \text{in } L^m(\mathbb{P}) \text{ for } f \in C^\gamma(\mathbb{R}^d).$$

Furthermore, the shifted stochastic sewing lemma will be crucially applied in Chapters 2 and 3. The reader however can skip the proof of Theorem 1.1.1 (Section 1.2 and Section 1.4) without any problem for further reading. The result of Section 1.3 will be used in Chapter 3.

Chapter 2: Level crossings of fractional Brownian motions

In this section, we provide a summary of Chapter 2, which is based on collaborative work with Purba Das, Rafał Łochowski, and Nicolas Perkowski.

We consider a fractional Brownian motion B^H with a Hurst parameter $H \in (0, 1)$. It is known that the following convergence holds:

$$\lim_{n\to\infty} \sum_{k=0}^{2^n-1} \left| B^H_{\frac{k+1}{2^n}T} - B^H_{\frac{k}{2^n}T} \right|^{1/H} = \mathbb{E}[|B_1^H|^{1/H}]T, \quad \text{a.s.}$$

One of the key objectives of Chapter 2 is to investigate the $(1/H)$-variations along *Lebesgue partitions*, which are random partitions defined by the level crossings of B^H. To construct these partitions, we start with $T_0^n := 0$ and recursively define the stopping times T_k^n by

$$T_k^n := \inf\{t > T_{k-1}^n : |B_t^H - B^H_{T_{k-1}^n}| = 2^{-n}\}.$$

The reader can refer to Figure 2 for an illustration. Each nth Lebesgue partition consists of intervals of the form $[T_{k-1}^n, T_k^n]$ for $k \in \mathbb{N}$ satisfying $T_k^n \leq T$. The main objective is to establish the convergence of $(1/H)$-variations along these Lebesgue partitions. In particular, we aim to prove the existence of the limit:

$$\lim_{n\to\infty} 2^{-n/H} \#\{k : T_k^n \leq T\}, \tag{4}$$

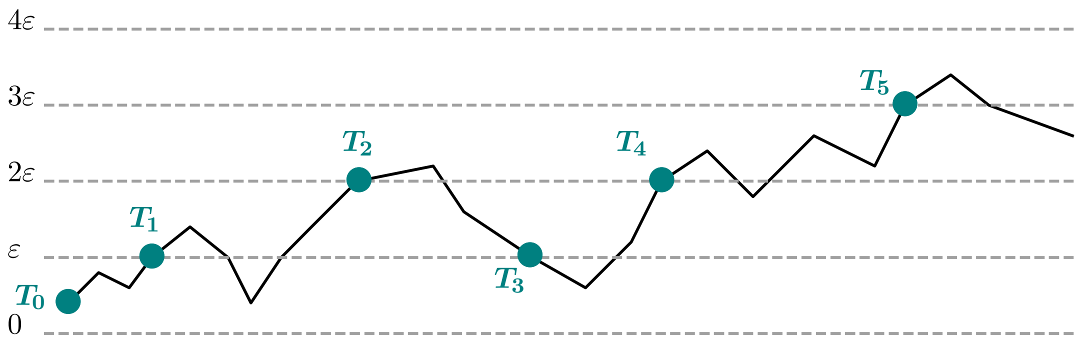


Figure 2: Stopping times T_k^n.

where $\#$ denotes the cardinality.

Additionally, we can count level crossings specifically around the level 0 and investigate the convergence of the number of these crossings towards the local time at 0.

The convergence of (4) and its local time counterpart is well-known for the Brownian case, where the proof relies on martingale or Markovian properties, such as Itô's formula (see, e.g., [RY99]). However, such martingale or Markovian arguments are not applicable when $H \neq 1/2$. Chapter 2 presents a completely different strategy to prove (4) for $H \in (0, 1)$ and its local time counterpart for $H < 1/2$. This novel approach resolves a conjecture posed in [CP19].

To demonstrate our strategy, we denote by $K_{s,t}(\varepsilon, w)$ the number of ε-level crossings of the process w in the interval $[s, t]$. Then the limit (4) is equal to

$$\lim_{n\to\infty} 2^{-n/H} K_{0,T}(2^{-n}, B).$$

The key observation is that the family $(K_{s,t}(\varepsilon, B))_{0\leq s<t\leq T}$ is superadditive and almost subadditive. This leads to the approximation

$$K_{0,T}(2^{-n}, B) \approx \sum_{[s,t]\in\pi_n} K_{s,t}(2^{-n}, B),$$

where π_n is a partition of $[0, T]$ with a mesh size of order $2^{-n/H}$. Hence, we can approximate $K_{0,T}(2^{-n}, B)$ by a stochastic Riemann sum, which can then be estimated by the shifted stochastic sewing. To verify the conditions of the shifted stochastic sewing, the computations will be carried out in the spirit of Picard [Pic08]. The strategy of proving the convergence to the local time follows similar arguments, but due to the lack of stationarity, technicalities dramatically increase.

Our result also poses a very interesting conjecture on whether the limit (4) is equal to that of $(1/H)$-variations along deterministic dyadic partitions. For $H = 1/2$, they are equal. Numerical simulation suggests that they are different for $H \neq 1/2$. If this were indeed true, it would be an interesting manifestation of non-Markovianity.

Chapter 3: Probabilistic Young and rough differential equations

This is the summary of Chapter 3, based on joint work with Avi Mayorcas. In this chapter, we focus on the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t^H, \quad X_0 = x \in \mathbb{R}^{d_1}, \tag{5}$$

where σ is a map valued in the space of $d_1 \times d_2$ matrices, and B^H is a d_2-dimensional fractional Brownian motion with Hurst parameter $H \in (1/3, 1)$. The differential equation is interpreted either as Young's differential equation ($H > 1/2$) or as Lyons' rough differential equation [Lyo98] ($H < 1/2$). Our main result is on the pathwise uniqueness of (5) under a low regularity assumption on σ. Rather than directly stating the result, we provide a background review leading up to this outcome.

For $H = 1/2$, we often apply Itô's theory to study (5); we will discuss the alternative theory of Lyons later. In Itô's theory, there are a few notions of solutions and their uniqueness, among which the most relevant to us is the notion of *pathwise uniqueness*. It says that two solutions, adapted to some filtration making the driving Brownian motion martingale, must be indistinguishable. Hence, pathwise uniqueness is a *probabilistic* concept of uniqueness (despite its name). It is a classical result, as can be found in all textbooks of stochastic calculus, that pathwise uniqueness holds for (5) with $H = 1/2$ provided that σ is Lipschitz. The proof is a consequence of Itô's isometry: for an adapted process Y we have

$$\mathbb{E}\Big[\Big|\int_0^T Y_r \mathrm{d}B_r^{1/2}\Big|^2\Big] = \mathbb{E}\Big[\int_0^T |Y_r|^2 \mathrm{d}r\Big].$$

Itô's isometry is due to the martingale property of the Brownian motion. Since B^H, $H \neq 1/2$, is not a martingale (nor Markovian), Itô's theory is not available for $H \neq 1/2$. Lack of probabilistic tools naturally motivates us to study the SDE (5) *pathwisely*. Based on Young's integration theory, Lyons [Lyo94] showed that the (deterministic) differential equation

$$\mathrm{d}x_t = f(x_t)\mathrm{d}y_t \tag{6}$$

driven by a path y of finite p-variation with $p \in [1,2)$ has a unique solution provided that f is α-Hölder with $\alpha > p$. Furthermore, [Lyo98] extended the result for $p \in [2,\infty)$, provided that we are additionally given "iterated integrals"

$$\int_s^t \int_s^{r_1} \mathrm{d}y_{r_2}\mathrm{d}y_{r_1}, \int_s^t \int_s^{r_1} \int_s^{r_2} \mathrm{d}y_{r_3}\mathrm{d}y_{r_2}\mathrm{d}y_{r_1}, \ldots,$$

satisfying certain analytic and algebraic conditions. The tuple of y and its (sufficient number of) iterated integrals is called a rough path of y. Later, Coutin and Qian [CQ02] proved that the fractional Brownian motion B^H, with $H > 1/4$, can be naturally lifted to a rough path. Since B^H has finite p-variation for any $p > 1/H$, we see that (5) has a unique solution provided that $\sigma \in C^\gamma$ with $\gamma > 1/H$ and $H \in (1/4, 1)$.

We remark two important differences in Itô's probabilistic theory and Lyons' pathwise theory. One is that the former considers uniqueness among solutions adapted to a given filtration, while the latter considers uniqueness among all solutions satisfying (6), which do not need to be adapted. In other words, the notion of uniqueness is stronger in the pathwise theory, referred to as *path-by-path uniqueness*, following the works of Davie [Dav07; Dav08]. The other difference lies in the regularity assumption on σ. When $H = 1/2$, Itô's theory assumes that σ is only Lipschitz, while Lyons' theory assumes that $\sigma \in C^\gamma$ with $\gamma > 2$. In summary, Itô's theory requires less regularity assumption on σ at the cost of a weaker notion of uniqueness.

Although Itô's theory is not available for $H \neq 1/2$, pathwise uniqueness is a well-defined notion in this setting. Now it is natural to ask if we can prove pathwise uniqueness of (5) for $\sigma \in C^\gamma$ with $\gamma < 1/H$. Our main result in this chapter answers the question affirmatively. That is, under the uniformly elliptic condition ($\sigma\sigma^{\mathrm{T}}$ is non-degenerate), we prove pathwise uniqueness under $\sigma \in C^\gamma$ with $\gamma > \max\{1/(2H), (1-H)/H\}$, as shown in Figure 3.

The proof follows Lê's strategy [Lê20] to prove pathwise uniqueness. In fact, pathwise uniqueness is deduced from a sharp estimate on the stochastic integral

$$\int f(X_r)\mathrm{d}B_r,$$

where X is a path controlled by B and f is a map of low regularity. Specifically, we can define the stochastic integral for $f \in C^\gamma$ with $\gamma > 1/(2H) - 1$. The estimate is proven using stochastic sewing techniques, including the shifted stochastic sewing.

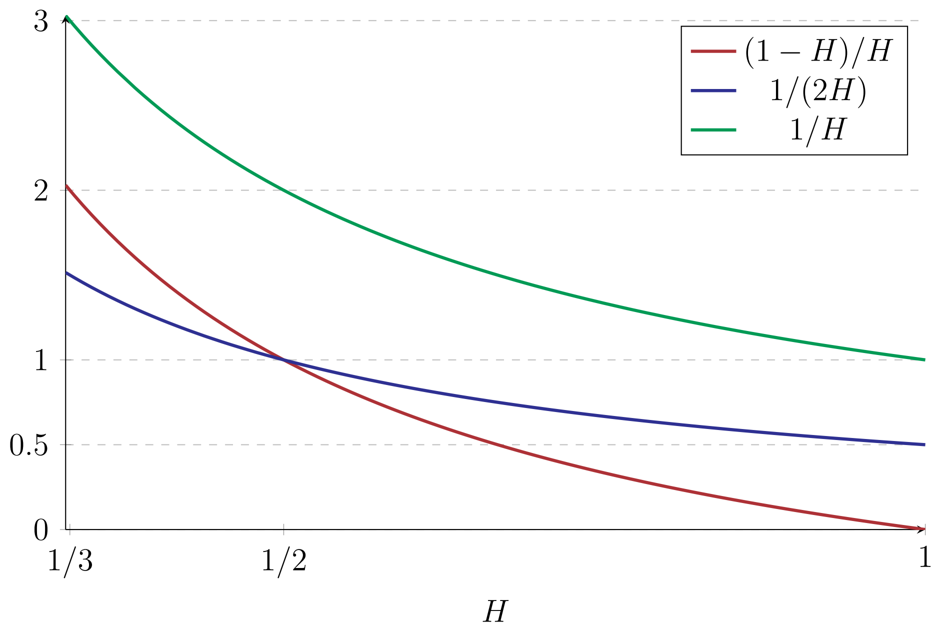


Figure 3: Some graphs of H related to the main result of Chapter 3. Pathwise theory covers $\sigma \in C^\gamma$ with $\gamma > 1/H$ (green), while the result of Chapter 3 says that pathwise uniqueness holds if $\gamma > 1/(2H)$ (blue) and if $\gamma > (1-H)/H$ (red).

Chapter 4: Regularization by fractional noise for integrable drift

This is the summary of Chapter 4, based on joint work with Oleg Butkovsky and Khoa Lê. This chapter is relatively independent as it does not rely on the shifted stochastic sewing. The main result focuses on the strong well-posedness of the fractional SDE

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}B_t^H. \tag{7}$$

The well-posedness is straightforward when b is smooth, but our interest lies in the case of non-smooth b. Here, we consider an integrable drift $b \in L^q([0,T]; L^p(\mathbb{R}^d))$, and our goal is to determine the condition on p, q, d, H for strong well-posedness. The topic discussed in this chapter falls within the highly active field of *regularization by noise*, which will be further explored and reviewed in the following paragraphs.

Ill-posed differential equations sometimes regain well-posedness by introducing noise. For instance, the differential equation $\mathrm{d}X_t = \sqrt{|X_t|}\mathrm{d}t$ may have multiple solutions, while the stochastic differential equation (SDE) $\mathrm{d}X_t = \sqrt{|X_t|}\mathrm{d}t + \mathrm{d}B_t^{1/2}$ has a unique strong solution (strongly well-posed). This phenomenon is known as *regularization by noise*. Recently, there has been growing interest in understanding this phenomenon beyond the Brownian setting; among them is regularization by fractional noise.

In their recent work [GG23], Galeati and Gerencsér introduced the notion of subcriticality for fractional SDEs. Subcriticality refers to the domination of fractional noise under small scales. If X solves the SDE (7), then the scaled process $X_t^{(\lambda)} = \lambda^{-H} X_{\lambda t}$ solves the SDE

$$\mathrm{d}X_t^{(\lambda)} = \lambda^{1-H-\frac{dH}{p}-\frac{1}{q}} b^{(\lambda)}(t, X_t^{(\lambda)})\mathrm{d}t + \mathrm{d}B_t^{(\lambda)},$$

where

$$b^{(\lambda)}(t,x) = \lambda^{\frac{dH}{p}+\frac{1}{q}} b(\lambda t, \lambda^H x), \quad B_t^{(\lambda)} = \lambda^{-H} B_{\lambda t}.$$

It is worth noting that $\|b^{(\lambda)}\|_{L_t^q L_x^p} = \|b\|_{L_t^q L_x^p}$ and $B^{(\lambda)}$ has the same law as B. The domination of the noise at small scales implies that the order of the drift term is smaller than that of the driving noise as λ approaches 0. This leads to the condition

$$1 - H - \frac{dH}{p} - \frac{1}{q} > 0. \tag{8}$$

Therefore, the condition (8) is natural for the solution theory of (7). In fact, the celebrated result by Krylov and Röckner [KR05] proves strong well-posedness for $H = 1/2$ under (8). The main result of Chapter 4 addresses the strong well-posedness of (7) in the

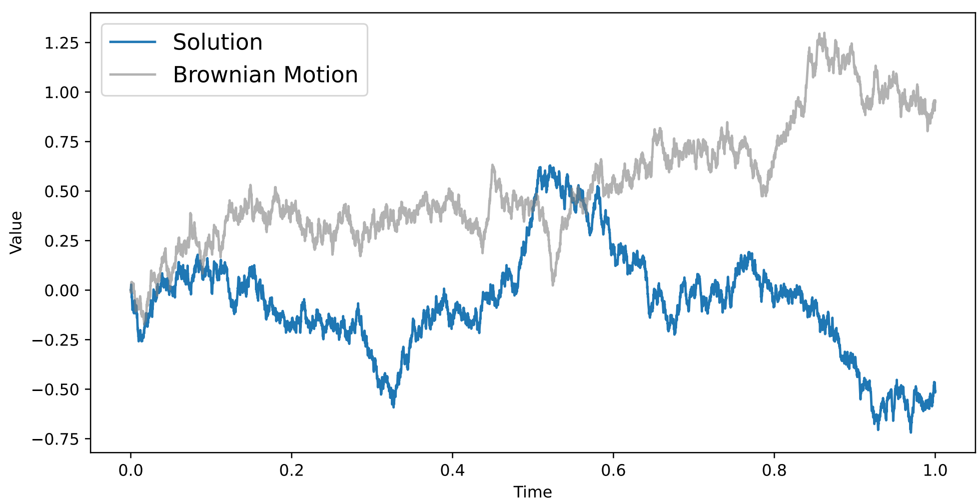


Figure 4: Simulation of the SDE $\mathrm{d}X_t = \sqrt{|X_t|}\mathrm{d}t + \mathrm{d}B_t^{1/2}$ with $X_0 = 0$. The differential equation $\mathrm{d}x_t = \sqrt{|x_t|}\mathrm{d}t$ with $x_0 = 0$ has multiple solutions, but the SDE has a unique strong solution. Heuristically, the solution of the SDE must behave like Brownian motion, which enforces a unique way for the solution to escape the singularity at 0.

fractional case $H < 1/2$. Specifically, we prove strong well-posedness for $H < 1/2$, $p \geq \max\{(1-H)^{-1}, 2dH\}$, and (8). We also consider the case where $p < (1-H)^{-1}$ with an additional condition. Additionally, we establish the stability of (7) with respect to the initial condition and the drift b. Our arguments are based on the stochastic sewing with control [Lê23] and the notion of processes of vanishing mean oscillation introduced in another seminal work by Lê [Lê22].

Reading guide

$$
\begin{tikzcd}
\boxed{\text{Section 1.1}} \arrow[d] \arrow[dr] & \boxed{\text{Section 1.3}} \arrow[d] & \\
\boxed{\text{Chapter 2}} & \boxed{\text{Chapter 3}} & \boxed{\text{Chapter 4}}
\end{tikzcd}
$$

The relation of each chapter is depicted in the above diagram. The statement of the shifted stochastic sewing (Theorem 1.1.1) appears in Section 1.1, and the result will be used in Chapters 2 and 3. However, the reader can skip its rather involved proof (Sections 1.2 and 1.4) for further reading. We remark that Chapter 4 is essentially independent of the preceding chapters.

Introduction of each chapter includes a section on notation specific to that chapter. Below, we collect the most frequently used notations:

- We use the notation $A := B$ to indicate that A is defined by B.

- The symbol $\mathbb{N}$ represents the set of natural numbers $\{1, 2, \ldots\}$, $\mathbb{Q}$ represents the set of rational numbers, and $\mathbb{R}$ represents the set of real numbers.

- We denote by $\mathbf{1}_A$ the indicator function for the set A.

- We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the underlying probability space, which is often implicit. The symbol $\mathbb{E}$ denotes the expectation. We write $\mathbb{E}[\cdot|\mathcal{G}]$ for the conditional expectation given $\mathcal{G}$. We set

$$\|F\|_{L^m(\mathbb{P})} := \Big(\int_\Omega |F(\omega)|^m \mathrm{d}\mathbb{P}(\omega) \Big)^{\frac{1}{m}}$$

  with usual convention for $m = \infty$.

- The d-dimensional fractional Brownian motion with Hurst parameter $H \in (0,1)$ is represented as $B^H = (B^{H,i})_{i=1}^d$. The components of B^H are independent. In Chapter 2, it takes values in $\mathbb{R}$ ($d = 1$), and in Chapter 3, it takes values in $\mathbb{R}^{d_2}$ ($d = d_2$). We typically use the symbol W to denote the Brownian motion.

- For a given map $f\colon [0,T] \to \mathbb{R}^d$, we write $f_{s,t} := f_t - f_s$.

- The notation $A \lesssim B$ signifies that there exists a constant C depending only on irrelevant parameters such that $A \le CB$. If we want to emphasize the dependency on $\alpha, \beta, \ldots$, we write $A \lesssim_{\alpha,\beta,\ldots} B$. We often write $C = C(\alpha, \beta, \ldots)$ to emphasize that the constant C depends on $\alpha, \beta, \ldots$.

I hope that you enjoy the reading :)

Bibliography

[CP19] R. Cont and N. Perkowski. "Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity". *Trans. Amer. Math. Soc. Ser. B* 6 (2019), pp. 161–186.

[CQ02] L. Coutin and Z. Qian. "Stochastic analysis, rough path analysis and fractional Brownian motions". *Probability Theory and Related Fields* 122.1 (2002), pp. 108–140.

[Dav07] A. M. Davie. "Uniqueness of Solutions of Stochastic Differential Equations". *International Mathematics Research Notices* 2007 (Dec. 2007), rnm124.

[Dav08] A. M. Davie. "Differential equations driven by rough paths: an approach via discrete approximation". *Appl. Math. Res. Express. AMRX* (2008), Art. ID abm009, 40.

[FH20] P. K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Cham: Springer International Publishing, 2020.

[GG23] L. Galeati and M. Gerencsér. *Solution theory of fractional SDEs in complete subcritical regimes*. 2023. arXiv: 2207.03475 `[math.PR]`.

[Ger22] M. Gerencsér. "Regularisation by regular noise". *Stoch. PDE: Anal. Comp.* (2022).

[Gub04] M. Gubinelli. "Controlling rough paths". *J. Funct. Anal.* 216.1 (2004), pp. 86–140.

[Hör67] L. Hörmander. "Hypoelliptic second-order differential equations". *Acta Mathematica* 119.1 (1967), pp. 147–171.

[Kol40] A. Kolmogorov. "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum". *C.R. (Doklady) Acad. Sci. URSS (N.S.)* 26 (1940), pp. 115–118.

[KR05] N. V. Krylov and M. Röckner. "Strong solutions of stochastic equations with singular time dependent drift". *Probability Theory and Related Fields* 131.2 (2005), pp. 154–196.

[Lê20] K. Lê. "A stochastic sewing lemma and applications". *Electron. J. Probab.* 25 (2020), Paper No. 38, 55.

[Lê22] K. Lê. *Quantitative John–Nirenberg inequality for stochastic processes of bounded mean oscillation*. 2022. arXiv: 2210.15736 `[math.PR]`.

[Lê23] K. Lê. "Stochastic sewing in Banach spaces". *Electronic Journal of Probability* 28.none (Jan. 2023), pp. 1–22.

[Lyo94] T. Lyons. “Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young”. *Math. Res. Lett.* 1.4 (1994), pp. 451–464.

[Lyo98] T. J. Lyons. “Differential equations driven by rough signals”. *Rev. Mat. Iberoamericana* 14.2 (1998), pp. 215–310.

[Mal78] P. Malliavin. “Stochastic calculus of variations and hypoelliptic operators”. In: *Proc. Inter. Symp. on Stoch. Diff. Equations, Kyoto 1976*. Wiley, 1978, pp. 195–263.

[Man82] B. B. Mandelbrot. *The fractal geometry of nature*. W. H. Freeman and Co., San Francisco, Calif., 1982, pp. v+460.

[MV68] B. B. Mandelbrot and J. W. Van Ness. “Fractional Brownian Motions, Fractional Noises and Applications”. *SIAM Review* 10.4 (1968), pp. 422–437.

[Nou12] I. Nourdin. *Selected aspects of fractional Brownian motion*. Vol. 4. Bocconi & Springer Series. Springer, Milan; Bocconi University Press, Milan, 2012, pp. x+122.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*. Second. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006, pp. xiv+382.

[Pic08] J. Picard. “A tree approach to p-variation and to integration”. *Ann. Probab.* 36.6 (2008), pp. 2235–2279.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Third. Vol. 293. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.

[You36] L. C. Young. “An inequality of the Hölder type, connected with Stieltjes integration”. *Acta Math.* 67.1 (1936), pp. 251–282.

[Zäh98] M. Zähle. “Integration with respect to fractal functions and stochastic calculus. I”. *Probab. Theory Related Fields* 111.3 (1998), pp. 333–374.

Chapter 1

Shifted stochastic sewing lemma

We give an extension of Lê's stochastic sewing lemma [Lê20]. The stochastic sewing lemma proves convergence in $L^m(\mathbb{P})$ of Riemann type sums $\sum_{[s,t]\in\pi} A_{s,t}$ for an adapted two-parameter stochastic process A, under certain conditions on the moments of $A_{s,t}$ and of conditional expectations of $A_{s,t}$ given $\mathcal{F}_s$. Our extension replaces the conditional expectation given $\mathcal{F}_s$ by that given $\mathcal{F}_v$ for $v < s$, and it allows us to make use of asymptotic decorrelation properties between $A_{s,t}$ and $\mathcal{F}_v$ by including a singularity in $(s-v)$. As a first application, we prove the convergence of Itô or Stratonovich approximations of stochastic integrals along fractional Brownian motions under low regularity assumptions. Further applications can be found in the following chapters.

This chapter is based on joint work with Nicolas Perkowski.

Keywords and phrases. stochastic sewing lemma, fractional Brownian motion, stochastic integrals.
MSC 2020. 60G22, 60H05.

Contents

1.1 Introduction and the main theorem 18

1.2 Proof of the main theorem 22

1.3 Integration along fractional Brownian motions 28

1.4 Proofs of technical results 38

Bibliography 47

1.1 Introduction and the main theorem

In analysis and probability theory, we often consider the convergence of sums

$$\sum_{[s,t]\in\pi} A_{s,t}. \tag{1.1}$$

Here π is a partition of an interval $[0,T]$, and we consider the limit of

$$|\pi| := \max_{[s,t]\in\pi} |t-s| \to 0.$$

For instance, if $A_{s,t} := f(s)(t-s)$, then we consider a Riemann sum approximation of $\int_0^T f(s)\mathrm{d}s$, and if $A_{s,t} := X_s(W_t - W_s)$, where W is a Brownian motion and X is an adapted process, then we consider the Itô approximation of the stochastic integral $\int_0^T X_r \mathrm{d}W_r$.

Gubinelli [Gub04], inspired by Lyons' results on almost multiplicative functionals in the theory of rough paths [Lyo98], showed that if

$$\delta A_{s,u,t} := A_{s,t} - A_{s,u} - A_{u,t}, \quad 0 \le s < u < t \le T, \tag{1.2}$$

satisfies $|\delta A_{s,u,t}| \lesssim |t-s|^{1+\varepsilon}$ for some $\varepsilon > 0$, then the sums (1.1) converge. This result is now called the *sewing lemma*, named so in the work of Feyel and de La Pradelle [FL06]. This lemma is so powerful that many applications and many extensions are known. For instance, it can be used to define rough integrals, see [Gub04] and the monograph [FH20] of Friz and Hairer.

When $(A_{s,t})_{s\le t}$ is random and when we want to prove the convergence of the sums (1.1), the above sewing lemma is often not sufficient. For instance, if $A_{s,t} := (W_t - W_s)^2$, the sums converge in $L^m(\mathbb{P})$, $m < \infty$, to the quadratic variation of the Brownian motion. However, we only have

$$\|\delta A_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim_m |t-s|,$$

and hence we cannot apply the sewing lemma.

Lê [Lê20] proved a stochastic version of the sewing lemma (*stochastic sewing lemma*): if a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is given such that

- $A_{s,t}$ is $\mathcal{F}_t$-measurable and
- for some $\varepsilon_1, \varepsilon_2 > 0$ and $m \in [2,\infty)$, we have for every $s < u < t$,

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]\|_{L^m(\mathbb{P})} \lesssim |t-s|^{1+\varepsilon_2}, \tag{1.3}$$

$$\|\delta A_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim |t-s|^{\frac{1}{2}+\varepsilon_1}, \tag{1.4}$$

then the sums (1.1) converge in $L^m(\mathbb{P})$. If $A_{s,t} := (W_t - W_s)^2$, then we have $\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s] = 0$ and (1.4) is satisfied with $\varepsilon_1 = \frac{1}{2}$. Therefore, we can prove the convergence of (1.1) in $L^m(\mathbb{P})$. The stochastic sewing lemma has been already shown to be very powerful in the original work [Lê20] of Lê, and an increasing number of papers are appearing that take advantage of the lemma.

However, there are situations where Lê's stochastic sewing lemma seems insufficient. For instance, consider

$$A_{s,t} := |B_t - B_s|^{\frac{1}{H}}, \tag{1.5}$$

where B is a fractional Brownian motion with Hurst parameter $H \in (0,1)$. It is well known that the sums (1.1) converge to $c_H T$ in $L^m(\mathbb{P})$. Although we have the estimate (1.4), we fail to obtain the estimate (1.3) unless $H = \frac{1}{2}$.

To get an idea on how Lê's stochastic sewing lemma should be modified for this problem, observe the following trivial fact:

$$\mathbb{E}[\delta A_{s,u,t}] = 0.$$

This suggests that we consider estimates that interpolate $\mathbb{E}[\delta A_{s,u,t}]$ and $\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s]$. In fact, we can obtain the following estimates:

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v]\|_{L^m(\mathbb{P})} \lesssim_H \left(\frac{t-s}{s-v}\right)^{1-H}(t-s), \quad 0 \le v < s < u < t \le T. \tag{1.6}$$

We can prove (1.6) for instance by applying Picard's result [Pic08, Lemma A.1] on the asymptotic independence of fractional Brownian increments. This discussion motivates the following main theorem of this chapter.

Theorem 1.1.1 (shifted stochastic sewing)**.** *Suppose that we have a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and a family of $\mathbb{R}^d$-valued random variables $(A_{s,t})_{0\le s\le t\le T}$ such that $A_{s,s}=0$ for every $s\in[0,T]$ and such that $A_{s,t}$ is $\mathcal{F}_t$-measurable. We define $\delta A_{s,u,t}$ by (1.2). Furthermore, suppose that there exist constants*

$$m\in[2,\infty),\quad \Gamma_1,\Gamma_2,M\in[0,\infty),\quad \alpha,\beta_1,\beta_2\in[0,\infty)$$

such that the following conditions are satisfied.

- *For every $0\le t_0<t_1<t_2<t_3\le T$, we have*

$$\|\mathbb{E}[\delta A_{t_1,t_2,t_3}|\mathcal{F}_{t_0}]\|_{L^m(\mathbb{P})}\le \Gamma_1(t_1-t_0)^{-\alpha}(t_3-t_1)^{\beta_1},\quad \text{if } M(t_3-t_1)\le t_1-t_0, \tag{1.7}$$

$$\|\delta A_{t_0,t_1,t_2}\|_{L^m(\mathbb{P})}\le \Gamma_2(t_2-t_0)^{\beta_2}. \tag{1.8}$$

- *We have*

$$\beta_1>1,\quad \beta_2>\frac{1}{2},\quad \beta_1-\alpha>\frac{1}{2}. \tag{1.9}$$

Then, there exists a unique, up to modifications, $\mathbb{R}^d$-valued stochastic process $(\mathcal{A}_t)_{t\in[0,T]}$ with the following properties.

- *$\mathcal{A}_0=0$, $\mathcal{A}_t$ is $\mathcal{F}_t$-measurable and $\mathcal{A}_t$ belongs to $L^m(\mathbb{P})$.*
- *There exist non-negative constants C_1, C_2 and C_3 such that*

$$\|\mathbb{E}[\mathcal{A}_{t_2}-\mathcal{A}_{t_1}-A_{t_1,t_2}|\mathcal{F}_{t_0}]\|_{L^m(\mathbb{P})}\le C_1|t_1-t_0|^{-\alpha}|t_2-t_1|^{\beta_1}, \tag{1.10}$$

$$\|\mathcal{A}_{t_2}-\mathcal{A}_{t_1}-A_{t_1,t_2}\|_{L^m(\mathbb{P})}\le C_2|t_2-t_1|^{\beta_1-\alpha}+C_3|t_2-t_1|^{\beta_2}, \tag{1.11}$$

  where $t_2-t_1\le M^{-1}(t_1-t_0)$ is assumed for the inequality (1.10).

In fact, we can choose C_1, C_2 and C_3 so that

$$C_1\lesssim_{\beta_1}\Gamma_1,\quad C_2\lesssim_{\alpha,\beta_1,\beta_2,M}\kappa_{m,d}\Gamma_1,\quad C_3\lesssim_{\alpha,\beta_1,\beta_2,M}\kappa_{m,d}\Gamma_2,$$

where $\kappa_{m,d}$ is the constant of the Burkholder-Davis-Gundy inequality, see (1.13). Furthermore, for $\tau\in[0,T]$, if we set

$$A^{\pi}_{\tau}:=\sum_{[s,t]\in\pi}A_{s,t},\quad \text{where } \pi \text{ is a partition of } [0,\tau],$$

then the family $(A^{\pi}_{\tau})_{\pi}$ converges to $\mathcal{A}_{\tau}$ in $L^m(\mathbb{P})$ as $|\pi|$ tends to 0.

Remark 1.1.2. The proof shows that if

$$1 + \alpha - \beta_1 < 2\alpha\beta_2 - \alpha, \tag{1.12}$$

then we have $C_2 \lesssim_{\alpha,\beta_1,\beta_2,M} \Gamma_1$, and we can omit the factor $\kappa_{m,d}$. This is similar to [Lê20], where C_2 also does not depend on $\kappa_{m,d}$. If $\alpha = 0$ and $M = 0$, Theorem 1.1.1 recovers Lê's stochastic sewing lemma [Lê20, Theorem 2.1]. If $\alpha = 0$ and $M > 0$, it recovers a lemma [Ger22, Lemma 2.2] by Gerencsér. The version of Gerencsér is often called the shifted stochastic sewing, and we continue to call Theorem 1.1.1 the same way.

Remark 1.1.3. The proof shows that there exists $\varepsilon = \varepsilon(\alpha, \beta_1, \beta_2) > 0$ such that

$$\|\mathcal{A}_\tau - A^\pi_\tau\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M,m,d,T} (\Gamma_1 + \Gamma_2)|\pi|^\varepsilon$$

for every $\tau \in [0, T]$ and every partition π of $[0, \tau]$.

Remark 1.1.4. As in another work [Lê23] of Lê, it should be possible to extend Theorem 1.1.1 so that the stochastic process $(A_{s,t})_{s,t\in[0,T]}$ takes values in a certain Banach space.

Remark 1.1.5. A multidimensional version of the sewing lemma is the *reconstruction theorem* [Hai14, Theorem 3.10] of Hairer. A stochastic version of the reconstruction theorem was obtained by Kern [Ker21]. It could be possible to extend Theorem 1.1.1 in the multidimensional setting, but we will not pursue it here.

The proof of Theorem 1.1.1 is given in Section 1.2. If $A_{s,t}$ is given by (1.5), then we can apply Theorem 1.1.1 with

$$\alpha = 1 - H, \quad \beta_1 = 2 - H, \quad \beta_2 = 1.$$

However, the application of Theorem 1.1.1 goes beyond this simple problem of $\frac{1}{H}$-variation of the fractional Brownian motion. Indeed, in Section 1.3 we prove the convergence of Itô and Stratonovich approximations to the stochastic integrals

$$\int_0^T f(B_s)\mathrm{d}B_s \quad \text{and} \quad \int_0^T f(B_s) \circ \mathrm{d}B_s$$

with $H > \frac{1}{2}$ in Itô's case and with $H > \frac{1}{6}$ in Stratonovich's case, under rather general assumptions on the regularity of f, in fact $f \in C^2_b(\mathbb{R}^d, \mathbb{R}^d)$ works for all $H > \frac{1}{6}$.

More interesting applications can be found in the following chapters. In Chapter 2, we will prove the convergence of the level-crossing counting of the fractional Brownian motion to its local time, and in Chapter 3, we will consider fractional Young and rough differential equations with irregular noise coefficients.

Notation

We write $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ and $\mathbb{N} := \{1, 2, \dots\}$. Given a function $f : [S, T] \to \mathbb{R}^d$, we write $f_{s,t} := f_t - f_s$. We denote by $\kappa_{m,d}$ the best constant of the discrete Burkholder-Davis-Gundy (BDG) inequality for $\mathbb{R}^d$-valued martingale differences [BDG72]. Namely, if we are given a filtration $(\mathcal{F}_n)_{n=1}^{\infty}$ and a sequence $(X_n)_{n=1}^{\infty}$ of $\mathbb{R}^d$-valued random variables such that X_n is $\mathcal{F}_n$-measurable for every $n \geq 1$ and $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0$ for every $n \geq 2$, then

$$\|\sum_{n=1}^{\infty} X_n\|_{L^m(\mathbb{P})} \leq \kappa_{m,d} \|\sum_{n=1}^{\infty} X_n^2\|_{L^{\frac{m}{2}}(\mathbb{P})}^{\frac{1}{2}}. \tag{1.13}$$

Rather than (1.13), we mostly use the inequality

$$\|\sum_{n=1}^{\infty} X_n\|_{L^m(\mathbb{P})} \leq \kappa_{m,d} \Big(\sum_{n=1}^{\infty} \|X_n\|_{L^m(\mathbb{P})}^2 \Big)^{\frac{1}{2}} \tag{1.14}$$

for $m \geq 2$, which follows from (1.13) by Minkowski's inequality. We write $A \lesssim B$ or $A = O(B)$ if there exists a non-negative constant C such that $A \leq CB$. To emphasize the dependence of C on some parameters $a, b, \dots$, we write $A \lesssim_{a,b,\dots} B$.

1.2 Proof of the main theorem

The overall strategy of the proof is the same as that of the original work [Lê20] of Lê. Namely, we combine the argument of the deterministic sewing lemma ([Gub04], [FL06] and Yaskov [Yas18]) with the discrete BDG inequality [BDG72]. However, the proof of Theorem 1.1.1 requires more labor at technical level. Some proofs will be postponed to Appendix 1.4.

As in [Lê20], the following lemma, which originates from [Yas18], will be needed. It allows us to replace general partitions by dyadic partitions.

Lemma 1.2.1 ([Lê20, Lemma 2.14])**.** *Under the setting of Theorem 1.1.1, let*

$$0 \leq t_0 < t_1 < \dots < t_{N-1} < t_N \leq T.$$

Then, we have

$$A_{t_0,t_N} - \sum_{i=1}^{N} A_{t_{i-1},t_i} = \sum_{n \in \mathbb{N}_0} \sum_{i=0}^{2^n-1} R_i^n, \tag{1.15}$$

where

$$R_i^n := \delta A_{s_1^{n,i}, s_2^{n,i}, s_3^{n,i}} + \delta A_{s_1^{n,i}, s_3^{n,i}, s_4^{n,i}}, \tag{1.16}$$

and

$$n \in \mathbb{N}_0, \quad i \in \{0, 1, \dots, 2^n - 1\}, \quad s_j^{n,i} \in [t_0 + \frac{i(t_N - t_0)}{2^n}, t_0 + \frac{(i+1)(t_N - t_0)}{2^n}],$$

and where $R_i^n = 0$ *for all sufficiently large* n.

The next two lemmas correspond to the estimates [Lê20, (2.50) and (2.51)] respectively.

Lemma 1.2.2. *Under the setting of Theorem 1.1.1, let*

$$0 \le s < t_0 < t_1 < \cdots < t_{N-1} < t_N \le T, \quad t_N - t_1 \le \frac{t_0 - s}{M}.$$

Then,

$$\|\mathbb{E}[A_{t_0,t_N} - \sum_{i=1}^{N} A_{t_{i-1},t_i} | \mathcal{F}_s]\|_{L^m(\mathbb{P})} \lesssim_{\beta_1} \Gamma_1 |t_0 - s|^{-\alpha} |t_N - t_0|^{\beta_1}.$$

Proof. In view of the decomposition (1.15), the triangle inequality gives

$$\|\mathbb{E}[A_{t_0,t_N} - \sum_{i=1}^{N} A_{t_{i-1},t_i} | \mathcal{F}_s]\|_{L^m(\mathbb{P})} \le \sum_{n \in \mathbb{N}_0} \sum_{i=0}^{2^n - 1} \|\mathbb{E}[R_i^n | \mathcal{F}_s]\|_{L^m(\mathbb{P})}.$$

By (1.7) and (1.16),

$$\|\mathbb{E}[R_i^n | \mathcal{F}_s]\|_{L^m(\mathbb{P})} \le 2\Gamma_1 (t_0 - s)^{-\alpha} (2^{-n} |t_N - t_0|)^{\beta_1} = 2\Gamma_1 2^{-n\beta_1} |t_0 - s|^{-\alpha} |t_N - t_0|^{\beta_1}.$$

Therefore, recalling $\beta_1 > 1$ from (1.9), the claim follows. $\square$

Lemma 1.2.3. *Under the setting of Theorem 1.1.1, let*

$$0 \le t_0 < t_1 < \cdots < t_{N-1} < t_N \le T.$$

Then,

$$\|A_{t_0,t_N} - \sum_{i=1}^{N} A_{t_{i-1},t_i}\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M} \kappa_{m,d} \Gamma_1 |t_N - t_0|^{\beta_1 - \alpha} + \kappa_{m,d} \Gamma_2 |t_N - t_0|^{\beta_2}.$$

Under (1.12), we can replace $\kappa_{m,d}\Gamma_1$ *by* Γ_1.

Proof under (1.12). This lemma is the most important technical ingredient for the proof of Theorem 1.1.1. To simplify the proof, here we assume (1.12), i.e. that the additional technical condition $1+\alpha-\beta_1 < 2\alpha\beta_2-\alpha$ holds. The proof in the general setting will be given in Appendix 1.4.

We again use the representation (1.15). We fix a large $n \in \mathbb{N}$ and set $\mathcal{F}^n_k := \mathcal{F}_{t_0+\frac{k}{2^n}(t_N-t_0)}$. Fix an integer $L = L_n \in [M+1, 2^n]$, which will be chosen later. We have

$$\sum_{i=0}^{2^n-1} R^n_i = \sum_{l=0}^{L-1} \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \Big(R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\mathbf{1}_{\{j\ge 1\}}\Big) + \sum_{l=0}^{L-1} \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]. \quad (1.17)$$

We estimate the first term of (1.17). By the BDG inequality together with Minkowski's inequality (see (1.14)), we have

$$\begin{aligned}
&\| \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \Big(R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\mathbf{1}_{\{j\ge 1\}}\Big)\|^2_{L^m(\mathbb{P})} \\
&\qquad \le \kappa^2_{m,d} \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \|R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\mathbf{1}_{\{j\ge 1\}}\|^2_{L^m(\mathbb{P})} \\
&\qquad \le 4\kappa^2_{m,d} \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \|R^n_{Lj+l}\|^2_{L^m(\mathbb{P})}.
\end{aligned}$$

Using (1.8) and (1.16) and noting that we include more terms in the sum by requiring $j \le 2^n/L$ only instead of $Lj+l \le 2^n-1$, we get

$$\sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \|R^n_{Lj+l}\|^2_{L^m(\mathbb{P})} \le 4\Gamma_2^2 2^{-n(2\beta_2-1)} L^{-1} |t_N-t_0|^{2\beta_2}.$$

Therefore,

$$\|\sum_{l=0}^{L-1} \sum_{\substack{j\ge 0:\\ Lj+l<2^n}} \Big(R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\mathbf{1}_{\{j\ge 1\}}\Big)\|_{L^m(\mathbb{P})} \lesssim \kappa_{m,d}\Gamma_2 L^{\frac{1}{2}} 2^{-n(\beta_2-\frac{1}{2})} |t_N-t_0|^{\beta_2}.$$

We next estimate the second term of (1.17). The triangle inequality yields

$$\|\sum_{l=0}^{L-1}\sum_{\substack{j\geq 0:\\ Lj+l<2^n}}\mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\|_{L^m(\mathbb{P})}\leq\sum_{l=0}^{L-1}\sum_{\substack{j\geq 0:\\ Lj+l<2^n}}\|\mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\|_{L^m(\mathbb{P})}.$$

By (1.7),

$$\|\mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\|_{L^m(\mathbb{P})}\leq\Gamma_1(L-1)^{-\alpha}2^{-(\beta_1-\alpha)n}|t_N-t_0|^{\beta_1-\alpha}.$$

Therefore,

$$\|\sum_{l=0}^{L-1}\sum_{\substack{j\geq 0:\\ Lj+l<2^n}}\mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}]\|_{L^m(\mathbb{P})}\lesssim_\alpha\Gamma_1L^{-\alpha}2^{-(\beta_1-\alpha-1)n}|t_N-t_0|^{\beta_1-\alpha}.$$

In conclusion,

$$\|\sum_{i=0}^{2^n-1}R^n_i\|_{L^m(\mathbb{P})}\lesssim_\alpha\Gamma_1L^{-\alpha}2^{-(\beta_1-\alpha-1)n}|t_N-t_0|^{\beta_1-\alpha}+\kappa_{m,d}\Gamma_2L^{\frac{1}{2}}2^{-n(\beta_2-\frac{1}{2})}|t_N-t_0|^{\beta_2}. \tag{1.18}$$

We wish to choose $L=L_n$ so that (1.18) is summable with respect to n. We therefore set $L_n:=\lfloor 2^{\delta n}\rfloor$, where

$$\alpha\delta+\beta_1-\alpha-1>0,\quad 0<\delta<\min\{2\beta_2-1,1\}. \tag{1.19}$$

Such a δ exists exactly under the additional technical assumption (1.12), namely if $1+\alpha-\beta_1<2\alpha\beta_2-\alpha$. Then, (1.18) yields

$$\|\sum_{n:2^{n\delta}\geq M+2}\sum_{i=0}^{2^n-1}R^n_i\|_{L^m(\mathbb{P})}\lesssim_{\alpha,\beta_1,\beta_2}\Gamma_1|t_N-t_0|^{\beta_1-\alpha}+\kappa_{m,d}\Gamma_2|t_N-t_0|^{\beta_2}.$$

To estimate the contribution coming from the small n with $2^{n\delta}<M+2$, we apply (1.8) which yields

$$\|\sum_{i=0}^{2^n-1}R^n_i\|_{L^m(\mathbb{P})}\leq 2\Gamma_2\sum_{i=0}^{2^n-1}2^{-n\beta_2}|t_N-t_0|^{\beta_2}=\Gamma_2 2^{1+n(1-\beta_2)}|t_N-t_0|^{\beta_2}.$$

Thus, we conclude

$$\|\sum_{n\in\mathbb{N}_0}\sum_{i=0}^{2^n-1}R^n_i\|_{L^m(\mathbb{P})}\lesssim_{\alpha,\beta_1,\beta_2,M}\Gamma_1|t_N-t_0|^{\beta_1-\alpha}+\kappa_{m,d}\Gamma_2|t_N-t_0|^{\beta_2},$$

where the fact $\kappa_{m,d}\geq 1$ is used. □

Lemma 1.2.4. *Under the setting of Theorem 1.1.1, let π, π' be partitions of $[0,T]$ such that π refines π'. Suppose that we have*

$$\min_{[s,t]\in\pi'} |s-t| \geq \frac{|\pi'|}{3}. \tag{1.20}$$

Then, there exists $\varepsilon \in (0,1)$ such that

$$\|A_T^{\pi'} - A_T^{\pi}\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M,m,d,T} (\Gamma_1 + \Gamma_2)|\pi'|^{\varepsilon}.$$

Sketch of the proof. Here we give a sketch of the proof under (1.12). The complete proof is given in Appendix 1.4. The argument is similar to Lemma 1.2.3.

Write

$$\pi' =: \{0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T\}$$

and

$$\{\, [s,t] \in \pi : t_j \leq s < t \leq t_{j+1} \,\} =: \{t_j = t_0^j < t_1^j < \cdots < t_{N_j-1}^j < t_{N_j}^j = t_{j+1}\}.$$

We set $L := \lfloor |\pi'|^{-\delta} \rfloor$, where δ satisfies (1.19). We set

$$Z_j^l := A_{t_{jL+l}, t_{jL+l+1}} - \sum_{k=1}^{N_{jL+l}} A_{t_{k-1}^{jL+l}, t_k^{jL+l}}.$$

As in Lemma 1.2.3, we consider the decomposition $A_T^{\pi'} - A_T^{\pi} = A + B$, where

$$A := \sum_{l<L} \sum_{j : Lj < N-l} \Big\{ Z_j^l - \mathbb{E}[Z_j^l | \mathcal{F}_{t_{(j-1)L+l+1}}] \Big\}, \quad B := \sum_{l<L} \sum_{j : Lj < N-l} \mathbb{E}[Z_j^l | \mathcal{F}_{t_{(j-1)L+l+1}}].$$

Then, we estimate A by using the BDG inequality and Lemma 1.2.3, and B by using the triangle inequality and Lemma 1.2.2. $\square$

Remark 1.2.5. In the setting of Lemma 1.2.4, assume that the adapted process $(\mathcal{A}_t)_{t\in[0,T]}$ satisfies (1.10) and (1.11). Then we obtain for some $\varepsilon > 0$:

$$\|A_T^{\pi'} - \mathcal{A}_T\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M,m,d,T} (\Gamma_1 + \Gamma_2)|\pi'|^{\varepsilon}.$$

Indeed, it suffices to replace $A_{t_{jL+l}, t_{jL+l+1}}$ by $\mathcal{A}_{t_{jL+l}, t_{jL+l+1}}$ in the previous proof.

Lemma 1.2.6. *Let π be a partition of $[0,T]$. Then, there exists a partition π' of $[0,T]$ such that π refines π', $|\pi'| \leq 3|\pi|$ and*

$$\min_{[s,t]\in\pi'} |t-s| \geq 3^{-1}|\pi'|.$$

Proof. We write $\pi = \{0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T\}$. We set $k_0 := -1$ and for $l \in \mathbb{N}$ we inductively set

$$k_l := \inf\{\, j > k_{l-1} : t_{j+1} - t_{k_{l-1}+1} \geq |\pi|\,\}, \quad \text{where } \inf\varnothing := N.$$

Set $L := \sup\{\, l : k_l < N\,\}$. Then, we define

$$s_j := \begin{cases} t_{k_j+1} & \text{if } j < L, \\ t_N & \text{if } j = L. \end{cases}$$

By construction, $\pi' = \{s_j\}_{j=1}^L$ satisfies the claimed properties: $s_{j+1} - s_j \leq 2|\pi|$ if $j < L-2$, and $s_L - s_{L-1} \leq 3|\pi|$, so $|\pi'| \leq 3|\pi|$; moreover, $\min_{[s,t]\in\pi'} |t-s| \geq |\pi| \geq 3^{-1}|\pi'|$. ☐

Proof of Theorem 1.1.1. We will not write down dependence on $\alpha, \beta_1, \beta_2, M, m, d, T$. We first prove the convergence of $(A^\pi_\tau)_\pi$. Without loss of generality, we assume $\tau = T$. Let π_1, π_2 be partitions of $[0,T]$. By Lemma 1.2.6, there exist partitions π_1', π_2' such that for $j \in \{1,2\}$ the partition π_j refines π_j', $|\pi_j'| \leq 3|\pi_j|$ and

$$\min_{[s,t]\in\pi_j'} |t-s| \geq 3^{-1}|\pi_j'|.$$

Lemma 1.2.4 shows that for some $\varepsilon > 0$ we have

$$\|A^{\pi_j}_T - A^{\pi_j'}_T\|_{L^m(\mathbb{P})} \lesssim (\Gamma_1 + \Gamma_2)|\pi_j|^\varepsilon.$$

Therefore, by the triangle inequality,

$$\|A^{\pi_1}_T - A^{\pi_2}_T\|_{L^m(\mathbb{P})} \lesssim \|A^{\pi_1'}_T - A^{\pi_2'}_T\|_{L^m(\mathbb{P})} + (\Gamma_1 + \Gamma_2)(|\pi_1|^\varepsilon + |\pi_2|^\varepsilon). \tag{1.21}$$

Let π refine both π_1' and π_2'. Lemma 1.2.4 implies that

$$\|A^{\pi_1'}_T - A^{\pi_2'}_T\|_{L^m(\mathbb{P})} \leq \|A^{\pi_1'}_T - A^{\pi}_T\|_{L^m(\mathbb{P})} + \|A^{\pi_1'}_T - A^{\pi}_T\|_{L^m(\mathbb{P})} \lesssim (\Gamma_1 + \Gamma_2)|\pi_1|^\varepsilon + |\pi_2|^\varepsilon. \tag{1.22}$$

The estimates (1.21) and (1.22) show

$$\|A^{\pi_1}_T - A^{\pi_2}_T\|_{L^m(\mathbb{P})} \lesssim (\Gamma_1 + \Gamma_2)(|\pi_1|^\varepsilon + |\pi_2|^\varepsilon).$$

Thus, $\{A^\pi_T\}_\pi$ forms a Cauchy net in $L^m(\mathbb{P})$. We denote the limit by $\mathscr{A}_T$.

We next prove that $(\mathscr{A}_t)_{t\in[0,T]}$ satisfies (1.10) and (1.11). Let $t_0 < t_1 < t_2$ be such that $M(t_2 - t_1) \leq t_1 - t_0$. Let $\pi_n = \{t_1 + k2^{-n}(t_2 - t_1) : k = 0, \ldots, 2^n\}$ be the nth dyadic partition of $[t_1, t_2]$, and we write

$$A^n_{t_1,t_2} := \sum_{[s,t]\in\pi_n} A_{s,t}.$$

We have

$$\mathbb{E}[\mathscr{A}_{t_1,t_2} - A_{t_1,t_2}|\mathcal{F}_{t_0}] = \lim_{n\to\infty} \mathbb{E}[A^n_{t_1,t_2} - A_{t_1,t_2}|\mathcal{F}_{t_0}] \quad \text{in } L^m(\mathbb{P}). \tag{1.23}$$

By Lemma 1.2.2,

$$\|\mathbb{E}[A_{t_1,t_2} - A^n_{t_1,t_2}|\mathcal{F}_{t_0}]\|_{L^m(\mathbb{P})} \lesssim_{\beta_1} \Gamma_1|t_1-t_0|^{-\alpha}|t_2-t_1|^{\beta_1}.$$

In this estimate, we can replace $A^n_{t_1,t_2}$ by $\mathscr{A}_{t_1,t_2}$ in view of (1.23). Similarly, by Lemma 1.2.3, we obtain

$$\|\mathscr{A}_{t_1,t_2} - A_{t_1,t_2}\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M} \kappa_{m,d}\Gamma_1|t_2-t_1|^{\beta_1-\alpha} + \kappa_{m,d}\Gamma_2|t_2-t_1|^{\beta_2}.$$

Under (1.12), we can replace $\kappa_{m,d}\Gamma_1$ by Γ_1.

Finally, let us prove the uniqueness of $\mathcal{A}$. Let $(\tilde{\mathcal{A}}_t)_{t\in[0,T]}$ be another adapted process satisfying $\tilde{\mathcal{A}}_0 = 0$, (1.10) and (1.11). It suffices to show $\mathcal{A}_T = \tilde{\mathcal{A}}_T$. Let π_n be the nth dyadic partition of $[0,T]$. By Remark 1.2.5 we have

$$\|\mathcal{A}_T - \tilde{\mathcal{A}}_T\|_{L^m(\mathbb{P})} \le \|\mathcal{A}_T - A^{\pi^n}_T\|_{L^m(\mathbb{P})} + \|A^{\pi^n}_T - \tilde{\mathcal{A}}_T\|_{L^m(\mathbb{P})} \lesssim 2^{-n\varepsilon}T^{\varepsilon}.$$

Since $n \in \mathbb{N}$ is arbitrary we must have $\mathcal{A}_T = \tilde{\mathcal{A}}_T$. $\square$

1.3 Integration along fractional Brownian motions

The goal of this section is to prove the convergence of Itô and Stratonovich approximations of

$$\int_0^t f(B_s)\mathrm{d}B_s \quad \text{and} \quad \int_0^t f(B_s)\circ \mathrm{d}B_s$$

along a multidimensional fractional Brownian motion B with Hurst parameter H, using Theorem 1.1.1. For Itô's case, we let $H \in (\frac{1}{2}, 1)$ and for Stratonovich's case, we let $H \in (\frac{1}{6}, \frac{1}{2})$.

Let us recall the fractional Brownian motion.

Definition 1.3.1. Let $H \in (0,1)$. The *fractional Brownian motion* with Hurst parameter H is a centered Gaussian process $B^H = (B^{H,i})_{i=1}^d$ such that $B^H_0 = 0$, the components $B^{H,1}, B^{H,2}, \dots, B^{H,d}$ are independent and identically distributed, and

$$\mathbb{E}[(B^{H,i}_t - B^{H,i}_s)^2] = c_H(t-s)^{2H}, \quad c_H := \frac{3/2-H}{2H}\mathfrak{B}(2-2H, H+1/2)$$

with $\mathfrak{B}$ being the usual Beta function.

In particular, we have

$$\mathbb{E}[B_s^i B_t^i] = \frac{c_H}{2}(t^{2H} + s^{2H} - |t-s|^{2H}). \tag{1.24}$$

We will use the *Mandelbrot–van Ness representation* ([MV68])

$$B_t = \int_{\mathbb{R}} \mathcal{K}_H(t,s)\mathrm{d}W_s, \tag{1.25}$$

where

$$\mathcal{K}(t,s) := \mathcal{K}_H(t,s) := (t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}},$$

and $W = (W_t)_{t\in\mathbb{R}}$ is a two-sided $\mathbb{R}^d$-valued Brownian motion. Regarding the expression of the constant c_H, see [Pic11, Appendix B].

We denote by $(\mathcal{F}_t)_{t\in\mathbb{R}}$ the filtration generated by W. An advantage of the representation (1.25) is that given $v < s$, we have the decomposition

$$B_s = \int_{-\infty}^{v} \mathcal{K}(s,r)\mathrm{d}W_r + \int_{v}^{s} \mathcal{K}(s,r)\mathrm{d}W_r,$$

where the second term $\int_v^s \mathcal{K}(s,r)\mathrm{d}W_r$ is independent of $\mathcal{F}_v$. Later we will need to estimate the correlation of

$$\int_v^s \mathcal{K}(s,r)\mathrm{d}W_r, \quad s > v.$$

We note that for $s \le t$

$$\mathbb{E}[\int_v^s \mathcal{K}(s,r)\mathrm{d}W_r^i \int_v^t \mathcal{K}(t,r)\mathrm{d}W_r^j] = \delta_{ij}\int_v^s \mathcal{K}(s,r)\mathcal{K}(t,r)\mathrm{d}r.$$

Lemma 1.3.2. *Let $H \neq \frac{1}{2}$. Let $0 \le v < s \le t$ be such that $t-s \le s-v$. Then,*

$$\int_v^s \mathcal{K}(s,r)\mathcal{K}(t,r)\mathrm{d}r$$
$$= \frac{1}{2H}(s-v)^{2H} + \frac{1}{2}(s-v)^{2H-1}(t-s) - \frac{c_H}{2}(t-s)^{2H} + g_H(v,s,t)$$

where we have

$$|g_H(v,s,t)| \lesssim_H (s-v)^{2H-2}(t-s)^2$$

uniformly over such v, s, t.

Proof. See Appendix 1.4. □

We apply Theorem 1.1.1 to construct a stochastic integral

$$\int_0^T f(B_s)\mathrm{d}B_s, \quad H \in (1/2, 1)$$

as the limit of Riemann type approximations. An advantage of the stochastic sewing lemma is that we do not need any regularity of f. We denote by $L^\infty(\mathbb{R}^d, \mathbb{R}^d)$ the space of bounded measurable maps from $\mathbb{R}^d$ to $\mathbb{R}^d$. We write

$$x \cdot y := \sum_{i=1}^d x^i y^i, \quad x = (x^i)_{i=1}^d, \; y = (y^i)_{i=1}^d$$

for the inner product of $\mathbb{R}^d$.

Proposition 1.3.3. *Let $H \in (1/2, 1)$ and $f \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$. Then, for any $\tau \in [0, T]$ and $m \in [2, \infty)$, the sequence*

$$\sum_{[s,t]\in\pi} f(B_s) \cdot (B_t - B_s), \quad \textit{where } \pi \textit{ is a partition of } [0, \tau],$$

converges in $L^m(\mathbb{P})$ for every $m < \infty$ as $|\pi| \to 0$. Furthermore, if we denote the limit by $\int_0^\tau f(B_r)\mathrm{d}B_r$ and if we write

$$\int_s^t f(B_r)\mathrm{d}B_r := \int_0^t f(B_r)\mathrm{d}B_r - \int_0^s f(B_r)\mathrm{d}B_r,$$

then for every $0 \le s < t \le T$,

$$\Big\| \int_s^t f(B_r)\mathrm{d}B_r \Big\|_{L^m(\mathbb{P})} \lesssim_{d,H,m} \|f\|_{L^\infty(\mathbb{R}^d)} |t-s|^H \tag{1.26}$$

Remark 1.3.4. We can replace $f(B_s)$ by $f(B_u)$ for any $u \in [s, t]$. It is well known that the sums converge to the Young integral if $f \in C^\gamma(\mathbb{R})$ with $\gamma > H^{-1}(1-H)$. Yaskov [Yas18, Theorem 3.7] proves that the sums converge in some $L^p(\mathbb{P})$-space if f is of bounded variation.

Remark 1.3.5. We can actually improve the estimate (1.26). In fact, we can replace the norm $\|f\|_{L^\infty(\mathbb{R}^d)}$ by $\|f\|_{C^{\frac{1}{2H}-1+\varepsilon}(\mathbb{R}^d)}$ for any positive ε, see Theorem 3.3.2.

Proof. We will not write down dependence on d, H and m. We will apply Theorem 1.1.1 with $A_{s,t} := f(B_s) \cdot (B_t - B_s)$. Let $m \ge 2$. We have

$$\|A_{s,t}\|_{L^m(\mathbb{P})} \lesssim \|f\|_{L^\infty} |t-s|^H.$$

To estimate conditional expectations, let $0 \le v < s < t$ be such that $t - s \le s - v$ and set

$$Y_s := \int_{-\infty}^{v} \mathcal{K}(s,r)\mathrm{d}W_r, \quad \tilde{B}_s := \int_{v}^{s} \mathcal{K}(s,r)\mathrm{d}W_r.$$

We write $y_s := Y_s$, if conditioned under $\mathcal{F}_v$. Namely, we write for instance

$$\mathbb{E}[g(y_s, \tilde{B}_s)] := \mathbb{E}[g(Y_s, \tilde{B}_s)|\mathcal{F}_v] = \mathbb{E}[g(y, \tilde{B}_s)]|_{y=Y_s}.$$

For $k \in \mathbb{N}_0^d$, we denote by $X^{:k:} =: X_1^{k_1} \dots X_d^{k_d}$: the kth Wick power of $X = (X_1, \dots, X_d)$. We are going to compute $\mathbb{E}[A_{s,t}|\mathcal{F}_v]$. Conditionally on $\mathcal{F}_v$, we have the Wiener chaos expansion

$$f(B_s) = f(y_s + \tilde{B}_s) = \sum_{k \in \mathbb{N}_0^d} a_k(s) \tilde{B}_s^{:k:}.$$

Although it is abuse of notation, for $i \in \{1, \dots, d\}$ we write $a_i(s) := a_{\boldsymbol{e}_i}$, where $\boldsymbol{e}_i$ is ith unit vector in $\mathbb{R}^d$. Note that

$$\begin{aligned} &a_0(s) = \mathbb{E}[f(y_s + \tilde{B}_s)], \\ &a_i(s) = \mathbb{E}[(\tilde{B}_s^i)^2]^{-1} \mathbb{E}[f(y_s + \tilde{B}_s)\tilde{B}_s^i] \overset{\text{Lem. 1.3.2}}{=} 2H(s-v)^{-2H} \mathbb{E}[f(y_s + \tilde{B}_s)\tilde{B}_s^i]. \end{aligned}$$

Then, by the orthogonality of the Wiener chaos decomposition,

$$\mathbb{E}[A_{s,t}|\mathcal{F}_v] = a_0(s) \cdot Y_{s,t} + \sum_{i=1}^{d} a_i(s) \cdot \mathbb{E}[\tilde{B}_s^i \tilde{B}_{s,t}].$$

Hence, for $u \in (s,t)$,

$$\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v] = A^0_{s,u,t} + \sum_{i=1}^{d} A^i_{s,u,t},$$

where

$$\begin{aligned} A^0_{s,u,t} :=& a_0(s) \cdot Y_{s,t} - a_0(s) \cdot Y_{s,u} - a_0(u) \cdot Y_{u,t} = (a_0(s) - a_0(u)) \cdot Y_{u,t}, \\ A^i_{s,u,t} :=& a_i(s) \cdot \mathbb{E}[\tilde{B}_s^i \tilde{B}_{s,t}] - a_i(s) \cdot \mathbb{E}[\tilde{B}_s^i \tilde{B}_{s,u}] - a_i(u) \cdot \mathbb{E}[\tilde{B}_u^i \tilde{B}_{u,t}] \\ =& [a_i(s) \cdot \boldsymbol{e}_i] \mathbb{E}[\tilde{B}_s^i \tilde{B}_{s,t}^i] - [a_i(s) \cdot \boldsymbol{e}_i] \mathbb{E}[\tilde{B}_s^i \tilde{B}_{s,u}^i] - [a_i(u) \cdot \boldsymbol{e}_i] \mathbb{E}[\tilde{B}_u^i \tilde{B}_{u,t}^i]. \end{aligned}$$

We first estimate $A^0_{s,u,t}$, for which we begin with estimating $a_0(s) - a_0(u)$. We set

$$F(m, \sigma) := \mathbb{E}[f(m + \sigma X)], \quad m \in \mathbb{R}^d, \ \sigma \in (0, \infty),$$

where X has the standard normal distribution in $\mathbb{R}^d$. Note that

$$a_0(s) = F(Y_s, (2H)^{-\frac{1}{2}}(s-v)^H)$$

and similarly for $a_0(u)$. we have

$$\begin{aligned}
\partial_{m^i} F(m,\sigma) &= \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+2}} \int_{\mathbb{R}^d} x^i e^{-\frac{|x|^2}{2\sigma^2}} f(x+m)\mathrm{d}x, \\
\partial_\sigma F(m,\sigma) &= \frac{-d}{(2\pi)^{\frac{d}{2}}\sigma^{d+1}} \int_{\mathbb{R}^d} f(m+x) e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{d}x \\
&\quad + \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+3}} \int_{\mathbb{R}^d} |x|^2 f(m+x) e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{d}x.
\end{aligned}$$

Therefore,

$$|\partial_m F(m,\sigma)| + |\partial_\sigma F(m,\sigma)| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \sigma^{-1}.$$

This yields

$$\begin{aligned}
&|a_0(s) - a_0(u)| \\
&\quad \le |F(Y_s, (2H)^{-\frac{1}{2}}(s-v)^H) - F(Y_u, (2H)^{-\frac{1}{2}}(s-v)^H)| \\
&\qquad + |F(Y_u, (2H)^{-\frac{1}{2}}(s-v)^H) - F(Y_u, (2H)^{-\frac{1}{2}}(u-v)^H)| \\
&\quad \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}|Y_{s,u}| + \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}(|u-v|^H - |s-v|^H) \\
&\quad \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}|Y_{s,u}| + \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-1}(t-s).
\end{aligned}$$

Therefore,

$$|A^0_{s,u,t}| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}|Y_{s,u}||Y_{u,t}| + \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-1}(t-s)|Y_{u,t}|. \tag{1.27}$$

The random variable $Y_{s,u}$ is Gaussian and

$$\begin{aligned}
\mathbb{E}[|Y_{s,u}|^2] &= d\int_{-\infty}^{v} (\mathcal{K}(s,r) - \mathcal{K}(u,r))^2 \mathrm{d}r = d\int_{s-v}^{\infty} ((u-s+r)^{H-\frac{1}{2}} - r^{H-\frac{1}{2}})^2 \mathrm{d}r \\
&\lesssim (u-s)^2 \int_{s-v}^{\infty} r^{2H-3}\mathrm{d}r \lesssim (s-v)^{2H-2}(u-s)^2. \qquad (1.28)
\end{aligned}$$

We have a similar estimate for $Y_{u,t}$. Therefore,

$$\|A^0_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{H-2}(t-s)^2 \quad \text{if } t-s \le v-s.$$

Now we move to estimate $A^i_{s,u,t}$. By Lemma 1.3.2, we have

$$\begin{aligned}\mathbb{E}[\tilde{B}^i_s \tilde{B}^i_{s,t}] &= \int_v^s \mathcal{K}(s,r)\mathcal{K}(t,r)\mathrm{d}r - \int_v^s \mathcal{K}(s,r)\mathcal{K}(s,r)\mathrm{d}r \\ &= \frac{1}{2}(s-v)^{2H-1}(t-s) + O((t-s)^{2H}).\end{aligned}$$

Therefore, if we write $a^i_i(s) := a_i(s) \cdot \boldsymbol{e}_i$,

$$\begin{aligned}A^i_{s,u,t} = \frac{1}{2}\big[a^i_i(s)(s-v)^{2H-1} - a^i_i(u)(u-v)^{2H-1}\big](t-u) \\ + O((|a^i_i(s)| + |a^i_i(u)|)|t-s|^{2H}).\end{aligned}$$

If we set

$$G_i(m,\sigma) := \sigma^{-1}\mathbb{E}[f^i(m+\sigma X)X^i], \quad m \in \mathbb{R}^d,\ \sigma \in (0,\infty),$$

then $a^i_i(s) = G_i(Y_s, (2H)^{-\frac{1}{2}}(s-v)^H)$ and similarly for $a^i_i(u)$. Since

$$G_i(m,\sigma) = (2\pi)^{-\frac{d}{2}}\sigma^{-d-2}\int_{\mathbb{R}^d} f^i(y)(y^i - m^i)e^{-\frac{|y-m|^2}{2\sigma^2}}\mathrm{d}y,$$

we have

$$\begin{aligned}(2\pi)^{\frac{d}{2}}\sigma^2\partial_{m^j}G_i(m,\sigma) &= \int_{\mathbb{R}^d} f^i(m+\sigma x)[-\delta_{ij} + x^i x^j]e^{-\frac{|x|^2}{2}}\mathrm{d}x \\ (2\pi)^{\frac{d}{2}}\sigma^2\partial_{\sigma}G_i(m,\sigma) &= \int_{\mathbb{R}^d} f^i(m+\sigma x)x^i[-(d+2) + |x|^2]e^{-\frac{|x|^2}{2}}\mathrm{d}x.\end{aligned}$$

Therefore,

$$\begin{gathered}|G_i(m,\sigma)| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}\sigma^{-1}, \\ |\partial_m G_i(m,\sigma)| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}\sigma^{-2}, \quad |\partial_\sigma G_i(m,\sigma)| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}\sigma^{-2}\end{gathered}$$

and thus

$$|a^i_i(s)| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H},$$

$$\begin{aligned}|a^i_i(s) - a^i_i(u)| &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-2H}\big(|Y_{s,u}| + (u-v)^H - (s-v)^H\big) \\ &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-2H}\big(|Y_{s,u}| + (s-v)^{H-1}(u-s)\big).\end{aligned}$$

This yields

$$\begin{aligned}|A^i_{s,u,t}| \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}\big[(s-v)^{-1}(t-s)|y_{s,u}| + (s-v)^{H-2}(t-s)^2 \\ + (s-v)^{H-2}(t-s)^2 + (s-v)^{-H}(t-s)^{2H}\big]\end{aligned}$$

and

$$\begin{aligned}\|A^i_{s,u,t}\|_{L^m(\mathbb{P})} &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)}\big[(s-v)^{H-2}(t-s)^2 + (s-v)^{-H}(t-s)^{2H}\big] \\ &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}(t-s)^{2H} \end{aligned} \tag{1.29}$$

if $t-s \le s-v$.

Therefore, by (1.27) and (1.29),

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v]\|_{L^m(\mathbb{P})} \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}(s-v)^{-H}(t-s)^{2H}$$

if $t-s \le s-v$. Hence, $(A_{s,t})$ satisfies the assumption of Theorem 1.1.1 with

$$\alpha = H, \quad \beta_1 = 2H, \quad \beta_2 = H, \quad M = 1. \qquad \square$$

Next, we consider the case $H \in (\frac{1}{6}, \frac{1}{2})$. The following result reproduces [Nou12, Theorem 3.5], with a more elementary proof and with improvement of the regularity of f. More precisely, the cited result requires $f \in C^6$ while here $f \in C^\gamma$ with $\gamma > \frac{1}{2H} - 1$ is sufficient and thus in particular $f \in C^2$ works for all $H \in (\frac{1}{6}, \frac{1}{2})$. We denote by $C^\gamma(\mathbb{R}^d, \mathbb{R}^d)$ the space of γ-Hölder maps from $\mathbb{R}^d$ to $\mathbb{R}^d$, with the norm

$$\|f\|_{C^\gamma} := \|f\|_{L^\infty(\mathbb{R}^d)} + \sup_{x \neq y} \frac{|f(x)-f(y)|}{|x-y|^\gamma}$$

if $\gamma \in (0,1)$ and

$$\|f\|_{C^\gamma} := \|f\|_{L^\infty(\mathbb{R}^d)} + \sum_{i=1}^d \|\partial_i f\|_{C^{\gamma-1}}$$

if $\gamma \in (1,2)$.

Proposition 1.3.6. *Let $H \in (\frac{1}{6}, \frac{1}{2})$, $\gamma > \frac{1}{2H} - 1$ and $f \in C^\gamma(\mathbb{R}^d, \mathbb{R}^d)$. If $H \le \frac{1}{4}$ and $d > 1$, assume furthermore that*

$$\partial_i f^j = \partial_j f^i, \quad \forall i,j \in \{1,\dots,d\}. \tag{1.30}$$

Then, for every $m \in [2,\infty)$ and $\tau \in [0,T]$, the family of Stratonovich approximations

$$\sum_{[s,t]\in\pi} \frac{f(B_s)+f(B_t)}{2} \cdot B_{s,t}, \quad \text{where } \pi \text{ is a partition of } [0,\tau],$$

converges in $L^m(\mathbb{P})$ as $|\pi| \to 0$. Moreover, if we denote the limit by $\int_0^\tau f(B_r) \circ \mathrm{d}B_r$ and if we write

$$\int_s^t f(B_r)\circ \mathrm{d}B_r := \int_0^t f(B_r)\circ \mathrm{d}B_r - \int_0^s f(B_r)\circ\mathrm{d}B_r,$$

then for every $0 \le s < t \le T$ *we have*

$$\Big\| \int_s^t f(B_r) \circ \mathrm{d}B_r - \frac{f(B_s) + f(B_t)}{2} \cdot B_{s,t} \Big\|_{L^m(\mathbb{P})} \lesssim_{d,H,m,\gamma} \|f\|_{C^\gamma} |t-s|^{(\gamma+1)H}.$$

Proof. We will not write down dependence on d, H, m and γ. We can assume

$$\gamma < \mathbf{1}_{\{H > \frac{1}{4}\}} + 2\mathbf{1}_{\{H \le \frac{1}{4}\}}.$$

We will apply Theorem 1.1.1 with

$$A_{s,t} := (f(B_s) + f(B_t)) \cdot B_{s,t}.$$

We first claim

$$\|\delta A_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim \|f\|_{C^\gamma} |t-s|^{(\gamma+1)H}. \tag{1.31}$$

Observe

$$\delta A_{s,u,t} = f(B)_{u,t} \cdot B_{s,u} + f(B)_{u,s} \cdot B_{u,t}.$$

If $H > \frac{1}{4}$, the claim (1.31) follows from the estimates

$$|f(B)_{u,t}| \le \|f\|_{C^\gamma} |B_{u,t}|^\gamma, \quad |f(B)_{u,s}| \le \|f\|_{C^\gamma} |B_{u,s}|^\gamma.$$

If $H \le \frac{1}{4}$, then $\gamma > 1$ and we have

$$\delta A_{s,u,t} = \Big(f(B)_{u,t} - \sum_{j=1}^d \partial_j f(B_u) B^j_{u,t} \Big) \cdot B_{s,u} + \Big(f(B)_{u,s} - \sum_{j=1}^d \partial_j f(B_u) B^j_{u,s} \Big) \cdot B_{u,t},$$

where (1.30) is used. Then, the claim (1.31) follows again from the Hölder estimate of f. Note that the condition $\gamma > \frac{1}{2H} - 1$ is equivalent to $(\gamma + 1)H > \frac{1}{2}$.

The rest of the proof consists of estimating the conditional expectation $\mathbb{E}[\delta A_{s,u,t} | \mathcal{F}_v]$. Let $t - s \le s - v$. We will use the same notation as in the proof of Proposition 1.3.3. We have

$$\mathbb{E}[\delta A_{s,u,t} | \mathcal{F}_v] = D^0_{s,u,t} + \sum_{i=1}^d D^i_{s,u,t},$$

where

$$\begin{aligned} D^0_{s,u,t} :=& (a_0(s) + a_0(t)) \cdot Y_{s,t} - (a_0(s) + a_0(u)) \cdot Y_{s,u} - (a_0(u) + a_0(t)) \cdot Y_{u,t} \\ =& (a_0(t) - a_0(u)) \cdot Y_{s,u} + (a_0(s) - a_0(u)) \cdot Y_{u,t} \end{aligned} \tag{1.32}$$

and

$$D^i_{s,u,t} := \mathbb{E}[(a^i_i(s)\tilde{B}^i_s + a^i_i(t)\tilde{B}^i_t)\tilde{B}^i_{s,t}|\mathcal{F}_v] \\ - \mathbb{E}[(a^i_i(s)\tilde{B}^i_s + a^i_i(u)\tilde{B}^i_u)\tilde{B}^i_{s,u}|\mathcal{F}_v] - \mathbb{E}[(a^i_i(u)\tilde{B}^i_u + a^i_i(t)\tilde{B}^i_t)\tilde{B}^i_{u,t}|\mathcal{F}_v]$$

We first estimate $D^0_{s,u,t}$. Suppose that $H > \frac{1}{4}$. Recall

$$\begin{aligned}\partial_{m^i} F(m,\sigma) &= \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+2}} \int_{\mathbb{R}^d} x^i e^{-\frac{|x|^2}{2\sigma^2}}[f(x+m) - f(m)]\mathrm{d}x, \\ \partial_\sigma F(m,\sigma) &= \frac{-d}{(2\pi)^{\frac{d}{2}}\sigma^{d+1}} \int_{\mathbb{R}^d} [f(m+x) - f(m)]e^{-\frac{|x|^2}{2\sigma^2}}\mathrm{d}x \\ &\qquad + \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+3}} \int_{\mathbb{R}^d} |x|^2[f(m+x) - f(m)]e^{-\frac{|x|^2}{2\sigma^2}}\mathrm{d}x.\end{aligned}$$

Therefore,

$$|\partial_{m^i} F(m,\sigma)| + |\partial_\sigma F(m,\sigma)| \lesssim \|f\|_{C^\gamma}\sigma^{\gamma-1}.$$

This yields

$$|D^0_{s,u,t}| \lesssim \|f\|_{C^\gamma}\left[(s-v)^{(\gamma-1)H}|Y_{s,u}||Y_{u,t}| + (s-v)^{\gamma H-1}(t-s)(|Y_{s,u}| + |Y_{u,t}|)\right]. \tag{1.33}$$

Therefore, by (1.28),

$$\|D^0_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma+1)H-2}(t-s)^2. \tag{1.34}$$

Now suppose that $H \leq \frac{1}{4}$. To simplify notation, we write $I(m,\sigma) := F(m,(2H)^{-\frac{1}{2}}\sigma)$. Since (1.30) gives $\partial_{m^i} I^j = \partial_{m^j} I^i$ for every i,j, we have

$$\begin{aligned}D^0_{s,u,t} =& [I(Y_s,(u-v)^H) - I(Y_u,(u-v)^H) - \sum_{i=1}^d \partial_{m^i} I(Y_u,(u-v)^H)Y^i_{u,s}] \cdot Y_{u,t} \\ &+ [I(Y_t,(u-v)^H) - I(Y_u,(u-v)^H) - \sum_{i=1}^d \partial_{m^i} I(Y_u,(u-v)^H)Y^i_{u,t}] \cdot Y_{s,u} \\ &+ [I(Y_s,(s-v)^H) - I(Y_s,(u-v)^H)] \cdot Y_{u,t} \\ &+ [I(Y_t,(t-v)^H) - I(Y_t,(u-v)^H)] \cdot Y_{s,u}.\end{aligned}$$

Since

$$\partial_{m^i}\partial_{m^j} F(m,\sigma) = \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+2}} \int_{\mathbb{R}^d} x^i e^{-\frac{|x|^2}{2\sigma^2}}[\partial_j f(x+m) - \partial_j f(m)]\mathrm{d}x,$$

we have

$$|I(Y_s,(u-v)^H) - I(Y_u,(u-v)^H) - \sum_{i=1}^{d} \partial_{m^i} I(Y_u,(u-v)^H) Y^i_{u,s}| \\ \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma-2)H}|Y_{s,u}|^2.$$

Notice

$$\partial_\sigma F(m,\sigma) = \frac{-d}{(2\pi)^{\frac{d}{2}}\sigma^{d+1}} \int_{\mathbb{R}^d} [f(m+x) - f(m) - \sum_{i=1}^{d} \partial_i f(m) x^i] e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{d}x \\ + \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{d+3}} \int_{\mathbb{R}^d} |x|^2 [f(m+x) - f(m) - \sum_{i=1}^{d} \partial_i f(m) x^i] e^{-\frac{|x|^2}{2\sigma^2}} \mathrm{d}x.$$

Therefore,

$$|\partial_\sigma F(m,\sigma)| \lesssim \|f\|_{C^\gamma}\sigma^{\gamma-1}.$$

This yields

$$|I(Y_s,(s-v)^H) - I(Y_s,(u-v)^H)| \lesssim \|f\|_{C^\gamma}(s-v)^{\gamma H-1}(t-s).$$

Hence, we obtain the estimate (1.34) when $H \leq \frac{1}{4}$.

We move to estimate $D^i_{s,u,t}$. By using the identity,

$$\mathbb{E}[(\tilde{B}^i_a + \tilde{B}^i_b)\tilde{B}^i_{a,b}] = \mathbb{E}[(\tilde{B}^i_b)^2] - \mathbb{E}[(\tilde{B}^i_a)^2],$$

we obtain

$$D^i_{s,u,t} = (a^i_i(t) - a^i_i(u))\mathbb{E}[\tilde{B}^i_t\tilde{B}^i_{s,t}] + (a^i_i(s) - a^i_i(u))\mathbb{E}[\tilde{B}^i_s\tilde{B}^i_{s,t}] \\ - (a^i_i(s) - a^i_i(u))\mathbb{E}[\tilde{B}^i_s\tilde{B}^i_{s,u}] - (a^i_i(t) - a^i_i(u))\mathbb{E}[\tilde{B}^i_t\tilde{B}^i_{u,t}]. \quad (1.35)$$

Since the other terms can be estimated similarly, we only estimate $(a^i_i(t) - a^i_i(u))\mathbb{E}[\tilde{B}^i_t\tilde{B}^i_{s,t}]$. By Lemma 1.3.2,

$$|\mathbb{E}[\tilde{B}_t\tilde{B}_{s,t}]| \lesssim |t-s|^{2H}.$$

Now we estimate $|a^i_i(t) - a^i_i(u)|$. Recall $a^i_i(s) = G_i(Y_s,(2H)^{-\frac{1}{2}}(s-v)^H)$,

$$(2\pi)^{\frac{d}{2}}\sigma^2\partial_{m^j}G_i(m,\sigma) = -\delta_{ij}\int_{\mathbb{R}^d}[f^i(m+\sigma x) - f^i(m)]e^{-\frac{|x|^2}{2}}\mathrm{d}x \\ + \int_{\mathbb{R}^d}[f^i(m+\sigma x) - f^i(m)]x^i x^j e^{-\frac{|x|^2}{2}}\mathrm{d}x,$$

$$(2\pi)^{\frac{d}{2}}\sigma^2\partial_\sigma G(m,\sigma) = -(d+2)\int_{\mathbb{R}^d}[f^i(m+\sigma x)-f^i(m)]x^i e^{-\frac{|x|^2}{2}}\mathrm{d}x \\ + \int_{\mathbb{R}^d}[f^i(m+\sigma x)-f^i(m)]x^i|x|^2 e^{-\frac{|x|^2}{2}}\mathrm{d}x.$$

If $H \le \frac{1}{4}$, we can replace $f^i(m+\sigma x)-f^i(m)$ by

$$f^i(m+\sigma x)-f^i(m)-\sum_{k=1}^{d}\partial_k f^i(m)\sigma x^k.$$

Therefore,

$$|\partial_{m^j}G_i(m,\sigma)|+|\partial_\sigma G_i(m,\sigma)| \lesssim \|f\|_{C^\gamma}\sigma^{\gamma-2}.$$

This yields

$$|a_i^i(t)-a_i^i(u)| \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma-2)H}(|Y_{u,t}|+(s-v)^{H-1}(t-s))$$

and hence

$$\|a_i^i(t)-a_i^i(u)\|_{L^m(\mathbb{P})} \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma-1)H-1}(t-s).$$

Therefore, we obtain

$$\|D^i_{s,u,t}\|_{L^m(\mathbb{P})} \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma-1)H-1}(t-s)^{1+2H}. \tag{1.36}$$

By (1.34) and (1.36), we conclude

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v]\|_{L^m(\mathbb{P})} \lesssim \|f\|_{C^\gamma}[(s-v)^{(1+\gamma)H-2}(t-s)^2+(s-v)^{(\gamma-1)H-1}(t-s)^{1+2H}] \\ \lesssim \|f\|_{C^\gamma}(s-v)^{(\gamma-1)H-1}(t-s)^{1+2H}$$

if $t-s \le s-v$. Therefore, we can apply Theorem 1.1.1 with

$$\alpha = 1-(\gamma-1)H, \quad \beta_1 = 1+2H, \quad \beta_2 = (\gamma+1)H, \quad M=1. \qquad \square$$

1.4 Proofs of technical results

Proofs of Lemma 1.2.3 and Lemma 1.2.4

Proof of Lemma 1.2.3 without (1.12). Let us first recall our previous strategy under (1.12). We used Lemma 1.2.1 to write

$$A_{t_0,t_N}-\sum_{i=1}^{N}A_{t_{i-1},t_i} = \sum_{n\in\mathbb{N}_0}\sum_{i=0}^{2^n-1}R_i^n. \tag{1.37}$$

Then, we decomposed

$$\sum_{i=0}^{2^n-1} R_i^n = \sum_{l=0}^{L-1} \sum_{j=0}^{2^n/L} \Big(R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l} | \mathcal{F}^n_{L(j-1)+l+1}] \mathbf{1}_{\{j \geq 1\}} \Big) \\ + \sum_{l=0}^{L-1} \sum_{j=1}^{2^n/L} \mathbb{E}[R^n_{Lj+l} | \mathcal{F}^n_{L(j-1)+l+1}], \quad (1.38)$$

where $\mathcal{F}^n_k := \mathcal{F}_{t_0 + \frac{k}{2^n}(t_N - t_0)}$. We estimated the first term of (1.38) by the BDG inequality and (1.8):

$$\| \sum_{l=0}^{L-1} \sum_{j=0}^{2^n/L} \Big(R^n_{Lj+l} - \mathbb{E}[R^n_{Lj+l} | \mathcal{F}^n_{L(j-1)+l+1}] \mathbf{1}_{\{j \geq 1\}} \Big) \|_{L^m(\mathbb{P})} \\ \lesssim \kappa_{m,d} \Gamma_2 L^{\frac{1}{2}} 2^{-n(\beta_2 - \frac{1}{2})} |t_N - t_0|^{\beta_2}. \quad (1.39)$$

In the proof under (1.12), we estimated the second term of (1.38) by the triangle inequality and (1.7):

$$\| \sum_{l=0}^{L-1} \sum_{j=1}^{2^n/L} \mathbb{E}[R^n_{Lj+l} | \mathcal{F}^n_{L(j-1)+l+1}] \|_{L^m(\mathbb{P})} \lesssim_\alpha \Gamma_1 L^{-\alpha} 2^{-(\beta_1 - \alpha - 1)n} |t_N - t_0|^{\beta_1 - \alpha}. \quad (1.40)$$

Then, we chose L so that both (1.39) and (1.40) are summable with respect to n, for which to be possible, we had to assume (1.12).

In order to remove the assumption (1.12), let us think again why we did the decomposition (1.38). This is because we do not want to apply the simplest estimate, namely the triangle inequality, since the condition (1.7) implies that $(A_{s,t})_{[s,t] \in \pi}$ are not so correlated. This point of view teaches us that, to estimate

$$\sum_{l=0}^{L-1} \sum_{j=1}^{2^n/L} \mathbb{E}[R^n_{Lj+l} | \mathcal{F}^n_{L(j-1)+l+1}],$$

we should not simply apply the triangle inequality. That is, we should again apply the decomposition as in (1.38).

To carry out our new strategy, set

$$S_j^{(1),l} := R^n_{Lj+l}, \quad \mathcal{G}_j^{(1),l} := \mathcal{F}^n_{L(j-1)+l+1}, \qquad j \in \mathbb{N}.$$

We use the convention $\mathbb{E}[X|\mathcal{G}_j^{(1),l}] = 0$ for $j \leq 0$. Then,

$$\begin{aligned}\sum_{l=0}^{L-1}\sum_{j=1}^{2^n/L}\mathbb{E}[R^n_{Lj+l}|\mathcal{F}^n_{L(j-1)+l+1}] &= \sum_{l=0}^{L-1}\sum_{j=1}^{L^{-1}2^n}\mathbb{E}[S_j^{(1),l}|\mathcal{G}_j^{(1),l}] \\ &= \sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}\sum_{j=0}^{L^{-2}2^n}\mathbb{E}[S^{(1),l_1}_{jL+l_2}|\mathcal{G}^{(1),l_1}_{jL+l_2}]. \end{aligned} \tag{1.41}$$

By setting

$$S_j^{(2),l_1,l_2} := S^{(1),l_1}_{jL+l_2}, \quad \mathcal{G}_j^{(2),l_1,l_2} := \mathcal{G}^{(1),l_1}_{(j-1)L+l_2},$$

the quantity (1.41) equals to

$$\sum_{l_1=0}^{L}\sum_{l_2=0}^{L}\sum_{j=0}^{L^{-2}2^n}\left(\mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_{j+1}^{(2),l_1,l_2}] - \mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_j^{(2),l_1,l_2}]\right) \\ + \sum_{l_1=0}^{L}\sum_{l_2=0}^{L}\sum_{j=0}^{L^{-2}2^n}\mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_j^{(2),l_1,l_2}]$$

The $L^m(\mathbb{P})$-norm of the first term can be estimated by the BDG inequality: it is bounded by

$$2\kappa_{m,d}\sum_{l_1,l_2\leq L}\Big(\sum_{j\leq L^{-2}2^n}\|\mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_{j+1}^{(2),l_1,l_2}]\|^2_{L^m(\mathbb{P})}\Big)^{\frac{1}{2}}. \tag{1.42}$$

By (1.7), we have

$$\|\mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_{j+1}^{(2),l_1,l_2}]\|_{L^m(\mathbb{P})} \leq \Gamma_1(L2^{-n}|t_N-t_0|)^{-\alpha}(2^{-n}|t_N-t_0|)^{\beta_1}.$$

Therefore, the quantity (1.42) is bounded by

$$2\kappa_{m,d}\Gamma_1 L^{1-\alpha}2^{-n(\beta_1-\alpha-\frac{1}{2})}|t_N-t_0|^{\beta_1-\alpha}.$$

As the reader may realize, we will repeat the same argument for

$$\sum_{l_1=0}^{L}\sum_{l_2=0}^{L}\sum_{j=1}^{L^{-2}2^n}\mathbb{E}[S_j^{(2),l_1,l_2}|\mathcal{G}_{j-1}^{(2),l_1,l_2}]$$

and continue. To state more precisely, set inductively,

$$S_j^{(k),l_1,\dots,l_k} := S^{(k-1),l_1,\dots,l_{k-1}}_{Lj+l_k}, \quad \mathcal{G}_j^{(k),l_1,\dots,l_k} := \mathcal{G}^{(k-1),l_1,\dots,l_{k-1}}_{L(j-1)+l_k}, \quad j \in [1, L^{-k}2^n] \cap \mathbb{N}. \tag{1.43}$$

We claim that, if $L^k \leq 2^n$, we have

$$\begin{aligned}\|\sum_{i=0}^{2^n-1} R_i^n\|_{L^m(\mathbb{P})} &\leq 2\kappa_{m,d}\Gamma_2 L^{\frac{1}{2}} 2^{-n(\beta_2-\frac{1}{2})}|t_N - t_0|^{\beta_2} \\ &\quad + 2\kappa_{m,d}\Gamma_1\Big(\sum_{j=1}^{k-1} L^{\frac{j}{2}-(j-1)\alpha}\Big)L^{\frac{1}{2}}2^{-n(\beta_1-\alpha-\frac{1}{2})}|t_N-t_0|^{\beta_1-\alpha} \\ &\quad + \|\sum_{l_1,\ldots,l_k\leq L}\sum_{j\leq L^{-k}2^n}\mathbb{E}[S_j^{(k),l_1,\ldots,l_k}|\mathcal{G}_j^{(k),l_1,\ldots,l_k}]\|_{L^m(\mathbb{P})}.\end{aligned}$$

The proof of the claim is based on induction. The case $k=1$ and $k=2$ is obtained. Suppose that the claim is correct for $k \geq 2$, and consider the case $k+1$. Again, decompose

$$\begin{aligned}&\sum_{l_1,\ldots,l_k\leq L}\sum_{j\leq L^{-k}2^n}\mathbb{E}[S_j^{(k),l_1,\ldots,l_k}|\mathcal{G}_j^{(k),l_1,\ldots,l_k}] \\ &\quad = \sum_{l_1,\ldots,l_k,l_{k+1}\leq L}\sum_{j\leq L^{-(k+1)}2^n}\big(\mathbb{E}[S_j^{(k+1),l_1,\ldots,l_k}|\mathcal{G}_{j+1}^{(k+1),l_1,\ldots,l_k,l_{k+1}}] \\ &\qquad\qquad - \mathbb{E}[S_j^{(k+1),l_1,\ldots,l_k}|\mathcal{G}_j^{(k+1),l_1,\ldots,l_k,l_{k+1}}]\big) \\ &\qquad + \sum_{l_1,\ldots,l_k,l_{k+1}\leq L}\sum_{j\leq L^{-(k+1)}2^n}\mathbb{E}[S_j^{(k+1),l_1,\ldots,l_k}|\mathcal{G}_j^{(k+1),l_1,\ldots,l_k,l_{k+1}}].\end{aligned}$$

To prove the claim, it suffices to estimate the first sum in the right-hand side. By the BDG inequality, its $L^m(\mathbb{P})$-norm is bounded by

$$2\kappa_{m,d}\sum_{l_1,\ldots,l_k,l_{k+1}\leq L}\Big(\sum_{j\leq L^{-(k+1)}2^n}\|\mathbb{E}[S_j^{(k+1),l_1,\ldots,l_k,l_{k+1}}|\mathcal{G}_{j+1}^{(k+1),l_1,\ldots,l_k,l_{k+1}}]\|^2_{L^m(\mathbb{P})}\Big)^{1/2}. \quad (1.44)$$

By (1.7),

$$\|\mathbb{E}[S_j^{(k+1),l_1,\ldots,l_k,l_{k+1}}|\mathcal{G}_{j+1}^{(k+1),l_1,\ldots,l_k,l_{k+1}}]\|_{L^m(\mathbb{P})} \leq \Gamma_1(L^k 2^{-n}|t_N-t_0|)^{-\alpha}(2^{-n}|t_N-t_0|)^{\beta_1}. \quad (1.45)$$

Therefore, the quantity (1.44) is bounded by

$$2\kappa_{m,d}\Gamma_1 L^{\frac{1}{2}}L^{(\frac{1}{2}-\alpha)k}2^{-n(\beta_1-\alpha-\frac{1}{2})}|t_N-t_0|^{(\beta_1-\alpha)}$$

and the claim follows.

Now let us estimate

$$\|\sum_{l_1,\ldots,l_k\leq L}\sum_{j\leq L^{-k}2^n}\mathbb{E}[S_j^{(k),l_1,\ldots,l_k}|\mathcal{G}_j^{(k),l_1,\ldots,l_k}]\|_{L^m(\mathbb{P})} \quad (1.46)$$

by the triangle inequality:

$$\begin{aligned}\|\sum_{l_1,\dots,l_k\le L}\sum_{j\le L^{-k}2^n}&\mathbb{E}[S_j^{(k),l_1,\dots,l_k}|\mathcal{G}_j^{(k),l_1,\dots,l_k}]\|_{L^m(\mathbb{P})}\\&\le\sum_{l_1,\dots,l_k\le L}\sum_{j\le L^{-k}2^n}\|\mathbb{E}[S_j^{(k),l_1,\dots,l_k}|\mathcal{G}_j^{(k),l_1,\dots,l_k}]\|_{L^m(\mathbb{P})}.\end{aligned}$$

By (1.7) (or essentially the estimate (1.45)),

$$\|\mathbb{E}[S_j^{(k),l_1,\dots,l_k}|\mathcal{G}_j^{(k),l_1,\dots,l_k}]\|_{L^m(\mathbb{P})}\le\Gamma_1(L^k2^{-n}|t_N-t_0|)^{-\alpha}(2^{-n}|t_N-t_0|)^{\beta_1}$$

and hence the quantity (1.46) is bounded by

$$\Gamma_1L^{-\alpha k}2^{-n(\beta_1-\alpha-1)}|t_N-t_0|^{\beta_1-\alpha}$$

In conclusion, we obtained for $L^k\le 2^n$,

$$\begin{aligned}\|\sum_{i=0}^{2^n-1}R_i^n\|_{L^m(\mathbb{P})}&\lesssim\kappa_{m,d}\Gamma_2L^{\frac12}2^{-n(\beta_2-\frac12)}|t_N-t_0|^{\beta_2}\\&+\kappa_{m,d}\Gamma_1f_k(L)2^{-n(\beta_1-\alpha-\frac12)}|t_N-t_0|^{\beta_1-\alpha}+\Gamma_1L^{-\alpha k}2^{-n(\beta_1-\alpha-1)}|t_N-t_0|^{\beta_1-\alpha},\end{aligned}\tag{1.47}$$

where

$$f_k(L)=\begin{cases}L^{\frac{k}{2}-\alpha(k-1)}\mathbf{1}_{\{k\ge2\}}, & \text{if } \alpha<\frac12,\\(k-1)L^{\frac12}, & \text{if } \alpha\ge\frac12.\end{cases}\tag{1.48}$$

We wish to choose L and k so that (1.47) is summable with respect to n.

- Assume $\alpha<\frac12$. For fixed $k\ge2$, we choose L so that

  $$L^{\frac{k}{2}-\alpha(k-1)}2^{-n(\beta_1-\alpha-\frac12)}+L^{-\alpha k}2^{-n(\beta_1-\alpha-1)}$$

  is minimized. Namely, we set $L:=\lfloor 2^{\frac{n}{k}}\rfloor$. Then,

  $$\|\sum_{i=0}^{2^n-1}R_i^n\|_{L^m(\mathbb{P})}\lesssim_\alpha\kappa_{m,d}\Gamma_22^{\frac{n}{k}}2^{-n(\beta_2-\frac12)}|t_N-t_0|^{\beta_2}+\Gamma_1\kappa_{m,d}2^{-n(\beta_1-1-\frac{\alpha}{k})}.$$

  Now we set k so that $\frac1k<\min\{2\beta_2-1,\frac{\beta_1-1}{\alpha}\}$.

- If $\alpha \geq \frac{1}{2}$, we set $L = \lfloor 2^{\delta n} \rfloor$, where

$$0 < \delta < 2\min\{\beta_2, \beta_1 - \alpha_1\} - 1.$$

Then, we choose k so that

$$\alpha k \delta + \beta_1 - \alpha - 1 > 0.$$

We also need to ensure $k\delta \leq 1$, but this is possible since $\frac{1+\alpha-\beta_1}{\alpha} < 1$.

In any case, we note that (1.47) is summable with respect to n and

$$\|A_{t_0,t_N} - \sum_{i=1}^{N} A_{t_{i-1},t_i}\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,M} \kappa_{m,d}\Gamma_2|t_N - t_0|^{\beta_2} + \kappa_{m,d}\Gamma_1|t_N - t_0|^{\beta_1-\alpha_1}. \quad \square$$

Proof of Lemma 1.2.4. The proof is similar to Lemma 1.2.3. Write

$$\pi' =: \{0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T\}$$

and

$$\{\, [s,t] \in \pi : t_j \leq s < t \leq t_{j+1} \,\} =: \{t_j = t_0^j < t_1^j < \cdots < t_{N_j-1}^j < t_{N_j}^j = t_{j+1}\}.$$

By (1.20), we have $N \leq 3|\pi'|^{-1}T$. We fix a parameter L, which will be chosen later, and set

$$Z_j^{(1),l} := A_{t_{jL+l},t_{jL+l+1}} - \sum_{k=1}^{N_{jL+l}} A_{t_{k-1}^{jL+l},t_k^{jL+l}}, \quad \mathcal{H}_j^{(1),l} := \mathcal{F}_{t_{(j-1)L+l+1}}.$$

Inductively, we set

$$Z_j^{(k),l_1,\dots,l_k} := Z_{jL+l_k}^{(k-1),l_1,\dots,l_{k-1}}, \quad \mathcal{H}_j^{(k),l_1,\dots,l_k} := \mathcal{H}_{(j-1)L+l_k}^{(k-1),l_1,\dots,l_{k-1}}.$$

As in Lemma 1.2.3, for each $k \in \mathbb{N}$, we consider the decomposition

$$A_T^{\pi'} - A_T^{\pi} = A + B$$

where

$$A := \sum_{p=1}^{k} \sum_{l_1,\dots,l_p \leq L} \sum_{j \leq NL^{-p}} \Big\{ \mathbb{E}[Z_j^{(p),l_1,\dots,l_p} | \mathcal{H}_{j+1}^{(p),l_1,\dots,l_p}] - \mathbb{E}[Z_j^{(p),l_1,\dots,l_p} | \mathcal{H}_j^{(p),l_1,\dots,l_p}] \Big\}$$

$$B := \sum_{l_1,\dots,l_k \leq L} \sum_{j \leq NL^{-k}} \mathbb{E}[Z_j^{(k),l_1,\dots,l_k} | \mathcal{H}_j^{(k),l_1,\dots,l_k}].$$

By the BDG inequality and the Cauchy-Schwarz inequality,

$$\|A\|_{L^m(\mathbb{P})} \lesssim \kappa_{m,d} \sum_{p=1}^{k} L^{\frac{p}{2}} \Big(\sum_{l_1,\dots,l_p \le L} \sum_{j \le NL^{-p}} \|\mathbb{E}[Z_j^{(p),l_1,\dots,l_p} | \mathcal{H}_{j+1}^{(p),l_1,\dots,l_p}]\|_{L^m(\mathbb{P})}^2 \Big)^{\frac{1}{2}}.$$

By Lemma 1.2.3,

$$\|Z_j^{(1),l}\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2} \Gamma_1 |t_{jL+l+1} - t_{jL+l}|^{\beta_1-\alpha} + \kappa_{m,d}\Gamma_2 |t_{jL+l+1} - t_{jL+l}|^{\beta_2}.$$

For $p \ge 2$, by Lemma 1.2.2 and (1.20),

$$\|\mathbb{E}[Z_j^{(p),l_1,\dots,l_p} | \mathcal{H}_{j+1}^{(p),l_1,\dots,l_p}]\|_{L^m(\mathbb{P})} \lesssim_{\beta_1} \Gamma_1 L^{-(p-1)\alpha} |\pi'|^{-\alpha} |\pi'|^{\beta_1}.$$

Therefore, we obtain

$$\|A\|_{L^m(\mathbb{P})} \lesssim_{\alpha,\beta_1,\beta_2,m,d,T} L^{\frac{1}{2}} (\Gamma_1 |\pi'|^{\beta_1-\alpha-\frac{1}{2}} + \Gamma_2 |\pi'|^{\beta_2-\frac{1}{2}}) + \Gamma_1 f_k(L) |\pi'|^{\beta_1-\alpha-\frac{1}{2}}, \quad (1.49)$$

where $f_k(L)$ is defined by (1.48).

We move to estimate B. By Lemma 1.2.2 and (1.20),

$$\|\mathbb{E}[Z_j^{(k),l_1,\dots,l_k} | \mathcal{H}_j^{(k),l_1,\dots,l_k}]\|_{L^m(\mathbb{P})} \lesssim_{\beta_1} \Gamma_1 L^{-\alpha k} |\pi'|^{\beta_1-\alpha}.$$

Therefore,

$$\|B\|_{L^m(\mathbb{P})} \lesssim_{\beta_1,T} \Gamma_1 L^{-\alpha k} |\pi'|^{\beta_1-\alpha-1}. \quad (1.50)$$

Combining (1.49) and (1.50), we obtain

$$\begin{aligned} &\|A_T^{\pi'} - A_T^{\pi}\|_{L^m(\mathbb{P})} \\ &\lesssim_{\alpha,\beta_1,\beta_2,m,d,T} L^{\frac{1}{2}} (\Gamma_1 |\pi'|^{\beta_1-\alpha-\frac{1}{2}} + \Gamma_2 |\pi'|^{\beta_2-\frac{1}{2}}) + \Gamma_1 f_k(L) |\pi'|^{\beta_1-\alpha-\frac{1}{2}} + \Gamma_1 L^{-\alpha k} |\pi'|^{\beta_1-\alpha-1}. \end{aligned}$$

By choosing L and k as in the proof of Lemma 1.2.3 (replace 2^n by $|\pi'|^{-1}$), we complete the proof. □

Proof of Lemma 1.3.2

Let $d = 1$. For $u > v$, we set

$$B_u^{(1)} := \int_{-\infty}^{v} \mathcal{K}(u,r) \mathrm{d}W_r, \quad B_u^{(2)} := \int_{v}^{u} \mathcal{K}(u,r) \mathrm{d}W_r$$

so that $B_u - B(0) = B_u^{(1)} + B_u^{(2)}$ and $B_\cdot^{(1)}$ and $B_\cdot^{(2)}$ are independent. Then, we have

$$\mathbb{E}[B_s^{(2)} B_t^{(2)}] = \int_v^s \mathcal{K}(s,r)\mathcal{K}(t,r)\mathrm{d}r,$$

and by (1.24), we have

$$\frac{c_H}{2}(s^{2H} + t^{2H} - |t-s|^{2H}) = \mathbb{E}[B_s^{(1)} B_t^{(1)}] + \mathbb{E}[B_s^{(2)} B_t^{(2)}],$$

and thus, we will estimate $\mathbb{E}[B_s^{(1)} B_t^{(1)}]$. We have

$$\mathbb{E}[B_s^{(1)} B_t^{(1)}] = \int_0^\infty \left[(s+r)^{H-1/2} - r^{H-1/2}\right]\left[(t+r)^{H-1/2} - r^{H-1/2}\right]\mathrm{d}r + \int_0^v (s-r)^{H-1/2}(t-r)^{H-1/2}\mathrm{d}r. \quad (1.51)$$

By [Pic11, Theorem 33], the first term of (1.51) equals to

$$(c_H - (2H)^{-1})s^{2H} + \int_0^\infty \left[(s+r)^{H-1/2} - r^{H-1/2}\right]\left[(t+r)^{H-1/2} - (s+r)^{H-1/2}\right]\mathrm{d}r. \quad (1.52)$$

Since

$$(t+r)^{H-1/2} - (s+r)^{H-1/2} = (H-1/2)(s+r)^{H-3/2}(t-s) + O((s+r)^{H-5/2}(t-s)^2),$$

the second term of (1.52) equals to

$$s^{2H-1}(t-s)(H-1/2)\int_0^\infty \left[(1+r)^{H-1/2} - r^{H-1/2}\right](1+r)^{H-3/2}\mathrm{d}r + O(s^{2H-2}(t-s)^2).$$

By [Pic11, Theorem 33],

$$(H-1/2)\int_0^\infty \left[(1+r)^{H-1/2} - r^{H-1/2}\right](1+r)^{H-3/2}\mathrm{d}r = -\frac{1}{2} + Hc_H.$$

Similarly, the second term of (1.51) equals to

$$\frac{1}{2H}(s^{2H} - (s-v)^{2H}) + \frac{t-s}{2}(s^{2H-1} - (s-v)^{2H-1}) + O((s-v)^{2H-2}(t-s)^2).$$

Therefore, $\mathbb{E}[B_s^{(1)} B_t^{(1)}]$ equals to

$$c_H s^{2H} + Hc_H s^{2H-1}(t-s) - \frac{1}{2H}(v-s)^{2H} - \frac{1}{2}(s-v)^{2H-1}(t-s) + O((s-v)^{2H-2}(t-s)^2).$$

Since

$$\frac{c_H}{2}(s^{2H} + t^{2H} - |t-s|^{2H}) - c_H s^{2H} + Hc_H s^{2H-1}(t-s) - \frac{1}{2H}(v-s)^{2H} = -\frac{c_H}{2}(t-s)^{2H} + O((s-v)^{2H-2}(t-s)^2),$$

the proof is complete. ∎

Bibliography

[BDG72] D. L. Burkholder, B. J. Davis, and R. F. Gundy. “Integral inequalities for convex functions of operators on martingales”. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*. 1972, pp. 223–240.

[FL06] D. Feyel and A. de La Pradelle. “Curvilinear integrals along enriched paths”. *Electron. J. Probab.* 11 (2006), no. 34, 860–892.

[FH20] P. K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Cham: Springer International Publishing, 2020.

[Ger22] M. Gerencsér. “Regularisation by regular noise”. *Stoch. PDE: Anal. Comp.* (2022).

[Gub04] M. Gubinelli. “Controlling rough paths”. *J. Funct. Anal.* 216.1 (2004), pp. 86–140.

[Hai14] M. Hairer. “A theory of regularity structures”. *Invent. Math.* 198.2 (2014), pp. 269–504.

[Ker21] H. Kern. *A stochastic reconstruction theorem*. 2021. arXiv: 2107.03867 [math.PR].

[Lê20] K. Lê. “A stochastic sewing lemma and applications”. *Electron. J. Probab.* 25 (2020), Paper No. 38, 55.

[Lê23] K. Lê. “Stochastic sewing in Banach spaces”. *Electronic Journal of Probability* 28.none (Jan. 2023), pp. 1–22.

[Lyo98] T. J. Lyons. “Differential equations driven by rough signals”. *Rev. Mat. Iberoamericana* 14.2 (1998), pp. 215–310.

[MV68] B. B. Mandelbrot and J. W. Van Ness. “Fractional Brownian Motions, Fractional Noises and Applications”. *SIAM Review* 10.4 (1968), pp. 422–437.

[Nou12] I. Nourdin. *Selected aspects of fractional Brownian motion*. Vol. 4. Bocconi & Springer Series. Springer, Milan; Bocconi University Press, Milan, 2012, pp. x+122.

[Pic08] J. Picard. “A tree approach to p-variation and to integration”. *Ann. Probab.* 36.6 (2008), pp. 2235–2279.

[Pic11] J. Picard. “Representation Formulae for the Fractional Brownian Motion”. In: *Séminaire de Probabilités XLIII*. Ed. by C. Donati-Martin, A. Lejay, and A. Rouault. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 3–70.

[Yas18] P. Yaskov. "Extensions of the sewing lemma with applications". *Stochastic Process. Appl.* 128.11 (2018), pp. 3940–3965.

Chapter 2

Level crossings of fractional Brownian motions

We prove that the number of level crossings of the fractional Brownian motion, after normalization, converges to its local time. This resolves a conjecture posed in [CP19], and our result can be viewed as an extension of [Lem83] for the fractional Brownian motion. We also prove the convergence of the $(1/H)$-variation, where H is the Hurst parameter, along random partitions defined by level crossings. This result raises an interesting conjecture, which seems to capture non-Markovianity of the fractional Brownian motion.

This chapter is based on joint work with Purba Das, Rafał Łochowski and Nicolas Perkowski.

Keywords and phrases. fractional Brownian motion, level crossings, local time, excursions, stochastic sewing lemma.
MSC 2020. 60G22, 60J55

Contents

2.1 Introduction **50**

2.1.1 Main results 52

2.2 Variations along Lebesgue partitions **56**

2.2.1 Elementary results 56

2.2.2 Convergence of the variations 61

2.3 Local time via Lebesgue partitions **66**

2.3.1 Heuristics 66

2.3.2 Convergence to the local time 68

2.3.3 Lemiuex type result 83

2.4 A fractional excursion measure **86**

Bibliography **93**

2.1 Introduction

Level crossings of stochastic processes have been studied since the classical works of Kac [Kac43] and Rice [Ric45]. Depending on whether the process is smooth or rough, the study of its level crossings rely on different methods. As for the smooth case, which is not the scope of this chapter, the reader can refer to the survey article [Kra06] and the textbook [AW09].

By far the most prominent example of rough stochastic processes is the Brownian motion. The first work on level crossings of the Brownian motion is attributed to Lévy [Lév48] , who characterized its local time as a limit of the counting of level crossings. More precisely, for a given process w, setting

$$U_{0,t}(\varepsilon, w) := \#\big\{(u,v) : 0 \leq u < v \leq t,\ \ w_u = 0, w_v = \varepsilon, \forall r \in (u,v)\ w_r \in (0,\varepsilon)\big\},$$

we have for the Brownian motion W and $a \in \mathbb{R}$,

$$\lim_{\varepsilon \to 0} \varepsilon U_{0,t}(\varepsilon, W - a) = L_t(a)$$

almost surely, where $L_t(a)$ is the local time at time t at the level a. This result can now be found in standard textbooks such as [IM74], [RY99] and [MP10], and it can be generalized for semimartingales [El 78] and for Markov processes [FT83].

There are many rough stochastic processes that are neither semimartingale nor Markovian. Among them is the *fractional Brownian motion* B^H, a Gaussian process parametrized by $H \in (0,1)$. (Precisely, B^H is neither semimartingale nor Markovian for $H \neq \frac{1}{2}$, and for $H = \frac{1}{2}$ it is the Brownian motion.) The process B^H is known to have the local time. In view of Lévy's result on the Brownian local time, it is very natural to ask if an analogous result holds for the fractional Brownian local time. So far, no complete answer is obtained. This is surprising, considering the age of Lévy's result and that of the fractional Brownian motion.

There are some works related to the question on the level crossing characterization of the fractional Brownian local time. For instance, [Aza90; AW96] show that the number of zeroes for some smoothed fractional Brownian motion converges in suitable sense to the local time. We note that this question gets attention in the pathwise stochastic calculus [DOS18; CP19; Łoc+21; Kim22; ACX20] as well as in some applied literatures [FHW94; Kru98].

Constructing the local time via level crossings is not only a natural problem, but also it can lead to a significant implication on the path property of the process. This was first observed by the brilliant thesis [Lem83] of Lemieux. Therein he proved the existence of a measurable set Ω_W such that $\mathbb{P}(W \in \Omega_W) = 1$ (recall that W is Brownian motion) and for every $w \in \Omega_W$, $a \in \mathbb{R}$ and $t \geq 0$, the limit

$$\lim_{\varepsilon \to 0} \varepsilon U_{0,t}(\varepsilon, w - a)$$

exists. Hence, the existence of the limit of the normalized level-crossing counting is a path property. This result explains why such construction of the local time receives attention in the pathwise stochastic calculus. It is worth noting that Lemieux proved the result for a large class of semimartingales.

Lemieux's result has a remarkable consequence on *pathwise quadratic variation*, calculated as a limit of sum square increments where the increments are taken along partitions of a fixed interval with vanishing mesh. The precise definition of the pathwise quadratic variation is as follows: given a sequence $\boldsymbol{\pi}$ of partitions π_n ($n \in \mathbb{N}$) with vanishing mesh, the pathwise quadratic variation $[w]_{\boldsymbol{\pi}}$ of a process w is defined by

$$[w]_{\boldsymbol{\pi}} := \lim_{n \to \infty} \sum_{[s,t] \in \pi_n} |w_t - w_s|^2$$

whenever the limit exists. In general, the pathwise quadratic variation (even when it exists) may depend on the choice of a sequence of partitions [DOS18].

Hence, an obvious question is if, given a stochastic process X, there are any (big) class P of partition sequences such that almost surely for any $\boldsymbol{\pi}, \boldsymbol{\pi}' \in \mathrm{P}$ we have $[X]_{\boldsymbol{\pi}} = [X]_{\boldsymbol{\pi}'}$. The classical works [Lév40; Lév48] of Lévy show that for any refining partition sequence

$\boldsymbol{\pi}$ of $[0, t]$ with vanishing mesh we have

$$\mathbb{P}([W]_{\boldsymbol{\pi}} = t) = 1. \tag{2.1}$$

Dudley [Dud73] proved that if the sequence $\boldsymbol{\pi} = (\pi_n)$ satisfies $|\pi_n| = o(1/\log n)$ then (2.1) holds. He even showed the optimality of the decaying condition. Among the works on the quadratic variation of the Brownian motion, most relevant to us is Lemieux's result, which proves that the Brownian motion has a measure zero set outside which any quadratic variation of the Brownian motion along any uniform *Lebesgue partition* (defined at the beginning of Section 2.1.1) of $[0, t]$ with vanishing mesh is equal to t. We remark that, unlike Dudley's result, there is no decaying condition on partition sequences in Lemieux's result.

In this chapter, we extend Levy's construction of the local time and Lemieux's result for fractional Brownian motions.

2.1.1 Main results

We write $B = B^H$ for a fractional Brownian motion (fBm) with $B_0 = 0$ and with Hurst parameter $H \in (0, 1)$. Given a partition π, we write

$$|\pi| := \max_{[s,t]\in\pi} |t - s|.$$

To define the Lebesgue partition, let us introduce some definitions. Given a process w and a positive constant ε, we set $T_0(\varepsilon, w) := 0$ and inductively

$$T_n(\varepsilon, w) := \inf\{t > T_{n-1}(\varepsilon, w) : w_t \in \varepsilon\mathbb{Z} \setminus \{w_{T_{n-1}(\varepsilon,w)}\}\}. \tag{2.2}$$

(If $T_{n-1} = +\infty$, we set $T_n := +\infty$.) Note that we do not assume $w_0 = 0$. See Figure 2.1 for a graphics. We denote by $K_{s,t}(\varepsilon, w)$ the number of ε-level crossings in $[s, t]$, that is[1]

$$K_{s,t}(\varepsilon, w) := \#\{n \in \mathbb{N} \setminus \{1\} : T_n(\varepsilon, w_{s+\cdot}) \le t\} + \mathbf{1}_{\{w_s\in\varepsilon\mathbb{Z}\}}\mathbf{1}_{\{T_1(\varepsilon,w_{s+\cdot})\le t\}}. \tag{2.3}$$

Remark 2.1.1. Our notation for the total number of ε-level crossings is slightly different from [Lem83]. $K_{s,t}(\varepsilon, w)$ in (2.3) represents $K_s^t(w, \varepsilon\mathbb{Z})$ in Lemieux's notation.

The partition

$$\{[T_{n-1}(\varepsilon, w), T_n(\varepsilon, w)] : n \in \mathbb{N}, T_n(\varepsilon, w) \le t\} \tag{2.4}$$

[1]Our convention is $\mathbb{N} = \{1, 2, 3, \ldots\}$. In particular $0 \notin \mathbb{N}$.

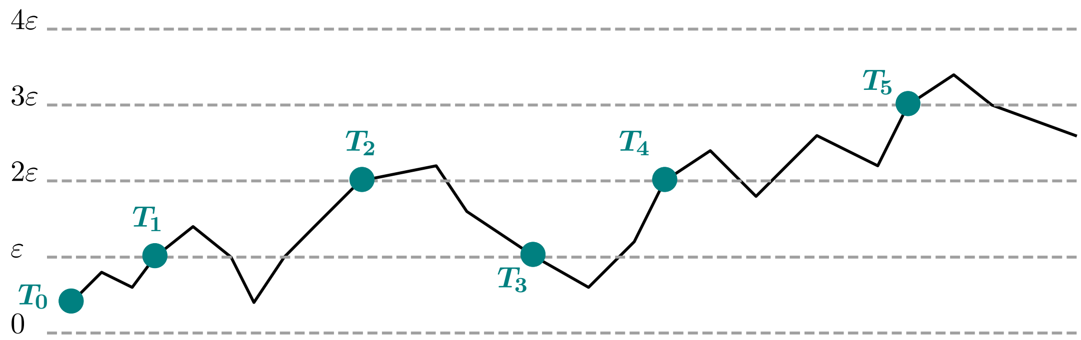


Figure 2.1: Visualisation of $T_n(\varepsilon, w)$.

is called a *Lebesgue partition*. We observe

$$\sum_{n:T_n(\varepsilon,w)\leq t} |w_{T_n(\varepsilon,w)} - w_{T_{n-1}(\varepsilon,w)}|^{\frac{1}{H}} = \varepsilon^{1/H} K_{0,t}(\varepsilon, w) + |w_{T_1(\varepsilon,w)} - w_0|^{\frac{1}{H}}.$$

Note that $|w_{T_1(\varepsilon,w)} - w_0| \leq \varepsilon$. Therefore, the study of the $1/H$-variation along a sequence of Lebesgue partitions is equivalent to that of $K_{0,t}(\varepsilon, w)$ as $\varepsilon \downarrow 0$.

.

For $\rho \in \mathbb{R}$ and a process w, the process $w + \rho$ is defined by $(w + \rho)_t := w_t + \rho$. Our first main result is on the $1/H$-variation along Lebesgue partitions of the fractional Brownian motion B.

Theorem 2.1.2 (Convergence of the variation along Lebesgue partitions)**.** *For every $H \in (0, 1)$, there exists a positive constant $\mathfrak{c}_H$ with the following property. Let $\rho \in \mathbb{R}$, $T \in (0, \infty)$ and $(\varepsilon_n)_{n=1}^{\infty}$ be a sequence of positive numbers such that $\varepsilon_n = O(n^{-\eta})$ for some $\eta > 0$. Then, we have*

$$\lim_{n\to\infty} \varepsilon_n^{1/H} K_{0,T}(\varepsilon_n, B^H + \rho) = \mathfrak{c}_H T, \quad \textit{almost surely.} \tag{2.5}$$

Theorem 2.1.2 concerns level crossings at all levels. We can also consider level crossings at a specific level. For $s < t$, we set $\Delta_{s,t} := \{(u, v) : s \leq u < v \leq t\}$. For each $\varepsilon \in (0, 1)$ and $w \in C([0, T]; \mathbb{R})$, we consider the number of upcrossings by setting

$$U_{s,t}(\varepsilon, w) := \#\big\{(u, v) \in \Delta_{s,t} : w_u = 0, w_v = \varepsilon, \forall r \in (u, v)\ w_r \in (0, \varepsilon)\big\}. \tag{2.6}$$

See Figure 2.2. In the case of the Brownian motion ($H = \frac{1}{2}$), it is well-known (e.g. [IM74,

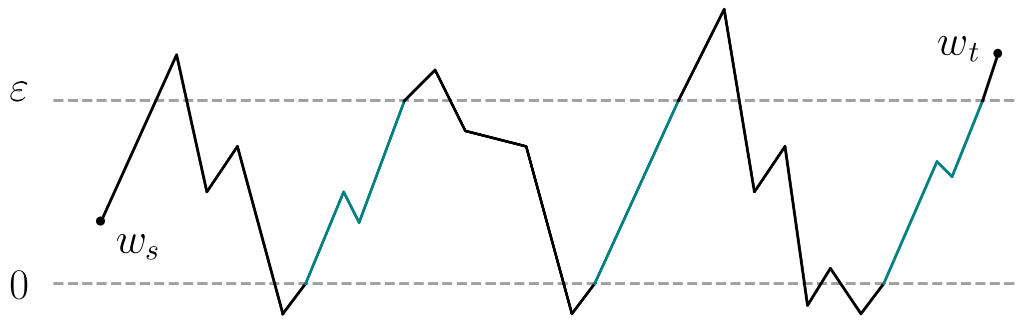


Figure 2.2: Visualisation of $U_{s,t}(\varepsilon, w)$. In the picture, $U_{s,t}(\varepsilon, w) = 3$.

Section 2.2], [RY99, Chapter VI], [MP10, Section 6]) that we have

$$\lim_{\varepsilon \to 0} \varepsilon U_{0,t}(\varepsilon, B^{1/2} - a) = \frac{1}{2} L_t^{1/2}(a) \quad \text{almost surely,} \tag{2.7}$$

where $L^{1/2}$ is the local time of the Brownian motion $B^{1/2}$. This representation of the local time is extended for semimartingales by Karoui [El 78].

We recall that the notion of the local time exists for general $H \in (0, 1)$.

Definition 2.1.3. We denote by $(L_t(a))_{t \geq 0} = (L_t^H(a))_{t \geq 0}$ the local time of B^H at the level a. That is, L is a unique random field satisfying the following occupation density formula:

$$\int_0^t f(B_r) \mathrm{d}r = \int_{\mathbb{R}} f(a) L_t(a) \mathrm{d}a, \quad \forall t \geq 0, \forall f \in C_c^\infty(\mathbb{R}).$$

As for the existence of L, see e.g. [GH80].

Our next main result is to prove analogue of (2.7) for $H < \frac{1}{2}$.

Theorem 2.1.4 (Local time via level crossings)**.** *Let $H < \frac{1}{2}$, $a \in \mathbb{R}$, $T \in (0, \infty)$. Then we have*

$$\lim_{\varepsilon \to 0, \varepsilon > 0} \varepsilon^{\frac{1}{H} - 1} U_{0,T}(\varepsilon, B^H - a) = \frac{\mathfrak{c}_H}{2} L_T^H(a) \quad \textit{almost surely.}$$

More strongly, we have the following result.

Theorem 2.1.5 (Lemieux type result)**.** *Let $H < \frac{1}{2}$. Then, there exists a measurable set $\Omega_H \subseteq C([0, \infty); \mathbb{R})$ with the following property.*

- *We have* $\mathbb{P}(B^H \in \Omega_H) = 1$.
- *For every* $w \in \Omega_H$, $t \in [0, \infty)$ *and* $a \in \mathbb{R}$, *the limit*

$$L_t(a, w) := \lim_{\varepsilon \to 0, \varepsilon > 0} \varepsilon^{\frac{1}{H} - 1} U_{0,t}(\varepsilon, w - a)$$

exists and finite. The limit $(L_t(a, w))_{a \in \mathbb{R}}$ *satisfies the occupation density formula*

$$\int_0^t f(w_r) \mathrm{d}w_r = \int_{\mathbb{R}} f(a) L_t(a, w) \mathrm{d}a$$

for every $t \geq 0$ *and continuous* f. *Furthermore,*

$$\lim_{\varepsilon \to 0, \varepsilon > 0} \varepsilon^{\frac{1}{H}} K_{0,t}(\varepsilon, B - a) = \mathfrak{c}_H t.$$

Proof. See Theorems 2.3.21 and 2.3.22. □

Remark 2.1.6. As noted in Remark 1.1.3, there exists a $\varepsilon \in (0, 1)$ such that for every $m \in (0, \infty)$ we have

$$\Big\| \sum_{[u,v] \in \pi} |B_v - B_u|^{\frac{1}{H}} - \mathbb{E}\big[|B_1|^{\frac{1}{H}}\big] t \Big\|_{L^m(\mathbb{P})} \lesssim_{m,t} |\pi|^{\varepsilon}$$

for any deterministic partition π of $[0, t]$. Therefore, by the Borel–Cantelli lemma, for any sequence $(\pi_n)_{n=1}^{\infty}$ of partitions with

$$|\pi_n| = O(n^{-\delta}), \quad \delta \in (0, \infty), \tag{2.8}$$

we have $\lim_{n \to \infty} \sum_{[u,v] \in \pi} |B_v - B_u|^{\frac{1}{H}} = \mathbb{E}[|B_1|^{\frac{1}{H}}] t$ almost surely. Unlike Theorem 2.1.5, we need the decaying condition (2.8). In view of [Dud73], the condition (2.8) is not optimal; finding the optimal condition seems open.

Conjecture

There is an interesting aspect on the constant $\mathfrak{c}_H$. For the Brownian motion, the quadratic variation along any deterministic partition almost surely matches with the quadratic variation along any uniform Lebesgue type partitions. That is,

$$\mathfrak{c}_{\frac{1}{2}} = \mathbb{E}[(B_1^{1/2})^2]. \tag{2.9}$$

It is tempting to guess that such relation holds for $H \neq \frac{1}{2}$ as well. Indeed, such conjecture is stated in [CP19]. However, the identity (2.9) is due to Markovianity of the Brownian motion. Therefore, for $H \neq \frac{1}{2}$, there is no reason to believe that $\mathfrak{c}_H$ and $\mathbb{E}[|B_1^H|^{\frac{1}{H}}]$ are equal. Motivated by the simulation shown in Figures 2.3 and 2.4, we propose the following remarkable conjecture.

Conjecture 2.1.7 (The constant $\mathfrak{c}_H$)**.** For the fractional Brownian motion with $H \neq \frac{1}{2}$, we conjecture that the $\frac{1}{H}$ variation of the fractional Brownian motion along deterministic partitions differs from the $\frac{1}{H}$ variation of the fractional Brownian motion along uniform Lebesgue partitions. To be more precise, we conjecture

$$\begin{cases} \mathfrak{c}_H > \mathbb{E}[|B_1^H|^{1/H}] & \text{if } H < \frac{1}{2}, \\ \mathfrak{c}_H < \mathbb{E}[|B_1^H|^{1/H}] & \text{if } H > \frac{1}{2}. \end{cases}$$

If this is indeed the case, the constant $\mathfrak{c}_H$ captures non-Markovianity of the fractional Brownian motion.

Notation

Given a path $f\colon [0,T] \to \mathbb{R}^d$, we write $f_{s,t} := f_t - f_s$ and we denote by $\dot{f}$ the derivative $\frac{\mathrm{d}f}{\mathrm{d}t}$. We write $A \lesssim B$ if there exists a positive constant C, depending only on unimportant parameters, such that $A \leq CB$. If we want to emphasize the dependency on parameters $\alpha, \beta, \ldots$, then we write $A \lesssim_{\alpha,\beta,\ldots} B$. **In this chapter we will not write down dependency on H.**

2.2 Variations along Lebesgue partitions

The goal of this section is to prove Theorem 2.1.2. We begin observing elementary results on the counting K of level crossings, defined by (2.3).

2.2.1 Elementary results

Recall the definition of the fractional Brownian motion B from Definition 1.3.1 and the Mandelbrot–van Ness representation (1.25), which will be used throughout the chapter.

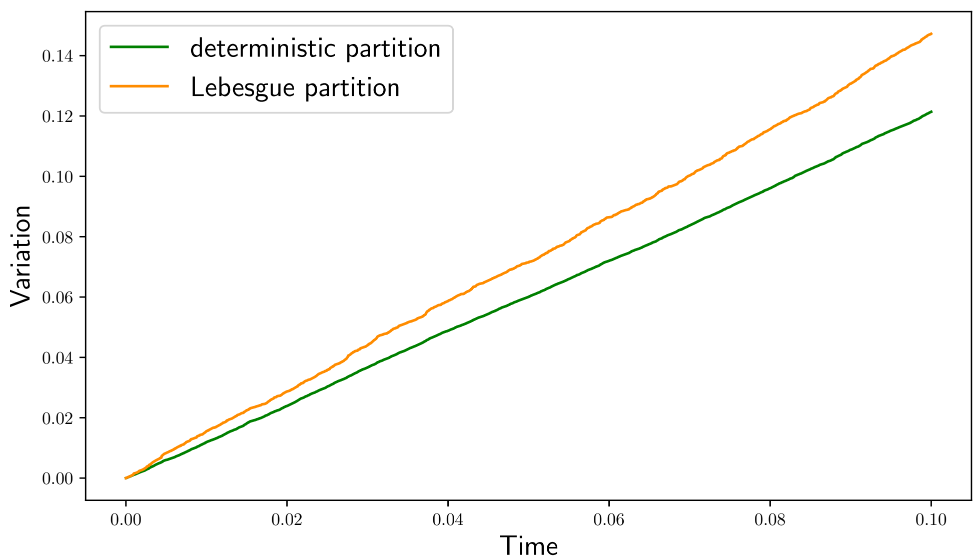
Lemma 2.2.1 (scaling of K)**.** *For $\lambda \in (0,\infty)$, we have*

$$(K_{s,t}(\varepsilon, B+\rho))_{s<t,\varepsilon>0,\rho\in\mathbb{R}} \overset{\mathrm{d}}{=} (K_{\lambda^{1/H}s,\lambda^{1/H}t}(\lambda\varepsilon, B+\lambda\rho))_{s<t,\varepsilon>0,\rho\in\mathbb{R}}.$$

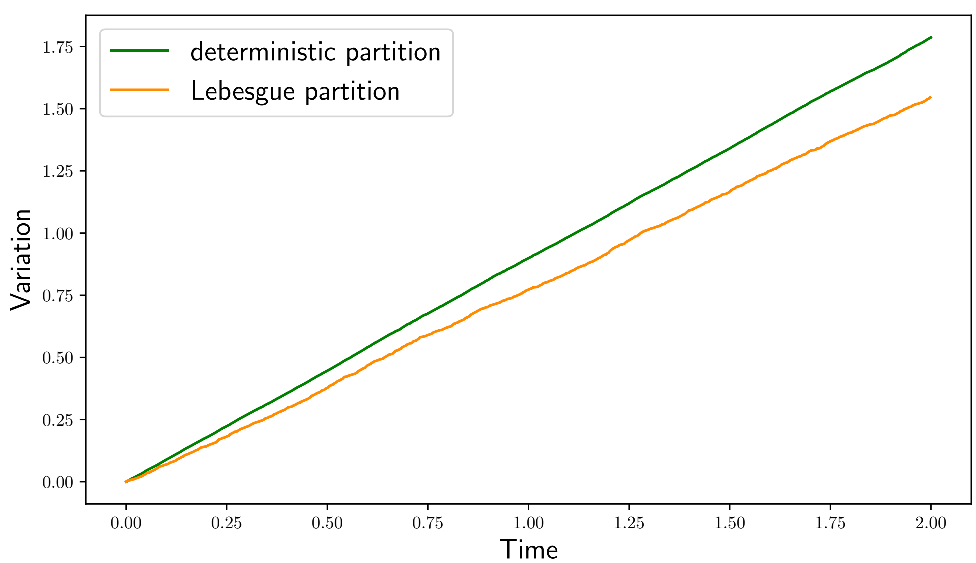
[2]We simulate the variation

$$V_t := \sum_{[u,v]\in\pi^{\#},v\leq t} |B_v - B_u|^{\frac{1}{H}}$$

up to time T with $\# \in \{\text{deterministic}, \text{Lebesgue}\}$. The fractional Brownian motion is discretized with step size T/n. We have $\pi^{\text{deterministic}} = \{kT/n\}_{k=1}^n$ and $\pi^{\text{Lebesgue}} = \{T_k(\varepsilon, B)\}_k$. For $H = 0.4$, it is simulated with $T = 0.1$, $n = 30000$, $\varepsilon = 0.015$. For $H = 0.6$, it is simulated with $T = 2$, $n = 30000$, $\varepsilon = 0.013$.



(a) $H = 0.4$



(b) $H = 0.6$

Figure 2.3: Comparison between the variation along a deterministic uniform partition and that along a Lebesgue partition. [2]

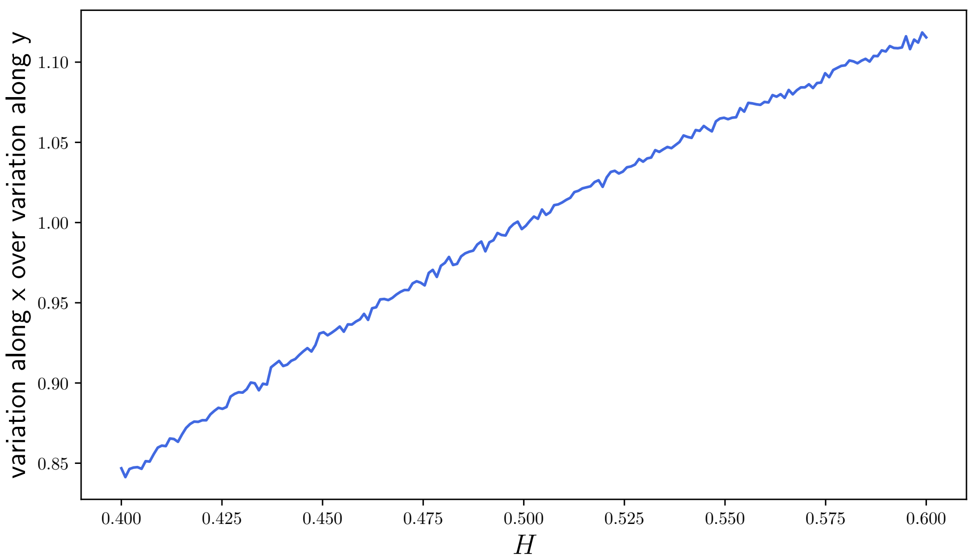


Figure 2.4: Plotting $\mathbb{E}[|B_1^H|^{1/H}]/\mathfrak{c}_H$ for $0.4 \leq H \leq 0.6$. The oscillation should be due to simulation error. The graph is expected to be increasing, with $\mathbb{E}[|B_1^{1/2}|^2]/\mathfrak{c}_{1/2} = 1$.

Proof. We set $B^{(\lambda)}_t := \lambda B_{\lambda^{-1/H}t}$. Note that $B^{(\lambda)} \overset{\mathrm{d}}{=} B$ and observe that

$$K_{s,t}(\varepsilon, B_t + \rho) = K_{s,t}(\lambda\varepsilon, \lambda(B + \rho)) = K_{\lambda^{1/H}s,\lambda^{1/H}t}(\lambda\varepsilon, B^{(\lambda)} + \lambda\rho). \qquad \square$$

Lemma 2.2.2 (superadditivity of K)**.** *Let $r < s < t$ and w be a process. Then,*

$$K_{r,s}(\varepsilon, w) + K_{s,t}(\varepsilon, w) \leq K_{r,t}(\varepsilon, w) \leq K_{r,s}(\varepsilon, w) + K_{s,t}(\varepsilon, w) + 1.$$

Proof. Recalling the definition of T_n from (2.2), we set

$$N := \max\{n \in \mathbb{N} \cup \{0\} : T_n(\varepsilon, w_{r+\cdot}) < s\}.$$

- If $w_{T_N(\varepsilon,w_{r+\cdot})} = w_{T_1(\varepsilon,w_{s+\cdot})}$, then $K_{r,t}(\varepsilon, w) = K_{r,s}(\varepsilon, w) + K_{s,t}(\varepsilon, w)$.
- If $w_{T_N(\varepsilon,w)} \neq w_{T_1(\varepsilon,w_{s+\cdot})}$, then $K_{r,t}(\varepsilon, w) = K_{r,s}(\varepsilon, w) + K_{s,t}(\varepsilon, w) + 1$. $\square$

For our arguments, the following variants of K will appear.

Definition 2.2.3. We set

$$\bar{K}_{s,t}(\varepsilon, w) := \varepsilon^{-1} \int_{-\varepsilon/2}^{\varepsilon/2} K_{s,t}(\varepsilon, w + \rho)\mathrm{d}\rho, \quad J_{s,t}(\varepsilon, w) := \sup_{\rho \in \mathbb{R}} K_{s,t}(\varepsilon, w + \rho).$$

Note that we have the obvious inequality $\bar{K}_{s,t}(\varepsilon, w) \leq J_{s,t}(\varepsilon, w)$.

The advantage of $\bar{K}$ is that in addition to the superadditivity, it is stationary.

Lemma 2.2.4 (scaling, superadditivity and stationarity of $\bar{K}$)**.** *Let $r < s < t$ and let w be a process.*

(i) For $\lambda > 0$, we have

$$(\bar{K}_{s,t}(\varepsilon, B))_{s<t,\varepsilon>0} \overset{\mathrm{d}}{=} (\bar{K}_{\lambda^{1/H}s,\lambda^{1/H}t}(\lambda\varepsilon, B))_{s<t,\varepsilon>0}.$$

(ii) We have

$$\bar{K}_{r,s}(\varepsilon, w) + \bar{K}_{s,t}(\varepsilon, w) \leq \bar{K}_{r,t}(\varepsilon, w). \leq \bar{K}_{r,s}(\varepsilon, w) + \bar{K}_{s,t}(\varepsilon, w) + 1.$$

(iii) We have $\bar{K}_{s,t}(\varepsilon, w) = \bar{K}_{0,t-s}(\varepsilon, w_{s+\cdot} - w_s)$. In particular,

$$\bar{K}_{s,t}(\varepsilon, B) \overset{\mathrm{d}}{=} \bar{K}_{0,t-s}(\varepsilon, B).$$

Proof. The claim (a) follows from Lemma 2.2.1 and the claim (b) follows from Lemma 2.2.2. For the claim (c), we observe that for every $\rho \in \mathbb{R}$ we have

$$\bar{K}_{s,t}(\varepsilon, w) = \bar{K}_{s,t}(\varepsilon, w + \rho) = \bar{K}_{0,t-s}(\varepsilon, w_{s+\cdot} + \rho)$$

In particular, we choose $\rho := -w_s$. $\square$

Lemma 2.2.5. *For every $p, t, \varepsilon \in (0, \infty)$ we have $\mathbb{E}[J_{0,t}(\varepsilon, B)^p] < \infty$.*

Proof. For $\alpha \in (0, H)$, we set

$$[\![B]\!]_{C^\alpha([0,t])} := \sup_{0 \le r < s \le t} \frac{|B_s - B_r|}{(s-r)^\alpha}.$$

By the Kolmogorov continuity theorem, we have

$$\mathbb{E}[[\![B]\!]^p_{C^\alpha([0,t])}] < \infty. \tag{2.10}$$

We set

$$\delta := \left\lceil \varepsilon^{-\frac{1}{\alpha}} (1 + [\![B]\!]_{C^\alpha([0,1])})^{\frac{1}{\alpha}} \right\rceil^{-1}.$$

Suppose that there exists n such that

$$k\delta \le T_n(\varepsilon, B + \rho) < T_{n+1}(\varepsilon, B + \rho) \le (k+1)\delta \quad \text{with } T_{n+1}(\varepsilon, B + \rho) \le t.$$

Then,

$$\begin{aligned} \varepsilon = |B_{T_{n+1}(\varepsilon, B+\rho)} - B_{T_n(\varepsilon, B+\rho)}| &\le [\![B]\!]_{C^\alpha([0,t])} \delta^\alpha \\ &< \varepsilon [\![B]\!]_{C^\alpha([0,1])} (1 + [\![B]\!]_{C^\alpha([0,1])})^{-1} < \varepsilon, \end{aligned}$$

which is a contradiction. Thus, we must have

$$\#\{n : k\delta \le T_n(\varepsilon, B + \rho) \le (k+1)\delta\} \le 1 \quad \text{for each } k$$

and

$$J_{0,t}(\varepsilon, B) = \sup_{\rho \in \mathbb{R}} K_{0,t}(\varepsilon, B + \rho) \le \delta^{-1} = \left\lceil \varepsilon^{-\frac{1}{\alpha}} (1 + [\![B]\!]_{C^\alpha([0,1])})^{\frac{1}{\alpha}} \right\rceil, \tag{2.11}$$

which is $L^p(\mathbb{P})$-integrable by (2.10). $\square$

In view of Lemma 2.2.4, the family $(\mathbb{E}[\bar{K}_{0,t}(1,B)])_{t\geq 0}$ satisfies

$$\mathbb{E}[\bar{K}_{0,s+t}(1,B)] \geq \mathbb{E}[\bar{K}_{0,s}(1,B)] + \mathbb{E}[\bar{K}_{0,t}(1,B)].$$

Therefore, we have

$$\mathfrak{c}_H := \lim_{t\to\infty} \frac{1}{t}\mathbb{E}[\bar{K}_{0,t}(1,B)] = \sup_{t>0} \frac{1}{t}\mathbb{E}[\bar{K}_{0,t}(1,B)].$$

The constant $\mathfrak{c}_H$ coincides with the one from Theorem 2.1.2. The following lemma shows that the constant $\mathfrak{c}_H$ is non-trivial.

Lemma 2.2.6. *We have $\mathfrak{c}_H \in (0,\infty)$.*

Proof. To see $\mathfrak{c}_H > 0$, we observe

$$\mathfrak{c}_H \geq \mathbb{E}[\bar{K}_{0,1}(1,B)] \geq \mathbb{P}(B_1 \geq 2) > 0.$$

To see $\mathfrak{c}_H < \infty$, we note by Lemma 2.2.4 that $(\bar{K}_{s,t}+1)$, $s<t$ is subadditive. Therefore,

$$\mathfrak{c}_H \leq \mathbb{E}[\bar{K}_{0,1}(1,B)] + 1 \leq \mathbb{E}[J_{0,1}(1,B)] + 1,$$

which is finite by Lemma 2.2.5. $\square$

Remark 2.2.7. By the subadditivity, we have

$$\frac{\mathbb{E}[\bar{K}_{0,t}(1,B)]}{t} \leq \mathfrak{c}_H \leq \frac{\mathbb{E}[\bar{K}_{0,t}(1,B)]+1}{t}.$$

In particular,

$$\left|\mathfrak{c}_H - \frac{\mathbb{E}[\bar{K}_{0,t}]}{t}\right| \leq t^{-1}. \tag{2.12}$$

2.2.2 Convergence of the variations

The aim of this section is to prove Theorem 2.1.2. The following is the first observation.

Lemma 2.2.8. *Let $\zeta \geq 1$ and $v < s < t$. We set $\varepsilon := (\frac{t-s}{\zeta})^H$. Then, if $\frac{t-s}{s-v}$ is sufficiently small, we have*

$$\|\mathbb{E}[\bar{K}_{s,t}(\varepsilon,B)|\mathcal{F}_v] - \mathbb{E}[\bar{K}_{0,\zeta}(1,B)]\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \left(\frac{t-s}{s-v}\right)^{1-H}.$$

Lemma 2.2.8 is an easy consequence of the following result.

Lemma 2.2.9 (asymptotic independence, [Pic08, Lemma A.1])**.** *Let $0 \le v < s < t$. Let F and G be respectively measurable with respect to*

$$\sigma(B_r : r < v) \quad \textit{and} \quad \sigma(B_{t'} - B_{s'} : s \le s' < t' \le t),$$

and suppose that $F, G \in L^p(\mathbb{P})$ with $p \in (1, \infty)$. If $(t-s)(s-v)^{-1}$ is sufficiently small, then we have

$$|\mathbb{E}[FG] - \mathbb{E}[F]\mathbb{E}[G]| \lesssim_p \left(\frac{t-s}{s-v}\right)^{1-H} \|F\|_{L^p(\mathbb{P})} \|G\|_{L^p(\mathbb{P})}.$$

In particular,

$$\|\mathbb{E}[F|\mathcal{F}_v] - \mathbb{E}[F]\|_{L^p(\mathbb{P})} \lesssim_p \left(\frac{t-s}{s-v}\right)^{1-H} \|F\|_{L^p(\mathbb{P})}. \tag{2.13}$$

Proof of Lemma 2.2.8. By Lemma 2.2.4, the random variable $\bar{K}_{s,t}(\varepsilon, B)$ is measurable with respect to $\sigma(B_r - B_s : s \le r \le t)$. The estimate (2.13) implies

$$\|\mathbb{E}[\bar{K}_{s,t}(\varepsilon, B)|\mathcal{F}_v] - \mathbb{E}[\bar{K}_{s,t}(\varepsilon, B)]\|_{L^p(\mathbb{P})} \lesssim \left(\frac{t-s}{s-v}\right)^{1-H} \|\bar{K}_{s,t}(\varepsilon, B)\|_{L^p(\mathbb{P})}.$$

By the stationarity and the scaling (Lemma 2.2.4),

$$\bar{K}_{s,t}(\varepsilon, B) \overset{\mathrm{d}}{=} \bar{K}_{0,\zeta}(1, B)$$

and the claim follows. □

We recall the Mandelbrot–van Ness representation (1.25). The next lemma is a consequence of Girsanov's theorem.

Lemma 2.2.10. *Let $v < s < t$, $\varepsilon \in (0,1)$, $\rho, \rho' \in [-\varepsilon/2, \varepsilon/2]$ and $y : [v,t] \to \mathbb{R}$ be a deterministic path. We set*

$$\tilde{B}^v_r := \int_v^r (r-u)^{H-1/2} \mathrm{d}W_u, \quad v \le r \le t. \tag{2.14}$$

Setting

$$a_H := \frac{1}{2}\Big(\frac{1}{\Gamma(H+1/2)\Gamma(3/2-H)}\Big)^2,$$

we have the bound

$$\begin{aligned} &|\mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho)] - \mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho')]| \\ &\quad \lesssim e^{a_H|\rho-\rho'|^2(s-v)^{-2}(t-v)^{2-2H}} \\ &\qquad \times \mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho)^2]^{\frac{1}{2}} |\rho - \rho'|(s-v)^{-1}(t-v)^{1-H}. \end{aligned}$$

Proof. The proof is inspired by [Pic08, Theorem A.1]. Let $\delta := \rho' - \rho$ and

$$h_r := \begin{cases} (s-v)^{-1}(r-v)\delta & \text{if } v \leq r \leq s, \\ \delta & \text{if } s \leq r. \end{cases}$$

Note that the functions $r \mapsto \tilde{B}^v_r + y_r + \rho'$ and $r \mapsto \tilde{B}^v_r + y_r + h_r + \rho$ are equal on the interval $[s,t]$. Thus,

$$K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho') = K_{s,t}(\varepsilon, \tilde{B}^v + y + h + \rho).$$

We claim

$$h_r = \int_v^r (r-u)^{H-1/2} \mathrm{d}g_u,$$

where for $r > v$,

$$g_r := \frac{\delta\{(r-v)^{3/2-H} - (r - r\wedge s)^{3/2-H}\}}{\Gamma(H+1/2)\Gamma(3/2-H)(3/2-H)(s-v)}.$$

Indeed,

$$\begin{aligned} \dot{g}_r &:= \frac{\mathrm{d}g_r}{\mathrm{d}r} \\ &= \frac{1}{\Gamma(H+1/2)\Gamma(3/2-H)} \frac{\delta}{s-v}\{(r-v)^{1/2-H} - (r-s)^{1/2-H}\mathbf{1}_{\{r>s\}}\} \end{aligned}$$

and

$$\begin{aligned} \int_v^r (r-u)^{H-1/2}(u-v)^{1/2-H}\mathrm{d}u &= \int_0^{r-v} (r-v-u)^{H-1/2}u^{1/2-H}\mathrm{d}u \\ &= (r-v)\int_0^1 (1-u)^{H-1/2}u^{1/2-H}\mathrm{d}u \\ &= \Gamma(H+1/2)\Gamma(3/2-H)(r-v). \end{aligned}$$

Therefore,

$$\int_v^r (r-u)^{H-1/2}\mathrm{d}g_u = \frac{\delta}{s-v}\{(r-v) - (r-s)\mathbf{1}_{\{r>s\}}\} = h_r.$$

If we set

$$F(w) := K_{s,t}\Big(\varepsilon, \int_v^{\cdot} (\cdot - u)^{H-1/2}\mathrm{d}w_u + y + \rho\Big),$$

then $K_{s,t}(1, \tilde{B}^v + y + \rho') = F(W + g)$ and by Girasnov's theorem (or the Cameron-Martin theorem)

$$\mathbb{E}[F(W+g)] = \mathbb{E}\Big[e^{\int_v^t \dot{g}_r \mathrm{d}W_r - \frac{1}{2}\int_v^t |\dot{g}_r|^2 \mathrm{d}r} F(W)\Big].$$

Thus,

$$\begin{aligned}\mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho')] - \mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho)]& \\ = \mathbb{E}\Big[\Big\{e^{\int_v^t \dot{g}_r \mathrm{d}W_r - \frac{1}{2}\int_v^t |\dot{g}_r|^2 \mathrm{d}r} - 1\Big\} K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho)\Big]&.\end{aligned}$$

By the Cauchy-Schwarz inequality, it is bounded by

$$\mathbb{E}\Big[\Big(e^{\int_v^t \dot{g}_r \mathrm{d}W_r - \frac{1}{2}\int_v^t |\dot{g}_r|^2 \mathrm{d}r} - 1\Big)^2\Big]^{1/2} \mathbb{E}[K_{s,t}(\varepsilon, \tilde{B}^v + y + \rho)^2]^{\frac{1}{2}}.$$

Since $\int_v^t \dot{g}_r \mathrm{d}W_r$ is centered Gaussian with variance $\int_v^t |\dot{g}_r|^2 \mathrm{d}r$,

$$\begin{aligned}&\mathbb{E}\Big[\Big(e^{\int_v^t \dot{g}_r \mathrm{d}W_r - \frac{1}{2}\int_v^t |\dot{g}_r|^2 \mathrm{d}r} - 1\Big)^2\Big] \\ &\qquad = e^{\int_v^t |\dot{g}_r|^2 \mathrm{d}r} - 1 \\ &\qquad \le \int_v^t |\dot{g}_r|^2 \mathrm{d}r \times e^{\int_v^t |\dot{g}_r|^2 \mathrm{d}r} \\ &\qquad \lesssim e^{2a_H |\rho - \rho'|^2 (s-v)^{-2}(t-v)^{2-2H}} |\rho - \rho'|^2 (s-v)^{-2}(t-v)^{2-2H},\end{aligned}$$

which completes the proof. □

Proof of Theorem 2.1.2. In view of the scaling, we may suppose that $T = 1$. The proof takes advantage of Theorem 1.1.1, combined with lemmas prepared above.

Step 1, lower bound. Let $\zeta \ge 1$, and let $\pi_{\varepsilon,\zeta}$ be the partition of $[0,1]$ with identical mesh size $\zeta\varepsilon^{\frac{1}{H}}$. By the superadditivity (Lemma 2.2.2),

$$\varepsilon^{\frac{1}{H}} K_{0,1}(\varepsilon, B + \rho) \ge \sum_{[s,t]\in\pi_{\varepsilon,\zeta}} \varepsilon^{\frac{1}{H}} K_{s,t}(\varepsilon, B + \rho) = \zeta^{-1} \sum_{[s,t]\in\pi_{\varepsilon,\zeta}} A_{s,t},$$

where $A^1_{s,t} := K_{s,t}((\frac{t-s}{\zeta})^H, B + \rho)(t-s)$. Furthermore, we set

$$A^2_{s,t} := \bar{K}_{s,t}\Big(\Big(\frac{t-s}{\zeta}\Big)^H, B\Big)(t-s), \quad A^3_{s,t} := \mathbb{E}[\bar{K}_{0,\zeta}(1, B)](t-s).$$

We see that $A_{s,t} := A^1_{s,t} - A^3_{s,t}$ satisfies the condition of Theorem 1.1.1. Indeed, by scaling we have

$$\|K_{s,t}(\varepsilon, B + \rho)\|_{L^p(\mathbb{P})} + \|\bar{K}_{s,t}(\varepsilon, B + \rho)\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} 1$$

and hence

$$\|A_{s,t}\|_{L^p(\mathbb{P})} \le \|A^1_{s,t}\|_{L^p(\mathbb{P})} + \|A^3_{s,t}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} (t-s).$$

Since

$$K_{s,t}(\varepsilon, B+\rho) - \bar{K}_{s,t}(\varepsilon, B) = \varepsilon^{-1} \int_{-\varepsilon/2}^{\varepsilon/2} \{K_{s,t}(\varepsilon, B+\rho) - K_{s,t}(\varepsilon, \rho')\} \mathrm{d}\rho',$$

by Lemma 2.2.10, we have

$$\|\mathbb{E}[A^1_{s,t} - A^2_{s,t} | \mathcal{F}_v]\|_p \lesssim_{p,\zeta} \left(\frac{t-s}{s-v}\right)^H (t-s).$$

By Lemma 2.2.8,

$$\|\mathbb{E}[A^2_{s,t} - A^3_{s,t} | \mathcal{F}_v]\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \left(\frac{t-s}{s-v}\right)^{1-H} (t-s).$$

Therefore,

$$\|\mathbb{E}[A_{s,t} | \mathcal{F}_v]\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \left(\frac{t-s}{s-v}\right)^{\min\{H,1-H\}} (t-s),$$

and we indeed see that $(A_{s,t})_{s<t}$ satisfies the conditions of Theorem 1.1.1.

Consequently, we obtain

$$\varepsilon^{\frac{1}{H}} K_{0,1}(\varepsilon, B) \ge \frac{\mathbb{E}[\bar{K}_{0,\zeta}(1, B)]}{\zeta} - R_{\varepsilon,\zeta},$$

where

$$\|R_{\varepsilon,\zeta}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \varepsilon^{\delta}$$

for some δ depending only on H. By the Borel–Cantelli lemma, if $\varepsilon_n = O(n^{-\eta})$ for some $\eta > 0$, then $R_{\varepsilon_n,\zeta} \to 0$ a.s. This implies

$$\liminf_{n\to\infty} \varepsilon_n^{\frac{1}{H}} K_{0,1}(\varepsilon_n, B+\rho) \ge \frac{\mathbb{E}[\bar{K}_{0,\zeta}(1, B)]}{\zeta} \quad \text{a.s.}$$

Since ζ is arbitrary, the lower bound is obtained.

Step 2, upper bound. Since $(K_{s,t}(\varepsilon, B+\rho)+1)_{s<t}$ is subadditive, we obtain

$$\varepsilon^{\frac{1}{H}} K_{0,1}(\varepsilon, B) \le \frac{\mathbb{E}[\bar{K}_{0,\zeta}(1, B)]}{\zeta} + \frac{1}{\zeta} + R_{\varepsilon,\zeta},$$

and we similarly obtain the upper bound. $\square$

2.3 Local time via Lebesgue partitions

In this section, we are interested in level crossings at a specific level. That is, we prove Theorem 2.1.4, and as an application we prove Lemieux type result Theorem 2.1.5. A more quantitative version of Theorem 2.1.4 is the following. Recall the definition of $U_{s,t}(\varepsilon, w)$ from (2.6), which counts the number of upcrossings from 0 to ε in the interval $[s,t]$.

Theorem 2.3.1 (quantitative bound on U)**.** *Let $H < \frac{1}{2}$, $a \in \mathbb{R}$ and $(\varepsilon_n)_{n=1}^{\infty}$ be a sequence of positive numbers tending to 0. Then, if $\varepsilon_n = O(n^{-\eta})$ for some $\eta > 0$, we have*

$$\lim_{n\to\infty} \varepsilon_n^{\frac{1}{H}-1} U_{0,1}(\varepsilon_n, B^H - a) = \frac{\mathfrak{c}_H}{2} L_1^H(a) \quad \textit{a.s.}$$

In fact, we have a quantitative bound

$$\left| \varepsilon^{\frac{1}{H}-1} U_{0,1}(\varepsilon, B^H - a) - \frac{\mathfrak{c}_H}{2} L_1^H(a) \right| \leq \frac{1}{\zeta} L_1^H(a) + \mathcal{R}_{\varepsilon,\zeta,a},$$

for all $\varepsilon \in (0,1)$ and $\zeta \in (1,\infty)$, where there exists a $\kappa > 0$ such that for every $p \in (0,\infty)$ we have

$$\|\mathcal{R}_{\varepsilon,\zeta,a}\|_{L^p(\mathbb{P})} \leq C_{p,\zeta} \varepsilon^{\kappa}$$

with $C_{p,\zeta}$ independent of a.

The proof of Theorem 2.3.1 is somewhat similar to that of Theorem 2.1.2. Indeed, it is based on the super(sub)-additivity, Girsanov's theorem and the shifted stochastic sewing. However, a major difficulty here is that we cannot use any ergodic theorem. This leads to more involved technical arguments. Instead of directly going to the proof, in the next section we heuristically explain our strategy.

2.3.1 Heuristics

Herein we explain our heuristic strategy to prove Theorem 2.3.1. We set

$$A_{s,t} := U_{s,t}((t-s)^H, B - a)(t-s)^{1-H}.$$

In view of Theorem 1.1.1, our goal is to show

$$\mathbb{E}[A_{s,t} | \mathcal{F}_v] \approx \frac{\mathfrak{c}_H}{2} \mathbb{E}[L_{s,t}(a) | \mathcal{F}_v]. \tag{2.15}$$

Indeed, once the estimate (2.15) is proven, the rest of the argument is similar to the proof of Theorem 2.1.2.

We thus explain heuristically how to prove (2.15). For simplicity, we set $a = 0$, and we write $\varepsilon := (t-s)^H$. (Strictly speaking, we actually introduce another parameter ζ going to infinity and set $\varepsilon := (\frac{t-s}{\zeta})^H$, but for simplicity here we set $\zeta = 1$.) Let us introduce another parameter $u \in (v, s)$ (in mind $t - s \ll s - u \ll u - v$), and, recalling the Mandelbrot–van Ness representation, for $r \in [s, t]$ we decompose

$$\begin{aligned} B_r &= \int_{-\infty}^{v} \mathcal{K}(r,\theta)\mathrm{d}W_\theta + \int_{v}^{u} \mathcal{K}(r,\theta)\mathrm{d}W_\theta + \int_{u}^{r} \mathcal{K}(r,\theta)\mathrm{d}W_\theta \\ &=: X_r + Y_r + Z_r. \end{aligned}$$

In the interval $[s, t]$ the smooth processes X and Y do not change much compared to Z. Therefore, we can freeze time of X and Y (Lemma 2.3.8):

$$\mathbb{E}[U_{s,t}(\varepsilon, B)|\mathcal{F}_v] \approx \mathbb{E}[U_{s,t}(\varepsilon, X_s + Y_s + Z)|\mathcal{F}_v].$$

But we see

$$\mathbb{E}[U_{s,t}(\varepsilon, X_s + Y_s + Z)|\mathcal{F}_v] = \mathbb{E}[U_{s,t}(\varepsilon, x + Y_s + Z)]|_{x=X_s},$$

and the Gaussian change of variable to Y yields

$$\mathbb{E}[U_{s,t}(\varepsilon, x + Y_s + Z)] = e^{-\frac{1}{2}(\frac{x}{\sigma_Y})^2}\mathbb{E}\Big[e^{\frac{xY_s}{\sigma_Y^2}} U_{s,t}(\varepsilon, Y_s + Z)\Big],$$

where σ_Y is the variance of Y_s (Lemma 2.3.10).

For $U_{s,t}(\varepsilon, Y_s + Z)$ to be positive, Y_s must be around 0. (In other words, if Y_s is far away from 0, the process Z must move quite a lot, which is costly.) Therefore (Lemma 2.3.11),

$$\begin{aligned} \mathbb{E}\Big[e^{\frac{xY_s}{\sigma_Y^2}} U_{s,t}(\varepsilon, Y_s + Z)\Big] &\approx \mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)] \\ &\approx \mathbb{E}\Big[U_{s,t}(\varepsilon, Y + Z)\Big]. \end{aligned}$$

As $v \ll u \ll s$, we have $\sigma_Y \approx \sigma_{Y+Z}$, with σ_{Y+Z} the variance of $Y + Z$ (Lemma 2.3.14). In the end, we have (Lemma 2.3.7)

$$\mathbb{E}[U_{s,t}(\varepsilon, B)|\mathcal{F}_v] \approx \mathbb{E}\Big[U_{s,t}(1, Y + Z)\Big]e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}.$$

We now observe (Lemma 2.3.17)

$$\begin{aligned} \mathbb{E}\Big[\int_s^t \delta_0(B_r)\mathrm{d}r\Big|\mathcal{F}_v\Big] &\approx \int_s^t \mathbb{E}[\delta_0(B_s)|\mathcal{F}_v]\mathrm{d}r \\ &= \frac{1}{\sqrt{2\pi}\sigma_{Y+Z}} e^{-\frac{X_s^2}{2\sigma_{Y+Z}^2}}(t-s). \end{aligned}$$

It is not obvious, but in Lemma 2.3.16 we prove

$$\sqrt{2\pi}\sigma_{Y+Z}(t-s)^{-H}\mathbb{E}[U_{s,t}(\varepsilon, Y+Z)] \approx \frac{\mathfrak{c}_H}{2}.$$

Now we see (2.15). With this heuristic argument in mind, we move to a rigorous proof in the next section.

2.3.2 Convergence to the local time

Estimates on level crossings

The following process will appear in our argument.

Definition 2.3.2. We denote by $\mathcal{B} = \mathcal{B}^H$ the *Riemann-Liouville process*

$$\mathcal{B}_t := \int_0^t \mathcal{K}(t,r)\mathrm{d}W_r.$$

In view of the Mandelbrot–van Ness representation (1.25), we have

$$B_t = \int_{-\infty}^0 \mathcal{K}(t,r)\mathrm{d}W_r + \mathcal{B}_t.$$

We begin with three elementary lemmas.

Lemma 2.3.3 (scaling of U)**.** *We have the following scaling property: for $\lambda > 0$,*

$$(U_{s,t}(\varepsilon, \mathcal{B}+\rho))_{s<t,\varepsilon>0,\rho\in\mathbb{R}} \stackrel{\mathrm{d}}{=} (U_{\lambda^{1/H}s,\lambda^{1/H}t}(\lambda\varepsilon, \mathcal{B}+\lambda\rho))_{s<t,\varepsilon>0,\rho\in\mathbb{R}}.$$

A similar result holds with $\mathcal{B}$ replaced by B.

Proof. As in Lemma 2.2.1, it follows from the scaling property of $\mathcal{B}$. □

Definition 2.3.4. We set

$$\bar{U}_{s,t}(\varepsilon, w) := U_{s,t}(\varepsilon, w) + \mathbf{1}_{\{w_s\in(0,\varepsilon)\}}.$$

Lemma 2.3.5 (sub/super-additivity of U)**.** *For $s < u < t$ we have*

$$U_{s,t}(\varepsilon, w) \geq U_{s,u}(\varepsilon, w) + U_{u,t}(\varepsilon, w), \quad \bar{U}_{s,t}(\varepsilon, w) \leq \bar{U}_{s,u}(\varepsilon, w) + \bar{U}_{u,t}(\varepsilon, w).$$

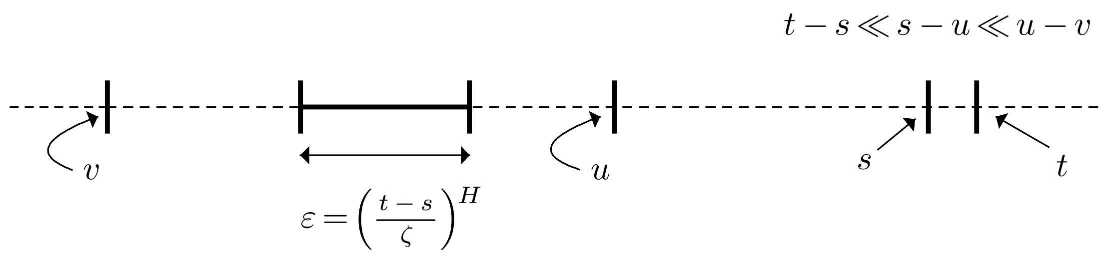


Figure 2.5: Parameters for Lemma 2.3.7

Proof. We have

$$U_{s,t}(\varepsilon, w) = U_{s,u}(\varepsilon, w) + U_{u,t}(\varepsilon, w) + 1$$

if there exist $a < u < b$ such that $w_a = 0, w_b = \varepsilon$ and $w_r \in (0, \varepsilon)$ for all $r \in (a, b)$, and otherwise

$$U_{s,t}(\varepsilon, w) = U_{s,u}(\varepsilon, w) + U_{u,t}(\varepsilon, w).$$ □

Lemma 2.3.6. *There exists a positive constant $\kappa = \kappa(H)$ such that for $a \in \mathbb{R}$, $\rho \in (0, \infty)$, $s < t$ and $p \in (0, \infty)$ we have*

$$\|U_{s,t}(\rho, B - a)\|_{L^p(\mathbb{P})} \lesssim_{H,p,\rho} (t - s)^{\kappa}.$$

A similar estimate holds with B replaced by $\mathcal{B}$.

Proof. This follows from the obvious inequality $U_{s,t}(\rho, B-a) \leq J_{s,t}(\rho, B)$ and the estimate (2.11). □

We introduce some notation which will be used throughout Section 2.3.2. Fix $\zeta \geq 1$. At the very end, we let $\zeta \to \infty$. (The parameter ζ corresponds to the parameter m in [Dur19, Theorem 6.4.1].) We fix $v < u < s < t$ with $t - s \ll s - u \ll u - v$ and set $\varepsilon := (\frac{t-s}{\zeta})^H$, as shown in Figure 2.5. We set

$$X_r := \int_{-\infty}^{v} \mathcal{K}(r, \theta) \mathrm{d}W_\theta - a, \quad Y_r := \int_{v}^{u} \mathcal{K}(r, \theta) \mathrm{d}W_\theta, \quad Z_r := \int_{u}^{r} \mathcal{K}(r, \theta) \mathrm{d}W_\theta$$

for $r \in [s, t]$. Then

$$\mathbb{E}[U_{s,t}(\varepsilon, B - a) | \mathcal{F}_v] = \mathbb{E}[U_{s,t}(\varepsilon, x + Y + Z)]|_{x=X}.$$

Finally, we write

$$\sigma_Y^2 := \mathbb{E}[Y_s^2] = \frac{1}{2H}\{(s-v)^{2H} - (s-u)^{2H}\}, \tag{2.16}$$

$$\sigma_{Y+Z}^2 := \mathbb{E}[(Y_s+Z_s)^2] = \frac{1}{2H}(s-v)^{2H}. \tag{2.17}$$

In the spirit of the shifted stochastic sewing (Theorem 1.1.1), we will estimate

$$\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_v].$$

The goal of this section is to prove the following estimate.

Lemma 2.3.7. *Let $H < \frac{1}{2}$ and $p \in (1,\infty)$. We further let $v < s < t$ and $\zeta \in [1,\infty)$, and set $\varepsilon := (\frac{t-s}{\zeta})^H$. For $\kappa \in (0,1)$, if $\frac{t-s}{s-v}$ is sufficiently small, we have*

$$\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_v] = \frac{\mathbb{E}[\bar{K}_{0,\zeta}(1,B)]}{2\sqrt{2\pi}\sigma_{Y+Z}} e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}\Big(\frac{t-s}{\zeta}\Big)^H + R_{v,s,t}, \tag{2.18}$$

where

$$\|R_{v,s,t}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta,\kappa} \Big(\frac{t-s}{s-v}\Big)^{(2-\kappa)H(1-H)}(t-s)^{-\kappa H}.$$

The proof of Lemma 2.3.7 will be built on several technical lemmas. For the sake of the next lemma, we recall the Riemann-Liouville operator (e.g. [Pic11])

$$I_\alpha f(r) := \frac{1}{\Gamma(\alpha)}\int_s^r (r-\theta)^{\alpha-1} f(\theta)\mathrm{d}\theta, \quad \text{for } \alpha > 0.$$

If f is Lipschitz with $f_s = 0$ and $\alpha \in (-1, 0]$, we set

$$I_\alpha f(r) := \frac{1}{\Gamma(1+\alpha)}\int_s^r (r-\theta)^{\alpha}\dot{f}(\theta)\mathrm{d}\theta.$$

The family $(I_\alpha)_{\alpha>-1}$ has the semigroup property $I_\alpha I_\beta = I_{\alpha+\beta}$.

Lemma 2.3.8. *For $p \in (1,\infty)$ and $\kappa \in (1-p^{-1}, 1)$, there exists a positive constant c such that if $(t-s)(s-u)^{-1}$ is sufficiently small, then*

$$\|\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_u] - \mathbb{E}[U_{s,t}(\varepsilon, X_s+Y_s+Z)|\mathcal{F}_u]\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \|U_{s,t}(\varepsilon, B-a)\|_{L^1(\mathbb{P})}^{1-\kappa} e^{-ca^2}\Big(\frac{t-s}{s-u}\Big)^{(1-H)}.$$

Proof. The proof is similar to [Pic08, Lemma A.1]. We have

$$\mathbb{E}[U_{s,t}(\varepsilon, X_s + Y_s + Z)|\mathcal{F}_u] = \mathbb{E}[U_{s,t}(\varepsilon, x_s + y_s + Z)]|_{x,=X,y=Y}$$

and

$$\mathbb{E}[U_{s,t}(\varepsilon, x_s + y_s + Z)] = \mathbb{E}[U_{s,t}(\varepsilon, x + y + w + Z)]$$

where

$$w_r := -(x_r + y_r - x_s - y_s).$$

Therefore, since $w_s = 0$,

$$w = I_1 \dot{w} = I_{H-\frac{1}{2}} I_{\frac{3}{2}-H} \dot{w}$$

and

$$w_r + \int_s^r \mathcal{K}(r,\theta) \mathrm{d}W_\theta = \int_s^r \mathcal{K}(r,\theta) \mathrm{d}\Big(W_\theta + c_1 \big(I_{\frac{3}{2}-H}\dot{w}\big)_\theta\Big)$$

for some constant c_1 depending only on H.

By Girsanov's theorem,

$$\mathbb{E}[U_{s,t}(\varepsilon, x + y + w + Z)] = \mathbb{E}\Big[U_{s,t}(\varepsilon, x + y + Z) \\ \times \exp\Big(c_1 \int_s^t \frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w} \mathrm{d}W_\theta - \frac{c_1^2}{2} \int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}\Big|^2 \mathrm{d}\theta\Big)\Big].$$

Therefore, if $p^{-1} + q^{-1} = 1$, by Hölder's inequality,

$$|\mathbb{E}[U_{s,t}(\varepsilon, x + y + Z)] - \mathbb{E}[U_{s,t}(\varepsilon, x_s + y_s + Z)]| \\ \lesssim \mathbb{E}\Big[\Big(\exp\Big(c_1 \int_s^t \frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w} \mathrm{d}W_\theta - \frac{c_1^2}{2} \int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}\Big|^2 \mathrm{d}\theta\Big) - 1\Big)^q\Big]^{\frac{1}{q}} \\ \times \mathbb{E}[U_{s,t}(\varepsilon, x + y + Z)^p]^{\frac{1}{p}}.$$

Since the random variable

$$\int_s^t \frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}(\theta) \mathrm{d}W_\theta$$

is Gaussian, we have

$$\mathbb{E}\Big[\Big(\exp\Big(c_1 \int_s^t \frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w} \mathrm{d}W_\theta - \frac{c_1^2}{2} \int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}\Big|^2 \mathrm{d}\theta\Big) - 1\Big)^q\Big]^{\frac{1}{q}} \\ \lesssim_q \Big(\int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}\Big|^2 \mathrm{d}\theta\Big)^{\frac{1}{2}} \exp\Big(C_q \int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}\dot{w}\Big|^2 \mathrm{d}\theta\Big).$$

Hence, by setting

$$S_{s,t} := \int_s^t \Big|\frac{\mathrm{d}}{\mathrm{d}\theta} I_{\frac{3}{2}-H}(\dot{X}+\dot{Y})\Big|^2 \mathrm{d}\theta,$$

we have

$$\begin{aligned}\|\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_u] - \mathbb{E}[U_{s,t}(\varepsilon, X_s+Y_s+Z)|\mathcal{F}_u]\|_{L^p(\mathbb{P})} \\ &\lesssim \mathbb{E}\Big[\mathbb{E}_u[U_{s,t}(\varepsilon, B-a)^p] S_{s,t}^{\frac{p}{2}} e^{pC_q S_{s,t}}\Big]^{\frac{1}{p}} \\ &\leq \|U_{s,t}(\varepsilon, B-a)\|_{L^{p_1}(\mathbb{P})} \Big\| S_{s,t}^{\frac{1}{2}} e^{C_q S_{s,t}} \Big\|_{L^{q_1}(\mathbb{P})},\end{aligned}$$

where $p_1^{-1} + q_1^{-1} = p^{-1}$. Choose p_2 so that $p_1^{-1} = (1-\kappa) + \kappa p_2^{-1}$ (since $\kappa > 1 - p^{-1}$, this is possible by choosing p_1 close to p). By interpolation,

$$\|U_{s,t}(\varepsilon, B-a)\|_{L^{p_1}(\mathbb{P})} \leq \|U_{s,t}(\varepsilon, B-a)\|_{L^1(\mathbb{P})}^{1-\kappa} \|U_{s,t}(\varepsilon, B-a)\|_{L^{p_2}(\mathbb{P})}^{\kappa}.$$

We have

$$\begin{aligned}\|U_{s,t}(\varepsilon, B-a)\|_{L^{p_2}(\mathbb{P})} &\leq \mathbb{P}(\|B\|_{L^\infty([0,1])} \geq a)^{\frac{1}{2}} \|U_{s,t}(\varepsilon, B-a)\|_{L^{2p_2}(\mathbb{P})} \\ &\lesssim e^{-c_2 a^2} \|U_{s,t}(\varepsilon, B-a)\|_{L^{2p_2}(\mathbb{P})}.\end{aligned}$$

The scaling property (Lemma 2.3.3) gives

$$\|U_{s,t}(\varepsilon, B-a)\|_{L^{2p_2}(\mathbb{P})} = \|U_{s/(t-s),t/(t-s)}(\zeta^{-H}, B-(t-s)^{-H}a)\|_{L^{2p_2}(\mathbb{P})}.$$

By Lemma 2.3.6,

$$\|U_{s/(t-s),t/(t-s)}(\zeta^{-H}, B-(t-s)^{-H}a)\|_{L^{2p_2}(\mathbb{P})} \lesssim_{p_2,\zeta} 1.$$

It remains to see

$$\Big\| S_{s,t}^{\frac{1}{2}} e^{C_q S_{s,t}} \Big\|_{L^{q_1}(\mathbb{P})} \lesssim \Big(\frac{t-s}{s-u}\Big)^{1-H}, \quad \text{if } \frac{t-s}{s-u} \text{ is sufficiently small.}$$

This was essentially proven in [Pic08, Lemma A.1] (our $S_{s,t}$ corresponds to L therein). □

Remark 2.3.9. We note that a similar reasoning shows that for $p < p_1 < \infty$ if $\frac{t-s}{s-u}$ is sufficiently small, we have

$$\|U_{s,t}(\varepsilon, Y_s+Z)\|_{L^p(\mathbb{P})} \lesssim_{\zeta,p,p_1} \|U_{s,t}(\varepsilon, Y+Z)\|_{L^{p_1}(\mathbb{P})}.$$

Lemma 2.3.10. *Recall σ_Y from (2.16). We have the estimate*

$$\mathbb{E}[U_{s,t}(\varepsilon, X_s + Y_s + Z)|\mathcal{F}_v] = e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}\Big[e^{\frac{X_s Y_s}{\sigma_Y^2}} U_{s,t}(\varepsilon, Y_s + Z)\Big|\mathcal{F}_v\Big].$$

Proof. Since X, Y and Z are independent,

$$\mathbb{E}[U_{s,t}(\varepsilon, X_s + Y_s + Z)|\mathcal{F}_v] = \mathbb{E}[\mathbb{E}[U_{s,t}(\varepsilon, x_s + y_s + Z)]|_{(x,y)=(X,Y)}|X].$$

If we set $F(\eta) := \mathbb{E}[U_{s,t}(\varepsilon, \eta + Z)]$ ($\eta \in \mathbb{R}$), then

$$\begin{aligned}
\mathbb{E}[F(x_s + Y_s)] &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} F(x_s + \sigma_Y\eta)e^{-\frac{\eta^2}{2}}\mathrm{d}\eta \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} F(\sigma_Y\eta)e^{-\frac{1}{2}(\eta-\sigma_Y^{-1}x_s)^2}\mathrm{d}\eta \\
&= e^{-\frac{1}{2}(\frac{x_s}{\sigma_Y})^2}\mathbb{E}\Big[e^{\frac{Y_s x_s}{\sigma_Y^2}} F(Y_s)\Big].
\end{aligned}$$

The claim thus follows. □

Lemma 2.3.11. *For every $p_1 \in (1, \infty)$, if $\frac{t-s}{s-u}$ is sufficiently small, then*

$$\begin{aligned}
&\Big|\mathbb{E}\Big[e^{\frac{x_s Y_s}{\sigma_Y^2}} U_{s,t}(\varepsilon, Y_s + Z)\Big] - \mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)]\Big| \\
&\qquad\qquad \lesssim_{H,p_1} \frac{|x_s|(t-u)^H}{\sigma_Y^2}e^{c(\frac{|x_s|(t-u)^H}{\sigma_Y^2})^2}\|U_{s,t}(\varepsilon, Y + Z)\|_{L^{p_1}(\mathbb{P})}
\end{aligned}$$

with c depending only on H and p_1.

Proof. For $U_{s,t}(\varepsilon, Y_s + Z)$ to be non-zero, we must have $\inf_{r\in[s,t]}|Y_s + Z_r| = 0$. Therefore,

$$\begin{aligned}
&\mathbb{E}\Big[e^{\frac{x_s Y_s}{\sigma_Y^2}} U_{s,t}(\varepsilon, Y_s + Z)\Big] - \mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)] \\
&\qquad\qquad = \mathbb{E}\Big[\Big(e^{\frac{x_s Y_s}{\sigma_Y^2}} - 1\Big)U_{s,t}(\varepsilon, Y_s + Z)\mathbf{1}_{\{\|Z\|_{L^\infty([s,t])}\geq|Y_s|\}}\Big].
\end{aligned}$$

Using the inequality

$$|e^\lambda - 1| \leq e^{|\lambda|}|\lambda|, \quad \lambda \in \mathbb{R},$$

we estimate

$$\begin{aligned}
&\Big|\mathbb{E}\Big[\Big(e^{\frac{x_s Y_s}{\sigma_Y^2}} - 1\Big)U_{s,t}(\varepsilon, Y_s + Z)\mathbf{1}_{\{\|Z\|_{L^\infty([s,t])}\geq|Y_s|\}}\Big]\Big| \\
&\qquad\qquad \leq \frac{|x_s|}{\sigma_Y^2}\mathbb{E}\Big[e^{\frac{|x_s|\|Z\|_{L^\infty([s,t])}}{\sigma_Y^2}}\|Z\|_{L^\infty([s,t])}U_{s,t}(\varepsilon, Y_s + Z)\Big],
\end{aligned}$$

which, by Hölder's inequality, is bounded by

$$\mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)^{p_1}]^{\frac{1}{p_1}} \mathbb{E}[\|Z\|^{p_2}_{L^\infty([s,t])}]^{\frac{1}{p_2}} \mathbb{E}\Big[e^{\frac{p_3|x_s|\|Z\|_{L^\infty([s,t])}}{\sigma_Y^2}}\Big]^{\frac{1}{p_3}},$$

where $p_1, p_2, p_3 \in (1, \infty)$ satisfy

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1.$$

By Remark 2.3.9, if $\frac{t-s}{s-u}$ is sufficiently small, we have

$$\mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)^{p_1}]^{\frac{1}{p_1}} \lesssim_{p_1} \mathbb{E}[U_{s,t}(\varepsilon, Y + Z)^{p_4}]^{\frac{1}{p_4}}, \quad p_4 := p_1^2.$$

Recalling $\mathcal{B}$ from Definition 2.3.2, the scaling property yields

$$\mathbb{E}[\|Z\|^{p_2}_{L^\infty([s,t])}]^{\frac{1}{p_2}} \leq \mathbb{E}[\|Z\|^{p_2}_{L^\infty([u,t])}]^{\frac{1}{p_2}} = (t-u)^H \mathbb{E}[\|\mathcal{B}\|^{p_2}_{L^\infty([0,1])}]^{\frac{1}{p_2}}$$

and similarly

$$\mathbb{E}\Big[e^{\frac{p_3|x_s|\|Z\|_{L^\infty([s,t])}}{\sigma_Y^2}}\Big] \leq \mathbb{E}\Big[e^{\frac{p_3|x_s|(t-u)^H}{\sigma_Y^2}\|\mathcal{B}\|_{L^\infty([0,1])}}\Big].$$

Since $\|\mathcal{B}\|_{L^\infty([0,1])}$ has a Gaussian tail by Fernique's theorem, there exists a constant c depending only on H such that

$$\mathbb{E}\Big[e^{\frac{p_3|x_s|(t-u)^H}{\sigma_Y^2}\|\mathcal{B}\|_{L^\infty([0,1])}}\Big] \lesssim e^{c(\frac{p_3|x_s|(t-u)^H}{\sigma_Y^2})^2}.$$

Now the claim is proved. $\square$

Lemma 2.3.12. *For every $p_1, p_2 \in (1, \infty)$ we have*

$$\|U_{s,t}(\varepsilon, Y + Z)\|_{L^{p_1}(\mathbb{P})} \lesssim_{\zeta, p_1, p_2} \left(\frac{t-s}{s-v}\right)^{\frac{H}{p_1 p_2}}.$$

Proof. By the scaling,

$$\|U_{s,t}(\varepsilon, Y + Z)\|_{L^{p_1}(\mathbb{P})} = \|U_{s-v,t-v}(\varepsilon, \mathcal{B})\|_{L^{p_1}(\mathbb{P})} = \|U_{\frac{s-v}{t-s}, \frac{t-v}{t-s}}(\zeta^{-H}, \mathcal{B})\|_{L^{p_1}(\mathbb{P})}.$$

We set $k_1 := \frac{s-v}{t-s}$ and $k_2 := \frac{t-v}{t-s}$, and note that $k_2 - k_1 = 1$. We observe

$$\begin{aligned}
&\|U_{k_1,k_2}(\zeta^{-H}, \mathcal{B})\|_{L^{p_1}(\mathbb{P})}^{p_1} \\
&\quad = \sum_{a\in\mathbb{Z}} \mathbb{E}\Big[U_{k_1,k_2}(\zeta^{-H}, \mathcal{B})^{p_1} \mathbf{1}_{\{\mathcal{B}_{k_1}\in(a-1,a]\}}\Big] \\
&\quad = \sum_{a\in\mathbb{Z}} \mathbb{E}\Big[U_{k_1,k_2}(\zeta^{-H}, \mathcal{B})^{p_1} \mathbf{1}_{\{\mathcal{B}_{k_1}\in(a-1,a]\}} \mathbf{1}_{\{\max_{r\in[k_1,k_2]}|\mathcal{B}_r - \mathcal{B}_{k_1}|\geq|a|-1\}}\Big] \\
&\quad \leq \sum_{a\in\mathbb{Z}} \mathbb{E}[U_{k_1,k_2}(\zeta^{-H}, \mathcal{B})^{p_1 q_2}]^{\frac{1}{q_2}} \mathbb{P}(\max_{r\in[k_1,k_2]} |\mathcal{B}_r - \mathcal{B}_{k_1}| \geq |a| - 1)^{\frac{1}{q_2}} \\
&\qquad\qquad \times \mathbb{P}(\mathcal{B}_{k_1} \in (a-1, a])^{\frac{1}{p_2}},
\end{aligned}$$

where

$$\frac{1}{p_2} + \frac{1}{2q_2} = 1.$$

By Lemma 2.3.6,

$$\mathbb{E}[U_{k_1,k_2}(\zeta^{-H}, \mathcal{B})^{p_1 q_2}]^{\frac{1}{q_2}} \lesssim_{p_1,p_2,\zeta} 1.$$

Since

$$\|\mathcal{B}_{r_1} - \mathcal{B}_{r_2}\|_{L^2(\mathbb{P})}^2 \leq \|B_{r_1} - B_{r_2}\|_{L^2(\mathbb{P})}^2 \lesssim |r_1 - r_2|^{2H},$$

by Kolmogorov's continuity theorem and Fernique's theorem we obtain

$$\mathbb{P}(\max_{r\in[k_1,k_2]} |\mathcal{B}_r - \mathcal{B}_{k_1}| \geq |a| - 1) \lesssim e^{-ca^2}.$$

Finally, we see $\mathbb{P}(\mathcal{B}_{k_1} \in (a-1, a]) \lesssim k_1^{-H}$, and the claim follows. □

Lemma 2.3.13. *For every $p_1 \in (1, \infty)$, if $\frac{t-s}{s-u}$ is sufficiently small, we have*

$$|\mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)] - \mathbb{E}[U_{s,t}(\varepsilon, Y + Z)]| \lesssim_{\zeta,p_1} \left(\frac{t-s}{s-v}\right)^{H/p_1} \left(\frac{t-s}{s-u}\right)^{1-H}.$$

Proof. As in Lemma 2.3.8, we get

$$|\mathbb{E}[U_{s,t}(\varepsilon, Y_s + Z)] - \mathbb{E}[U_{s,t}(\varepsilon, Y + Z)]| \lesssim_{\zeta,p_1} \|U_{s,t}(\varepsilon, Y + Z)\|_{L^{p_1}} \left(\frac{t-s}{s-u}\right)^{1-H}.$$

We then apply Lemma 2.3.12. □

Lemma 2.3.14. *If $\frac{s-u}{s-v} \leq \frac{1}{2}$, for every $p_1 \in (1,\infty)$, we have*

$$\left| e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2} - e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2} \right| \lesssim_{p_1} e^{-\frac{1}{2p_1}(\frac{X_s}{\sigma_{Y+Z}})^2} \Big(\frac{s-u}{s-v}\Big)^{2H}.$$

Proof. Recalling (2.16), we have

$$|\sigma_{Y+Z}^{-2} - \sigma_Y^{-2}| = \sigma_{Y+Z}^{-2}\sigma_Y^{-2}\frac{(s-u)^{2H}}{2H} \lesssim \sigma_{Y+Z}^{-2}\Big(\frac{s-u}{s-v}\Big)^{2H}.$$

Using the inequality $1-e^{-\lambda} \leq \lambda$ for $\lambda \geq 0$, we observe

$$\begin{aligned} e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2} - e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2} &= e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}\Big\{1 - e^{-\frac{X_s^2}{2}(\sigma_Y^{-2} - \sigma_{Y+Z}^{-2})}\Big\} \\ &\lesssim e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2} X_s^2(\sigma_Y^{-2} - \sigma_{Y+Z}^{-2}) \\ &\lesssim e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}\sigma_{Y+Z}^{-2}X_s^2\Big(\frac{s-u}{s-v}\Big)^{2H}. \end{aligned}$$

Since $\sup_{\lambda\geq 0}\lambda e^{-\varepsilon\lambda} < \infty$ for every $\varepsilon > 0$, we obtain the claimed estimate. $\square$

Lemma 2.3.15. *Let $p \in (1,\infty)$. For $\kappa \in (1-p^{-1}, 1)$ and $p_1 \in (1,2)$, if $\frac{t-s}{s-u}$ and $\frac{t-u}{u-v}$ are sufficiently small, we have*

$$\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_v] = \mathbb{E}[U_{s-v,t-v}(\varepsilon,\mathcal{B})]e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2} + R^1_{v,u,s,t} + R^2_{v,u,s,t},$$

where

$$\|R^1_{v,s,t}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta,\kappa} \mathbb{E}[U_{s,t}(\varepsilon, B-a)]^{1-\kappa}e^{-ca^2}\Big(\frac{t-s}{s-u}\Big)^{1-H}$$

with c being a constant depending only on H,κ,p, and almost surely

$$\begin{aligned} |R^2_{v,u,s,t}| \lesssim_{p_1,\zeta} e^{-\frac{1}{2p_1}(\frac{X_s}{\sigma_{Y+Z}})^2}\frac{(t-u)^H}{\sigma_{Y+Z}}\Big(\frac{t-s}{s-v}\Big)^{H/p_1} \\ + e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\Big(\frac{t-s}{s-v}\Big)^{H/p_1}\Big(\frac{t-s}{s-u}\Big)^{1-H}. \end{aligned}$$

Proof. In view of Lemma 2.3.10, we decompose

$$\mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_v] = R_1 + R_2 + R_3 + R_4 + R_5,$$

where

$$\begin{aligned}
R_1 &:= \mathbb{E}[U_{s,t}(\varepsilon, B-a)|\mathcal{F}_v] - \mathbb{E}[U_{s,t}(\varepsilon, X_s+Y_s+Z)|\mathcal{F}_v],\\
R_2 &:= e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}\Big[e^{\frac{X_sY_s}{\sigma_Y^2}}U_{s,t}(\varepsilon, Y_s+Z)\Big|\mathcal{F}_v\Big] - e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y_s+Z)],\\
R_3 &:= e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y_s+Z)] - e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y+Z)],\\
R_4 &:= e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y+Z)] - e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y+Z)],\\
R_5 &:= e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}\mathbb{E}[U_{s,t}(\varepsilon, Y+Z)] = \mathbb{E}[U_{s-v,t-v}(\varepsilon, \mathcal{B})]e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2}.
\end{aligned}$$

By Lemma 2.3.8 ,

$$\|R_1\|_{L^p(\mathbb{P})} \lesssim_{H,p,\zeta} \mathbb{E}[U_{s,t}(\varepsilon, B-a)]^{1-\kappa}e^{-ca^2}\Big(\frac{t-s}{s-u}\Big)^{1-H}.$$

By Lemma 2.3.13,

$$|R_3| \lesssim_{\zeta,p_1} e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\Big(\frac{t-s}{s-v}\Big)^{H/p_1}\Big(\frac{t-s}{s-u}\Big)^{1-H}.$$

To estimate R_2, by Lemma 2.3.11,

$$|R_2| \lesssim e^{-\frac{1}{2}(\frac{X_s}{\sigma_Y})^2}\frac{|X_s|(t-u)^H}{\sigma_Y^2}e^{c(\frac{|X_s|(t-u)^H}{\sigma_Y^2})^2}\|U_{s,t}(\varepsilon, Y+Z)\|_{L^{p_1}(\mathbb{P})}.$$

If $\frac{t-u}{u-v}$ is sufficiently small, we have

$$c\frac{(t-u)^H}{\sigma_Y} \le \frac{1}{2} - \frac{1}{2p_1},$$

hence

$$|R_2| \lesssim e^{-\frac{1}{2p_1}(\frac{X_s}{\sigma_Y})^2}\frac{|X_s|(t-u)^H}{\sigma_Y^2}\|U_{s,t}(\varepsilon, Y+Z)\|_{L^{p_1}(\mathbb{P})}.$$

Using the estimate $\sup_{\lambda\ge 0}\lambda e^{-\varepsilon\lambda^2} < \infty$ and Lemma 2.3.12, we get

$$\begin{aligned}
|R_2| &\lesssim e^{-\frac{1}{2p_1^2}(\frac{X_s}{\sigma_Y})^2}\frac{(t-u)^H}{\sigma_Y}\Big(\frac{t-s}{s-v}\Big)^{H/p_1^2}\\
&\lesssim e^{-\frac{1}{2p_1^2}(\frac{X_s}{\sigma_{Y+Z}})^2}\frac{(t-u)^H}{\sigma_{Y+Z}}\Big(\frac{t-s}{s-v}\Big)^{H/p_1^2}.
\end{aligned}$$

Finally, we estimate R_4. By Lemma 2.3.14, we get

$$|R_4| \lesssim_{p_1} \|U_{s,t}(\varepsilon, Y+Z)\|_{L^1(\mathbb{P})} e^{-\frac{1}{2p_1}(\frac{X_s}{\sigma_{Y+Z}})^2} \Big(\frac{s-u}{s-v}\Big)^{2H}$$

By Lemma 2.3.12, we obtain

$$\begin{aligned} |R_4| &\lesssim e^{-\frac{1}{2p_1}(\frac{X_s}{\sigma_{Y+Z}})^2} \Big(\frac{s-u}{s-v}\Big)^{2H} \Big(\frac{t-s}{s-v}\Big)^{H/p_1^2} \\ &\lesssim e^{-\frac{1}{2p_1^2}(\frac{X_s}{\sigma_{Y+Z}})^2} \frac{(t-u)^H}{\sigma_{Y+Z}} \Big(\frac{t-s}{s-v}\Big)^{H/p_1^2}. \end{aligned}$$

Therefore, we set $R^1_{v,u,s,t} := R_1$ and $R^2_{v,u,s,t} := R_2 + R_3 + R_4$. □

As a final ingredient in the proof of Lemma 2.3.7, we estimate $\mathbb{E}[U_{s-v,t-v}(\varepsilon, \mathcal{B})]$.

Lemma 2.3.16. *For every $\kappa \in (0,1)$ we have the estimate*

$$\begin{aligned} &\Big|\sqrt{2\pi}\zeta^H (t-s)^{-H} \sigma_{Y+Z} \mathbb{E}[U_{s-v,t-v}(\varepsilon, \mathcal{B})] - \frac{1}{2}\mathbb{E}[\bar{K}_0^\zeta(1,B)]\Big| \\ &\qquad \lesssim_{\kappa,\zeta} (t-s)^{-\kappa H} \Big(\frac{t-s}{s-u}\Big)^{1-H} + \Big(\frac{t-s}{s-u}\Big)^{-H} \Big(\frac{t-s}{s-v}\Big)^{(1-\kappa)H}. \end{aligned}$$

Proof. By taking the expectation in (2.18) and integrating over $\mathbb{R}$ with respect to a, we get the claimed estimate. Indeed, by the scaling (Lemma 2.3.3) and the stationarity of $\bar{K}$ (Lemma 2.2.4),

$$\begin{aligned} \int_{\mathbb{R}} \mathbb{E}[U_{s,t}(\varepsilon, B-a)]\mathrm{d}a &= \int_{\mathbb{R}} \mathbb{E}[U_{\frac{s}{t-s}\zeta, \frac{t}{t-s}\zeta}(1, B - \zeta^H (t-s)^{-H} a)]\mathrm{d}a \\ &= \zeta^{-H}(t-s)^H \int_{\mathbb{R}} \mathbb{E}[U_{\frac{s}{t-s}\zeta, \frac{t}{t-s}\zeta}(1, B-a)]\mathrm{d}a \\ &= \frac{(t-s)^H}{2\zeta^H} \mathbb{E}[\bar{K}_{0,\zeta}(1,B)]. \end{aligned} \tag{2.19}$$

We see

$$\int_{\mathbb{R}} e^{-\frac{1}{2}(\frac{X_s - a}{\sigma_{Y+Z}})^2} \mathrm{d}a = \sqrt{2\pi}\sigma_{Y+Z} \lesssim (s-v)^H.$$

Therefore,

$$\begin{aligned} &\Big|\sqrt{2\pi}\sigma_{Y+Z} \mathbb{E}[U_{s-v,t-v}(\varepsilon, \mathcal{B})] - \frac{1}{2}\mathbb{E}[\bar{K}_0^\zeta(1,B)]\Big(\frac{t-s}{\zeta}\Big)^H\Big| \\ &\lesssim_{\zeta,\kappa} \Big(\frac{t-s}{s-u}\Big)^{1-H} \int_{\mathbb{R}} \mathbb{E}[U_{s,t}(\varepsilon, B-a)]^{1-\kappa} e^{-ca^2}\mathrm{d}a + (t-u)^H \Big(\frac{t-s}{s-v}\Big)^{(1-\kappa)H} \\ &\qquad + (s-v)^H \Big(\frac{t-s}{s-v}\Big)^{(1-\kappa)H} \Big(\frac{t-s}{s-u}\Big)^{1-H}. \end{aligned}$$

By Jensen's inequality and (2.19),

$$\int_{\mathbb{R}} \mathbb{E}[U_{s,t}(\varepsilon, B-a)]^{1-\kappa} e^{-ca^2} \mathrm{d}a \lesssim \Big(\int_{\mathbb{R}} \mathbb{E}[U_{s,t}(\varepsilon, B-a)] e^{-ca^2} \mathrm{d}a \Big)^{1-\kappa}$$
$$\lesssim_\zeta (t-s)^{(1-\kappa)H}.$$

The claimed estimate is now established. $\square$

Proof of Lemma 2.3.7. By Lemma 2.3.15 and Lemma 2.3.16, we have

$$\Big\| U_{s,t}(\varepsilon, B-a) - \frac{1}{2}\mathbb{E}[\bar{K}_0^\zeta(1,B)] \frac{1}{\sqrt{2\pi}\sigma_{Y+Z}} e^{-\frac{1}{2}(\frac{X_s}{\sigma_{Y+Z}})^2} \Big(\frac{t-s}{\zeta}\Big)^H \Big\|_{L^p(\mathbb{P})}$$
$$\lesssim \Big(\frac{t-s}{s-u}\Big)^{1-H} (t-s)^{-\kappa} + \Big(\frac{t-s}{s-u}\Big)^{-H} \Big(\frac{t-s}{s-v}\Big)^{H(2-\kappa)}$$

if $\frac{t-s}{s-u}$ and $\frac{t-u}{u-v}$ are sufficiently small. To optimize, we choose u so that

$$\frac{t-s}{s-u} = \Big(\frac{t-s}{s-v}\Big)^{(2-\kappa)H}.$$

Note that, as $H < \frac{1}{2}$, the exponent $(2-\kappa)H$ is less than 1. Therefore, if $\frac{t-s}{s-v}$ is sufficiently small, then $\frac{t-s}{s-u}$ and $\frac{t-u}{u-v}$ are sufficiently small as well. This gives the claimed bound. $\square$

Estimates on the local time

The following is the last technical ingredient for Theorem 2.3.1.

Lemma 2.3.17. *Let $H < \frac{1}{2}$. We set*

$$\begin{aligned}\tilde{A}_{s,t} :=& \mathbb{E}[\delta_0(B_s - a) | \mathcal{F}_{s-(t-s)}](t-s) \\ =& \sqrt{\frac{H}{\pi}} e^{-\frac{H}{(t-s)^{2H}} \mathbb{E}[B_s - a | \mathcal{F}_{s-(t-s)}]^2} (t-s)^{1-H}.\end{aligned} \tag{2.20}$$

Then, there exists a $\delta > 0$ such that for any $p < \infty$ and for any partition π of $[0,1]$,

$$\Big\| L_1(a) - \sum_{[s,t]\in\pi} \tilde{A}_{s,t} \Big\|_{L^p(\mathbb{P})} \lesssim_p |\pi|^\delta.$$

Proof. We use the shifted stochastic sewing (Theorem 1.1.1). To this end, it suffices to check

$$\|L_{s,t}(a)\|_{L^p(\mathbb{P})} \lesssim (t-s)^{1-H}, \quad \|\tilde{A}_{s,t}\|_{L^p(\mathbb{P})} \lesssim (t-s)^{1-H} \tag{2.21}$$

and

$$\|\mathbb{E}[L_{s,t}(a) - \tilde{A}_{s,t}|\mathcal{F}_v]\|_{L^p(\mathbb{P})} \lesssim (s-v)^{-1-H}(t-s)^2, \quad s-v \le t-s. \tag{2.22}$$

The estimates (2.21) are trivial, hence we focus on the estimate (2.22).

We have

$$\mathbb{E}[L_{s,t}(a) - \tilde{A}_{s,t}|\mathcal{F}_v] = \sqrt{\frac{H}{\pi}} \int_s^t \Big\{ e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r - a|\mathcal{F}_v]^2}(r-v)^{-H} \\ - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_s - a|\mathcal{F}_v]^2}(s-v)^{-H} \Big\} \mathrm{d}r.$$

For simplification, we replace $B-a$ by B. We decompose the integrand as $R_1 + R_2 + R_3$, where

$$\begin{aligned}
R_1 &:= e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(r-v)^{-H} - e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(s-v)^{-H},\\
R_2 &:= e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(s-v)^{-H} - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(s-v)^{-H},\\
R_3 &:= e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(s-v)^{-H} - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_s|\mathcal{F}_v]^2}(s-v)^{-H}.
\end{aligned}$$

Since

$$0 \le (s-v)^{-H} - (r-v)^{-H} \lesssim (s-v)^{-H-1}(r-s),$$

we have

$$|R_1| \lesssim (s-v)^{-H-1}(t-s).$$

We observe

$$\begin{aligned}
&e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2} - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}\\
&\quad = e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}\big(1 - e^{-H((s-v)^{-2H}-(r-v)^{-2H})\mathbb{E}[B_r|\mathcal{F}_v]^2}\big)\\
&\quad \lesssim e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}\mathbb{E}[B_r|\mathcal{F}_v]^2((s-v)^{-2H} - (r-v)^{-2H})\\
&\quad \lesssim e^{-\frac{H}{(r-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}\mathbb{E}[B_r|\mathcal{F}_v]^2(s-v)^{-2H-1}(r-s)\\
&\quad \lesssim (r-v)^{2H}(s-v)^{-2H-1}(r-s).
\end{aligned}$$

Hence,

$$|R_2| \lesssim (s-v)^{-1-H}(t-s).$$

Finally, we estimate R_3. Suppose that $\mathbb{E}[B_r|\mathcal{F}_v]^2 \le \mathbb{E}[B_s|\mathcal{F}_v]^2$. Then,

$$\Big| e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2} - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_s|\mathcal{F}_v]^2} \Big| \\ \le e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}\frac{H}{(s-v)^{2H}}(\mathbb{E}[B_s|\mathcal{F}_v]^2 - \mathbb{E}[B_r|\mathcal{F}_v]^2).$$

Since

$$e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2}(s-v)^{-H}|\mathbb{E}[B_r|\mathcal{F}_v]| \lesssim 1,$$

we obtain

$$\begin{aligned}&\left|e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_r|\mathcal{F}_v]^2} - e^{-\frac{H}{(s-v)^{2H}}\mathbb{E}[B_s|\mathcal{F}_v]^2}\right| \\ &\qquad \lesssim (s-v)^{-H}|\mathbb{E}[B_s|\mathcal{F}_v] - \mathbb{E}[B_r|\mathcal{F}_v]| + (s-v)^{-2H}|\mathbb{E}[B_s|\mathcal{F}_v] - \mathbb{E}[B_r|\mathcal{F}_v]|^2.\end{aligned}$$

A similar estimate holds if $\mathbb{E}[B_r|\mathcal{F}_v]^2 \geq \mathbb{E}[B_s|\mathcal{F}_v]^2$. Therefore, it remains to note

$$\|\mathbb{E}[B_s|\mathcal{F}_v] - \mathbb{E}[B_r|\mathcal{F}_v]\|_{L^p(\mathbb{P})} \lesssim_p (s-v)^{H-1}(t-s). \qquad \square$$

Concluding estimates

Now we can finish the proof of Theorem 2.3.1. Let π be a partition of $[0,1]$. By Lemma 2.3.5,

$$\varepsilon^{\frac{1}{H}-1}U_{0,1}(\varepsilon, B-a) \geq \sum_{[s,t]\in\pi} \varepsilon^{\frac{1}{H}-1}U_{s,t}(\varepsilon, B-a), \tag{2.23}$$

$$\varepsilon^{\frac{1}{H}-1}U_{0,1}(\varepsilon, B-a) \leq \sum_{[s,t]\in\pi} \varepsilon^{\frac{1}{H}-1}\bar{U}_{s,t}(\varepsilon, B-a). \tag{2.24}$$

Lemma 2.3.18. *Let $H < \frac{1}{2}$, $p \in [2,\infty)$, $\varepsilon \in (0,1)$ and $\zeta \in [1,\infty)$. Then, we have*

$$\varepsilon^{\frac{1}{H}-1}U_{0,1}(\varepsilon, B-a) \geq \frac{1}{2\zeta}\mathbb{E}[\bar{K}_{0,\zeta}(1,B)]L_1(a) - R_\varepsilon,$$

where for some δ depending only on H we have

$$\|R_\varepsilon\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \varepsilon^\delta.$$

Proof. We define $\tilde{A}$ by (2.20), and we set

$$\hat{A}_{s,t} := U_{s,t}(\zeta^{-H}(t-s)^H, B-a)\Big(\frac{t-s}{\zeta}\Big)^{1-H}.$$

By Lemma 2.3.6, we have

$$\|\hat{A}_{s,t}\|_{L^p(\mathbb{P})} \lesssim (t-s)^{1-H}.$$

By Lemma 2.3.7,

$$\begin{aligned}\mathbb{E}[\hat{A}_{s,t}|\mathcal{F}_v] &= \frac{1}{2\zeta}\mathbb{E}[\bar{K}_0^{\zeta}(1,B)]\sqrt{\frac{H}{\pi(s-v)^{2H}}}e^{-\frac{H\mathbb{E}[B_s|\mathcal{F}_v]^2}{(s-v)^{2H}}}(t-s)+R_{v,s,t} \\ &= \frac{1}{2\zeta}\mathbb{E}[\bar{K}_0^{\zeta}(1,B)]\mathbb{E}[\tilde{A}_{s,t}|\mathcal{F}_v]+R_{v,s,t},\end{aligned}$$

where

$$\|R_{v,s,t}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta,\kappa} \left(\frac{t-s}{s-v}\right)^{(2-\kappa)H(1-H)}(t-s)^{1-(1+\kappa)H}.$$

for any $\kappa \in (0,1)$. Since $H < \frac{1}{2}$, choosing κ sufficiently small, we can suppose that

$$1-(1+\kappa)H > \frac{1}{2}, \quad 1-(1+\kappa)H+(2-\kappa)H(1-H) > 1.$$

Hence, by Theorem 1.1.1, with some $\delta = \delta(H)$,

$$\Big\|\sum_{[s,t]\in\pi}\Big(\hat{A}_{s,t}-\frac{1}{2\zeta}\mathbb{E}[\bar{K}_0^{\zeta}(1,B)]\tilde{A}_{s,t}\Big)\Big\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} |\pi|^{\delta}.$$

In particular, considering a partition of size $\zeta\varepsilon^{\frac{1}{H}}$, the claim follows in view of (2.23) and Lemma 2.3.17.

□

Lemma 2.3.19. *Let $H < \frac{1}{2}$, $p \in [2,\infty)$, $\varepsilon \in (0,1)$ and $\zeta \in [1,\infty)$. Then, we have*

$$\varepsilon^{\frac{1}{H}-1}U_{0,1}(\varepsilon, B-a) \leq \frac{1}{2\zeta}(\mathbb{E}[\bar{K}_{0,\zeta}(1,B)]+1)L_1(a)+\bar{R}_{\varepsilon},$$

where for some δ depending only on H we have

$$\|\bar{R}_{\varepsilon}\|_{L^p(\mathbb{P})} \lesssim_{p,\zeta} \varepsilon^{\delta}.$$

Proof. In view of (2.24), the proof is similar to Lemma 2.3.18. □

Proof of Theorem 2.3.1. It readily follows from Lemma 2.3.18 and Lemma 2.3.19, and the estimate (2.12). □

2.3.3 Lemiuex type result

Theorem 2.3.1 tells that for any $a \in \mathbb{R}$ and for any $\boldsymbol{\varepsilon} = (\varepsilon_n)_{n=1}^{\infty}$ with polynomial decay, there exists a measurable set $\Omega_{a,\boldsymbol{\varepsilon}}$ such that $\mathbb{P}(B \in \Omega_{a,\boldsymbol{\varepsilon}}) = 1$ and for every $w \in \Omega_{a,\boldsymbol{\varepsilon}}$ the limit

$$\lim_{n\to\infty} \varepsilon_n^{\frac{1}{H}-1} U_{0,t}(\varepsilon_n, w-a)$$

exists for every $t \geq 0$. As observed by Lemieux [Lem83], the quantitative estimate in Theorem 2.3.1 implies that we can take $\Omega_{a,\boldsymbol{\varepsilon}}$ uniformly over a and $\boldsymbol{\varepsilon}$. Furthermore, we can remove the polynomial decaying condition.

We begin with the following lemma.

Lemma 2.3.20. *We define the grid*

$$G_k := \{ik^{-7} : i \in \mathbb{Z}, |i| \leq k^8\}, \quad k \in \mathbb{N}.$$

We then have

$$\lim_{k\to\infty} \max_{x\in G_k} \left| k^{-6(\frac{1}{H}-1)} U_{0,t}(k^{-6}, B^H - x) - \frac{\mathfrak{c}_H}{2} L_t^H(x) \right| = 0 \quad \textit{a.s.}$$

Proof. In the notation of Theorem 2.3.1, we have

$$\max_{x\in G_k} \left| k^{-6(\frac{1}{H}-1)} U_{0,t}(k^{-6}, B^H - x) - \frac{\mathfrak{c}_H}{2} L_t^H(x) \right| \leq \zeta^{-1} \sup_{x\in\mathbb{R}} L_t(x) + \max_{x\in G_k} R_{k,\zeta,x}.$$

Since $x \mapsto L_t(x)$ is continuous and $L_t(\cdot)$ is supported on

$$\{x \in \mathbb{R} : |x| \leq \|B\|_{L^\infty([0,t])}\},$$

we see that $\sup_{x\in\mathbb{R}} L_t(x) < \infty$ a.s. By Theorem 2.3.1,

$$\|\max_{x\in G_k} R_{k,\zeta,x}\|_p^p \leq \sum_{x\in G_k} \|R_{k,\zeta,x}\|_p^p \lesssim_{p,\zeta} k^{-p\delta+8},$$

where δ is independent of p. Since p can be arbitrarily large, the Borel–Cantelli lemma implies that almost surely we have

$$\lim_{k\to\infty} \max_{x\in G_k} R_{k,\zeta,x} = 0$$

and

$$\limsup_{k\to\infty} \max_{x\in G_k} \left| k^{-6(\frac{1}{H}-1)} U_{0,t}(k^{-6}, B^H - x) - \frac{\mathfrak{c}_H}{2} L_t^H(x) \right| \leq \zeta^{-1} \sup_{x\in\mathbb{R}} L_t(x).$$

Since ζ is arbitrary, we complete the proof. $\square$

Theorem 2.3.21. *Let $H \in (0, \frac{1}{2})$. Almost surely, we have*

$$\lim_{\varepsilon \to 0} \sup_{a \in \mathbb{R}} \left| \varepsilon^{\frac{1}{H}-1} U_{0,t}(\varepsilon, B - a) - \frac{\mathfrak{c}_H}{2} L_t(a) \right| = 0 \quad \forall t \geq 0.$$

Proof. The proof follows [Lem83, Theorem II.2.4]. Due to the monotonicity of U and L, we can fix a time $t \geq 0$. By Lemma 2.3.20, we can find an $\Omega_1 \subseteq \Omega$ with $\mathbb{P}(\Omega_1) = 1$ such that for any $\delta \in (0,1)$ and $\omega \in \Omega_1$ there exists an $N = N(\delta, \omega)$ with the following inequalities:

$$(k-1)^{-6} - k^{-7} > k^{-6} \quad \forall k \geq N, \tag{2.25}$$

$$\|B(\omega)\|_{L^\infty([0,t])} < N - 1, \tag{2.26}$$

$$\sup_{k \geq N} \max_{x \in G_k} \left| k^{-6(\frac{1}{H}-1)} U_{0,t}(k^{-6}, B^H(\omega) - x) - \frac{\mathfrak{c}_H}{2} L_t^H(x)(\omega) \right| < \delta. \tag{2.27}$$

The argument below holds on the event Ω_1. For $\varepsilon \leq (N+1)^{-6}$, there exists a unique $m = m_\varepsilon \geq N+1$ such that

$$(m+1)^{-6} < \varepsilon^{\frac{1}{H}-1} \leq m^{-6}.$$

If $|x| \geq N-1$, then by (2.26) we have $L_t^H(x) = 0$. On the other hand, if $|x| < N-1$, then we define

$$x_k := \max_{y \in G_k} \{y \leq x\}$$

for all $k \geq N$. Since $x < x_{m-1} + (m-1)^{-7}$, we have

- $x_{m-1} \leq x < x + \varepsilon < x_{m-1} + (m-1)^{-7}$ and
- $x < x_{m+2} + (m+2)^{-7} < x_{m+2} + (m+2)^{-7} + (m+2)^{-6} < x + \varepsilon$,

where (2.25) is applied in the inequality of the second item. Hence, defining the two sets I_{m-1} and $\bar{I}_{m+2}$ as

$$I_{m-1} := \left[x_{m-1}, x_{m-1} + (m-1)^{-7}\right], \quad \bar{I}_{m+2} := \left[\bar{x}_{m+2}, \bar{x}_{m+2} + (m+2)^{-6}\right],$$

where $\bar{x}_{m+2} := x_{m+2} + (m+2)^{-7}$, we have the inclusions

$$\bar{I}_{m+2} \subseteq [x, x+\varepsilon] \subseteq I_{m-1}. \tag{2.28}$$

Now we move to the bound on U. We first observe the monotonicity of U:

$$U_{0,t}(\varepsilon_1, B - x_1) \leq U_{0,t}(\varepsilon_2, B - x_2)$$

provided that $[x_2, x_2 + \varepsilon_2] \subseteq [x_1, x_1 + \varepsilon_1]$. The relation (2.28) thus yields

$$U_{0,t}((m-1)^{-7}, B - x_{m-1}) \leq U_{0,t}(\varepsilon, B - x) \leq U_{0,t}((m+2)^{-7}, B - \bar{x}_{m+2}).$$

Hence,

$$\sup_{x \in \mathbb{R}} \left| \varepsilon^{\frac{1}{H}-1} U_{0,t}(\varepsilon, B - x) - \frac{\mathfrak{c}_H}{2} L_t(x) \right| \leq A_\varepsilon + \bar{A}_\varepsilon + \frac{\mathfrak{c}_H}{2} \sup_{x,y:|x-y| \leq 2\varepsilon} |L_t(x) - L_t(y)|,$$

where

$$A_\varepsilon := \sup_{x \in R_{m_\varepsilon - 1}} \left| \varepsilon^{\frac{1}{H}-1} U_{0,t}((m_\varepsilon - 1)^{-6}, B^H - x) - \frac{\mathfrak{c}_H}{2} L_t(x) \right|,$$
$$\bar{A}_\varepsilon := \sup_{x \in R_{m_\varepsilon + 2}} \left| \varepsilon^{\frac{1}{H}-1} U_{0,t}((m_\varepsilon + 2)^{-6}, B^H - x) - \frac{\mathfrak{c}_H}{2} L_t(x) \right|.$$

By (2.27),

$$\limsup_{\varepsilon \to 0} A_\varepsilon + \limsup_{\varepsilon \to 0} \bar{A}_\varepsilon \leq 2\delta.$$

Since δ is arbitrary, we conclude the proof. □

Analogous to $U_{s,t}$, we denote by $D_{s,t}$ the total number of downcrossings

$$D_{s,t}(\varepsilon, w) := \#\big\{(u,v) \in \Delta_{s,t} : w_v = 0, w_u = \varepsilon, \forall r \in (u,v)\ w_r \in (0,\varepsilon)\big\}. \tag{2.29}$$

By definition, we have the identity

$$K_{0,t}(\varepsilon, w) = \sum_{x \in \varepsilon\mathbb{Z}} \{U_{0,t}(\varepsilon, w - x) + D_{0,t}(\varepsilon, w - x)\}.$$

Furthermore, since the total number of upcrossings and that of downcrossings can differ by at most 1, almost surely we have

$$\lim_{\varepsilon \to 0} \sup_{x \in \mathbb{R}} \left| \varepsilon^{\frac{1}{H}-1} D_{0,t}(\varepsilon, B - x) - \frac{\mathfrak{c}_H}{2} L_t(x) \right| = 0 \quad \forall t \geq 0,$$

or

$$\lim_{\varepsilon \to 0} \sup_{x \in \mathbb{R}} \left| \varepsilon^{\frac{1}{H}-1} (U_{0,t}(\varepsilon, B - x) + D_{0,t}(\varepsilon, B - x)) - \mathfrak{c}_H L_t(x) \right| = 0 \quad \forall t \geq 0, \tag{2.30}$$

Theorem 2.3.22. *Let $H \in (0, \frac{1}{2})$. Almost surely, we have*

$$\lim_{\varepsilon \to 0} \sup_{\rho \in \mathbb{R}} |\varepsilon^{\frac{1}{H}} K_{0,t}(\varepsilon, B - \rho) - \mathfrak{c}_H t| = 0, \quad \forall t \geq 0.$$

Proof. We observe

$$\begin{aligned}
&\sup_\rho |\varepsilon^{1/H} K_{0,t}(\varepsilon, B-\rho) - \mathfrak{c}_H t| \\
&\quad = \sup_\rho \Big| \varepsilon^{1/H} K_{0,t}(\rho, B-\rho) - \mathfrak{c}_H \int_{\mathbb{R}} L_t^H(a) \mathrm{d}a \Big| \\
&\quad = \sup_\rho \Big| \sum_{x \in \rho + \varepsilon\mathbb{Z}} \int_x^{x+\varepsilon} \big\{ \varepsilon^{\frac{1}{H}-1} (U_{0,t}(\varepsilon, B-x) + D_{0,t}(\varepsilon, B-x)) - \mathfrak{c}_H L_t(a) \big\} \mathrm{d}a \Big| \\
&\quad \lesssim_{\|B\|_{L^\infty([0,t])}} \sup_x \big\{ \varepsilon^{\frac{1}{H}-1} (U_{0,t}(\varepsilon, B-x) + D_{0,t}(\varepsilon, B-x)) - \mathfrak{c}_H L_t(x) \big\} \\
&\qquad\qquad + \sup_{x,y : |x-y| \le \varepsilon} |L_t(x) - L_t(y)|.
\end{aligned}$$

In view of (2.30) and the uniform continuity of $L_t(\cdot)$, the claim follows. □

2.4 A fractional excursion measure

Instead of counting level crossings, we can similarly count excursions. By so doing, we can define a natural notion of an excursion measure of the fractional Brownian motion.

Definition 2.4.1. Let w be a continuous path and $s < t$. The set $\{r \in [s,t] : w_r > 0\}$ is open in $[s,t]$. Therefore, we can write

$$\{r \in [s,t] : w(r) > 0\} = \bigcup_{\lambda \in \bar{\Lambda}_{s,t}(w)} I_\lambda(w),$$

where $I_\lambda(w)$, $\lambda \in \bar{\Lambda}_{s,t}(w)$, are disjoint intervals of the form $[s,a)$, (a,b), $(b,t]$ or $[s,t]$. We set

$$\Lambda_{s,t}(w) := \{\lambda \in \bar{\Lambda}_{s,t}(w) : w(\inf I_\lambda(w)) = w(\sup I_\lambda(w)) = 0\}.$$

Note that removed intervals are of the form $[s,a)$, $(b,t]$ or $[s,t]$. For $\lambda \in \Lambda_{s,t}(w)$ we define $e^\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$e^\lambda(t) := w((t + \inf I_\lambda) \wedge \sup I_\lambda).$$

We say that e^λ is an *excursion* of w (see Figure 2.6). For a Borel subset Γ of $C(\mathbb{R}_+, \mathbb{R}_+)$ we set

$$\mathfrak{m}_{s,t}(\Gamma, w) := \#\{\lambda \in \Lambda_{s,t}(w) : e^\lambda \in \Gamma\}.$$

Note that $\mathfrak{m}(\cdot, w)$ is a measure on $C(\mathbb{R}_+, \mathbb{R}_+)$ and that $\mathfrak{m}(\cdot, w)$ is supported on

$$\mathcal{E} := \{w \in C(\mathbb{R}_+, \mathbb{R}_+) : w(0) = 0, \sigma(w) > 0, \forall t \ge \sigma(w), w_t = 0\} \setminus \{0\},$$

where $\sigma(w) := \inf\{t > 0 : w_t = 0\}$.

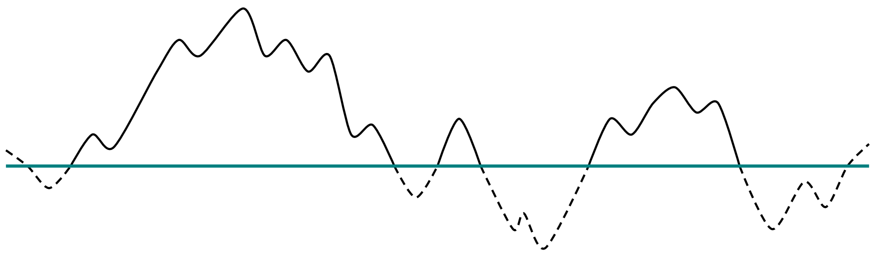


Figure 2.6: Excursions

The family $(\mathfrak{m}_{s,t}(\Gamma, w))_{s<t}$ is similar to $(U_{s,t}(\varepsilon, w))_{s<t}$. As in Lemma 2.3.5, we can prove the following.

Lemma 2.4.2. *The two-parameter family* $(\mathfrak{m}_{s,t}(\Gamma, w))_{s<t}$ *is superadditive: for* $r < s < t$ *we have*

$$\mathfrak{m}_{r,t}(\Gamma, w) \geq \mathfrak{m}_{r,s}(\Gamma, w) + \mathfrak{m}_{s,t}(\Gamma, w).$$

Definition 2.4.3. We set

$$\mathfrak{M}_{s,t}(\Gamma, w) := \int_{\mathbb{R}} \mathfrak{m}_{s,t}(\Gamma, w - \rho) \mathrm{d}\rho.$$

Lemma 2.4.4. *Regarding* $\mathfrak{M}$, *we have the following.*

(i) The family $(\mathfrak{M}(\Gamma, w))_{s<t}$ *is superadditive.*

(ii) (stationarity) For $s < t$ *we have* $\mathfrak{M}_{s,t}(\Gamma, B) \stackrel{\mathrm{d}}{=} \mathfrak{M}_{0,t-s}(\Gamma, B)$.

Proof. The proof is similar to Lemma 2.2.4. □

Definition 2.4.5. For $\Gamma \in \mathcal{B}(\mathcal{E})$ and $t \in (0, \infty)$ we consider

$$a(t) := \mathbb{E}[\mathfrak{M}_{0,t}(\Gamma, B)].$$

Then, by Lemma 2.4.4 we have $a(s + t) \geq a(s) + a(t)$. Therefore, the following limit exists:

$$\hat{P}(\Gamma) := \hat{P}^H(\Gamma) := \lim_{t\to\infty} \frac{1}{t} a(t) = \sup_{t\in(0,\infty)} \frac{1}{t} a(t) \in [0, \infty].$$

Now we would like to show that $\hat{P}$ is non-trivial. We begin with the lower bound. We set

$$\mathcal{E}_\delta := \{e \in \mathcal{E} : \max_t e(t) \geq \delta\}.$$

Lemma 2.4.6. *We have* $\hat{P}(\mathcal{E}_\delta) > 0$.

Proof. By the superadditivity it suffices to show $\mathbb{E}[\mathfrak{M}_{0,1}(\Gamma, B)] > 0$. Then we note

$$\mathbb{E}[\mathfrak{M}_{0,1}(\Gamma, B)] \geq \mathbb{P}(\max_{t\in[0,1]} B(t) \geq \delta, B(1) < 0) > 0. \qquad \square$$

To prove the upper bound $\hat{P}(\mathcal{E}_\delta) < \infty$, we introduce the following.

Definition 2.4.7. For $\Gamma \in \mathcal{B}(\mathcal{E})$ and a continuous path w, we set

$$\mathfrak{n}_{s,t}(\Gamma, w) := \mathfrak{m}_{s,t}(\Gamma, w) + \mathbf{1}_{\{\min_{r\in[s,t]} w(r)\leq 0, w(t)>0\}},$$

and $\mathfrak{N}_{s,t}(\Gamma, w) := \int_{\mathbb{R}} \mathfrak{n}_{s,t}(\Gamma, w-\rho)\mathrm{d}\rho$.

Lemma 2.4.8. *Regarding* $\mathfrak{N}$*, we have the following.*

(i) The family $(\mathfrak{n}_{s,t}(\Gamma, w))_{s<t}$ *is subadditive.*

(ii) (stationarity) For $s < t$ *we have* $\mathfrak{N}_{s,t}(\Gamma, B) \stackrel{\mathrm{d}}{=} \mathfrak{N}_{0,t-s}(\Gamma, B)$.

Proof. We only prove (i), since the proof of (ii) is similar to Lemma 2.4.4-(ii). To prove (i), we separate cases.

- Suppose that $w(s) \leq 0$. In this case we have

$$\mathfrak{n}_{r,s}(\Gamma, w) = \mathfrak{n}_{r,s}(\Gamma, w) + \mathfrak{n}_{s,t}(\Gamma, w).$$

- Suppose that $w(s) > 0$.
  - If $\inf_{u\in[s,t]} w(s) > 0$ then $\mathfrak{n}_{s,t}(\Gamma, w) = 0$ and

$$\mathfrak{n}_{r,t}(\Gamma, w) = \mathfrak{n}_{r,s}(\Gamma, w) = \mathfrak{n}_{r,s}(\Gamma, w) + \mathfrak{n}_{s,t}(\Gamma, w).$$

    The case is similar if $\inf_{u\in[r,s]} w(u) > 0$.
  - If $\inf_{u\in[s,t]} w(u) \leq 0$ and $\inf_{u\in[r,s]} w(u) \leq 0$, let e be the excursion of w such that the starting point is less than s and the ending point is greater than s. Then,

$$\mathfrak{m}_{r,t}(\Gamma, w) = \mathfrak{m}_{r,s}(\Gamma, w) + \mathfrak{m}_{s,t}(\Gamma, w) + \mathbf{1}_{\{e\in\Gamma\}},$$

$$\mathfrak{n}_{r,t}(\Gamma, w) - \mathfrak{m}_{r,t}(\Gamma, w) = \mathfrak{n}_{s,t}(\Gamma, w) - \mathfrak{m}_{s,t}(\Gamma, w),$$
$$\mathfrak{n}_{r,s}(\Gamma, w) = \mathfrak{m}_{r,s}(\Gamma, w) + 1.$$

    Therefore, we have $\mathfrak{n}_{r,t}(\Gamma, w) + \mathbf{1}_{\{e\notin\Gamma\}} = \mathfrak{n}_{r,s}(\Gamma, w) + \mathfrak{n}_{s,t}(\Gamma, w)$. $\square$

Again, by the subadditive ergodic theorem, the limit

$$\tilde{P}(\Gamma) := \lim_{T\to\infty} \frac{1}{T}\mathbb{E}[\mathfrak{N}_{0,T}(\Gamma, B)] = \inf_{T>0} \frac{1}{T}\mathbb{E}[\mathfrak{N}_{0,T}(\Gamma, B)] \tag{2.31}$$

exists. We will see that $\hat{P}(\Gamma) = \tilde{P}(\Gamma)$. But before that we see the following.

Lemma 2.4.9. *We have* $\tilde{P}(\mathcal{E}_\delta) < \infty$.

Proof. By (2.31), it suffices to show $\mathbb{E}[\mathfrak{N}_{0,1}(\Gamma, B)] < \infty$. Then, the proof is similar to Lemma 2.2.5. □

Lemma 2.4.10. *We have* $\hat{P} = \tilde{P}$. *In particular,* $\hat{P}(\mathcal{E}_\delta) < \infty$ *for all* $\delta > 0$.

Proof. We have the inequalities

$$\mathfrak{m}_{0,T}(\Gamma, B-\rho) \le \mathfrak{n}_{0,T}(\Gamma, B-\rho) \le \mathfrak{m}_{0,T}(\Gamma, B-\rho) + \mathbf{1}_{\{|\rho|\le \|B\|_{L^\infty([0,T])}\}}.$$

Therefore,

$$\hat{P}(\Gamma) \le \tilde{P}(\Gamma) \le \hat{P}(\Gamma) + \limsup_{T\to\infty} \frac{2}{T}\mathbb{E}[\|B\|_{L^\infty([0,T])}].$$

By the scaling we have $\mathbb{E}[\|B\|_{L^\infty([0,T])}] = T^H \mathbb{E}[\|B\|_{L^\infty([0,1])}]$ and the claim follows. □

Now we show that $\hat{P}$ is a measure.

Lemma 2.4.11. *The map* $\mathcal{B}(\mathcal{E}) \ni \Gamma \mapsto \hat{P}(\Gamma) \in [0,\infty]$ *defines a* σ*-finite measure on* $\mathcal{E}$.

Proof. We have $\mathcal{E} = \bigcup_{n\in\mathbb{N}} \mathcal{E}_{1/n}$, and by Lemma 2.4.10 we have $\hat{P}(\mathcal{E}_{1/n}) < \infty$. Thus, it suffices to show that $\hat{P}$ is a measure. Let $(\Gamma_n)_{n=1}^\infty$ be disjoint subsets from $\mathcal{B}(\mathcal{E})$. Since

$$T \mapsto \frac{1}{T}\mathbb{E}\left[\mathfrak{M}(\Gamma, B|_{[0,T]})\right]$$

is non-decreasing, by the monotone convergence theorem,

$$\begin{aligned}
\hat{P}\Big(\bigcup_n \Gamma_n\Big) &= \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\mathfrak{M}_{0,T}\Big(\bigcup_n \Gamma_n, B\Big)\Big] \\
&= \lim_{T\to\infty} \sum_n \frac{1}{T}\mathbb{E}[\mathfrak{M}_{0,T}(\Gamma_n, B)] \\
&= \sum_n \hat{P}(\Gamma_n).
\end{aligned}$$

□

Definition 2.4.12. We call the σ-finite measure $\hat{P} = \hat{P}^H$ on $\mathcal{E}$ an excursion measure of the fractional Brownian motion B^H.

To see that $\hat{P}$ is a natural notion of an excursion measure of B, we show that it satisfies the self similar property (Proposition 2.4.13) and that for $H = \frac{1}{2}$ it coincides with the usual Brownian excursion measure (Proposition 2.4.16).

Proposition 2.4.13 (Self-similarity)**.** *For $\Gamma \in \mathcal{B}(\mathcal{E})$ and $\lambda \in (0, \infty)$ we set*

$$\lambda\Gamma := \{\lambda w : w \in \Gamma\}, \quad \Gamma_\lambda := \{w \in \mathcal{E} : w_{\lambda^{-1}\cdot} \in \Gamma\}.$$

Then we have

$$\hat{P}(\lambda\Gamma) = \lambda^{1-1/H}\hat{P}(\Gamma_{\lambda^{1/H}}).$$

Proof. We observe

$$\begin{aligned}
\hat{P}(\lambda\Gamma) &= \lim_{T\to\infty} \frac{1}{T}\int_{\mathbb{R}} \mathbb{E}[\mathfrak{m}_{0,T}(\lambda\Gamma, B-\rho)]\mathrm{d}\rho \\
&= \lim_{T\to\infty} \frac{1}{T}\int_{\mathbb{R}} \mathbb{E}[\mathfrak{m}_{0,T}(\Gamma, \lambda^{-1}(B-\rho))]\mathrm{d}\rho \\
&= \lim_{T\to\infty} \frac{1}{T}\int_{\mathbb{R}} \mathbb{E}[\mathfrak{m}_{0,T}(\Gamma, \lambda^{-1}B-\eta)]\lambda\mathrm{d}\eta & \lambda^{-1}\rho = \eta \\
&= \lambda\lim_{T\to\infty} \frac{1}{T}\int_{\mathbb{R}} \mathbb{E}[\mathfrak{m}_{0,T}(\Gamma, B_{\lambda^{-1/H}\cdot}-\eta)]\mathrm{d}\eta & \lambda^{-1}B \overset{d}{=} B_{\lambda^{-1/H}\cdot} \\
&= \lambda\lim_{T\to\infty} \frac{1}{T}\int_{\mathbb{R}} \mathbb{E}[\mathfrak{m}_{0,\lambda^{-1/H}T}(\Gamma_{\lambda^{-1/H}}, B-\eta)]\mathrm{d}\eta \\
&= \lambda^{1-1/H}\hat{P}(\Gamma_{\lambda^{-1/H}}).
\end{aligned}$$

□

Corollary 2.4.14. *For $\delta \in (0, \infty)$ we have $\hat{P}^H(\mathcal{E}_\delta) = \frac{\mathfrak{c}_H}{2}\delta^{1-1/H}$.*

Proof. By Proposition 2.4.13, we have

$$\hat{P}^H(\mathcal{E}_\delta) = \hat{P}^H(\delta\mathcal{E}_1) = \delta^{1-1/H}\hat{P}(\mathcal{E}_1).$$

It therefore remains to show $\hat{P}(\mathcal{E}_1) = \mathfrak{c}_H/2$. To this end, we observe

$$\mathfrak{m}_{0,T}(\mathcal{E}_1, B-\rho) \le U_{0,T}(1, B-\rho) \le \mathfrak{n}_{0,T}(\mathcal{E}_1, B-\rho). \tag{2.32}$$

Integrating (2.32) with respect to ρ over $\mathbb{R}$, we get

$$\mathfrak{M}_{0,T}(\mathcal{E}_1, B) \le \int_{\mathbb{R}} U_{0,T}(1, B-\rho)\mathrm{d}\rho \le \mathfrak{N}_{0,T}(\mathcal{E}_1, B).$$

Taking the expectation, dividing them by T and letting $T \to \infty$, we complete the proof. □

Corollary 2.4.15. *We set* $\sigma(w) := \inf\{s > 0 : w_s = 0\}$ *for* $w \in \mathcal{E}$*. Then* $\hat{P}(\sigma > t) = \hat{P}(\sigma > 1)t^{-(1-H)}$.

Proof. If we set $\Gamma := \{w \in \mathcal{E} : \sigma(w) > 1\}$, then using the notation in Proposition 2.4.13 we have $\lambda\Gamma = \Gamma$ and $\Gamma_\lambda = \{w \in \mathcal{E} : \sigma(w) > \lambda\}$. Therefore, by Proposition 2.4.13 we have

$$\hat{P}(\Gamma) = \lambda^{1-1/H}\hat{P}(\Gamma_{\lambda^{-1/H}}).$$

Setting $\lambda := t^{-H}$ we obtain the result. □

Next, we show that $\hat{P}^{\frac{1}{2}}$ coincides with the standard Brownian excursion measure. To distinguish it from the fractional Brownian motion, we write $W := B^{1/2}$ for the Brownian motion. A standard construction of its excursion measure $\check{P}$ for the Brownian motion is as follows (e.g. [Blu92]). Let $(L_t(x))_{t\geq 0}$ be the (occupation density) local time of W and let β be the inverse local time at 0:

$$\beta(t) := \inf\{s : L_s(0) \geq t\}.$$

Then, recalling notation from Definition 2.4.1, the measure $\check{P}$ is given by[3]

$$\check{P}(\Gamma) = \mathbb{E}\Big[\sum_{\lambda\in\Lambda_{0,\beta(1)}(W)} \mathbf{1}_{\{e^\lambda\in\Gamma\}}\Big].$$

A natural question is if $\hat{P} = \check{P}$ for W. The answer is affirmative:

Proposition 2.4.16. *For the Brownian motion* W*, we have* $\hat{P} = \check{P}$.

Proof. This is an easy consequence of the excursion formula ([Blu92, Eq.(3.27)]):

$$\mathbb{E}\Big[\sum_{\lambda\in\Lambda(W+x)} Z_{\inf I_\lambda(W+x)}\mathbf{1}_{\{e^\lambda\in\Gamma\}}\Big] = \check{P}(\Gamma)\mathbb{E}\Big[\int_0^\infty Z_s \mathrm{d}L_s(x)\Big], \tag{2.33}$$

where $(Z_s)_{s\geq 0}$ is an adapted non-negative process and $\Gamma \in \mathcal{B}(\mathcal{E})$. For $T \in (0,\infty)$, if $Z_s = \mathbf{1}_{\{s\leq T\}}$, we have

$$\mathbb{E}\Big[\sum_{\lambda:\inf I_\lambda(W+x)\leq T} \mathbf{1}_{\{e^\lambda\in\Gamma\}}\Big] = \check{P}(\Gamma)\mathbb{E}[L_T(x)]. \tag{2.34}$$

[3]The measure $\check{P}$ is different from the Brownian excursion measure constructed in [Blu92] by factor of $\sqrt{2}$. This is because the book [Blu92] considers the local time with normalized Laplace transform, which differs from the occupation density local time by factor of $\sqrt{2}$. More discussion can be found in [Blu92, Section VI.2.b].

Observing the inequalities

$$\mathfrak{m}_{0,T}(\Gamma, W+x) \leq \sum_{\lambda:\inf I_\lambda(W+x)\leq T} \mathbf{1}_{\{e^\lambda\in\Gamma\}} \leq \mathfrak{n}_{0,T}(\Gamma, W+x)$$

and the identity $\int_{\mathbb{R}} L_T(x)\mathrm{d}x = T$, we obtain

$$\frac{1}{T}\mathbb{E}\left[\mathfrak{M}_{0,T}(\Gamma, W)\right] \leq \check{P}(\Gamma) \leq \frac{1}{T}\mathbb{E}\left[\mathfrak{N}_{0,T}(\Gamma, W)\right].$$

Now the claim follows by Lemma 2.4.10. □

When $H < \frac{1}{2}$, the relation (2.34) holds asymptotically. That is:

Theorem 2.4.17. *Let $H < \frac{1}{2}$. Provided $\hat{P}(\Gamma) < \infty$, we have*

$$\lim_{T\to\infty} T^{-(1-H)}|\mathfrak{m}_{0,T}(\Gamma, B) - \hat{P}(\Gamma)L_T(0)| = 0 \quad \text{in } L^p(\mathbb{P}) \text{ for all } p < \infty.$$

In particular, if $\hat{P}(\Gamma_i) < \infty$ ($i = 1, \dots, n$), then

$$\lim_{T\to\infty} T^{-(1-H)}(\mathfrak{m}_{0,T}(\Gamma_1, B), \dots, \mathfrak{m}_{0,T}(\Gamma_n, B)) \\ = (\hat{P}(\Gamma_1), \dots, \hat{P}(\Gamma_n))L_1(0) \quad \text{in law.}$$

Proof. For $\Gamma \in \mathcal{B}(\mathcal{E})$ and $\lambda \in (0,\infty)$ we set

$$\lambda\Gamma := \{\lambda w : w \in \Gamma\}, \quad \Gamma_\lambda := \{w \in \mathcal{E} : w_{\lambda^{-1}\cdot} \in \Gamma\}$$

and $\Gamma^{(\lambda)} := \lambda^{-H}\Gamma_\lambda$, $B^{(\lambda)} := \lambda^{-H}B_{\lambda\cdot}$. Since

$$\mathfrak{m}_{0,T}(\Gamma, B) = \mathfrak{m}_{0,1}(\Gamma^{(T)}, B^{(T)}),$$

we see

$$(\mathfrak{m}_{0,T}(\Gamma, B), L_T(0)) \stackrel{\mathrm{d}}{=} (\mathfrak{m}_{0,1}(\Gamma^{(T)}, B), T^{1-H}L_1(0)). \tag{2.35}$$

As in Theorem 2.3.1, we have

$$|T^{-(1-H)}\mathfrak{m}_{0,1}(\Gamma^{(T)}, B) - \hat{P}(\Gamma)L_1(0)|_p \leq \delta_\zeta L_1(0) + R_{T,\zeta},$$

where δ_ζ is non-random with $\lim_{\zeta\to 0}\delta_\zeta = 0$ and $R_{T,\zeta}$ satisfies

$$\|R_{T,\zeta}\|_p \lesssim_{\zeta,p} T^{-\varepsilon}$$

for some $\varepsilon > 0$ independent of ζ and p. In view of (2.35), this proves the claim. □

Bibliography

[ACX20] A. Ananova, R. Cont, and R. Xu. “Excursion Risk” (2020).

[Aza90] J.-M. Azaïs. “Conditions for convergence of number of crossings to the local time. Application to stable processes with independent increments and to gaussian processes”. *Probability and Mathematical Statistics (Pol)* 11 (1990), pp. 19–36.

[AW96] J.-M. Azaïs and M. Wschebor. “Almost sure oscillation of certain random processes”. *Bernoulli* 2.3 (Sept. 1996), pp. 257–270.

[AW09] J.-M. Azaïs and M. Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009, pp. xii+393.

[Blu92] R. Blumenthal. *Excursions of Markov Processes*. Probability and Its Applications. Birkhäuser Boston, 1992.

[CP19] R. Cont and N. Perkowski. “Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity”. *Trans. Amer. Math. Soc. Ser. B* 6 (2019), pp. 161–186.

[DOS18] M. Davis, J. Obłój, and P. Siorpaes. “Pathwise stochastic calculus with local times”. *Ann. Inst. H. Poincaré Probab. Statist.* 54.1 (Feb. 2018), pp. 1–21.

[Dud73] R. M. Dudley. “Sample Functions of the Gaussian Process”. *The Annals of Probability* 1.1 (Feb. 1973), pp. 66–103.

[Dur19] R. Durrett. *Probability: Theory and Examples*. 5th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2019.

[El 78] N. El Karoui. “Sur les montées des semi-martingales”. In: *Temps locaux*. Astérisque 52-53. Sociétémathématique de France, 1978.

[FHW94] A. Feuerverger, P. Hall, and A. T. A. Wood. “Estimation Of Fractal Index And Fractal Dimension Of A Gaussian Process By Counting The Number Of Level Crossings”. *Journal of Time Series Analysis* 15.6 (Nov. 1994), pp. 587–606.

[FT83] B. Fristedt and S. J. Taylor. “Constructions of local time for a Markov process”. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 62.1 (1983), pp. 73–112.

[GH80] D. Geman and J. Horowitz. “Occupation Densities”. *The Annals of Probability* 8.1 (Feb. 1980), pp. 1–67.

[IM74] K. Itô and H. P. McKean Jr. *Diffusion processes and their sample paths*. Die Grundlehren der mathematischen Wissenschaften, Band 125. Second printing, corrected. Springer-Verlag, Berlin-New York, 1974, pp. xv+321.

[Kac43] M. Kac. “On the average number of real roots of a random algebraic equation”. *Bulletin of the American Mathematical Society* 49.4 (Apr. 1943), pp. 314–320.

[Kim22] D. Kim. “Local Times for Continuous Paths of Arbitrary Regularity”. *Journal of Theoretical Probability* 35.4 (2022), pp. 2540–2568.

[Kra06] M. F. Kratz. “Level crossings and other level functionals of stationary Gaussian processes”. *Probability Surveys* 3.none (Jan. 2006), pp. 230–288.

[Kru98] J. Krug. “Persistence of non-Markovian processes related to fractional Brownian motion”. In: vol. 4. 4. I Brazilian School in Probability (Rio de Janeiro, 1997). 1998, pp. 509–516.

[Lem83] M. Lemieux. “On the quadratic variation of semi-martingales”. MA thesis. University of British Columbia, 1983.

[Lév40] P. Lévy. “Le mouvement brownien plan”. *American Journal of Mathematics* 62 (1940), pp. 487–550.

[Lév48] P. Lévy. *Processus stochastiques et mouvement brownien*. Gauthier-Villars & Cie, Paris, 1948, pp. vi+438.

[Łoc+21] R. M. Łochowski, J. Obłój, D. J. Prömel, and P. Siorpaes. “Local times and Tanaka–Meyer formulae for càdlàg paths”. *Electronic Journal of Probability* 26.none (2021), pp. 1–29.

[MP10] P. Mörters and Y. Peres. *Brownian Motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2010.

[Pic08] J. Picard. “A tree approach to p-variation and to integration”. *Ann. Probab.* 36.6 (2008), pp. 2235–2279.

[Pic11] J. Picard. “Representation Formulae for the Fractional Brownian Motion”. In: *Séminaire de Probabilités XLIII*. Ed. by C. Donati-Martin, A. Lejay, and A. Rouault. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 3–70.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Third. Vol. 293. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.

[Ric45] S. O. Rice. “Mathematical analysis of random noise”. *The Bell System Technical Journal* 24.1 (1945), pp. 46–156.

Chapter 3

Young and rough differential equations driven by fractional Brownian motion

We consider Young or rough differential equations $\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t^H$ driven by fractional Brownian motion B^H with Hurst parameter $H \in (\frac{1}{3}, 1)$. When $H \neq \frac{1}{2}$, the equation is usually treated either by Young's theory or by Lyons' rough path theory, whereby we assume that $\sigma \in C^\gamma$ with $\gamma \geq \frac{1}{H}$. However, for $H = \frac{1}{2}$, Itô's theory covers the case where σ is Lipschitz. The aim of this chapter is to fill the gap, by proving that for any $H \in (\frac{1}{3}, 1)$ pathwise uniqueness holds for $\sigma \in C^\gamma$ with $\gamma > \max\{\frac{1}{2H}, \frac{1-H}{H}\}$, under a natural elliptic condition on σ. The result relies on new probabilistic estimates on stochastic integrals along fractional Brownian motions, proven by stochastic sewing techniques.

This chapter is based on joint work with Avi Mayorcas.

Keywords and phrases. Stochastic differential equations, fractional Brownian motion, regularization by noise, stochastic sewing, processes of vanishing mean oscillation, rough paths.
MSC 2020 – 60H10, 60H50, 60G22, 60L20.

Contents

3.1 Introduction **96**

3.2 Preliminaries **101**

3.2.1 Function spaces and heat semigroup estimates 102

3.2.2 SDE driven by fractional Brownian motion 104

3.2.3 Stochastic estimates 107

3.3 Young case **109**

3.3.1 Young integral with irregular integrand 110

3.3.2 Pathwise uniqueness 116

3.4 Rough case **121**

3.4.1 Rough integral with irregular integrand 123

3.4.2 Iterated integrals 130

3.4.3 Pathwise uniqueness 134

Bibliography **139**

3.1 Introduction

In this chapter we consider the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t^H, \quad X_0 = x \in \mathbb{R}^{d_1}, \tag{3.1}$$

where σ is a map valued in the space of $d_1 \times d_2$ matrices, and B^H is a d_2-dimensional fractional Brownian motion with Hurst parameter $H \in (1/3, 1)$. The differential equation is interpreted either as Young's differential equation ($H > 1/2$) or as Lyons' rough differential equation [Lyo98] ($H < 1/2$). We will review such SDEs in Section 3.2.2.

For $H = 1/2$, we often apply Itô's theory to study (3.1); we will discuss the alternative theory of Lyons later. In Itô's theory, there are a few notions of solutions and their uniqueness, among which the most relevant to us is the notion of *pathwise uniqueness*. It says that two solutions, adapted to some filtration making the driving Brownian motion martingale, must be indistinguishable. Hence, pathwise uniqueness is a *probabilistic* concept of uniqueness (despite it is called "pathwise" uniqueness). It is a classical result, as can be found in all textbooks of stochastic calculus, that pathwise uniqueness holds for (3.1) with $H = 1/2$

provided that σ is Lipschitz. The proof is a consequence of Itô's isometry: for an adapted process Y we have

$$\mathbb{E}\Big[\Big|\int_0^T Y_r \mathrm{d}B_r^{1/2}\Big|^2\Big] = \mathbb{E}\Big[\int_0^T |Y_r|^2 \mathrm{d}r\Big].$$

Itô's isometry is a consequence of the martingale property of the Brownian motion. Since B^H, $H \neq 1/2$, is not a martingale (nor Markovian), Itô's theory is not available for $H \neq 1/2$. Lack of probabilistic tools naturally motivates us to study the SDE (3.1) *pathwisely*. Starting from Young's integration theory [You36], Lyons [Lyo94] showed that the differential equation

$$\mathrm{d}x_t = f(x_t)\mathrm{d}y_t \tag{3.2}$$

driven by a path y of finite p-variation with $p \in [1, 2)$ has a unique solution provided that f is α-Hölder with $\alpha > p$. Since B^H has finite p-variation for any $p > 1/H$, we see that (3.1) has a unique solution provided that $\sigma \in C^\gamma$ with $\gamma > 1/H$. The reader can also refer to [NR02] or [Nua06, Section 5.3], where the approach is based on fractional stochastic calculus of Zähle [Zäh98]. The critical case $\gamma = 1/H$ is covered by [Dav08].

When $H \leq 1/2$, the path B^H is too irregular to apply Young's integration theory. The earliest work towards pathwise Itô calculus can be dated back to Föllmer [Föl81]. Afterwards, in his groundbreaking work [Lyo98], Lyons develops the theory of *rough paths*. The theory tells that we can make sense of (3.2) for y of finite p-variation for *any* $p < \infty$, provided that we are additionally given "iteraged integrals"

$$\int_s^t \int_s^{r_1} \mathrm{d}y_{r_2}\mathrm{d}y_{r_1}, \ \int_s^t \int_s^{r_1} \int_s^{r_2} \mathrm{d}y_{r_3}\mathrm{d}y_{r_2}\mathrm{d}y_{r_1}, \ldots,$$

satisfying certain analytic and algebraic conditions. The tuple of y and its (sufficient number of) iterated integrals is called a rough path of y. Furthermore, [Lyo98] proved well-posedness of (3.2) if $f \in C^\alpha$ with $\alpha > p$. Later, Coutin and Qian [CQ02] proved that the fractional Brownian motion B^H, with $H > 1/4$, can be naturally lifted to a rough path. This means that we can make sense of and prove well-posedness of (3.1) with $H \in (1/4, 1/2)$, provided that $\sigma \in C^\gamma$ with $\gamma > 1/H$.

We remark two important differences in Itô's probabilistic theory and Lyons' pathwise theory. One is that the former considers uniqueness among solutions adapted to a given filtration, while the latter considers that among all solutions satisfying (3.2) that do not need to be adapted. That is, notion of uniqueness is stronger in the pathwise theory. Such uniqueness is called *path-by-path uniqueness*, after the works of Davie [Dav07; Dav08]. The other difference is the regularity assumption on σ: when $H = 1/2$, Itô's theory assume that σ is only Lipschitz, while Lyons' theory assume that $\sigma \in C^\gamma$ with $\gamma \geq 2$ (the critical case $\gamma = 2$ is covered by [Dav08]). In summary, Itô's theory requires less regularity assumption on σ, at the cost of weaker notion of uniqueness.

Although Itô's theory is not available for $H \neq 1/2$, notion of pathwise uniqueness is well-defined in this setting (Definition 3.2.6). Now we find it natural to ask if we can prove pathwise uniqueness of (3.1) for $\sigma \in C^\gamma$ with $\gamma < 1/H$. Our main result of this chapter answers the question affirmatively.

Theorem 3.1.1. *Let $H \in (1/3, 1)$ and $\sigma \in C^\gamma$ with*

$$\gamma > \max\left\{\frac{1}{2H}, \frac{1-H}{H}\right\} = \begin{cases} \frac{1}{2H}, & \text{if } H > \frac{1}{2}, \\ \frac{1-H}{H}, & \text{if } H \leq \frac{1}{2}. \end{cases} \tag{3.3}$$

Furthermore, suppose that $\sigma\sigma^{\mathrm{T}}$ is uniformly elliptic, i.e. there exists a $K \in (1, \infty)$ such that

$$K^{-1} \leq \sigma(x)\sigma(x)^{\mathrm{T}} \leq K$$

for all $x \in \mathbb{R}^{d_1}$. Then pathwise uniqueness holds for (3.1).

Proof. See Theorem 3.3.3 for $H > 1/2$ and Theorem 3.4.7 for $H < 1/2$. (The case $H = 1/2$ is well known.) □

Remark 3.1.2. It is expected that a similar result holds for $H \in (1/4, 1/3]$. However, studying this case will surely increase the amount of technical computation, and in order to present our main ideas clearly, we restrict to the case $H > 1/3$.

Remark 3.1.3. By standard compactness argument, we can prove weak existence under $\gamma > \frac{1-H}{H}$. By the Yamada–Watanabe theorem (Proposition 3.2.7 below), weak uniqueness and strong existence hence hold under the assumptions of Theorem 3.1.1.

Before explaining the strategy of the proof, let us further explain connections of our result to recent literatures. After Lyons' groundbreaking works, pathwise approach has become central in stochastic analysis. Precisely, pathwise approach here means the approach that separates almost completely probabilistic argument (e.g. lifting to a rough path) and pathwise argument (e.g. analysis of differential equations driven by rough paths). It is worth mentioning that this pathwise approach is behind the breakthrough of singular SPDEs [Hai14; GIP15].

With flavor of pathwise approach, Catellier and Gubinelli [CG16] study fractional SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t^H \tag{3.4}$$

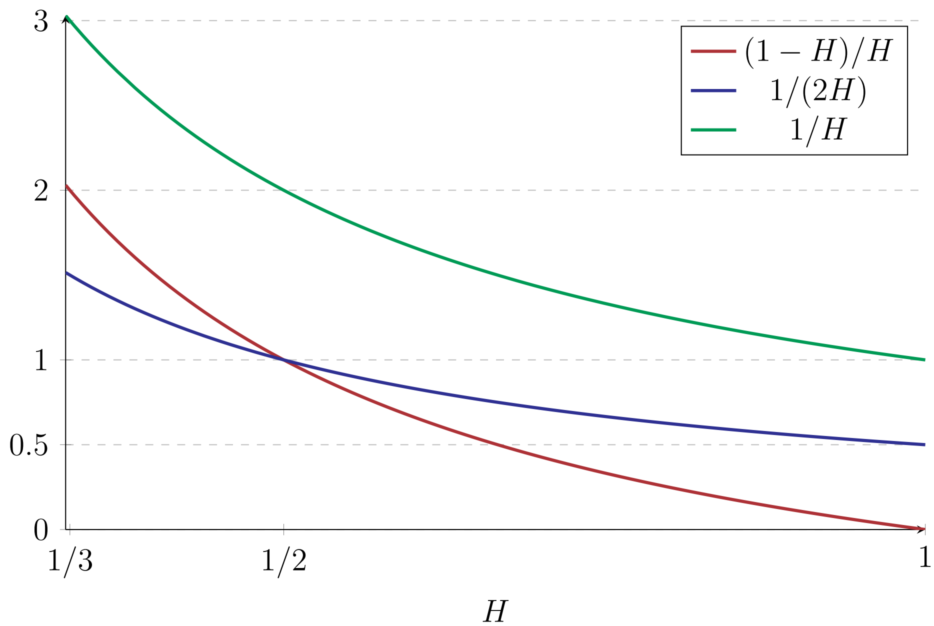


Figure 3.1: Some graphs of H from Theorem 3.1.1. Pathwise theory covers $\sigma \in C^\gamma$ with $\gamma \geq 1/H$ (green), while our result says that pathwise uniqueness holds if $\gamma > 1/(2H)$ (blue) and if $\gamma > (1-H)/H$ (red).

with irregular b, which can even be a distribution such as Dirac's delta function. Study of such SDEs belongs to the field called *regularization by noise*; the name comes from the fact that we can restore certain well-posedness of ill-posed ordinary differential equations by adding a noise. The argument of Catellier & Gubinelli is based on the nonlinear Young integration (cf. [Gal21]), and their work led to further developments, [GG22; HP21; GHM22; GH22; CD22] to name just a few.

On the other hand, Lê [Lê20] introduced a new method for regularization by fractional noise. Lê relates well-posedness of (3.4) to the stochastic integral

$$\int_s^t \nabla b(B_r^H) \mathrm{d}B_r^H,$$

which is then estimated by the *stochastic sewing lemma*, introduced in the same paper. (See [BNP20] for the approach based on Malliavin calculus.) The work [Lê20] is a landmark in that it is based on more probabilistic viewpoints than on pathwise viewpoints. However, it is worth noting that many ideas of [Lê20] are inspired by those developed in pathwise stochastic calculus; obviously the stochastic sewing lemma is inspired by Gubinelli's sewing lemma [Gub04]. Since the work [Lê20], the probabilistic approach based on the stochastic sewing has witnessed tremendous progress, see [BDG21; Ath+21; Ger22; FHL23; GG23; DG22; BLM23] and references therein.

Our main result is another contribution to the trend initiated by Lê. However, unlike the mentioned previous works, we are interested in the noise coefficient rather than the drift coefficient. We also mention the work [HTV22] of Hinz, Tölle and Viitasaari, where studied is the differential equation (3.2) with irregular f that can even be discontinuous. However, their argument is based on Doss transformation, which imposes some restrictions on the coefficient f, especially for uniqueness (e.g. [HTV22, Assumption 3.15]).

Strategy of the proof

Our argument is inspired by [Lê20]. Let us review here his strategy to prove the pathwise uniqueness. For simplicity, suppose that $H \in (1/2, 1)$. Let X and Y be two adapted solutions to (3.1). If $\sigma \in C^{1/H+\varepsilon}$,

$$\begin{aligned} X_t - Y_t &= \int_0^t \{\sigma(X_r) - \sigma(Y_r)\} \mathrm{d}B_r \\ &= \int_0^t \Big\{ \sum_{k=1}^{d_1} \int_0^1 (X_r^k - Y_r^k) \partial_k \sigma(\theta X_r + (1-\theta) Y_r) \mathrm{d}\theta \Big\} \mathrm{d}B_r \\ &= \sum_{k=1}^{d_1} \int_0^t (X_r^k - Y_r^k) \mathrm{d}V_r^k, \end{aligned} \tag{3.5}$$

where

$$V_t^k := \int_0^t \int_0^1 \partial_k \sigma(\theta X_r + (1-\theta) Y_r) \mathrm{d}\theta \mathrm{d}B_r. \tag{3.6}$$

Then, $Z := X - Y$ satisfies the linear Young SDE

$$Z_t = \int_0^t Z_r \mathrm{d}V_r,$$

by uniqueness of the Young SDE, we see $Z = 0$ or $X = Y$.

If σ is less regular, then the Young integration theory does not make sense of (3.6). However, we can still hope to give a natural meaning to (3.6) for such irregular σ, by taking advantage of randomness of the fractional Brownian motion. Indeed, our main ingredient is to give a probabilistic estimate on the integral

$$\int_s^t f(Y_r) \mathrm{d}B_r^H \tag{3.7}$$

for some irregular f, where Y is a path *controlled* by B (Definition 3.2.5). When $H = 1/2$, we can apply Itô's isometry to estimate (3.7). When $H \neq 1/2$, Itô's isometry is not available, and we replace it by the stochastic sewing estimate. For $H \in (1/2, 1)$, we apply Gerencser's shifted stochastic sewing [Ger22], and our choice of the germs resemble the one from [DG22]. For $H \in (1/3, 1/2)$, we apply the stochastic sewing *twice*; first by Lê's original stochastic sewing, and second by the fully shifted stochastic sewing (Theorem 1.1.1). See Theorem 3.3.2 ($H > 1/2$) and Theorem 3.4.4 ($H < 1/2$) for the precise estimates.

Notation

Throughout the chapter, we fix the dimensions d_1 (the dimension for the solution) and d_2 (the dimension for the fractional Brownian motion). Given a function f on some interval $[0, T]$, we write $f_{s,t} := f_t - f_s$ ($s < t$). Given a two parameter map $(A_{s,t})_{s<t}$, we set

$$\delta A_{s,u,t} := A_{s,t} - A_{s,u} - A_{u,t}, \quad s < u < t.$$

We denote by a^{T} the transpose of the matrix a. We write $A \lesssim_{\alpha,\beta,\dots} B$ if there exists a constant C, depending on irrelevant parameters $\alpha, \beta, \dots$ such that $A \leq CB$. We will ignore dependence of constants on d_1, d_2 and the Hurst parameter H.

3.2 Preliminaries

Here we review some estimates in function spaces and SDEs driven by fractional Brownian motion.

3.2.1 Function spaces and heat semigroup estimates

We follow the convention of [DGL21]. For $k \in \mathbb{N}$, we denote by C^k the space of $(k-1)$-times differentiable functions such that their $(k-1)$-derivatives are Lipschitz continuous, and we write

$$\|f\|_{C^k} := \sum_{l \in \mathbb{N}_0^{d}, |l| < k} \|\partial_l f\|_{L^\infty} + \sum_{l \in \mathbb{N}_0^{d_1}, |l| = k} \sup_{x,y \in \mathbb{R}^{d_1}} \frac{|\partial^l f(x) - \partial^l f(y)|}{|x-y|}.$$

For $\gamma \in (0,\infty) \setminus \mathbb{N}$, we denote by C^γ the space of γ-Hölder functions. That is, $f \in C^\gamma$ if

$$\|f\|_{C^\gamma} := \sum_{l \in \mathbb{N}_0^{d_1}, |l| < \lfloor \gamma \rfloor} \|\partial_l f\|_{L^\infty} + \sum_{l \in \mathbb{N}_0^{d_1}, |l| = \lfloor \gamma \rfloor} \sup_{x,y \in \mathbb{R}^{d}} \frac{|\partial^l f(x) - \partial^l f(y)|}{|x-y|^{\gamma - \lfloor \gamma \rfloor}}.$$

We set $C^0 := L^\infty$ and for $\alpha < 0$ we denote by C^γ the space of distributions f such that

$$\|f\|_{C^\gamma} := \sup_{t \in (0,1)} t^{-\gamma/2} \|\mathcal{P}_t f\|_{L^\infty},$$

where $\mathcal{P}_t$ is the heat semigroup (3.8) defined just below. If $\gamma \in \mathbb{R} \setminus \mathbb{N}_0$ the space C^γ coincides with the Hölder–Besov space constructed by Littlewood-Paley blocks [Tri92, Theorem 2.6.4].

We prepare two lemmas on heat semigroups. We denote by $\mathcal{M}$ the space of $d_1 \times d_1$ matrices and by $\mathcal{M}_+$ the space of $d_1 \times d_1$ positive-definite matrices. By identifying $\mathcal{M}$ with $\mathbb{R}^{(d_1)^2}$, we can equip $\mathcal{M}$ with the Euclidean norm $|\cdot|$. Given a $\Gamma \in \mathcal{M}_+$ we set

$$p_\Gamma(x) := \frac{1}{(2\pi)^{d/2} (\det \Gamma)^{1/2}} \exp\Big(-\frac{\langle x, \Gamma^{-1} x\rangle}{2}\Big)$$

and for a function f on $\mathbb{R}^{d_1}$ we set $\mathcal{P}_\Gamma f := p_\Gamma * f$. For $t \in (0,\infty)$ we simply write

$$\mathcal{P}_t := \mathcal{P}_{t I_{d_1}}, \tag{3.8}$$

where I_{d_1} is the $d_1 \times d_1$ identity matrix.

Lemma 3.2.1. *Let $K \in (1,\infty)$ and $\Gamma \in \mathcal{M}_+$ such that $K^{-1} \le \Gamma \le K$. Then, for $\alpha, \beta \in \mathbb{R}$ with $\beta \ge 0 \vee \alpha$ and $t \in (0,T)$ we have*

$$\|\mathcal{P}_{\Gamma t} f\|_{C^\beta} \lesssim_{K,\alpha,\beta,T} t^{-\frac{\beta-\alpha}{2}} \|f\|_{C^\alpha}.$$

Proof. We write
$$\Lambda_\Gamma f(x) := f(\sqrt{\Gamma}x). \tag{3.9}$$
Then $\mathcal{P}_{\Gamma t} f = \Lambda_{\Gamma^{-1}} \mathcal{P}_t \Lambda_\Gamma f$. Since
$$\|\Lambda_\Gamma f\|_{C^\gamma} \lesssim_K \|f\|_{C^\gamma} \lesssim_K \|\Lambda_\Gamma f\|_{C^\gamma},$$
we can assume that $\Gamma = I_d$. Then the claim follows from [DGL21, Lemma 2.3]. □

Lemma 3.2.2. *Let $K \in (1,\infty)$ and let $\Gamma_i \in \mathcal{M}_+$ $(i = 1,2)$ be such that $K^{-1} \le \Gamma_i \le K$. Then, for $\alpha, \beta \in \mathbb{R}$ with $\beta \ge \alpha \vee 0$ and $t \in (0,T)$ we have*
$$\|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f\|_{C^\beta} \lesssim_{K,\alpha,\beta,T} t^{-\frac{\beta-\alpha}{2}} |\Gamma_1 - \Gamma_2| \|f\|_{C^\alpha}. \tag{3.10}$$

Proof. We first claim
$$\|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f\|_{L^\infty} \lesssim_K |\Gamma_1 - \Gamma_2| \|f\|_{L^\infty}. \tag{3.11}$$
Indeed, we have
$$(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f(x) = \int_{\mathbb{R}^{d_1}} \{p_{\Gamma_1 t}(y) - p_{\Gamma_2 t}(y)\} f(x-y) \mathrm{d}y.$$
By [DGL21, Proposition 2.7],
$$|p_{\Gamma_1 t}(y) - p_{\Gamma_2 t}(y)| \lesssim_K |\Gamma_1 - \Gamma_2| (p_{\Gamma_1 t/2}(y) + p_{\Gamma_2 t/2}(y)), \tag{3.12}$$
which yields the estimate (3.11).

Next we prove the estimate (3.10) for $\alpha = \beta \notin \mathbb{N}$. The case for $\alpha = \beta = 0$ is proved by (3.11). Let $\{\Delta_j\}_{j=-1}^\infty$ be Littlewood–Paley blocks. Since Δ_j and $\mathcal{P}_{\Gamma_i t}$ commute, the estimate (3.11) yields
$$\|\Delta_j (\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f\|_{L^\infty} \lesssim_K |\Gamma_1 - \Gamma_2| (\|\Delta_j f\|_{L^\infty} + \|\Delta_j f\|_{L^\infty}).$$
Therefore,
$$\|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f\|_{C^\beta} \lesssim_K |\Gamma_1 - \Gamma_2| (\|f\|_{C^\beta} + \|f\|_{C^\beta}).$$

Next we prove the estimate (3.10) for $\alpha = \beta \in \mathbb{N}$. Since we only need the case $\alpha = \beta = 1$, we only prove that case. We have
$$\begin{aligned} &|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f(x) - (\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t}) f(y)| \\ &\qquad \le \int_{\mathbb{R}^{d_1}} |p_{\Gamma_1 t}(z) - p_{\Gamma_2 t}(z)| |f(x-z) - f(y-z)| \mathrm{d}z. \end{aligned}$$

By (3.12),

$$
\begin{aligned}
&|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t})f(x) - (\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t})f(y)| \\
&\qquad \lesssim_K |\Gamma_1 - \Gamma_2| \int_{\mathbb{R}^{d_1}} |p_{\Gamma_1 t/2}(z) + p_{\Gamma_2 t/2}(z)||f(x-z) - f(y-z)|\mathrm{d}z \\
&\qquad \leq 2|\Gamma_1 - \Gamma_2| \|f\|_{C^1}|x-y|,
\end{aligned}
$$

which prove the claim.

Finally, we prove the general case. Since

$$
\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t} = (\mathcal{P}_{\Gamma_1 t/2} - \mathcal{P}_{\Gamma_2 t/2})\mathcal{P}_{\Gamma_1 t/2} + \mathcal{P}_{\Gamma_2 t/2}(\mathcal{P}_{\Gamma_1 t/2} - \mathcal{P}_{\Gamma_2 t/2}),
$$

we have

$$
\begin{aligned}
&\|(\mathcal{P}_{\Gamma_1 t} - \mathcal{P}_{\Gamma_2 t})f\|_{C^\beta} \\
&\qquad\qquad \leq \|(\mathcal{P}_{\Gamma_1 t/2} - \mathcal{P}_{\Gamma_2 t/2})\mathcal{P}_{\Gamma_1 t/2}f\|_{C^\beta} + \|\mathcal{P}_{\Gamma_2 t/2}(\mathcal{P}_{\Gamma_1 t/2} - \mathcal{P}_{\Gamma_2 t/2})f\|_{C^\beta}.
\end{aligned}
$$

As for the first term, the estimate (3.10) with $\alpha = \beta$ and Lemma 3.2.1 imply

$$
\begin{aligned}
\|(\mathcal{P}_{\Gamma_1 t/2} - \mathcal{P}_{\Gamma_2 t/2})\mathcal{P}_{\Gamma_1 t/2}f\|_{C^\beta} &\lesssim_K |\Gamma_1 - \Gamma_2| \|\mathcal{P}_{\Gamma_1 t/2}f\|_{C^\beta} \\
&\lesssim_{K,\alpha,\beta,T} |\Gamma_1 - \Gamma_2| t^{-\frac{\beta-\alpha}{2}} \|f\|_{C^\alpha}.
\end{aligned}
$$

The estimate of the second term is similar. □

3.2.2 SDE driven by fractional Brownian motion

The goal of this section is to review the notion of solutions to (3.1). First we review the fractional Brownian motion. We define the kernel K_H by for $H > 1/2$

$$
K_H(t,s) := c_H s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{3}{2}} u^{H-\frac{1}{2}} \mathrm{d}u
$$

and for $H < 1/2$

$$
K_H(t,s) := c_H \Big[\Big(\frac{t}{s}\Big)^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}} - \Big(H - \frac{1}{2}\Big) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} \mathrm{d}u \Big].
$$

Then given a Brownian motion W the process

$$
K_H W(t) := \int_0^t K_H(t,s) \mathrm{d}W_s
$$

is a fractional Brownian motion with Hurst parameter H (recall Definition 1.3.1). The constant c_H is chosen to have the correct covariance. We refer the reader to [Nua06, Section 5.1.3] for more details. In fact, the operator K_H is bijective, hence given a fractional Brownian motion B^H we can recover a Brownian motion $K_H^{-1}B^H$. Recall that a Brownian motion W is called a $(\mathcal{F}_t)$-Brownian motion if it is a $(\mathcal{F}_t)$-martingale.

Definition 3.2.3. Given a complete filtered space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, an adapted process $(B_t^H)_{t\geq 0}$ is called a $(\mathcal{F}_t)$-*fractional Brownian motion* if the process $K_H^{-1}B^H$ is a $(\mathcal{F}_t)$-Brownian motion.

When $H < 1/2$, we need the theory of rough paths to interpret the SDE (3.1). Let us recall the rough path of B^H for $H < 1/2$, and the notion of controlled paths.

Definition 3.2.4 ([FH20, Chapter 10])**.** Let $H \in (1/3, 1/2)$ and $(B^{(n)})_{n=1}^{\infty}$ be a piecewise linear approximation of the fractional Brownian motion $B = B^H$. Then the canonical lift of $B^{(n)}$

$$(B^{(n)}, \mathbb{B}^{(n)}) := \Big(B^{(n)}, (\int_s^t B_{s,r}^{(n)} \otimes \mathrm{d}B_r^{(n)})_{s<t}\Big)$$

converges in the space of α-Hölder rough paths for any $\alpha \in (1/3, H)$ to some geometric rough path $(B, \mathbb{B})$, called the *canonical lift of the fractional Brownian motion* B.

Definition 3.2.5 ([FH20, Chapter 4])**.** Let $H \in (1/3, 1/2)$, $\alpha \in (1/3, H)$ and $\beta \in (1-\alpha, 1)$. A pair (X, X') of continuous paths is said to be *controlled by* B^H, and we write $(X, X') \in \mathscr{D}^\beta$, if

$$\|X, X'\|_{\mathscr{D}^\beta} := \|X'\|_{C^{\beta-\alpha}([0,T])} + \sup_{0\leq s<t\leq T} \frac{|X_{s,t} - X'_s B^H_{s,t}|}{|t-s|^\beta} < \infty. \tag{3.13}$$

The process X' is called a Gubinelli derivative. When no confusion is expected, we simply write $\|X\|_{\mathscr{D}^\beta}$. It was shown in [Gub04] and is now standard that for $(X, X') \in \mathscr{D}^\beta$ we can define the rough integral

$$\int_s^t X_r \mathrm{d}B_r^H := \lim_{\substack{\pi \text{ is a partition of } [s,t],\\ |\pi|\to 0}} \sum_{[u,v]\in\pi} X_u B^H_{u,v} + X'_u \mathbb{B}^H_{u,v}.$$

We assume that the coefficient σ of the SDE (3.1) belongs to the Hölder space C^γ with

$$\gamma > \frac{1-H}{H}. \tag{3.14}$$

The lower bound is necessary to make sense of the integral $\int \sigma(X_r)\mathrm{d}B_r^H$ as Young integral ($H > 1/2$) or rough integral ($H < 1/2$). If $H < 1/2$, we have $\gamma > 1$ and

$$|\sigma(X_t) - \sigma(X_s) - D\sigma(X_s)X_{s,t}| \leq \|\sigma\|_{C^\gamma}|X_{s,t}|^\gamma.$$

Therefore, provided $(X, X') \in \mathscr{D}^{\gamma\alpha}$, where $\alpha < H$ is sufficiently close to H, we have $(\sigma(X), \sigma'(X)X') \in \mathscr{D}^{\gamma\alpha}$. Thus, for any $H \in (1/3, 1)$, by writing $\boldsymbol{e}_i$ for the ith unit vector, the integral

$$\int_s^t \sigma(X_r)\mathrm{d}B_r^H := \sum_{\substack{1\le i\le d_1,\\ 1\le j\le d_2}} \Big(\int_s^t \sigma^{ij}(X_r)\mathrm{d}B_r^{H,j}\Big)\boldsymbol{e}_i$$

is well-defined either as Young integral ($H > 1/2$) or as rough integral ($H < 1/2$).

For $H > 1/2$ we say that a path X solves (3.1) if it satisfies the integral equation

$$X_t = x + \int_0^t \sigma(X_r)\mathrm{d}B_r^H, \quad \forall t. \tag{3.15}$$

Similarly, for $H < 1/2$ we say that a controlled path (X, X') solves (3.1) if it satisfies (3.15). In this case, the fundamental estimate of the rough integral yields

$$\Big|\int_s^t \sigma(X_r)\mathrm{d}B_r^H - \sigma(X_s)B_{s,t}^H - D\sigma(X_s)X_s'\mathbb{B}_{s,t}^H\Big| \\ \lesssim_H \|\sigma\|_{C^\gamma}\|X\|_{\mathscr{D}^{\gamma\alpha}}\|B^H\|_{C^\alpha}|t-s|^{(\gamma+1)\alpha},$$

hence $(X, \sigma(X)) \in \mathscr{D}^{\gamma\alpha}$. In particular, due to the uniqueness of the Gubinelli derivative for the fractional Brownian rough path [FH20, Chapter 6], we have $\sigma(X) = X'$. Therefore, without loss of generality, we say that X solves (3.1) if $(X, \sigma(X)) \in \mathcal{D}^{\gamma\alpha}$ and the pair solves (3.1).

We review the notion of solutions in the probabilistic setting. We fix α that is less than but sufficiently close to H.

Definition 3.2.6. Let $H > 1/2$. We say that a quintuple $(\Omega, (\mathcal{F}_t)_{t\ge0}, \mathbb{P}, B^H, X)$ is a *weak solution* to (3.1) if (B^H, X) are random variables defined on the filtered probability space $(\Omega, (\mathcal{F}_t), \mathbb{R})$, if B^H is a $(\mathcal{F}_t)$-fractional Brownian motion, if $X \in C^\alpha([0,T])$ is adapted to $(\mathcal{F}_t)$ and if X solves the Young differential equation (3.1). Given a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ and a $(\mathcal{F}_t)$-fractional Brownian motion B^H, we say that a $C^\alpha([0,T])$-valued random variable X defined on $(\Omega, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ is a *strong solution* if it solves (3.1) and if it is adapted to the natural filtration generated by B. We say that the *pathwise uniqueness* holds for (3.1) if, for any two adapted $C^\alpha([0,T])$-valued random process X and Y defined on a common filtered probability space that solve (3.1) driven by a common $(\mathcal{F}_t)$-fractional Brownian motion, we have $X = Y$ almost surely.

For $H < 1/2$, such notion of solutions naturally extend: we replace X and Y by controlled paths in $\mathscr{D}^{\gamma\alpha}$ and (3.1) is interpreted as rough differential equation.

As in the SDE driven by Brownian motion, we have the following Yamada–Watanabe theorem.

Proposition 3.2.7. *Regarding the SDE (3.1), if weak existence and pathwise uniqueness hold, then weak uniqueness and strong existence hold as well.*

Proof. This follows from a generalized Yamada-Watanabe theorem of Kurtz [Kur07]. □

A standard compactness argument allows us to construct a weak solution to (3.1) for $\sigma \in C^\gamma$ with $\gamma > (1-H)/H$. Hence, weak uniqueness and strong existence follow under the assumptions of Theorem 3.1.1.

3.2.3 Stochastic estimates

Here we provide technical estimates on integrals involving Gaussian processes. Skipping their proofs will not affect further reading.

Definition 3.2.8. We define the *Riemann–Liouville process* by

$$\mathcal{B}_t^H := \int_0^t (t-r)^{H-\frac{1}{2}} \mathrm{d}W_r.$$

In computations later, we use the Mandelbrot–van Ness representation [MV68]: for $t \geq 0$ we have

$$B_t^H = \int_{-\infty}^t \mathcal{K}_H(t,r)\mathrm{d}\tilde{W}_r, \quad \mathcal{K}_H(t,r) := (t-r)^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}}, \tag{3.16}$$

where $(\tilde{W}_t)_{t\in\mathbb{R}}$ is a two-sided Brownian motion. We note that the process

$$(B_{v+t}^H - \mathbb{E}[B_{v+t}^H | \sigma(\tilde{W}_r : r \leq v)])_{t\geq 0}$$

has the law of $\mathcal{B}$. Due to this relation, we need some computations involving $\mathcal{B}^H$.

As in B^H, throughout the chapter the process $\mathcal{B}^H$ takes values in $\mathbb{R}^{d_2}$. We remark that for $H \in (1/3, 1/2)$ the process $\mathcal{B}^H$ can be lifted to the second order geometric rough path (e.g. [FH20, Chap. 10]). Below, integrals with respect to $\mathcal{B}^H$ are understood as rough integral.

Lemma 3.2.9. *Let $H \in (1/3, 1)$, $g \in C^1(\mathbb{R}^{d_1}, \mathbb{R})$ and a be a $d_1 \times d_2$ matrix. We have the identity*

$$\mathbb{E}\Big[\int_s^t g(a\mathcal{B}_r^H)\mathrm{d}\mathcal{B}_r^{H,i}\Big] = \frac{1}{2}\sum_{j=1}^{d_1} a^{ji} \int_s^t r^{2H-1} \partial_j \mathcal{P}_{\frac{r^{2H}}{2H} aa^{\mathrm{T}}} g(0)\mathrm{d}r.$$

Proof. We will drop superscripts of H. We first note that, due to Gaussianity,

$$g(a\mathcal{B}_r) - c_r\mathcal{B}^i_r, \quad c_r := \mathbb{E}[(\mathcal{B}^i_r)^2]^{-1}\mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r]$$

is orthogonal to $\mathcal{B}^i$ in $L^2(\mathbb{P})$. Therefore,

$$\begin{aligned}\mathbb{E}\Big[\int_s^t g(a\mathcal{B}_r)\mathrm{d}\mathcal{B}^i_r\Big] &= \mathbb{E}\Big[\int_s^t c_r\mathcal{B}^i_r\mathrm{d}\mathcal{B}^i_r\Big] \\ &= \frac{1}{2}\mathbb{E}\Big[\int_s^t c_r\mathrm{d}(\mathcal{B}^i_r)^2\Big] = \frac{1}{2}\int_s^t c_r\mathrm{d}\mathbb{E}[(\mathcal{B}^i_r)^2]\end{aligned}$$

and we obtain

$$\mathbb{E}\Big[\int_s^t g(a\mathcal{B}_r)\mathrm{d}\mathcal{B}^i_r\Big] = \frac{1}{2}\int_s^t \mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r]\frac{\mathrm{d}}{\mathrm{d}r}\log\mathbb{E}[(\mathcal{B}^i_r)^2]\mathrm{d}r.$$

Since $\mathbb{E}[(\mathcal{B}^i_r)^2] = \frac{1}{2H}r^{2H}$, we obtain

$$\frac{1}{2}\int_s^t \mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r]\frac{\mathrm{d}}{\mathrm{d}r}\log\mathbb{E}[(\mathcal{B}^i_r)^2]\mathrm{d}r = H\int_s^t r^{-1}\mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r]\mathrm{d}r$$

or

$$\mathbb{E}\Big[\int_s^t g(a\mathcal{B}_r)\mathrm{d}\mathcal{B}^i_r\Big] = H\int_s^t r^{-1}\mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r]\mathrm{d}r.$$

To compute further, let N be the standard normal distribution on $\mathbb{R}^d$, so that

$$\mathbb{E}[g(a\mathcal{B}_r)\mathcal{B}^i_r] = \frac{r^H}{\sqrt{2H}}\mathbb{E}\Big[g\Big(\frac{r^H}{\sqrt{2H}}aN\Big)N^i\Big].$$

The Gaussian integration by parts yields

$$\mathbb{E}\Big[g\Big(\frac{r^H}{\sqrt{2H}}aN\Big)N^i\Big] = \frac{r^H}{\sqrt{2H}}\sum_{j=1}^{d_1}a^{ji}\mathbb{E}\Big[\partial_j g\Big(\frac{r^H}{\sqrt{2H}}aN\Big)\Big]$$

Therefore,

$$\begin{aligned}\mathbb{E}\Big[\int_s^t g(a\mathcal{B}_r)\mathrm{d}\mathcal{B}^i_r\Big] =& \frac{1}{2}\sum_{j=1}^{d_1}a^{ji}\int_s^t r^{2H-1}\mathbb{E}\Big[\partial_j g\Big(\frac{r^H}{\sqrt{2H}}aN\Big)\Big]\mathrm{d}r \\ =& \frac{1}{2}\sum_{j=1}^{d_1}a^{ji}\int_s^t r^{2H-1}\partial_j\mathcal{P}_{\frac{r^{2H}}{2H}aa^{\mathrm{T}}}g(0)\mathrm{d}r.\end{aligned}$$

□

Lemma 3.2.10. *Let* $g\colon \mathbb{R}^{d_2} \to \mathbb{R}$ *be such that* $g(\mathcal{B}_s) \in L^2(\mathbb{P})$*. Then we have*

$$\mathbb{E}[g(\mathcal{B}^H_s) \int_s^t \mathcal{B}^{H,i}_{s,r} \mathrm{d}\mathcal{B}^{H,j}_r] = \frac{1}{2}\mathbb{E}[g(\mathcal{B}^H_s)\mathcal{B}^{H,i}_{s,t}\mathcal{B}^{H,j}_{s,t}].$$

Proof. We drop scripts on H. If $i = j$, the identity is trivial, since the rough path of $\mathcal{B}$ is geometric. We therefore assume below that $i \neq j$. Since $i \neq j$, the random variable

$$\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r$$

belongs to the second order Wiener chaos. The Wiener chaos expansion yields

$$\mathbb{E}\Big[g(\mathcal{B}_s)\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r\Big] = \mathbb{E}[(\mathcal{B}^i_s\mathcal{B}^j_s)^2]^{-1}\mathbb{E}[g(\mathcal{B}_s)\mathcal{B}^i_s\mathcal{B}^j_s]\mathbb{E}\Big[\mathcal{B}^i_s\mathcal{B}^j_s\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r\Big].$$

Since $(\mathcal{B}^i, \mathcal{B}^j) \stackrel{\mathrm{d}}{=} (\mathcal{B}^j, \mathcal{B}^i)$, we have

$$\mathbb{E}\Big[\mathcal{B}^i_s\mathcal{B}^j_s\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r\Big] = \mathbb{E}\Big[\mathcal{B}^i_s\mathcal{B}^j_s\int_s^t \mathcal{B}^j_{s,r}\mathrm{d}\mathcal{B}^i_r\Big]$$

and

$$\begin{aligned}\mathbb{E}\Big[\mathcal{B}^i_s\mathcal{B}^j_s\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r\Big] &= \frac{1}{2}\mathbb{E}\Big[\mathcal{B}^i_s\mathcal{B}^j_s\Big(\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r + \int_s^t \mathcal{B}^j_{s,r}\mathrm{d}\mathcal{B}^i_r\Big)\Big] \\ &= \frac{1}{2}\mathbb{E}[\mathcal{B}^i_s\mathcal{B}^j_s\mathcal{B}^i_{s,t}\mathcal{B}^j_{s,t}].\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}\Big[g(\mathcal{B}_s)\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r\Big] &= \frac{1}{2}\mathbb{E}[(\mathcal{B}^i_s\mathcal{B}^j_s)^2]^{-1}\mathbb{E}[g(\mathcal{B}_s)\mathcal{B}^i_s\mathcal{B}^j_s]\mathbb{E}[\mathcal{B}^i_s\mathcal{B}^j_s\mathcal{B}^i_{s,t}\mathcal{B}^j_{s,t}] \\ &= \frac{1}{2}\mathbb{E}[g(\mathcal{B}_s)\mathcal{B}^i_{s,t}\mathcal{B}^j_{s,t}],\end{aligned}$$

and the proof is established. □

3.3 Young case

Throughout the section we fix the Hurst parameter $H \in (1/2, 1)$ and the final time $T \in (0, \infty)$. We simply write B for the fractional Brownian motion. Our goal is to prove pathwise uniqueness of the Young SDE

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t, \quad X_0 = x$$

for $\sigma \in C^\gamma$ with $\gamma > 1/(2H)$, or Theorem 3.1.1 for $H > 1/2$. As explained in Section 3.1, the key is to obtain a stochastic estimate on

$$\int_s^t f(X_r)\mathrm{d}B_r$$

for an irregular f and a path X controlled by B, which is the subject of the next section.

3.3.1 Young integral with irregular integrand

The following is a probabilistic notion of controlled path.

Definition 3.3.1. Let $\beta \in (0,1]$ and B be a $(\mathcal{F}_t)$-fractional Brownian motion. For $K \in (1,\infty)$ and $(\mathcal{F}_t)$-adapted processes (X, X'), valued in $\mathbb{R}^{d_1}$ and in the space of $d_1 \times d_2$ matrices respectively, we write $(X, X') \in \mathfrak{D}^{\beta,K}$ if

- we have
$$\sup_{0\le s<t\le T} \frac{||X'_{s,t}||_{L^p(\mathbb{P})}}{(t-s)^{\beta H}} < \infty \quad \forall p < \infty$$
and
$$R_{s,t} := X_{s,t} - X'_s B_{s,t}, \quad 0 \le s < t \le T$$
satisfies
$$\sup_{0\le s<t\le T} \frac{||R_{s,t}||_{L^p(\mathbb{P})}}{(t-s)^{(1+\beta)H}} < \infty, \quad \forall p < \infty;$$
- we have $K^{-1} \le X'_t(X'_t)^{\mathrm{T}} \le K$ for all $t \in [0,T]$.

We write

$$|||X|||_p := \sup_{0\le t\le T} \|X'_t\|_{L^p(\mathbb{P})} + \sup_{0\le s<t\le T} \frac{||X'_{s,t}||_{L^p(\mathbb{P})}}{(t-s)^{\beta H}} + \sup_{0\le s<t\le T} \frac{||R_{s,t}||_{L^p(\mathbb{P})}}{(t-s)^{(1+\beta)H}}.$$

The main goal of this section is the following.

Theorem 3.3.2. *Let $H \in (1/2,1)$, $\beta \in (\frac{1}{2H},1]$, $(X,X') \in \mathfrak{D}^{\beta,K}$, $f \in C^1(\mathbb{R}^{d_1},\mathbb{R})$, $p \in [2,\infty)$ and $\gamma \in (\frac{1}{2H}-1,0)$. Then for $s<t$ we have*

$$\Big\| \int_s^t f(X_r)\mathrm{d}B_r \Big\|_{L^p(\mathbb{P})} \lesssim_{K,\beta,p,\gamma} \|f\|_{C^\gamma} \big[(t-s)^{(1+\gamma)H} + |||X|||_{2p}(t-s)^{(1+\beta+\gamma)H}\big].$$

Proof. We decompose the proof into five steps.

Step 1. Firstly, we can assume that B has Mandelbrot–van Ness representation (3.16). Indeed, by Skorokhod's theorem (e.g. [Kal21, Theorem 8.17]), we can enlarge the original probability space so that we have $\tilde{W}$ and (3.16) holds. Set

$$\tilde{\mathcal{F}}_t := \sigma(X_s, X'_s, \tilde{W}_s : s \le t).$$

We claim that $\tilde{W}$ is a $(\tilde{\mathcal{F}}_t)$-Brownian motion. Indeed, fix $s \ge 0$, and set

$$Iw(t) := \tilde{c}_H \int_s^t (t-r)^{H-\frac{1}{2}} \mathrm{d}w_r, \quad t > s.$$

We observe the identity

$$I\tilde{W}_t = \int_s^t K_H(t,r)\mathrm{d}W_r + \int_0^s K_H(t,r)\mathrm{d}W_r - \int_{-\infty}^s \mathcal{K}(t,r)\mathrm{d}\tilde{W}_r.$$

As I corresponds to the fractional differential/integral operator of order $H - 1/2$, it is invertible. Since $(W_r)_{r\le s}$ is measurable with respect to $\mathcal{G}_s := \sigma(\tilde{W}_r : r \le s)$, we can write

$$\tilde{W}_t - \tilde{W}_s = F((W_r - W_s)_{s\le r\le t}, (\tilde{W}_r)_{r\le s})$$

with some measurable map F. Since $(W_r - W_s)_{s\le r\le t}$ is independent of $\mathcal{F}_s$, we see that $\tilde{W}_t - \tilde{W}_s$ is independent of $\mathcal{H}_s := \sigma(X_r, X'_r : r \le s)$ conditionally on $\mathcal{G}_s$. On the other hand, $\tilde{W}_t - \tilde{W}_s$ is independent of $\mathcal{G}_s$. Hence, by the chain rule (e.g. [Kal21, Theorem 8.12]), $W_t - W_s$ is independent of $\tilde{\mathcal{F}}_s = \mathcal{G}_s \vee \mathcal{H}_s$.

Below, although it is an abuse of notation, we write W and $\mathcal{F}_t$ for $\tilde{W}$ and $\tilde{\mathcal{F}}_t$. We write $\mathbb{E}_s[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_s]$.

Step 2. Our strategy is to apply Gerencsér's shifted stochastic sewing (Theorem 1.1.1 with $\alpha = 0$ and $v = s - (t-s)$) to the germ

$$A_{s,t} := \mathbb{E}_{s-(t-s)} \int_s^t f(X_{s-(t-s)} + X'_{s-(t-s)} B_{s-(t-s),r})\mathrm{d}B_r. \tag{3.17}$$

In this step, we will show that in $L^p(\mathbb{P})$

$$\int_s^t f(X_r)\mathrm{d}B_r = \lim_{\substack{\pi \text{ is a partition of } [s,t], \\ |\pi|\to 0}} \sum_{[s',t']\in\pi} A_{s',t'}. \tag{3.18}$$

We set $v := s - (t-s)$ and

$$A^0_{s,t} := \int_s^t f(X_v + X'_v B_{v,r})\mathrm{d}B_r.$$

We first see that

$$\int_s^t f(X_r)\mathrm{d}B_r = \lim_{\substack{\pi \text{ is a partition of } [s,t],\\ |\pi|\to 0}} \sum_{[s',t']\in\pi} A^0_{s',t'}. \tag{3.19}$$

Indeed, we have

$$\begin{aligned}\Big|\int_s^t f(X_r)\mathrm{d}B_r - A^0_{s,t}\Big| \le &\Big|\int_s^t f(X_r)\mathrm{d}B_r - f(X_s)B_{s,t}\Big| \\ &+ |A^0_{s,t} - f(X_v + X'_v B_{v,r})B_{s,t}| \\ &+ |f(X_s)B_{s,t} - f(X_v + X'_v B_{v,r})B_{s,t}|.\end{aligned}$$

By the fundamental estimate of Young's integral, we obtain

$$\begin{aligned}\Big\|\int_s^t f(X_r)\mathrm{d}B_r - f(X_s)B_{s,t}\Big\|_p &\lesssim \|f\|_{C^1}\|\!|X|\!\|_{2p}(t-s)^{(1+\beta)H},\\ \|A^0_{s,t} - f(X_v + X'_v B_{v,r})B_{s,t}\|_p &\lesssim \|f\|_{C^1}\|\!|X|\!\|_{2p}(t-s)^{(1+\beta)H}.\end{aligned}$$

In addition, since

$$|f(X_s) - f(X_v + X'_v B_{v,r})| \le \|f\|_{C^1}|R_{v,s}|,$$

we get

$$\|f(X_s)B_{s,t} - f(X_v + X'_v B_{v,r})B_{s,t}\|_p \le \|f\|_{C^1}\|\!|X|\!\|_{2p}(t-s)^{(2+\beta)H}.$$

Therefore, by the uniqueness part of the sewing lemma, the identity (3.19) is established.

To see (3.18), we note that

$$\|A_{s,t}\|_p \lesssim (t-s)^H, \quad \|A^0_{s,t}\|_p \lesssim (t-s)^H$$

and $\mathbb{E}_v[A_{s,t} - A^0_{s,t}] = 0$. Therefore, the identity (3.18) is proved by the uniqueness part of the stochastic sewing lemma.

Step 3. We set

$$Y^v_t := \int_{-\infty}^v \mathcal{K}(t,r)\mathrm{d}W_r, \quad \tilde{B}^v_t := \int_v^t \mathcal{K}(t,r)\mathrm{d}W_r.$$

Let $\tilde{B} = \tilde{B}^v$ for some $v < s$. By Lemma 3.2.9, for $g \in C^1(\mathbb{R}^{d_2}, \mathbb{R})$ and for a $d_1 \times d_2$ matrix a,

$$\mathbb{E}\Big[\int_s^t g(a\tilde{B}_r)\mathrm{d}\tilde{B}^i_r\Big] = \frac{1}{2}\sum_{j=1}^{d_1} a^{ji}\int_s^t (r-v)^{2H-1}\partial_j \mathcal{P}_{\frac{(r-v)^{2H}}{2H}aa^{\mathrm{T}}}g(0)\mathrm{d}r. \tag{3.20}$$

Step 4. Recall the germ $A_{s,t}$ from (3.17). Here we prove

$$\|A_{s,t}\|_p \lesssim \|f\|_{C^\gamma}(t-s)^{(1+\gamma)H}.$$

For $v := s-(t-s)$, we observe

$$\begin{aligned} A_{s,t} &= \int_s^t \mathbb{E}[f(x+x'(y_{v,r}+\tilde{B}_r))]|_{x=X_v,x'=X'_v,y=Y^v}\dot{Y}^v_r \mathrm{d}r \\ &\quad + \mathbb{E}\int_s^t f(x+x'(y_{v,r}+\tilde{B}_r))\mathrm{d}\tilde{B}_r\Big|_{x=X_v,x'=X'_v,y=Y^v} \\ &=: I_1+I_2. \end{aligned}$$

We compute

$$I_1 = \int_s^t \mathcal{P}_{\frac{(r-v)^{2H}}{2H}X'_v(X'_v)^{\mathrm{T}}} f(X_v+X'_v y_{v,r})\dot{Y}^v_r \mathrm{d}r$$

and by Lemma 3.2.1

$$\|\mathcal{P}_{\frac{(r-v)^{2H}}{2H}X'_v(X'_v)^{\mathrm{T}}} f\|_{L^\infty} \lesssim_K \|f\|_{C^\gamma}(r-v)^{\gamma H}.$$

Since

$$\dot{Y}^v_r = \Big(H-\frac{1}{2}\Big)\int_{-\infty}^v (r-u)^{H-\frac{3}{2}}\mathrm{d}W_u$$

and $\dot{Y}$ is Gaussian, we have

$$\|\dot{Y}^v_r\|_p \lesssim \|\dot{Y}^v_r\|_2 \lesssim (r-v)^{H-1}. \tag{3.21}$$

Therefore,

$$\|I_1\|_p \lesssim \|f\|_{C^\gamma}\int_s^t (r-v)^{\gamma H+H-1}\mathrm{d}r \lesssim \|f\|_{C^\gamma}(t-s)^{(1+\gamma)H}.$$

To estimate I_2, note by (3.20) that

$$I_2 = \frac{1}{2}\sum_{j=1}^{d_1}(X'_v)^{ji}\int_s^t (r-v)^{2H-1}\partial_j \mathcal{P}_{\frac{(r-v)^{2H}}{2H}X'_v(X'_v)^{\mathrm{T}}} f(X_v+X'_v Y^v_{v,r})\mathrm{d}r.$$

Since by Lemma 3.2.1

$$\|\mathcal{P}_{\frac{(r-v)^{2H}}{2H}X'_v(X'_v)^{\mathrm{T}}}f^i(X_v + X'_vY^v_{v,r}))\|_{C^1} \lesssim (r-v)^{H(\gamma-1)}\|f\|_{C^\gamma},$$

we obtain

$$\|I_2\|_p \lesssim \|f\|_{C^\gamma}\int_s^t (r-v)^{2H-1+H(\gamma-1)}\mathrm{d}r \lesssim \|f\|_{C^\gamma}(t-s)^{(\gamma+1)H}.$$

Step 5. We prove for $s < t$

$$\|\mathbb{E}_{s-(t-s)}\delta A_{s,u,t}\|_p \lesssim \|f\|_{C^\gamma}\|\!|X|\!\|_{2p}(t-s)^{(1+\beta+\gamma)H}, \quad u := s + \frac{t-s}{2}.$$

In view of Theorem 1.1.1, this step will complete the proof.

We set

$$s_1 := s-(t-s), \quad s_2 := s-(u-s), \quad s_3 := u-(t-u).$$

We have

$$\mathbb{E}_{s_1}\delta A_{s,u,t} = \mathbb{E}_{s_1}[I_3 + I_4],$$

where

$$I_3 := \mathbb{E}_{s_2}\int_s^u \{f(X_{s_1} + X'_{s_1}B_{s_1,r}) - f(X_{s_2} + X'_{s_2}B_{s_2,r})\}\mathrm{d}B_r,$$
$$I_4 := \mathbb{E}_{s_3}\int_u^t \{f(X_{s_1} + X'_{s_1}B_{s_1,r}) - f(X_{s_3} + X'_{s_3}B_{s_3,r})\}\mathrm{d}B_r.$$

Since the estimate of I_3 and that of I_4 are similar, we only estimate I_3. For $j = 1, 2$, we observe

$$\begin{aligned}\mathbb{E}_{s_2}\int_s^u f(X_{s_j} + X'_{s_j}B_{s_j,r})\mathrm{d}B_r =& \mathbb{E}_{s_2}\int_s^u f(X_{s_j} + X'_{s_j}B_{s_j,r})\dot{Y}^{s_2}_r\mathrm{d}r \\ &+ \mathbb{E}_{s_2}\int_s^u f(X_{s_j} + X'_{s_j}B_{s_j,r})\mathrm{d}\tilde{B}^{s_2}_r.\end{aligned}$$

The first term is equal to

$$\int_s^u \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_j}(X'_{s_j})^{\mathrm{T}}}f(X_{s_j} + X'_{s_j}(B_{s_j,s_2} + Y^{s_2}_{s_2,r}))\dot{Y}^{s_2}_r\mathrm{d}r$$

and by (3.20) the second term is equal to

$$\frac{1}{2}X'_{s_j}\int_s^u (r-s_2)^{2H-1}\nabla\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_j}(X'_{s_j})^{\mathrm{T}}}f(X_{s_j}+X'_{s_j}(B_{s_j,s_2}+Y^{s_2}_{s_2,r}))\mathrm{d}r.$$

Hence, $I_3 = I_5 + I_6$, where

$$\begin{aligned} I_5 := \int_s^u \big\{ & \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}}f(X_{s_1}+X'_{s_1}(B_{s_1,s_2}+Y^{s_2}_{s_2,r})) \\ & - \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_2}(X'_{s_2})^{\mathrm{T}}}f(X_{s_2}+X'_{s_2}Y^{s_2}_{s_2,r})\big\}\dot{Y}^{s_2}_r\mathrm{d}r, \end{aligned}$$

$$\begin{aligned} I_6 := \frac{1}{2}\int_s^u & (r-s_2)^{2H-1} \\ & \times \big\{X'_{s_1}\nabla\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}}f(X_{s_1}+X'_{s_1}(B_{s_1,s_2}+Y^{s_2}_{s_2,r})) \\ & - X'_{s_2}\nabla\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_2}(X'_{s_2})^{\mathrm{T}}}f(X_{s_2}+X'_{s_2}Y^{s_2}_{s_2,r})\big\}\mathrm{d}r. \end{aligned}$$

Since both estimates are similar, we only estimate I_5. We decompose $I_5 = I_7 + I_8$, where

$$\begin{aligned} I_7 := \int_s^u \big\{ & \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}}f(X_{s_1}+X'_{s_1}(B_{s_1,s_2}+Y^{s_2}_{s_2,r})) \\ & - \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}}f(X_{s_2}+X'_{s_2}Y^{s_2}_{s_2,r})\big\}\dot{Y}^{s_2}_r\mathrm{d}r, \end{aligned}$$

$$I_8 := \int_s^u \big\{\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}} - \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_2}(X'_{s_2})^{\mathrm{T}}}\big\}f(X_{s_2}+X'_{s_2}Y^{s_2}_{s_2,r})\dot{Y}^{s_2}_r\mathrm{d}r.$$

Estimating I_7, by Lemma 3.2.1,

$$\|\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}}f\|_{C^1} \lesssim (r-s_2)^{H(\gamma-1)}\|f\|_{C^\gamma}$$

and

$$|I_7| \lesssim \|f\|_{C^\gamma}\int_s^u (r-s_2)^{H(\gamma-1)}|\mathbb{E}_{s_2}[R_{s_1,r}-R_{s_2,r}]||\dot{Y}^{s_2}_r|\mathrm{d}r.$$

Therefore, using (3.21)

$$\begin{aligned} \|I_7\|_p &\lesssim \|f\|_{C^\gamma}\int_s^u (r-s_2)^{H(\gamma-1)}|\!|\!|X|\!|\!|_{2p}(r-s_1)^{(1+\beta)H}(r-s_2)^{H-1}\mathrm{d}r \\ &\lesssim \|f\|_{C^\gamma}|\!|\!|X|\!|\!|_{2p}(t-s)^{(1+\beta+\gamma)H}. \end{aligned}$$

Estimating I_8, by Lemma 3.2.2,

$$\begin{aligned}\big\|\{\mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_1}(X'_{s_1})^{\mathrm{T}}} - \mathcal{P}_{\frac{(r-s_2)^{2H}}{2H}X'_{s_2}(X'_{s_2})^{\mathrm{T}}}\}f\big\|_{L^\infty}& \\ \lesssim |X'_{s_1}(X'_{s_1})^{\mathrm{T}} - X'_{s_2}(X'_{s_2})^{\mathrm{T}}|(r-s_2)^{H\gamma}\|f\|_{C^\gamma}.&\end{aligned}$$

By (3.21)

$$\begin{aligned}\|I_8\|_p &\lesssim \|f\|_{C^\gamma}\|\!|X|\!\|_{2p}\int_s^t (r-s_2)^{H\gamma}(t-s)^{\beta H}(r-s_2)^{H-1}\mathrm{d}r \\ &\lesssim \|f\|_{C^\gamma}\|\!|X|\!\|_{2p}(t-s)^{(1+\beta+\gamma)H}.\end{aligned}$$

□

3.3.2 Pathwise uniqueness

Theorem 3.3.3. *Suppose that $\sigma \in C^\gamma$ with $\gamma > \frac{1}{2H}$. Furthermore, suppose that $\sigma\sigma^{\mathrm{T}}$ is uniformly elliptic, that is, there exists a positive K such that*

$$K^{-1} \le \sigma(y)\sigma^{\mathrm{T}}(y) \le K$$

for all $y \in \mathbb{R}^d$. Then, pathwise uniqueness holds for (3.1).

Before going to the proof of Theorem 3.3.3, we prepare the following lemma.

Lemma 3.3.4. *In the setting of Theorem 3.3.3, let X be a pathwise solution. Then, we have $(X, \sigma(X)) \in \mathfrak{D}^{\gamma,K}$.*

Proof. Let $\alpha < H$, but sufficiently close to H. The fundamental estimate of the Young integral gives

$$|X_{s,t} - \sigma(X_s)B_{s,t}| \le \|\sigma\|_{C^\gamma}\|X\|_{C^\alpha}\|B\|_{C^\alpha}(t-s)^{2\alpha}. \tag{3.22}$$

This implies

$$\|X\|_{C^\alpha([s,t])} \le 2\|\sigma\|_{L^\infty}\|B\|_{C^\alpha([s,t])} \tag{3.23}$$

provided $\|\sigma\|_{C^\gamma}\|B\|_{C^\alpha([0,T])}(t-s)^\alpha \lesssim_\alpha 1$. Therefore,

$$\big\|\|X\|_{C^\alpha([0,T])}\big\|_p \lesssim_{p,\|\sigma\|_{C^\gamma}} 1 \tag{3.24}$$

for all $p < \infty$. Combining (3.22) and (3.24), we get

$$\|X_{s,t}\|_p \lesssim_{p,\|\sigma\|_{C^\gamma}} (t-s)^H.$$

In particular, we have $\|\sigma(X)_{s,t}\|_p \lesssim_{p,\sigma} (t-s)^{\gamma H}$. To get an estimate of the remainder

$$R_{s,t} = X_{s,t} - \sigma(X_s)B_{s,t},$$

it suffices to apply the sewing lemma in $L^p(\mathbb{P})$ with germ

$$A_{s,t} := \sigma(X_s)B_{s,t}. \qquad \square$$

Proof of Theorem 3.3.3. Let $\alpha < H$, but sufficiently close to H. Let X and Y be two pathwise solutions. The fundamental estimate of the Young integral gives

$$\Big|\int_s^t \sigma(Z_r)\mathrm{d}B_r - \sigma(Z_s)B_{s,t}\Big| \lesssim \|\sigma\|_{C^\gamma}\|Z\|_{C^\alpha}\|B\|_{C^\alpha}(t-s)^{2\alpha} \tag{3.25}$$

for $Z \in \{X, Y\}$. Let $(\sigma_n)_{n=1}^\infty$ be a smooth approximation in C^γ of σ. The estimate (3.25) yields

$$\lim_{n\to\infty}\int_0^\cdot \sigma_n(Z_r)\mathrm{d}B_r = \lim_{n\to\infty}\int_0^\cdot \sigma(Z_r)\mathrm{d}B_r$$

uniformly in $[0,T]$. As the computation (3.5) shows, we have

$$\int_0^t \{\sigma_n(X_r) - \sigma_n(Y_r)\}\mathrm{d}B_r = \sum_k \int_0^t (X_r^k - Y_r^k)\mathrm{d}V_n^k(r),$$

where

$$V_n^k(r) := \int_0^1\int_0^r \partial_k\sigma_n(\theta X_u + (1-\theta)Y_u)\mathrm{d}B_u\mathrm{d}\theta.$$

Hence, we obtain

$$X_t - Y_t = \lim_{n\to\infty}\sum_k \int_0^t (X_r^k - Y_r^k)\mathrm{d}V_n^k(r). \tag{3.26}$$

Now our task is to prove the convergence of V_n. By Lemma 3.3.4, we have

$$\theta X_{s,t} + (1-\theta)Y_{s,t} = (\theta\sigma(X_s) + (1-\theta)\sigma(Y_s))B_{s,t} + R_{s,t}^\theta$$

with

$$\|R_{s,t}^\theta\|_p \lesssim_{p,\sigma} (t-s)^{(1+\gamma)H}.$$

One technical problem is that the Gubinelli derivative $\theta\sigma(X) + (1-\theta)\sigma(Y)$ might be degenerate. (In the end, we will show that $X = Y$, hence it must be non-degenerate; but a priori we do not know this.) Because of this technical problem, we proceed as follows.

Step 1. To clarify our argument, we first assume that

$$\sup_{x,y\in\mathbb{R}^d} |\sigma(x) - \sigma(y)| \le \frac{K^{-2}}{4}. \tag{3.27}$$

This condition implies that for all x and y we have

$$\begin{aligned}
&[\theta\sigma(x) + (1-\theta)\sigma(y)][\theta\sigma(x) + (1-\theta)\sigma(y)]^{\mathrm{T}} \\
&\quad =[\sigma(x) + (1-\theta)(\sigma(y) - \sigma(x))][\sigma(x) + (1-\theta)(\sigma(y) - \sigma(x))]^{\mathrm{T}} \\
&\quad \ge K^{-1} - \frac{K^{-1}}{2} - \frac{K^{-2}}{4} \ge \frac{K^{-1}}{4}.
\end{aligned}$$

Therefore, in combination of Lemma 3.3.4, we see that

$$(\theta X + (1-\theta)Y, \theta\sigma(X) + (1-\theta)\sigma(Y)) \in \mathfrak{D}^{\gamma,K/4}.$$

Theorem 3.3.2 yields

$$\Big\| \int_s^t \partial_k(\sigma_n - \sigma_m)(\theta X_r^1 + (1-\theta)X_r^2)\mathrm{d}B_r \Big\|_p \lesssim_{p,\gamma,K,\sigma} \|\sigma_n - \sigma_m\|_{C^\gamma}(t-s)^{\gamma H}.$$

By Kolmogorov's continuity theorem, we see that there exists a process V^k such that for any $\beta < \gamma H$,

$$\lim_{n\to\infty} \| \|V^k - V_n^k\|_{C^\beta}\|_p = 0.$$

Since $\gamma H > 1/2$, we may take $\beta > 1/2$. Therefore, recalling (3.26) we observe

$$X_t - Y_t = \sum_k \int_0^t (X_r^k - Y_r^k)\mathrm{d}V_r^k \quad \text{a.s.}$$

As this shows that $X - Y$ solves the linear Young differential equation

$$\mathrm{d}x_t = \sum_k x_t^k \mathrm{d}V_t^k, \quad x_0 = 0,$$

we see that $X - Y = 0$ a.s.

Step 2. Now we do not assume (3.27). The new ingredient is a stopping time argument. We choose $\varepsilon > 0$ so that

$$\sup_{x,y:|x-y|\leq\varepsilon} |\sigma(x) - \sigma(y)| \leq \frac{K^{-2}}{5},$$

and we set

$$T^{(1)} := \inf\{\, t \geq 0 : |X_t - x| \geq \varepsilon/2 \text{ or } |Y_t - x| \geq \varepsilon/2 \,\}$$

and inductively

$$T^{(i)} := \inf\{\, t \geq T^{(i-1)} : |X_t - X_{T^{(i-1)}}| \geq \varepsilon/2 \text{ or } |Y_t - Y_{T^{(i-1)}}| \geq \varepsilon/2 \,\}.$$

If $T^{(i)} \leq T$,

$$\begin{aligned}\frac{\varepsilon}{2} &= \max\{|X_{T^{(i)}} - X_{T^{(i-1)}}|, |Y_{T^{(i)}} - Y_{T^{(i-1)}}|\} \\ &\leq \max\{\|X\|_{C^\alpha([0,T])}, \|Y\|_{C^\alpha([0,T])}\}(T^{(i)} - T^{(i-1)})^\alpha.\end{aligned}$$

The a priori estimate (3.23) implies that

$$\max\{\|X\|_{C^\alpha}, \|Y\|_{C^\alpha}\} \lesssim_{\sigma,\|B\|_{C^\alpha([0,T])}} 1$$

and hence

$$T^{(i)} - T^{(i-1)} \gtrsim_{\sigma,\|B\|_{C^\alpha([0,T])}} 1. \tag{3.28}$$

uniformly over i, as long as $T^{(i)} \leq T$.

To see that $X = Y$ up to time $T^{(1)}$, let $\sigma^{(1)}$ be a γ-Hölder map such that $\sigma^{(1)} = \sigma$ in an ε-neighborhood of the initial condition x and such that

$$\sup_{x,y} |\sigma^{(1)}(x) - \sigma^{(1)}(y)| \leq \frac{K^{-2}}{4}.$$

For $Z \in \{X, Y\}$, we set

$$Z_t^{(1)} := x + \int_0^t \sigma^{(1)}(Z_r)\mathrm{d}r,$$

and

$$V_n^{(1),k}(t) := \int_0^t \int_0^1 \partial_k \sigma_n(\theta X_r^{(1)} + (1-\theta)Y_r^{(1)})\mathrm{d}\theta \mathrm{d}B_r.$$

Up to time $T^{(1)}$, we have

$$X_t - Y_t = \lim_{n\to\infty} \sum_k \int_0^t (X_r^k - Y_r^k) \mathrm{d}V_n^{(1),k}(r).$$

Due to our choice of $\sigma^{(1)}$, the argument of Step 1 shows that $V_n^{(1),k}$ converges to some $V^{(1),k}$ in C^β, and that $X = Y$ up to time $T^{(1)}$.

Step 3. Obviously we want to repeat the operation of Step 2, but now there is a small problem that $X_{T^{(1)}}$ is random. For this sake, let $(x^m)_{m\in\mathbb{N}}$ be a countable dense set of $\mathbb{R}^d$, and let σ^m be a γ-Hölder map such that $\sigma^m = \sigma$ in an ε-neighborhood of x^m and such that

$$\sup_{x,y} |\sigma^m(x) - \sigma^m(y)| \leq \frac{K^{-2}}{4}. \tag{3.29}$$

For each m and $Z \in \{X, Y\}$, we set

$$\Sigma^{Z,m}(t) := \begin{cases} \sigma(Z_t)\mathbf{1}_{\{t\leq T^{(1)}\}} + \sigma^m(Z_t)\mathbf{1}_{\{t>T^{(1)}\}}, & \text{if } |x^m - X_{T^{(1)}}| \leq \varepsilon/2, \\ \sigma(Z_{\min\{t,T^{(1)}\}}), & \text{otherwise }, \end{cases}$$

$$Z_t^{(2,m)} := x + \int_0^t \Sigma^{Z,m}(r)\mathrm{d}B_r,$$

$$V_n^{(2,m),k}(t) := \int_0^t \int_0^1 \partial_k\sigma_n(\theta X_r^{(2,m)} + (1-\theta)Y_r^{(2,m)})\mathrm{d}\theta\mathrm{d}B_r.$$

Notice that $Z^{(2,m)}$ is adapted. We claim that

$$|Z_{s,t}^{(2,m)} - \Sigma^{Z,m}(s)B_{s,t}| \lesssim (\|\sigma\|_{C^\gamma} + \|\sigma^m\|_{C^\gamma})\|Z\|_{C^\alpha}^\gamma \|B\|_{C^\alpha}(t-s)^{(1+\gamma)\alpha}. \tag{3.30}$$

Indeed, by the fundamental estimate of Young's integral, (3.30) is obvious if $t \leq T^{(1)}$ or if $s \geq T^{(1)}$. Suppose that $s < T^{(1)} < t$. If $|x_m - X_{T^{(1)}}| \leq \varepsilon/2$, then

$$|Z_{s,T^{(1)}}^{(2,m)} - \sigma(Z_s)B_{s,T^{(1)}}| \lesssim \|\sigma\|_{C^\gamma}\|Z\|_{C^\alpha}^\gamma\|B\|_{C^\alpha}(t-s)^{(1+\gamma)\alpha},$$

$$|Z_{T^{(1)},t}^{(2,m)} - \sigma^m(Z_{T^{(1)}})B_{T^{(1)},t}| \lesssim \|\sigma^m\|_{C^\gamma}\|Z\|_{C^\alpha}^\gamma\|B\|_{C^\alpha}(t-s)^{(1+\gamma)\alpha}.$$

Since $|x_m - X_{T^{(1)}}| = |x_m - Y_{T^{(1)}}| \leq \varepsilon/2$, we have

$$|\sigma^m(Z_{T^{(1)}}) - \sigma(Z_s)| = |\sigma(Z_{T^{(1)}}) - \sigma(Z_s)| \leq \|\sigma\|_{C^\gamma}\|Z\|_{C^\alpha}^\gamma(T^{(1)} - s)^{\gamma\alpha},$$

and the estimate (3.30) follows. The case where $|x_m - X_{T^{(1)}}| > \varepsilon/2$ is similar.

Now we check that the Gubinelli derivative $\Sigma^{\theta,m} := \theta\Sigma^{X,m} + (1-\theta)\Sigma^{Y,m}$ is non-degenerate. Indeed, for $t \le T^{(1)}$ we have

$$\Sigma^{\theta,m}(t)\Sigma^{\theta,m}(t)^{\mathrm{T}} = \sigma(X_t)\sigma(X_t)^{\mathrm{T}} \ge K^{-1},$$

and for $t > T^{(1)}$ the condition (3.29) implies that $\Sigma^{\theta,m}(t)\Sigma^{\theta,m}(t)^{\mathrm{T}} \ge K^{-1}/4$. Therefore, Theorem 3.3.2 and the Kolmogorov continuity theorem show that $V_n^{(2,m)}$ converges to some $V^{(2,m)}$ in $L^p(\mathbb{P})$. By the diagonalization argument, we may suppose that almost surely for every $m \in \mathbb{N}$ and $\delta \in (0, \gamma H)$ we have

$$\lim_{n\to\infty} \|V^{(2,m)} - V_n^{(2,m)}\|_{C^{\gamma H - \delta}} = 0.$$

We can find a random m so that $|x_m - X_{T^{(1)}}| < \varepsilon/2$. We then have

$$X_t - Y_t = \sum_k \int_0^t (X_r^k - Y_r^k)\mathrm{d}V^{(2,m),k}(r)$$

up to $t \le T^{(2)}$, hence $X = Y$ on $[0, T^{(2)}]$. It is now clear that this algorithm can be continued, and at some point we must have $T^{(i)} \ge T$ due to (3.28). ☐

3.4 Rough case

Throughout this section we fix $H \in (1/3, 1/2)$, and we will drop scripts on H. We always interpret the integral of the form

$$\int_s^t Y_r \mathrm{d}B_r$$

as the rough integral with respect to the canonical lift of B (Definition 3.2.4). We consider the rough differential equation

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t, \quad X_0 = x.$$

Given two adapted solutions X and Y, *if we pretend that σ and B are smooth*, then

$$X_t - Y_t = \sum_{k=1}^{d_1} \int_0^t (X_r^k - Y_r^k)\mathrm{d}G_r^k, \tag{3.31}$$

where

$$G_t^k := \int_0^t \Big[\int_0^1 \partial_k\sigma(\theta X_r + (1-\theta)Y_r)\mathrm{d}\theta\Big]\mathrm{d}B_r$$

Yet, as $H < 1/2$, we must appropriately interpret the integral in the right-hand side of (3.31). Since X and Y are controlled by B, it is natural to construct a joint rough path of B and $(G^k)_{k=1}^{d_1}$:

$$(B, G^k, \int B \otimes \mathrm{d}B, \int B \otimes \mathrm{d}G^k, \int G^k \otimes \mathrm{d}B, \int G^k \otimes \mathrm{d}G^l). \tag{3.32}$$

As in the Young case, the problem here is that σ is too irregular to make sense of the lift (3.32) by pathwise method. One obvious difference from the Young case is that we have to make sense of iterated integrals. However, even for constructing G^k, there is a new difficulty, which we explain now. Our task is to make sense of

$$\int_0^t f(X_r)\mathrm{d}B_r \tag{3.33}$$

for an irregular f, where X is controlled by B with Gubinelli derivative X'. A similar problem for $H > 1/2$ was already solved in Section 3.3 by considering the germ

$$A_{s,t} := \mathbb{E}_{s-(t-s)} \int_s^t f(X_{s-(t-s)} + X'_{s-(t-s)}B_{s-(t-s),r})\mathrm{d}B_r. \tag{3.34}$$

The very crucial advantage for the case $H > 1/2$ is that we can take conditional expectation on $\mathcal{F}_{s-(t-s)}$, which allows us to perform Gaussian computations. The reason why we can take the conditional expectation in this case is that the Hölder exponent of the process (3.33) is greater than $1/2$. For $H < 1/2$, this is not the case, and hence we are not allowed to consider the germ (3.34). To see further why taking the conditional expectation does not help, we can consider the case $H = 1/2$; if the integral is understood in Itô's sense, (3.34) is just 0, and this observation also implies that the germ (3.34) is not correct even for the Stratonovich integral, in view of the Itô–Stratonovich correction.

Therefore, in the rough case we have to consider the germ without conditional expectation:

$$A_{s,t} := \int_s^t f(X_s + X'_s B_{s,r})\mathrm{d}B_r. \tag{3.35}$$

To get $(1/2+\varepsilon)$-exponent, we must estimate $\delta A_{s,u,t}$ rather than just $A_{s,t}$. To this end, we apply the shifted stochastic sewing again. That is, we apply the stochastic sewing *twice* — first to estimate $\|\delta A_{s,u,t}\|_p$ by applying the fully shifted version, then to apply Lê's version with $(A_{s,t})_{s<t}$.

3.4.1 Rough integral with irregular integrand

Technical estimates

Let us recall the Riemann–Liouville process $\mathcal{B}$ from Definition 3.2.8. Here we give technical estimates involving $\mathcal{B}$, which will be necessary to estimate $\|\delta A_{s,u,t}\|_p$. Our computations here resemble Section 1.3.

Lemma 3.4.1. *Let $g\colon \mathbb{R}^{d_2} \to \mathbb{R}$ be a bounded measurable function. We set*

$$A^1_{s,t} := g(\mathcal{B}_s)\mathcal{B}^i_{s,t}\mathcal{B}^j_{s,t}, \quad A^2_{s,t} := g(\mathcal{B}_s)\int_s^t \mathcal{B}^i_{s,r}\mathrm{d}\mathcal{B}^j_r.$$

Then, for $p \in [2, \infty)$ and $v < s < t$ with $t - s \leq s - v$, we have

$$\|\mathbb{E}[A^1_{s,t} - 2A^2_{s,t}|\mathcal{F}_v]\|_p \lesssim_p \|g\|_{L^\infty(\mathbb{R}^{d_2})}\Big(\frac{t-s}{s-v}\Big)^{1-H}(t-s)^{2H}.$$

Proof. We set

$$Y_t := \int_0^v \mathcal{K}(t,r)\mathrm{d}W_r, \quad \tilde{\mathcal{B}}_t := \int_v^t \mathcal{K}(t,r)\mathrm{d}W_r.$$

Note that $Y = \mathbb{E}[\mathcal{B}|\mathcal{F}_v]$ and $\tilde{\mathcal{B}}$ is independent of $\mathcal{F}_v$. We have

$$A^1_{s,t} = g(Y_s + \tilde{\mathcal{B}}_s)[Y^i_{s,t}Y^j_{s,t} + Y^i_{s,t}\tilde{\mathcal{B}}^j_{s,t} + \tilde{\mathcal{B}}^i_{s,t}Y^j_{s,t} + \tilde{\mathcal{B}}^i_{s,t}\tilde{\mathcal{B}}^j_{s,t}].$$

To estimate the first three terms, observe that

$$\|Y_{s,t}\|_p \lesssim_p (s-v)^{H-1}(t-s), \quad \|\tilde{\mathcal{B}}_{s,t}\|_p \lesssim_p (t-s)^H.$$

Hence, we obtain

$$\|A^1_{s,t} - g(Y_s + \tilde{\mathcal{B}}_s)\tilde{\mathcal{B}}^i_{s,t}\tilde{\mathcal{B}}^j_{s,t}\|_p \lesssim_p \|g\|_{L^\infty(\mathbb{R}^{d_2})}\Big(\frac{t-s}{s-v}\Big)^{1-H}(t-s)^{2H}.$$

Similarly, $A^2_{s,t}$ equals to

$$g(Y_s + \tilde{B}_s)\Big[\int_s^t Y^i_{s,r}\dot{Y}^j_r\mathrm{d}r + \int_s^t Y^i_{s,r}\mathrm{d}\tilde{\mathcal{B}}^j_r + \int_s^t \tilde{\mathcal{B}}^i_{s,r}\dot{Y}^j_r\mathrm{d}r + \int_s^t \tilde{\mathcal{B}}^i_{s,r}\mathrm{d}\tilde{\mathcal{B}}^j_r\Big]$$

and

$$\Big\|A^2_{s,t} - g(Y_s + \tilde{B}_s)\int_s^t \tilde{\mathcal{B}}^i_{s,r}\mathrm{d}\tilde{\mathcal{B}}^j_r\Big\|_p \lesssim_p \|g\|_{L^\infty(\mathbb{R}^{d_2})}\Big(\frac{t-s}{s-v}\Big)^{1-H}(t-s)^{2H}.$$

To complete the proof, it remains to observe

$$\mathbb{E}\Big[g(Y_s+\tilde{\mathcal{B}}_s)\int_s^t \tilde{\mathcal{B}}^i_{s,r}\mathrm{d}\tilde{\mathcal{B}}^j_r\Big|\mathcal{F}_v\Big]=\frac{1}{2}\mathbb{E}[g(Y_s+\tilde{\mathcal{B}}_s)\tilde{\mathcal{B}}^i_{s,t}\tilde{\mathcal{B}}^j_{s,t}|\mathcal{F}_v],\tag{3.36}$$

which follows from Lemma 3.2.10. □

Lemma 3.4.2. *Let $H\in(1/3,1/2)$ and let E be a measurable space. Let ξ be a random variable valued in E and let η be a random path valued in $\mathbb{R}^d$ such that*

$$\|\eta_{s,t}\|_p\lesssim_p s^{-(1-H)}(t-s)$$

for all $p\in[1,\infty)$ and $s<t$ with $t-s\le s$. Suppose that $F\colon E\times\mathbb{R}^d\to\mathbb{R}$ satisfies

$$|F(x,y_1)-F(x,y_2)|\le M(x)|y_1-y_2|^\gamma$$

for all $x\in E$ and $y_1,y_2\in\mathbb{R}^d$, with $\gamma\in(\frac{1}{2H}-1,1]$. Finally, suppose that ξ,η are independent of $\mathcal{B}$. We set

$$A_{s,t}:=(F(\xi,\eta_s+\mathcal{B}_s)+F(\xi,\eta_t+\mathcal{B}_t))\mathcal{B}^i_{s,t}$$

and $\mathcal{F}_t:=\sigma(\xi,\eta,\mathcal{B}_r:r\le t)$. We then have

$$\|\delta A_{s,u,t}\|_p\lesssim_{p,\gamma}\|M(\xi)\|_{2p}(t-s)^{(1+\gamma)H},\tag{3.37}$$

$$\|\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v]\|_p\lesssim_{p,\gamma}\|M(\xi)\|_{2p}\Big(\frac{t-s}{s-v}\Big)^{1+(1-\gamma)H}(t-s)^{(1+\gamma)H}.\tag{3.38}$$

Proof. We essentially repeat the argument of Proposition 1.3.6. We first consider the estimate (3.37). By [LL22, Lemma 3.4], we see that

$$\|\eta_{s,t}\|_p\lesssim_p(t-s)^H$$

for all $s<t$. We compute

$$\delta A_{s,u,t}=F(\xi,\eta+\mathcal{B})_{u,t}\mathcal{B}^i_{s,u}-F(\xi,\eta+\mathcal{B})_{s,u}\mathcal{B}^i_{u,t},$$

and

$$|\delta A_{s,u,t}|\lesssim_\gamma M(\xi)\big[|\eta_{u,t}+\mathcal{B}_{u,t}|^\gamma|\mathcal{B}_{s,u}|+|\eta_{s,u}+\mathcal{B}_{s,u}|^\gamma|\mathcal{B}_{u,t}|\big].$$

Now it is easy to see the estimate (3.37).

We turn to the estimate (3.38). We set

$$Y_s := \eta_s + \int_0^v \mathcal{K}(s,r)\mathrm{d}W_r, \quad \tilde{\mathcal{B}}_s := \int_v^s \mathcal{K}(s,r)\mathrm{d}W_r$$

and

$$\begin{aligned} a_0(s) &:= \mathbb{E}[F(x, y_s + \tilde{\mathcal{B}}_s)]|_{x=\xi, y=Y}, \\ a_i(s) &:= \mathbb{E}[(\tilde{\mathcal{B}}^i_s)^2]^{-1}\mathbb{E}[F(x, y_s + \tilde{\mathcal{B}}_s)\tilde{B}^i_s]|_{x=\xi, y=Y}. \end{aligned}$$

We note that

$$\begin{aligned} \|Y_{s,t}\|_p &\lesssim_p s^{-(1-H)}(t-s) + (s-v)^{-(1-H)}(t-s) \\ &\leq (s-v)^{-(1-H)}(t-s). \end{aligned} \tag{3.39}$$

As in Proposition 1.3.6, especially (1.32) and (1.35), we have

$$\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_v] = D^0_{s,u,t} + D^i_{s,u,t},$$

where

$$D^0_{s,u,t} := (a_0(t) - a_0(u))Y^1_{s,u} + (a_0(s) - a_0(u))Y^i_{u,t}$$

and

$$\begin{aligned} D^i_{s,u,t} := (a_i(t) - a_i(u))\mathbb{E}[\tilde{B}^i_t\tilde{B}^i_{s,t}] + (a_i(s) - a_i(u))\mathbb{E}[\tilde{B}^i_s\tilde{B}^i_{s,t}] \\ - (a_i(s) - a_i(u))\mathbb{E}[\tilde{B}^i_s\tilde{B}^i_{s,u}] - (a_i(t) - a_i(u))\mathbb{E}[\tilde{B}^i_t\tilde{B}^i_{u,t}]. \end{aligned}$$

The estimate as in (1.33) gives

$$|D^0_{s,u,t}| \lesssim_\gamma M(\xi)\big[(s-v)^{(\gamma-1)H}|Y_{s,u}||Y_{u,t}| + (s-v)^{\gamma H-1}(t-s)(|Y_{s,u}| + |Y_{u,t}|)\big].$$

In view of (3.39), we obtain

$$\|D^0_{s,u,t}\|_p \lesssim_{p,\gamma} \|M(\xi)\|_{2p}(s-v)^{(\gamma+1)H-2}(t-s)^2.$$

Similarly, the estimate as in (1.36) gives

$$\|D^i_{s,u,t}\|_p \lesssim_{p,\gamma} \|M(\xi)\|_{2p}(s-v)^{(\gamma-1)H-1}(t-s)^{1+2H}.$$

Now the estimate (3.38) is proven. □

Lemma 3.4.3. *In the setting of Lemma 3.4.2, suppose further that for every x the map $F(x,\cdot)$ is locally C^2, that*

$$|\nabla F(x,y)| + |\nabla^2 F(x,y)| \leq N(x)(1+|y|)$$

for all x and y, and that

$$\sup_{r\in[s,t]} \|N(\xi)(1+|\eta_r|)\|_{2p} < \infty.$$

Then we have

$$\Big\| \int_s^t F(\xi,\eta_r+\mathcal{B}_r)\mathrm{d}\mathcal{B}_r - \frac{F(\xi,\eta_s+\mathcal{B}_s)+F(\xi,\eta_t+\mathcal{B}_t)}{2}\mathcal{B}_{s,t}\Big\|_p \\ \lesssim_{p,\gamma} \|M(\xi)\|_{2p}(t-s)^{(1+\gamma)H}.$$

Proof. The integral $\int_s^t F(\xi,\eta_r+\mathcal{B}_r)\mathrm{d}\mathcal{B}_r$ is understood as rough integral:

$$\int_s^t f(\mathcal{B}_r)\mathrm{d}\mathcal{B}_r = \lim_{|\pi|\to 0} \sum_{[s',t']\in\pi} A^1_{s',t'}, \tag{3.40}$$

where π is a partition of $[s,t]$ and

$$A^1_{s,t} := F(\xi,\eta_s+\mathcal{B}_s)\mathcal{B}_{s,t} + \nabla F(\xi,\eta_s+\mathcal{B}_s)\int_s^t \mathcal{B}_{s,r}\mathrm{d}\mathcal{B}_r.$$

(The operator ∇ acts only on the second variable.) In view of Theorem 1.1.1 and Lemma 3.4.2, it suffices to show

$$\int_s^t F(\xi,\eta_r+\mathcal{B}_r)\mathrm{d}\mathcal{B}_r = \lim_{|\pi|\to 0} \sum_{[s',t']\in\pi} A^2_{s',t'}, \tag{3.41}$$

where

$$A^2_{s,t} := \frac{F(\xi,\eta_s+\mathcal{B}_s)+F(\xi,\eta_t+\mathcal{B}_t)}{2}\mathcal{B}_{s,t}.$$

To this end, we apply the sewing technique. We have

$$\begin{aligned} A^2_{s,t} - A^1_{s,t} \\ &= \frac{F(\xi,\eta_t+\mathcal{B}_t)-F(\xi,\eta_s+\mathcal{B}_s)}{2}\mathcal{B}_{s,t} - \nabla F(\xi,\eta_s+\mathcal{B}_s)\int_s^t \mathcal{B}_{s,r}\mathrm{d}\mathcal{B}_r \\ &= \frac{1}{2}\nabla F(\xi,\eta_s+\mathcal{B}_s)\mathcal{B}_{s,t}\otimes\mathcal{B}_{s,t} - \nabla F(\xi,\eta_s+\mathcal{B}_s)\int_s^t \mathcal{B}_{s,r}\mathrm{d}\mathcal{B}_r + R_{s,t}, \end{aligned} \tag{3.42}$$

where

$$|R_{s,t}| \lesssim_{\omega,F} (t-s)^{3H}.$$

Since $H > 1/3$, we have

$$\Big| \sum_{[s',t']\in\pi} R_{s,t} \Big| \lesssim_{\omega,F} |\pi|^{3H-1}.$$

Setting $A^3_{s,t} := A^2_{s,t} - A^1_{s,t} - R_{s,t}$, we therefore see that

$$\lim_{|\pi|\to 0} \sum_{[s',t']\in\pi} (A^1_{s,t} - A^2_{s,t}) = \lim_{|\pi|\to 0} \sum_{[s',t']\in\pi} A^3_{s,t} \quad \text{a.s.}$$

To estimate the right-hand side, we will apply the uniqueness part of Theorem 1.1.1. By the representation (3.42), we easily obtain

$$\|A^3_{s,t}\|_p \lesssim \|N(\xi)(1 + |\eta_s + \mathcal{B}_s|)\|_{2p}(t-s)^{2H},$$

and note that the exponent $2H$ is greater than $1/2$. By Lemma 3.4.1,

$$\|\mathbb{E}[A^3_{s,t}|\mathcal{F}_v]\|_p \lesssim_p \|N(\xi)(1+|\eta_s|)\|_p \Big(\frac{t-s}{s-v}\Big)^{1-H}(t-s)^{2H}.$$

Hence, (3.41) is proven in view of (3.40) and Theorem 1.1.1. □

Main estimate

The following is the most important technical result of this section.

Theorem 3.4.4. *Let $H \in (1/3, 1/2)$, $(X, X') \in \mathfrak{D}^{1,K}$, $f \in C^2(\mathbb{R}^d)$, $\gamma \in (1/H - 2, 1)$ and $\varepsilon \in (0,1)$. Then, if $|t-s| \le 1$, we have*

$$\Big\| \int_s^t f(X_r)\mathrm{d}B_r - f(X_s)B_{s,t} \Big\|_p \lesssim_{p,K,\gamma,\varepsilon} \|f\|_{C^\gamma}(1 + |||X|||_{(1+\varepsilon)p})(t-s)^{(1+\gamma)H}.$$

Proof. We set

$$A_{s,t} := \int_s^t f(X_s + X'_s B_{s,r})\mathrm{d}B_r. \tag{3.43}$$

The integral is understood as rough integral: the map

$$r \mapsto f(X_s + X'_s B_{s,r})$$

is controlled by B with Gubinelli derivative $f'(X_s + X'_s B_{s,r})X'_s$. In particular, we have

$$|A_{s,t} - f(X_s)B_{s,t} - f'(X_s)X'_s\mathbb{B}_{s,t}| \lesssim (t-s)^{3\alpha}$$

for $\alpha \in (1/3, H)$, and hence

$$\int_s^t f(X_r)\mathrm{d}B_r = \lim_{|\pi|\to 0} \sum_{[s',t']\in\pi} A_{s',t'},$$

where π is a partition of $[s,t]$. As in Theorem 3.3.2, we apply the stochastic sewing to $A_{s,t}$. This time, the version of Lê is sufficient (that is $\alpha = 0$ and $v = s$ in Theorem 1.1.1). Furthermore, as valided in Step 1 of Theorem 3.3.2, we assume the Mandelbrot–van Ness representation.

Step 1. We estimate $\|\delta A_{s,u,t}\|_p$. We observe

$$\begin{aligned} \delta A_{s,u,t} &= \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{u,r})\}\mathrm{d}B_r \\ &= \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{u,r})\}\dot{Y}_r \mathrm{d}r \qquad (3.44) \\ &\quad + \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{u,r})\}\mathrm{d}\tilde{B}_r, \qquad (3.45) \end{aligned}$$

where

$$Y_t := \int_{-\infty}^u \mathcal{K}(t,r)\mathrm{d}W_r, \quad \tilde{B}_r := \int_u^t \mathcal{K}(t,r)\mathrm{d}W_r. \qquad (3.46)$$

The integral (3.44), by the triangle inequality, is bounded by

$$\|f\|_{C^\gamma} \int_u^t |X_s + X'_s B_{s,r} - (X_u + X'_u B_{u,r})|^\gamma |\dot{Y}_r| \mathrm{d}r.$$

Since

$$X_s + X'_s B_{s,r} - (X_u + X'_u B_{u,r}) = -(X_{s,u} - X'_s B_{s,u}) - X'_{s,u} B_{u,r},$$

and $\|\dot{Y}_r\|_p \lesssim (r-u)^{H-1}$, we obtain

$$\Big\| \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{s,r})\}\dot{Y}_r \mathrm{d}r \Big\|_p \lesssim |\!|\!| X |\!|\!|_q^\gamma \|f\|_{C^\gamma}(t-s)^{(1+2\gamma)H}.$$

To estimate (3.45), by Lemma 3.4.2 for $v \in \{s, u\}$ we have

$$\Big\| \int_u^t f(X_v + X'_v B_{v,r}) \mathrm{d}\tilde{B}_r - \frac{f(X_v + X'_v B_{v,u}) + f(X_v + X'_v B_{v,t})}{2} \tilde{B}_{u,t} \Big\|_p \lesssim_{\gamma,K} \|f\|_{C^\gamma} (t-s)^{(1+\gamma)H}.$$

Since

$$\Big| \frac{f(X_s + X'_s B_{s,u}) + f(X_s + X'_s B_{s,t})}{2} - \frac{f(X_u) + f(X_u + X'_u B_{u,t})}{2} \Big| \lesssim \|f\|_{C^\gamma} \big(|X_{s,u} - X'_s B_{s,u}|^\gamma + |X'_{s,u} B_{u,t}|^\gamma\big),$$

we see that the $L^p(\mathbb{P})$-norm of the integral (3.45) is bounded by (up to constant)

$$\|f\|_{C^\gamma} (t-s)^{(1+\gamma)H} + \|\|X\|\|_q^\gamma \|f\|_{C^\gamma} (t-s)^{(1+2\gamma)H}$$

Step 2. Next we estimate $\mathbb{E}[\delta A_{s,u,t} | \mathcal{F}_s]$, which is essentially done in Theorem 3.3.2. As in Step 1, Y and $\tilde{B}$ are defined by (3.46). We have

$$\begin{aligned} \mathbb{E}_u \delta A_{s,u,t} =& \mathbb{E}_u \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{s,r})\} \dot{Y}_r \mathrm{d}r \\ &+ \mathbb{E}_u \int_u^t \{f(X_s + X'_s B_{s,r}) - f(X_u + X'_u B_{s,r})\} \mathrm{d}\tilde{B}_r. \end{aligned}$$

For $v \in \{s, u\}$, we observe

$$\mathbb{E}_u \int_u^t f(X_v + X'_v B_{v,r}) \dot{Y}_r \mathrm{d}r = \int_u^t \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_v (X'_v)^{\mathrm{T}}} f(X_v + X'_v (B_{v,u} + Y_{u,r})) \dot{Y}_r \mathrm{d}r$$

and by Step 2 of the proof of Theorem 3.3.2 we observe

$$\begin{aligned} &\mathbb{E}_u \int_u^t f(X_v + X'_v B_{v,r}) \mathrm{d}\tilde{B}_r \\ &\quad = \frac{1}{2} X'_v \int_u^t (r-u)^{2H-1} \nabla \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_v (X'_v)^{\mathrm{T}}} f(X_v + X'_v (B_{v,u} + Y_{u,r})) \mathrm{d}r. \end{aligned}$$

Hence, $\mathbb{E}_u \delta A_{s,u,t} = J_1 + J_2 + J_3$, where

$$J_1 := \int_u^t \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_s (X'_s)^{\mathrm{T}}} \{f(X_s + X'_s (B_{s,u} + Y_{u,r})) - f(X_u + X'_u Y_{u,r})\} \dot{Y}_r \mathrm{d}r$$

$$J_2 := \int_u^t (\mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_s (X'_s)^{\mathrm{T}}} - \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_u (X'_u)^{\mathrm{T}}}) f(X_u + X'_u Y_{u,r}) \dot{Y}_r \mathrm{d}r$$

$$\begin{aligned} J_3 := \frac{1}{2} \int_u^t & (r-u)^{2H-1} \\ & \times \big\{ X'_s \nabla \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_s (X'_s)^{\mathrm{T}}} f(X_s + X'_s(B_{s,u} + Y_{u,r})) \\ & \qquad - X'_u \nabla \mathcal{P}_{\frac{(r-u)^{2H}}{2H} X'_u (X'_u)^{\mathrm{T}}} f(X_u + X'_u Y_{u,r}) \big\} \mathrm{d}r. \end{aligned}$$

Then the rest of the argument is identical to Step 4 of Theorem 3.3.2. (J_1 corresponds to I_7, J_2 to I_8 and J_3 to I_6.) In particular, we obtain

$$\|\mathbb{E}_s[\delta A_{s,u,t}]\|_{L^p(\mathbb{P})} \lesssim \|f\|_{C^\gamma} \|\!|X|\!\|_{(1+\varepsilon)p} (t-s)^{(2+\gamma)H}.$$

Step 3. By Theorem 1.1.1,

$$\begin{aligned} \Big\| \int_s^t f(X_r) \mathrm{d}B_r - A_{s,t} \Big\|_p \lesssim \|f\|_{C^\gamma} \big[& (t-s)^{(1+\gamma)H} \\ & + \|\!|X|\!\|_{(1+\varepsilon)p\gamma}^\gamma (t-s)^{(1+2\gamma)H} + \|\!|X|\!\|_{(1+\varepsilon)p} (t-s)^{(2+\gamma)H} \big]. \end{aligned}$$

By Lemma 3.4.3,

$$\Big\| A_{s,t} - \frac{f(X_s) + f(X_s + X'_s B_{s,t})}{2} B_{s,t} \Big\|_p \lesssim \|f\|_{C^\gamma} (t-s)^{(1+\gamma)H}.$$

It remains to observe

$$\Big\| \frac{f(X_s) + f(X_s + X'_s B_{s,t})}{2} B_{s,t} - f(X_s) B_{s,t} \Big\|_p \lesssim \|f\|_{C^\gamma} (t-s)^{(1+\gamma)H}. \qquad \square$$

3.4.2 Iterated integrals

The goal of this section is to estimate iterated integrals in (3.32). This turns out to be easy corollaries of Theorem 3.4.4.

Lemma 3.4.5. *Let* $H \in (1/3, 1/2)$, $(X, X') \in \mathfrak{D}^{1,K}$, $f \in C^2(\mathbb{R}^d)$, $\gamma \in (1/H - 2, 1)$ *and* $\varepsilon \in (0,1)$. *Then, if* $|t-s| \le 1$, *we have*

$$\Big\| \int_s^t f(X_r) B^i_{s,r} \mathrm{d}B^j_r - f(X_s) \mathbb{B}^{i,j}_{s,t} \Big\|_p \lesssim_{p,K,\varepsilon,\gamma} \|f\|_{C^\gamma} (1 + \|\!|X|\!\|_{(1+\varepsilon)p}) (t-s)^{(2+\gamma)H}.$$

Proof. We fix a $\tau \leq s$, and we set

$$\bar{A}_{s,t} := B^i_{\tau,s} \int_s^t f(X_r) \mathrm{d}B^j_r + f(X_s) \mathbb{B}^{i,j}_{s,t}.$$

Step 1. We show that $\bar{A}_{s,t}$ is a correct approximation. For this sake, we set

$$\bar{A}'_{s,t} := f(X_s) B^i_{\tau,s} B^j_{s,t} + f(X_s) \mathbb{B}^{i,j}_{s,t} + \sum_{k,l} B^i_{\tau,s} \partial_k f(X_s) (X'_s)^{kl} \mathbb{B}^{lj}_{s,t}.$$

By definition of the rough integral, we have

$$\int_s^t f(X_r) B^i_{\tau,r} \mathrm{d}B^j_r = \lim_{\substack{\pi \text{ is a partition of } [s,t], \\ |\pi| \to 0}} \sum_{[u,v] \in \pi} \bar{A}'_{u,v}.$$

Our goal is to prove $|\bar{A}_{s,t} - \bar{A}'_{s,t}| \lesssim_{f,\omega} (t-s)^{3H}$. We compute

$$\bar{A}_{s,t} - \bar{A}'_{s,t} = B^i_{\tau,s} \Big(\int_s^t f(X_r) \mathrm{d}B_r - f(X_s) B^j_{s,t} - \sum_{k,l} \partial_k f(X_s) (X'_s)^{kl} \mathbb{B}^{lj}_{s,t} \Big)$$

By the fundamental estimate of the rough integral, we have

$$\Big| \int_s^t f(X_r) \mathrm{d}B_r - f(X_s) B^j_{s,t} - \sum_{k,l} \partial_k f(X_s) (X'_s)^{kl} \mathbb{B}^{lj}_{s,t} \Big| \lesssim_{f,\omega} (t-s)^{3H},$$

hence $\bar{A}$ is a correct approximation.

Step 2. This time we do not need the stochastic sewing; the usual sewing in $L^p(\mathbb{P})$ suffices. We compute

$$\delta \bar{A}_{s,u,t} = -B^i_{s,u} \int_u^t f(X_r) \mathrm{d}B^j_r + f(X_s)(\mathbb{B}^{i,j}_{s,t} - \mathbb{B}^{i,j}_{s,u}) - f(X_u) \mathbb{B}^{i,j}_{u,t}.$$

Using Chen's relation

$$\mathbb{B}^{i,j}_{u,t} = \mathbb{B}^{i,j}_{s,t} - \mathbb{B}^{i,j}_{s,u} - B^i_{s,u} B^j_{u,t},$$

we obtain

$$\delta A_{s,u,t} = -B^i_{s,u} \Big(\int_u^t f(X_r) \mathrm{d}B^j_r - f(X_u) B^j_{u,t} \Big) - f(X)_{s,u} (\mathbb{B}^{i,j}_{s,t} - \mathbb{B}^{i,j}_{s,u}).$$

Since

$$\|B^i_{s,u}\|_p \lesssim (t-s)^H, \quad \|\mathbb{B}^{i,j}_{s,t}\|_p \lesssim (t-s)^{2H},$$
$$\|f(X)_{s,u}\|_p \lesssim \|f\|_{C^\gamma} \|\!|X|\!\|^\gamma_{p\gamma} (t-s)^{\gamma H},$$

and by Theorem 3.4.4

$$\Big\| \int_u^t f(X_r) \mathrm{d}B^j_r - f(X_u) B^j_{u,t} \Big\|_p \lesssim \|f\|_{C^\gamma}(1 + \|\!|X|\!\|_{(1+\varepsilon)p})(t-s)^{(1+\gamma)H},$$

we obtain

$$\|\delta A_{s,u,t}\|_p \lesssim \|f\|_{C^\gamma}(1 + \|\!|X|\!\|_{(1+\varepsilon)p})(t-s)^{(2+\gamma)H}.$$

As $(2+\gamma)H > 1$, the sewing lemma in $L^p(\mathbb{P})$ gives

$$\Big\| \int_s^t f(X_r) B^i_{s,r} \mathrm{d}B^j_r - \bar{A}_{s,t} \Big\|_p \lesssim_{p,K,\varepsilon,\gamma} \|f\|_{C^\gamma}(1 + \|\!|X|\!\|_{(1+\varepsilon)p})(t-s)^{(2+\gamma)H}.$$

It remains to set $\tau = s$. □

Lemma 3.4.6. *Let $H \in (1/3, 1/2)$, $(X, X'), (Y, Y') \in \mathfrak{D}^{1,K}$, $g, h \in C^2(\mathbb{R}^d)$, $\gamma \in (1/H - 2, 1)$ and $\varepsilon \in (0,1)$. Then, if $|t-s| \le 1$, we have*

$$\Big\| \int_s^t \Big(\int_s^{r_2} g(X_{r_1}) \mathrm{d}B^i_{r_1} \Big) h(Y_{r_2}) \mathrm{d}B^j_{r_2} - g(X_s) h(Y_s) \mathbb{B}^{i,j}_{s,t} \Big\|_p$$
$$\lesssim_{p,K,\varepsilon,\gamma} \|g\|_{C^\gamma} \|h\|_{C^\gamma}(1 + \|\!|X|\!\|_{2(1+\varepsilon)p})(1 + \|\!|Y|\!\|_{2(1+\varepsilon)p})(t-s)^{(2+\gamma)H}.$$

Proof. Let $\tau \le s$. This time our germ is

$$\tilde{A}_{s,t} := \Big(\int_\tau^s g(X_r) \mathrm{d}B^i_r \Big) \times \int_s^t h(Y_r) \mathrm{d}B^j_r + g(X_s) \int_s^t B^i_{s,r} h(Y_r) \mathrm{d}B^j_r.$$

Step 1. To see that $\tilde{A}_{s,t}$ is a correct approximation, we set

$$\tilde{A}'_{s,t} := \Big(\int_\tau^s g(X_r) \mathrm{d}B^i_r \Big) h(Y_s) B^j_{s,t}$$
$$+ g(X_s) h(Y_s) \mathbb{B}^{ij}_{s,t} + \sum_{k,l} \Big(\int_\tau^s g(X_r) \mathrm{d}B^i_r \Big) \partial_k h(Y_s) (Y'_s)^{kl} \mathbb{B}^{lj}_{s,t}.$$

By definition of the rough integral, we have

$$\int_s^t \Big(\int_\tau^{r_2} g(X_{r_1})\mathrm{d}B^i_{r_1}\Big)h(Y_{r_2})\mathrm{d}B^j_{r_2} = \lim_{\substack{\pi \text{ is a partition of } [s,t],\\ |\pi|\to 0}} \sum_{[u,v]\in\pi} \tilde{A}'_{u,v}.$$

Our goal is to prove $|\tilde{A}_{s,t} - \tilde{A}'_{s,t}| \lesssim_\omega (t-s)^{3H}$. We observe

$$\begin{aligned}
&\tilde{A}_{s,t} - \tilde{A}'_{s,t} \\
&\quad = \Big(\int_\tau^s g(X_r)\mathrm{d}B^i_r\Big)\Big(\int_s^t h(Y_r)\mathrm{d}B^j_r - h(Y_s)B^j_{s,t} - \sum_{k,l} \partial_k h(Y_s)(Y'_s)^{kl}\mathbb{B}^{lj}_{s,t}\Big) \\
&\qquad + g(X_s)\Big(\int_s^t B^i_{s,r} h(Y_r)\mathrm{d}B^j_r - h(Y_s)\mathbb{B}^{ij}_{s,r}\Big).
\end{aligned}$$

By the fundamental estimate of the rough integral, we have

$$\Big|\int_s^t h(Y_r)\mathrm{d}B^j_r - h(Y_s)B^j_{s,t} - \sum_{k,l}\partial_k h(Y_s)(Y'_s)^{kl}\mathbb{B}^{lj}_{s,t}\Big| \lesssim_\omega (t-s)^{3H},$$

$$\Big|\int_s^t B^i_{s,r}h(Y_r)\mathrm{d}B^j_r - h(Y_s)\mathbb{B}^{ij}_{s,r}\Big| \lesssim_\omega (t-s)^{3H}.$$

Hence, $\tilde{A}_{s,t}$ is a correct approximation.

Step 2. Again we do not need the stochastic sewing; the usual sewing in $L^p(\mathbb{P})$ suffices. We compute

$$\begin{aligned}
\delta\tilde{A}_{s,u,t} = -\Big(\int_s^u g(X_r)\mathrm{d}B^i_r - g(X_s)B^i_{s,u}\Big)\int_u^t h(Y_r)\mathrm{d}B^j_r \\
- g(X)_{s,u}\int_u^t h(Y_r)B^i_{u,r}\mathrm{d}B^j_r.
\end{aligned}$$

By Theorem 3.4.4 and Lemma 3.4.5,

$$\Big\|\int_s^u g(X_r)\mathrm{d}B^i_r - g(X_s)B^i_{s,u}\Big\|_{2p} \lesssim \|g\|_{C^\gamma}(1+|||X|||_{2(1+\varepsilon)p})(t-s)^{(1+\gamma)H},$$

$$\Big\|\int_u^t h(Y_r)\mathrm{d}B^j_r\Big\|_{2p} \lesssim \|h\|_{C^\gamma}(1+|||Y|||_{2(1+\varepsilon)p})(t-s)^{H},$$

$$\|g(X)_{s,u}\|_{2p} \lesssim \|g\|_{C^\gamma}|||X|||_{2p}(t-s)^{\gamma H},$$

$$\Big\|\int_u^t h(Y_r)B^i_{u,r}\mathrm{d}B^j_r\Big\|_{2p} \lesssim \|h\|_{C^\gamma}(1+|||Y|||_{2(1+\varepsilon)p})(t-s)^{2H}.$$

Since $(2+\gamma)H > 1$, we obtain

$$\Big\| \int_s^t \Big(\int_s^{r_2} g(X_{r_1}) \mathrm{d}B^i_{r_1} \Big) h(Y_{r_2}) \mathrm{d}B^j_{r_2} - \tilde{A}_{s,t} \Big\|_p \\ \lesssim \|g\|_{C^\gamma} \|h\|_{C^\gamma} (1 + \|\!|X|\!\|_{2(1+\varepsilon)p})(1 + \|\!|Y|\!\|_{2(1+\varepsilon)p})(t-s)^{(2+\gamma)H}.$$

Setting $\tau = s$, we get

$$\Big\| \int_s^t \Big(\int_s^{r_2} g(X_{r_1}) \mathrm{d}B^i_{r_1} \Big) h(Y_{r_2}) \mathrm{d}B^j_{r_2} - g(X_s) \int_s^t h(Y_r) B^i_{s,r} \mathrm{d}B^j_r \Big\|_p \\ \lesssim \|g\|_{C^\gamma} \|h\|_{C^\gamma} (1 + \|\!|X|\!\|_{2(1+\varepsilon)p})(1 + \|\!|Y|\!\|_{2(1+\varepsilon)p})(t-s)^{(2+\gamma)H}.$$

By Lemma 3.4.5,

$$\Big\| \int_s^t h(Y_r) B^i_{s,r} \mathrm{d}B^j_r - h(Y_s) \mathbb{B}^{i,j}_{s,t} \Big\|_p \lesssim \|h\|_{C^\gamma} (1 + \|\!|X|\!\|_{2(1+\varepsilon)p})(t-s)^{(2+\gamma)H},$$

and the proof is complete. □

3.4.3 Pathwise uniqueness

Our final result is the following.

Theorem 3.4.7. *Let $H \in (1/3, 1/2)$, $\gamma \in (\frac{1-H}{H}, 2)$ and $\sigma \in C^\gamma$. Suppose that $\sigma\sigma^{\mathrm{T}}$ is uniformly elliptic. Then, pathwise uniqueness holds for* (3.1).

Before going into the proof, we need some preparations. Let ρ be a smooth map from $\mathbb{R}^{d_1}$ to the space of $d_1 \times d_2$ matrices, and let (X, X'), (Y, Y') be paths controlled by B. We set

$$G^k_t := G[\rho]^k_t := \int_0^t \Big(\int_0^1 \partial_k \rho(\theta X_r + (1-\theta) Y_r) \mathrm{d}\theta \Big) \mathrm{d}B_r.$$

The integral is understood as the rough integral, since the path

$$t \mapsto \int_0^1 \partial_k \rho(\theta X_t + (1-\theta) Y_t) \mathrm{d}\theta$$

is controlled by B with Gubinelli derivative

$$t \mapsto \int_0^1 \nabla \partial_k \rho(\theta X_t + (1-\theta) Y_t)(\theta X'_t + (1-\theta) Y'_t) \mathrm{d}\theta.$$

We furthermore set

$$\mathcal{G}^k_{s,t} := \mathcal{G}[\rho]^k_{s,t} = \int_s^t \Big(\int_0^1 \partial_k \rho(\theta X_t + (1-\theta) Y_t) \mathrm{d}\theta \Big) B_{s,r} \otimes \mathrm{d}B_r.$$

We set $\mathbf{G}^k := (G^k, \mathcal{G}^k)$. Note that for $\alpha \in (1/3, H)$ we have

$$|||\mathbf{G}^k|||_\alpha := \sup_{0 \le s < t \le T} \Big\{ \frac{|G^k_{s,t}|}{(t-s)^\alpha} + \frac{|\mathcal{G}^k_{s,t}|}{(t-s)^{\gamma\alpha}} \Big\} < \infty, \tag{3.47}$$

for any $\gamma \in (\frac{1-H}{H}, 2)$. We choose α so close to H to have

$$(1+\gamma)\alpha > 1.$$

Finally, we observe the following *modified* Chen's relation

$$\mathcal{G}^k_{s,t} = \mathcal{G}^k_{s,u} + \mathcal{G}^k_{u,t} + B_{s,u} \otimes G^k_{u,t}. \tag{3.48}$$

More abstractly, we can consider any pair $\mathbf{G} = (G, \mathcal{G})$ satisfying the analytic condition (3.47) and the algebraic condition (3.48). Let Z be a path controlled by B. It is not difficult to see (simply by repeating the arguments in the usual rough path setting, e.g. [FH20, Theorem 4.10]) that the integral

$$\int_s^t Z_r \mathrm{d}\mathbf{G}_r := \lim_{|\pi| \to 0} \sum_{[u,v] \in \pi} (Z_u G_{u,v} + Z'_u \mathcal{G}_{u,v})$$

exists, and we have the quantitative estimate

$$\Big| \int_s^t Z_r \mathrm{d}\mathbf{G}_r - Z_s G_{s,t} - Z'_s \mathcal{G}_{s,t} \Big| \lesssim \|Z\|_{\mathcal{D}^{\gamma\alpha}} |||\mathbf{G}|||_\alpha (t-s)^{(1+\gamma)\alpha}. \tag{3.49}$$

Furthermore, as in [FH20, Theorem 4.17], we have the stability estimate: by setting

$$d_\alpha(\mathbf{G}, \bar{\mathbf{G}}) := \sup_{0 \le s < t \le T} \Big(\frac{|G_{s,t} - \bar{G}_{s,t}|}{(t-s)^\alpha} + \frac{|\mathcal{G}_{s,t} - \bar{\mathcal{G}}_{s,t}|}{(t-s)^{\gamma\alpha}} \Big), \tag{3.50}$$

and provided that

$$|||\mathbf{G}|||_\alpha + |||\bar{\mathbf{G}}|||_\alpha \le M$$

for some $M \ge 1$, we have

$$\Big| \int_s^t Z_r \mathrm{d}\mathbf{G}_r - \int_s^t Z_r \mathrm{d}\bar{\mathbf{G}}_r \Big| \lesssim_M \|Z\|_{\mathcal{D}^{\gamma\alpha}} d_\alpha(\mathbf{G}, \bar{\mathbf{G}}) (t-s)^\alpha. \tag{3.51}$$

To rigorously derive the identity (3.31), we use the following.

Lemma 3.4.8. *Let ρ be a smooth map from $\mathbb{R}^{d_1}$ to the space of $d_1 \times d_2$ matrices, and let $(X, X'), (Y, Y')$ be paths controlled by B. We then have*

$$\int_0^t \{\rho(X_r) - \rho(Y_r)\} \mathrm{d}B_r = \sum_{k=1}^{d_1} \int_0^t (X_r^k - Y_r^k) \mathrm{d}\mathbf{G}_r^k[\rho].$$

Proof. We set

$$\begin{aligned} A^1_{s,t} &:= \{\rho(X_s) - \rho(Y_s)\} B_{s,t} + \{\rho(X_s) X'_s - \rho(Y_s) Y'_s\} \mathbb{B}_{s,t}, \\ A^2_{s,t} &:= \sum_{k=1}^{d_1} (X_s^k - Y_s^k) G^k_{s,t} + \{(X'_s)^k - (Y'_s)^k\} \mathcal{G}^k_{s,t}, \end{aligned}$$

where $(Z'_s)^k$ is the kth row of the matrix Z'_s ($Z \in \{X, Y\}$). Note that

$$\begin{aligned} \int_0^t \{\rho(X_r) - \rho(Y_r)\} \mathrm{d}B_r &= \lim_{\substack{\pi \text{ is a partition of } [0,t], \\ |\pi| \to 0}} \sum_{[u,v] \in \pi} A^1_{u,v}, \\ \sum_{k=1}^{d_1} \int_0^t (X_r^k - Y_r^k) \mathrm{d}\mathbf{G}_r^k[\rho] &= \lim_{\substack{\pi \text{ is a partition of } [0,t], \\ |\pi| \to 0}} \sum_{[u,v] \in \pi} A^2_{u,v}. \end{aligned}$$

Therefore, it suffices to show $|A^1_{s,t} - A^2_{s,t}| \lesssim (t-s)^{3H}$. By the mean value theorem for integrals,

$$\rho(X_s) - \rho(Y_s) = \sum_{k=1}^{d_1} \int_0^1 \partial_k \rho(\theta X_s + (1-\theta) Y_s)(X_s^k - Y_s^k) \mathrm{d}\theta$$

and

$$\begin{aligned} &\nabla\rho(X_s) X'_s - \nabla\rho(Y_s) Y'_s \\ &\quad = \sum_{k=1}^{d_1} \Big\{ \int_0^1 (X_s^k - Y_s^k) \partial_k \nabla\rho(\theta X_s + (1-\theta) Y_s)(\theta X'_s + (1-\theta) Y'_s) \mathrm{d}\theta \\ &\qquad\qquad + \int_0^1 \nabla\rho(\theta X_s + (1-\theta) Y_s)(X'_s - Y'_s) \mathrm{d}\theta \Big\}. \end{aligned}$$

We compute

$$A^2_{s,t} - A^1_{s,t} = \sum_{k=1}^{d_1} \big[(X_s^k - Y_s^k) \Delta^{1,k}_{s,t} + \{(X'_s)^k - (Y'_s)^k\} \Delta^{2,k}_{s,t} \big],$$

where

$$\Delta^{1,k}_{s,t} := \int_s^t \Big(\int_0^1 \partial_k \rho(\theta X_r + (1-\theta)Y_r)\mathrm{d}\theta \Big)\mathrm{d}B_r - \Big(\int_0^1 \partial_k \rho(\theta X_s + (1-\theta)Y_s)\mathrm{d}\theta \Big) B_{s,t} - \Big(\int_0^1 \partial_k \nabla\rho(\theta X_s + (1-\theta)Y_s)(\theta X'_s + (1-\theta)Y'_s)\mathrm{d}\theta \Big) \mathbb{B}_{s,t}$$

and

$$\Delta^{2,k}_{s,t} := \int_s^t \Big(\int_0^1 \partial_k \rho(\theta X_r + (1-\theta)Y_r)\mathrm{d}\theta \Big) B_{s,r} \otimes \mathrm{d}B_r - \Big(\int_0^1 \partial_k \rho(\theta X_s + (1-\theta)Y_s)\mathrm{d}\theta \Big) \mathbb{B}_{s,t}.$$

By the fundamental estimate of the rough integral, we have

$$|\Delta^{1,k}_{s,t}| + |\Delta^{2,k}_{s,t}| \lesssim |t-s|^{3H},$$

and the proof is complete. □

Proof of Theorem 3.4.7. We assume that (3.27) holds, as the general case can be handled as in the proof of Theorem 3.3.3 by stopping time arguments. Let (X, Y) be two solutions to (3.1). Let $(\sigma_n)_{n=1}^\infty$ be a smooth approximation to σ. We have

$$X_t - Y_t = \lim_{n\to\infty} \int_0^t \{\sigma_n(X_r) - \sigma_n(Y_r)\}\mathrm{d}B_r.$$

By Lemma 3.4.8,

$$X_t - Y_t = \lim_{n\to\infty} \sum_{k=1}^{d_1} \int_0^t (X^k_r - Y^k_r)\mathrm{d}\mathbf{G}^k_r[\sigma_n].$$

By repeating the argument of Lemma 3.3.4, we can show that $(X, \sigma(X)) \in \mathfrak{D}^{1,K}$ and similarly for Y. Therefore, by Theorem 3.4.4, Lemma 3.4.5 and the Kolmogorov criterion for rough paths ([FH20, Theorem 3.1]), the sequence of the lifts $\mathbf{G}^k[\sigma_n]$ converges in $L^p(\mathbb{P})$ to some limit $\mathbf{G}^k$ in the metric d_α defined by (3.50). The stability estimate (3.51) yields

$$X_t - Y_t = \sum_{k=1}^{d_1} \int_0^t (X^k_r - Y^k_r)\mathrm{d}\mathbf{G}^k_r.$$

The remainder estimate (3.49) shows that $X-Y$ is controlled by $(G^k)_{k=1}^{d_1}$ with Gubinelli derivative $X-Y$. Moreover, by Lemma 3.4.6 and the Kolmogorov criterion for rough paths, the sequence of the rough paths

$$\hat{\mathbf{G}}[\sigma_n] := \Big((G^k[\sigma_n])_{k=1}^{d_1}, \Big(\int G^k[\sigma_n] \otimes \mathrm{d}G^l[\sigma_n]\Big)_{k,l=1}^{d_1}\Big)$$

converges to some limit $\hat{\mathbf{G}}$. We claim that

$$\int_0^t (X_r^k - Y_r^k)\mathrm{d}\mathbf{G}_r^k = \int_0^t (X_r^k - Y_r^k)\mathrm{d}\hat{\mathbf{G}}_r^k \quad \text{a.s.} \tag{3.52}$$

Indeed, by definition we have

$$\int_0^t (X_r^k - Y_r^k)\mathrm{d}\mathbf{G}_r^k - \int_0^t (X_r^k - Y_r^k)\mathrm{d}\hat{\mathbf{G}}_r^k = \lim_{\substack{\pi \text{ is a partition of } [0,t],\\ |\pi|\to 0}} \sum_{[u,v]\in\pi} \Delta_{u,v},$$

where

$$\begin{aligned}\Delta_{s,t} &:= (\sigma(X_s) - \sigma(Y_s))\mathcal{G}_{s,t}^k - \sum_{l=1}^{d_1}(X_s^l - Y_s^l)\int_s^t G_{s,r}^l \otimes \mathrm{d}G_r^k \\ &= \sum_{l=1}^{d_1}(X_s^l - Y_s^l)\Big\{\Big(\int_0^1 \partial_l\sigma(\theta X_s + (1-\theta)Y_s)\mathrm{d}\theta\Big)\mathcal{G}_{s,t}^k - \int_s^t G_{s,r}^l \otimes \mathrm{d}G_r^k\Big\}.\end{aligned}$$

By Lemma 3.4.5 and Lemma 3.4.6,

$$\Big\|\Big(\int_0^1 \partial_l\sigma(\theta X_s + (1-\theta)Y_s)\mathrm{d}\theta\Big)\mathcal{G}_{s,t}^k - \int_s^t G_{s,r}^l \otimes \mathrm{d}G_r^k\Big\|_p \lesssim (t-s)^{(1+\gamma)H},$$

and as $(1+\gamma)H > 1$ the identity (3.52) is established.

Hence, $X-Y$ solves the linear rough differential equation

$$\mathrm{d}Z_t = \sum_{k=1}^{d_1} Z_t^k \mathrm{d}\hat{\mathbf{G}}_t^k, \quad Z_0 = 0,$$

and its uniqueness implies that $X = Y$. $\square$

Bibliography

[Ath+21] S. Athreya, O. Butkovsky, K. Lê, and L. Mytnik. *Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation*. 2021. arXiv: 2011.13498 [math.PR].

[BNP20] D. Baños, T. Nilssen, and F. Proske. "Strong Existence and Higher Order Fréchet Differentiability of Stochastic Flows of Fractional Brownian Motion Driven SDEs with Singular Drift". *Journal of Dynamics and Differential Equations* 32.4 (2020), pp. 1819–1866.

[BDG21] O. Butkovsky, K. Dareiotis, and M. Gerencsér. "Approximation of SDEs: a stochastic sewing approach". *Probability Theory and Related Fields* 181.4 (2021), pp. 975–1034.

[BLM23] O. Butkovsky, K. Lê, and L. Mytnik. *Stochastic equations with singular drift driven by fractional Brownian motion*. 2023. arXiv: 2302.11937 [math.PR].

[CG16] R. Catellier and M. Gubinelli. "Averaging along irregular curves and regularisation of ODEs". *Stochastic Processes and their Applications* 126.8 (2016), pp. 2323–2366.

[CD22] R. Catellier and R. Duboscq. *Regularization by noise for rough differential equations driven by Gaussian rough paths*. 2022. arXiv: 2207.04251 [math.PR].

[CQ02] L. Coutin and Z. Qian. "Stochastic analysis, rough path analysis and fractional Brownian motions". *Probability Theory and Related Fields* 122.1 (2002), pp. 108–140.

[DG22] K. Dareiotis and M. Gerencsér. *Path-by-path regularisation through multiplicative noise in rough, Young, and ordinary differential equations*. 2022. arXiv: 2207.03476 [math.PR].

[DGL21] K. Dareiotis, M. Gerencsér, and K. Lê. *Quantifying a convergence theorem of Gyöngy and Krylov*. 2021. arXiv: 2101.12185 [math.PR].

[Dav07] A. M. Davie. "Uniqueness of Solutions of Stochastic Differential Equations". *International Mathematics Research Notices* 2007 (Dec. 2007), rnm124.

[Dav08] A. M. Davie. "Differential equations driven by rough paths: an approach via discrete approximation". *Appl. Math. Res. Express. AMRX* (2008), Art. ID abm009, 40.

[Föl81] H. Föllmer. “Calcul d’Ito sans probabilités”. *Séminaire de probabilités de Strasbourg* 15 (1981), pp. 143–150.

[FH20] P. K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Cham: Springer International Publishing, 2020.

[FHL23] P. K. Friz, A. Hocquet, and K. Lê. *Rough stochastic differential equations*. 2023. arXiv: 2106.10340 [math.PR].

[Gal21] L. Galeati. “Nonlinear Young Differential Equations: A Review”. *Journal of Dynamics and Differential Equations* (2021).

[GG23] L. Galeati and M. Gerencsér. *Solution theory of fractional SDEs in complete subcritical regimes*. 2023. arXiv: 2207.03475 [math.PR].

[GG22] L. Galeati and M. Gubinelli. “Noiseless regularisation by noise”. *Rev. Mat. Iberoam.* 38.2 (2022), pp. 433–502.

[GH22] L. Galeati and F. A. Harang. “Regularization of multiplicative SDEs through additive noise”. *The Annals of Applied Probability* 32.5 (Oct. 2022), pp. 3930–3963.

[GHM22] L. Galeati, F. A. Harang, and A. Mayorcas. “Distribution dependent SDEs driven by additive fractional Brownian motion”. *Probability Theory and Related Fields* (2022).

[Ger22] M. Gerencsér. “Regularisation by regular noise”. *Stochastics and Partial Differential Equations: Analysis and Computations* (2022).

[Gub04] M. Gubinelli. “Controlling rough paths”. *J. Funct. Anal.* 216.1 (2004), pp. 86–140.

[GIP15] M. Gubinelli, P. Imkeller, and N. Perkowski. “Paracontrolled distributions and singular PDEs”. *Forum of Mathematics, Pi* 3 (2015), e6.

[Hai14] M. Hairer. “A theory of regularity structures”. *Invent. Math.* 198.2 (2014), pp. 269–504.

[HP21] F. A. Harang and N. Perkowski. “C^∞ regularization of ODEs perturbed by noise”. *Stochastics and Dynamics* 21.08 (2021), p. 2140010.

[HTV22] M. Hinz, J. M. Tölle, and L. Viitasaari. *Variability of paths and differential equations with BV-coefficients*. 2022. arXiv: 2003.11698 [math.PR].

[Kal21] O. Kallenberg. *Foundations of modern probability*. Vol. 99. Probability Theory and Stochastic Modelling. Third edition. Springer, Cham, 2021.

[Kur07] T. G. Kurtz. “The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities”. *Electron. J. Probab* 12 (2007), pp. 951–965.

[Lê20] K. Lê. “A stochastic sewing lemma and applications”. *Electronic Journal of Probability* 25.none (Jan. 2020), pp. 1–55.

[LL22] K. Lê and C. Ling. *Taming singular stochastic differential equations: A numerical method*. 2022. arXiv: 2110.01343 [math.PR].

[Lyo94] T. Lyons. “Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young”. *Math. Res. Lett.* 1.4 (1994), pp. 451–464.

[Lyo98] T. J. Lyons. “Differential equations driven by rough signals”. *Rev. Mat. Iberoamericana* 14.2 (1998), pp. 215–310.

[MV68] B. B. Mandelbrot and J. W. Van Ness. “Fractional Brownian Motions, Fractional Noises and Applications”. *SIAM Review* 10.4 (1968), pp. 422–437.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*. Second. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006, pp. xiv+382.

[NR02] D. Nualart and A. Răşcanu. “Differential equations driven by fractional Brownian motion”. *Collect. Math.* 53.1 (2002), pp. 55–81.

[Tri92] H. Triebel. *Theory of function spaces. II*. Vol. 84. Monographs in Mathematics. Birkhäuser Verlag, Basel, 1992, pp. viii+370.

[You36] L. C. Young. “An inequality of the Hölder type, connected with Stieltjes integration”. *Acta Math.* 67.1 (1936), pp. 251–282.

[Zäh98] M. Zähle. “Integration with respect to fractal functions and stochastic calculus. I”. *Probab. Theory Related Fields* 111.3 (1998), pp. 333–374.

BIBLIOGRAPHY

Chapter 4

Strong regularization of differential equations with integrable drifts by fractional noise

We prove well-posedness and stability for stochastic differential equations with integrable time-dependent drift driven by additive fractional Brownian noise whose Hurst parameter is less than $1/2$. Our result can be considered as an extension of that from Krylov and Röckner [KR05] for Brownian motion. It holds under the entire subcritical regime observed earlier by Galeati and Gerencsér [GG23], and improves upon previous results of Nualart and Ouknine [NO03] for dimension one and of Lê [Lê20]. Our methods are built around Lyons' rough path theory, Girsanov's theorem, the stochastic sewing lemma, and the quantitative John–Nirenberg inequality for stochastic processes of vanishing mean oscillation.

This chapter is based on joint work with Oleg Butkovsky and Khoa Lê.

Keywords and phrases. Stochastic differential equations, fractional Brownian motion, regularization by noise, stochastic sewing, processes of vanishing mean oscillation, rough paths.
MSC 2020 – 60H10, 60H50, 60G22, 60L20.

Contents

4.1 Introduction **144**

4.1.1 Literatures and discussions 148

4.1.2 Strategies of proofs 149

4.2 Weak well-posedness **152**

4.2.1 Preliminaries 153

4.2.2 Weak solutions and stochastic estimates 156

4.3 Strong well-posedness **165**

4.3.1 Some technical estimates 166

4.3.2 Pathwise uniqueness 168

4.4 Stability **174**

4.4.1 VMO estimates 177

4.4.2 Lifting paths 181

4.4.3 Gronwall's argument in the probabilistic setting 182

Bibliography **187**

4.1 Introduction

We consider the stochastic differential equation (SDE)

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \mathrm{d}B_t^H, \quad X_0 = x \in \mathbb{R}^d, \tag{4.1}$$

where $b \in L_t^q L_x^p := L^q([0,T]; L^p(\mathbb{R}^d))$ and B^H is a fractional Brownian motion with Hurst parameter H on a fixed filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$. The aim of this chapter is to prove well-posedness of (4.1) and its stability under the following conditions:

$$p, q \in [1, \infty], \quad H \in \left(0, \frac{1}{2}\right), \tag{4.2a}$$

$$\frac{dH}{p} + \frac{1}{q} < 1 - H, \tag{4.2b}$$

$$p \geq 2dH \tag{4.2c}$$

and

$$\frac{dH}{p} < \frac{p^{-1} - q^{-1}}{p^{-1} - 2^{-1}}\Big(\frac{1}{2} - H\Big) \quad \text{if } p < (1-H)^{-1} \text{ and } q > 2. \tag{4.2d}$$

Theorem 4.1.1 (Well-posedness)**.** *Weak existence, weak uniqueness, pathwise uniqueness and strong existence for* (4.1) *hold under* (4.2)*.*

For a function $b \in L_t^q L_x^p$, M_b is the smallest constant M satisfying

$$\|b\mathbf{1}_{\{|b|\geq M\}}\|_{L_t^q L_x^p} \leq \mathfrak{c}, \tag{4.3}$$

for some constant $\mathfrak{c} = \mathfrak{c}(d, H, p, q, T)$, precisely defined in (4.60).

Theorem 4.1.2 (Stability)**.** *Suppose that the condition* (4.2) *is satisfied. For each* $i \in \{1, 2\}$*, let* $x^i \in \mathbb{R}^d$ *and* $b^i \in L_t^q L_x^p$*, and let* X^i *solve* (4.1) *with* (b, x) *replaced by* (b^i, x^i)*. Then we have the following stability estimates.*

(i) (Pathwise stability) *There exist some non-negative random variables* A *and* D *such that*

$$\begin{aligned}\|A\|_{L^m(\mathbb{P})} &\lesssim_{d,p,q,H,T,m,M_{b^1},M_{b^2}} 1 + \|b^1\|_{L_t^q L_x^p},\\ \|D\|_{L^m(\mathbb{P})} &\lesssim_{d,p,q,H,T,m,M_{b^1},M_{b^2}} \|b^1 - b^2\|_{L_t^q W_x^{-1,p}}\end{aligned}$$

for all $m \in (0, \infty)$*, and*

$$\sup_{t\in[0,T]} |X_t^1 - X_t^2| \leq e^A(|x^1 - x^2| + D) \quad \textit{almost surely}.$$

(ii) (Stability in moment norms) *For every* $m \in (0, \infty)$ *and*

$$\gamma \in \Big(0, 1 - H - \frac{dH}{p} - \frac{1}{q}\Big),$$

we have

$$\|\|X^1 - X^2\|_{C^\gamma}\|_{L^m(\mathbb{P})} \lesssim_{d,p,q,H,T,M_{b^1},M_{b^2},m,\gamma} |x^1 - x^2| + \|b^1 - b^2\|_{L_t^q W_x^{-1,p}}.$$

In the above theorem and throughout the article, C^γ is the space of $\mathbb{R}^d$-valued functions on $[0, T]$ with Hölder regularity γ and $W^{-1,p}$ is the Sobolev space on $\mathbb{R}^d$ of regularity -1 and integrability p.

Remark 4.1.3. In the publication version, we will use Theorem 4.1.2 to establish so called path-by-path uniqueness, as in [ALL23].

Remark 4.1.4. When $p > 2Hd$, we can choose $\theta > 1$ so that $(\bar{p}, \bar{q}) = (\theta^{-1}p, \theta^{-1}q)$ still satisfies (4.2). For any M, we have

$$\|b^i \mathbf{1}_{\{|b^i| \geq M\}}\|_{L_t^{\bar{q}} L_x^{\bar{p}}} \leq M^{1-\theta} \|b^i\|_{L_t^q L_x^p}^{\theta}.$$

It is evident that there exists a constant $M_i = M_i(p, q, H, d, \|b^i\|_{L_t^q L_x^p})$ such that the condition $\|b^i \mathbf{1}_{\{|b^i| \geq M\}}\|_{L_t^{\bar{q}} L_x^{\bar{p}}} \leq \mathfrak{c}$ becomes trivial. In this case, the stability estimates in Theorem 4.1.2 are uniform with respect to the size of $\max_i \|b^i\|_{L_t^q L_x^p}$.

When $p = 2Hd$, such choices are no longer possible and it is not expected to have uniform stability estimates with respect to the norms of b^1, b^2. This phenomenon also happens for SDEs with Brownian motion and drift $b \in L_x^d$, see [Kry21, Theorem 3.7].

On parameters

If $p \geq \max\{2dH, (1-H)^{-1}\}$, our results hold under the entire *subcritical regime* (4.2b) under which the fractional noise dominates at small scales. Indeed, following Galeati and Gerencsér in [GG23], if X solves the SDE (4.1), then the scaled process $X_t^{(\lambda)} := \lambda^{-H} X_{\lambda t}$ solves the SDE

$$\mathrm{d}X_t^{(\lambda)} = \lambda^{1-H-\frac{dH}{p}-\frac{1}{q}} b^{(\lambda)}(t, X_t^{(\lambda)})\mathrm{d}t + \mathrm{d}B_t^{(\lambda)}, \tag{4.4}$$

where

$$b^{(\lambda)}(t,x) := \lambda^{\frac{dH}{p}+\frac{1}{q}} b(\lambda t, \lambda^H x), \quad B_t^{(\lambda)} := \lambda^{-H} B_{\lambda t}.$$

We note that $\|b^{(\lambda)}\|_{L_t^q L_x^p} = \|b\|_{L_t^q L_x^p}$ and that $B^{(\lambda)}$ is equal to B in law. Domination of the noise at small scales entails that the order of the drift term is smaller than that of the driving noise as λ vanishes. This enforces that

$$1 - H - \frac{dH}{p} - \frac{1}{q} > 0,$$

which is exactly the last condition (4.2b).

Let us explain technical conditions (4.2c) and (4.2d). Both conditions are due to Girsanov's arguments. The condition (4.2d) ensures that the drift term

$$t \mapsto \int_0^t b_r(X_r)\mathrm{d}r$$

belongs to the Cameron–Martin space $\mathcal{H}^H$ of the fractional Brownian motion, and the condition (4.2c) ensures that the Cameron–Martin norm of the drift satisfies the strong Novikov's condition, that is

$$\mathbb{E}\Big[\exp\Big(\lambda\Big\|\int_0^\cdot b_r(X_r)\mathrm{d}r\Big\|_{\mathcal{H}^H}^2\Big)\Big] < \infty, \quad \forall\lambda \geq 0.$$

If $q = \infty$, which includes the case where b is time-independent, our condition (4.2) simplifies as follows. The condition (4.2d) becomes

$$H + \frac{Hd}{p} < \frac{1}{2} + \frac{Hd}{2}, \quad p < (1-H)^{-1}. \tag{4.5}$$

Hence, the condition (4.2) reduces to $H \in (0, \frac{1}{2})$ and

$$p \in \begin{cases} [1,\infty] \cap (2H/(1-H), \infty], & \text{if } d = 1, \\ [\max\{1, 4H\}, \infty] \setminus \{4H\}, & \text{if } d = 2, \\ [2dH, \infty], & \text{if } d \geq 3. \end{cases}$$

In particular, for $d = 1$ it allows $b \in L_x^1$ provided $H < \frac{1}{3}$.

Krylov and Röckner in [KR05] obtained similar results for the Brownian motion. Setting $H = \frac{1}{2}$ in our condition (4.2), we recover their condition

$$\frac{d}{2p} + \frac{1}{q} < \frac{1}{2}, \quad p \geq 2. \tag{4.6}$$

In this sense, our results can be viewed as an extension of their work when the driving process is a fractional Brownian with $H < 1/2$. In the case when $H > 1/2$, we expect that certain positive regularity, say in Sobolev scale, is necessary for uniqueness of strong solutions. However, this situation will not be considered herein.

Organization of the chapter

The rest of the chapter is organized as follows. In Section 4.1.1, we review some literatures and discuss their connections with our results. In Section 4.1.2, we explain our strategy to prove Theorem 4.1.1 and Theorem 4.1.2. Section 4.2 reviews basics of SDEs with fractional Brownian motion and prove weak well-posedness. In Section 4.3 we prove Theorem 4.1.1, and in Section 4.4 we prove Theorem 4.1.2.

4.1.1 Literatures and discussions

Ill-posed differential equations sometimes become well-posed by an addition of an irregular noise. This phenomenon is called *regularization by noise*. The study of regularization by Brownian noise has a long history and can be dated back to classical works of Zvonkin [Zvo74] and Veterennikov [Ver81]. In the seminal work [KR05] of Krylov and Röckner, the first result on strong well-posedness with $L_t^q L_x^p$-drift in arbitrary finite dimensions is obtained. More precisely, they prove strong well-posedness (pathwise uniqueness and strong existence) of (4.1) with $H = \frac{1}{2}$ under (4.6). The work [KR05] has inspired many others, notably Zhang's extension to multiplicative noise in [Zha10]. More recently, the critical case is addresses by Röckner and Zhao in [RZ23; RZ21] and independently by Krylov in [Kry21]. Other directions include stability and numerical approximations with sharp rates, as discussed in [LL22; GL23; Lê22].

Surprisingly, prior to [KR05], Nualart and Ouknine [NO03], building upon their earlier work [NO02], showed regularization by fractional noise for SDEs with integrable drifts although only in dimension one. Their arguments are based on Girsanov's theorem and the comparison principle, the latter being specific to dimension one. It is worth noting that the arguments from [NO03] also works for the case of driving Brownian noise. Yet, due to the use of the comparison principle, their arguments cannot translate to higher dimensional settings. Proving regularization by fractional noise in multi-dimensions requires more robust methods, which forego both one-dimensional and Markovian techniques.

Different approaches have been developed for regularization by fractional Brownian noise in multi-dimensions. Catellier and Gubinelli in [CG16], building upon [Dav07], introduce a path-by-path approach lying on the framework of nonlinear Young differential equations. Baños, Nilssen and Proske in [BNP20] employ tools from Malliavin calculus to construct strong solutions. The Lê in [Lê20] initiates the *stochastic sewing* techniques and proves the strong well-posedness of (4.1) provided that

$$\frac{dH}{p} + \frac{1}{q} < \frac{1}{2} - H, \quad p \geq 2, \quad q > 2. \tag{4.7}$$

Stochastic sewing techniques are also used in [GG23] to show strong well-posedness for equation (4.1) where the drift belongs to $L_t^q \mathbb{C}_x^\alpha$ for $\alpha \in \mathbb{R}$ and $q \in (1, \infty]$ satisfying

$$-H\alpha + \frac{1}{\min(q,2)} < 1 - H. \tag{4.8}$$

In view of the embedding $L^p(\mathbb{R}^d) \hookrightarrow \mathbb{C}^{-d/p}(\mathbb{R}^d)$, their results allow drifts in $L_t^q L_x^p$ with $p \in [1, \infty]$, $q \in (1, \infty]$ satisfying

$$\frac{dH}{p} + \frac{1}{\min(q,2)} < 1 - H. \tag{4.9}$$

Yet, this is still not optimal compared to (4.2) because for $q > 2$, the condition (4.9) renders to $\frac{dH}{p} < \frac{1}{2} - H$. The arguments from [GG23] are not applicable under (4.2). In fact, for $q > 2$, any similar strategy using the embedding $L^p(\mathbb{R}^d) \hookrightarrow \mathbb{C}^{-d/p}(\mathbb{R}^d)$ would eventually lead to the non-optimal condition (4.7). Therefore, our results are complementary to [GG23] and improve upon the previous related works.

Stability estimates of SDEs with irregular drift play an important intermediate role to derive other applications in distribution dependent SDEs [GH22], reflected SDEs [GM23], deriving sharp strong convergence rate of numerical approximations [DGL21; LL22; Lê22]. For Brownian noise, a typical approach to derive stability is the Zvonkin transformation, [DGL21; LL22; GL23]. For fractional Brownian noise, Zvonkin transformation is no longer available, nevertheless, one can utilize Hölder regularity and deduce stability results from that of Young–Lyons equations ([Lyo94]). This connection was observed in [Lê20] and developed further in [Ath+21; GH22; GG23; GM23]. In our situation, the relevant Hölder exponent vanishes, and we propose two different ways to capture regularity, one is to use pathwise variations and the other is to use stochastic mean oscillations in combination with Lyons' signature. Using pathwise variations, we can resort to stability properties of Young–Lyons equations once again to obtain some desirable stability estimates in pathwise sense (but not in any moment norm, because of some issues of exponential integrability). Using the latter method, we can resort to stability properties of (random) Lyons rough differential equations to obtain stability estimates in any positive moments.

When $H \in (\frac{1}{4}, \frac{1}{2})$, one may consider the multiplicative noise by the rough path theory of Lyons [Lyo98], with recent progress reported by Dareiotis and Gerencsér [DG22] and Catellier and Duboscq [CD22]. Yet another interesting program is to extend our results to stochastic partial differential equations as in [BM19; Ath+21].

4.1.2 Strategies of proofs

Weak existence

It is known that the validity of Girsanov's theorem implies existence of a weak solution. In [NO03; Lê20], using moment calculations, the authors verify Novikov's condition for Girsanov's theorem under the condition

$$\frac{dH}{p} + \frac{1}{q} < \frac{1}{2}, \quad p \geq 2.$$

Instead of moment methods, we apply the quantitative John–Nirenberg inequality for stochastic processes of *vanishing mean oscillation* (VMO) from [Lê22] (or alternatively, the quantitative Khasminski lemma from [LL22]) to establish exponential integrability for certain random integrals related to the Novikov's condition. This allows us to extend the arguments from [NO03; Lê20] under the regime (4.2).

Uniqueness and stability

We continue with the setup in Theorem 4.1.2. The difference $X^\Delta = X^1 - X^2$ of two solutions solves the equation

$$dX^\Delta = X^\Delta \mathrm{d}V + \mathrm{d}R, \quad X_0^\Delta = x^1 - x^2, \tag{4.10}$$

where

$$V_t = \int_0^t \int_0^1 \nabla b^1(r, \theta X_r^1 + (1-\theta)X_r^2)\mathrm{d}\theta \mathrm{d}r, \quad R_t = \int_0^t (b^1 - b^2)(r, X_r^2)\mathrm{d}r.$$

If b^1 has continuous bounded derivative, equation (4.10) is a perturbed ODE, whose known properties can be utilized to derive estimations on X^Δ. When b^1 is irregular, it is noted in [Lê20] that (4.10) can be interpreted as a Young–Lyons differential equation ([Lyo94]). The aforementioned work, however, treated (4.10) in the framework of Hölder continuous functions, which only works under the non-optimal condition (4.7). In contrast, we show that the process V has finite ρ-variation for a ρ which is less than but arbitrarily close to 2. This allows employment of Lyons' approach to (4.10) in [Lyo94] to derive pathwise stability and hence, pathwise uniqueness.

To show that V has the necessary variational regularity, we apply the stochastic sewing lemma with control. A change of measure links V directly to the process

$$U_t := \int_0^t \nabla b_r^1(B_r)\mathrm{d}r. \tag{4.11}$$

The stochastic sewing lemma relates the moment norm of $U_t - U_s$ to that of the germ

$$A_{s,t} := \mathbb{E}_s \int_s^t b_r^1(B_r)\mathrm{d}r = \mathbb{E}_s \int_s^t b_r^1(\mathbb{E}_s(B_r) + (B_r - \mathbb{E}_s B_r))\mathrm{d}r.$$

In [Lê20], moments of $A_{s,t}$ are estimated from above by constant multiples of

$$\|b^1\|_{L_t^q L_x^p}(t-s)^{1-H-\frac{Hd}{p}-\frac{1}{q}}$$

by utilizing the *local nondeterminism* property that $\mathrm{Var}(B_r - \mathbb{E}_s B_r) \gtrsim (r-s)^{2H}$. In our argument, we take into account the randomness of the $\mathcal{F}_s$-measurable part $\mathbb{E}_s B_r$ additionally, which leads to the following estimate with improved variational regularity

$$\|A_{s,t}\|_{L^p(\mathbb{P})} \lesssim s^{-\frac{dH}{p}} \|b^1\|_{L^q([s,t];L_x^p)}(t-s)^{1-H-\frac{1}{q}}.$$

As in [GG23], we make use of the fact that $(s,t) \mapsto \|b^1\|_{L^q([s,t];L^p_x)}$ is a control. The stochastic sewing lemma with control then yields that

$$\|U_t - U_s\|_{L^p(\mathbb{P})} \lesssim s^{-\frac{dH}{p}} \|b^1\|_{L^q([s,t];L^p_x)} (t-s)^{1-H-\frac{1}{q}}. \tag{4.12}$$

When regularity is measured by variations, this estimate has higher regularity counting, $(1-H)$ to be exact. This allows us to show that the path U given by (4.11) has finite ρ-variation for any $\rho > \frac{1}{1-H}$, and hence for some $\rho < 2$.

The estimate (4.54) does not imply stability in any moment norm because of the lack of exponential integrability of $\|V\|^{\rho}_{\rho\text{-var}}$ and $\|U\|^{\rho}_{\rho\text{-var}}$. Although similar problems involving exponential integrability have been resolved in literatures ([CLL13] uses a clever argument based on Gaussian concentration inequality, [GG23] uses a stochastic sewing argument built on Azuma–Hoeffding inequality), we are not able to apply these methods in our situation.

We thus adopt a different perspective, inspired by [FHL23; Ath+21], by estimating moments of X^{Δ} directly from equation (4.10). To employ this approach, one relies on the modulus of mean oscillation of the driving process V. In fact, we have

$$[V]_{\mathrm{VMO}^{\gamma}} := \sup_{s<t} \frac{\|\mathbb{E}[V_t - V_s|\mathcal{F}_s]\|_{L^\infty(\mathbb{P})}}{(t-s)^{\gamma}} < \infty, \quad \gamma := 1 - H - \frac{dH}{p} - \frac{1}{q}.$$

Under (4.2), γ is positive but can be arbitrarily small. Therefore, in view of Lyons' rough path theory, to obtain closed estimates for X^{Δ} from (4.10), we have to construct a rough path lift $\mathbb{V}$ on V and consider (4.10) as a rough differential equation. The construction and the estimate of such rough path lift (see Lemma 4.4.8) are based upon VMO-type estimate of the ρ-variation of V. We then apply the John–Nirenberg inequality (Proposition 4.2.8) and the sewing lemma to obtain moment estimates for the integral $\int_s^t X^{\Delta} \mathrm{d}V$. Combined with (4.10), we obtain that

$$\begin{aligned}\|X^{\Delta}_t - X^{\Delta}_s\|_{L^m(\mathbb{P})} \lesssim \|X^{\Delta}_s\|_{L^m(\mathbb{P})} (t-s)^{\gamma} + \|X^{\Delta}\|_{C^{\gamma}L^m(\mathbb{P})} (t-s)^{2\gamma} \\ + \|R_t - R_s\|_{L^m(\mathbb{P})}\end{aligned}$$

Another stochastic sewing argument in combination with John–Nirenberg inequality (Lemma 4.4.3) shows that

$$\|R_t - R_s\|_{L^m(\mathbb{P})} \lesssim \|b^1 - b^2\|_{L^q_t W^{-1,p}_x} (t-s)^{\gamma}.$$

From here, standard Gronwall argument is applied, which yields strong stability in moment norms.

Notation

Throughout the chapter we fix the dimension d, the Hurst parameter $H \in (0, \frac{1}{2})$, the integrability parameters $p, q \in [1, \infty)$ and the final time $T \in [0, \infty)$. Note that T is a fixed positive number, not a stopping time. We will write

$$\|b\|_{L^q([s,t];L^p_x)} := \Big(\int_s^t \|b_r\|^q_{L^p(\mathbb{R}^d)} \mathrm{d}r \Big)^{\frac{1}{q}}, \quad \|b\|_{L^q_t L^p_x} := \|b\|_{L^q([0,T];L^p)}.$$

Remark 4.1.5. We can assume without loss of generality that both p, q are finite. This allows us to approximate any element of $L^q_t L^p_x$ by smooth functions. Since we consider the fixed time interval $[0, T]$, the condition $q < \infty$ causes no harm. The case $p = \infty$ is already covered by [GG23]. Indeed, we can use the embedding $L^q_t L^\infty_x \hookrightarrow L^q_t C^\alpha_x$ for some $\alpha < 0$ such that (4.8) holds.

Given a path $f \colon [0, T] \to E$ we write $f_{s,t} := f_t - f_s$. Given a two-parameter maps $(A_{s,t})_{s<t}$ we write

$$\delta A_{s,u,t} := A_{s,t} - A_{s,u} - A_{u,t}. \tag{4.13}$$

We denote by $\|\cdot\|_{\rho\text{-var};[s,t]}$ the ρ-variation norm [FV10, Definition 5.1]. That is,

$$\|w\|_{\rho\text{-var};[s,t]} := \Big(\sup_{\pi} \sum_{[u,v]\in\pi} |w_{s,t}|^\rho \Big)^{\frac{1}{\rho}},$$

where the supremum is over all partitions π of the interval $[s, t]$. We simply write $\|\cdot\|_{\rho\text{-var}} := \|\cdot\|_{\rho\text{-var};[0,T]}$. The following inequalities hold: for $s < u < t$

$$(\|w\|^\rho_{\rho\text{-var};[s,u]} + \|w\|^\rho_{\rho\text{-var};[u,t]})^{\frac{1}{\rho}} \leq \|w\|_{\rho\text{-var};[s,t]} \leq \|w\|_{\rho\text{-var};[s,u]} + \|w\|_{\rho\text{-var};[u,t]}. \tag{4.14}$$

We write $\|\cdot\|_m := \|\cdot\|_{L^m(\mathbb{P})}$, $\mathbb{E}_s[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_s]$ (when the filtration $(\mathcal{F}_s)$ is obvious from the context) and

$$\|X|\mathcal{G}\|_m := \mathbb{E}[|X|^m|\mathcal{G}]^{\frac{1}{m}}, \quad \|\|X|\mathcal{G}\|_m\|_n := \|\|X|\mathcal{G}\|_m\|_n. \tag{4.15}$$

We write $A \lesssim B$ if there exists a constant $C \in (0, \infty)$ depending on irrelevant parameters such that $A \leq CB$. If we want to emphasize dependency on parameters, say $\alpha, \beta, \ldots$, we write $A \lesssim_{\alpha,\beta,\ldots} B$. **In this chapter, we will not write down dependency on** d, H, p, q.

4.2 Weak well-posedness

The goal of this section is to prove weak well-posedness for (4.1) and Girsanov's theorem for the solutions.

4.2.1 Preliminaries

Review of fractional SDEs

Recall the notion of $(\mathcal{F}_t)$-fractional Brownian motion from Definition 3.2.3.

Definition 4.2.1. We say that a quintuple

$$(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}, (B_t)_{t\in[0,T]}, (X_t)_{t\in[0,T]})$$

is a *weak solution* to (4.1) if the following conditions are satisfied.

(i) The triplet $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is a complete filtered probability space.

(ii) The process B is a $(\mathcal{F}_t)$-fractional Brownian motion.

(iii) The process X is $(\mathcal{F}_t)$-adapted and satisfies a.s.

$$t \mapsto b_t(X_t) \in L^1([0,T]), \quad K_H^{-1}\Big(\int_0^\cdot b_r(X_r)\mathrm{d}r\Big) \in H^1([0,T]).$$

(iv) We have

$$X_t = x + \int_0^t b_r(X_r)\mathrm{d}r + B_t \quad \forall t \in [0,T].$$

We say that *weak uniqueness* holds for (4.1) if the law of X is unique.

Remark 4.2.2. The technical second condition of (iii) is expected to be removed for the publication, by proving path-by-path uniqueness.

Definition 4.2.3. Given a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with $(\mathcal{F}_t)$-fractional Brownian motion B, we say that a process X is a *pathwise solution* to (4.1) if X satisfies the conditions (iii) and (iv) of Definition 4.2.1. We say that *pathwise uniqueness* for (4.1) holds if two pathwise solutions to (4.1) are indistinguishable. A pathwise solution is called a *strong solution* if it is adapted to the natural filtration generated by B. We say that *strong existence* holds if there exists a strong solution.

Remark 4.2.4. In view of a general version of the Yamada-Watanabe theorem proved by Kurtz [Kur07], to prove Theorem 4.1.1 it suffices to prove weak existence and pathwise uniqueness. Following the arguments of [Lê20, Theorem 6.1] with new ingredient of a VMO-type estimate from [Lê22], we will prove weak existence in Proposition 4.2.11. Pathwise uniqness will be proven in Section 4.3.

Girsanov's theorem

Our argument relies on the following Girsanov's theorem. Since the Cameron–Martin space of the Brownian motion is

$$\{\,\phi \in H^1([0,T]) : \phi_0 = 0\,\},$$

the Cameron–Martin space $\mathcal{H}^H$ of the fractional Brownian motion B^H is

$$\{\,\phi : \phi_0 = 0,\ K_H^{-1}\phi \in H^1([0,T])\,\}$$

with

$$\|\phi\|_{\mathcal{H}^H} := \Big(\int_0^T \Big|\frac{\mathrm{d}}{\mathrm{d}t}K_H^{-1}\phi\Big|^2(t)\mathrm{d}t\Big)^{\frac{1}{2}}.$$

For an adapted process ϕ belonging to $\mathcal{H}^H$ and a stopping time τ, we set

$$\xi_\tau(\phi) := \exp\Big\{-\int_0^\tau \frac{\mathrm{d}}{\mathrm{d}t}K_H^{-1}\phi(t)\mathrm{d}W_t - \frac{1}{2}\int_0^\tau \Big|\frac{\mathrm{d}}{\mathrm{d}t}K_H^{-1}\phi(t)\Big|^2\mathrm{d}t\Big\}. \tag{4.16}$$

Lemma 4.2.5. *For an adapted process ϕ with $\phi_0 = 0$, we define the measure $\tilde{\mathbb{P}}$ by $\mathrm{d}\tilde{\mathbb{P}} := \xi_T(\phi)\mathrm{d}\mathbb{P}$. If*

$$\mathbb{E}\exp\big(\lambda\|\phi\|^2_{\mathcal{H}^H}\big) < \infty, \quad \forall\lambda \in [0,\infty), \tag{4.17}$$

then $\tilde{\mathbb{P}}$ is a probability measure and the law of $B+\phi$ under $\tilde{\mathbb{P}}$ is that of B under $\mathbb{P}$. Furthermore, for any $\lambda \in \mathbb{R}$, we have

$$\mathbb{E}[\xi_T(\phi)^\lambda] \le \mathbb{E}[\exp(\lambda^2\|\phi\|^2_{\mathcal{H}^H})]^{\frac{1}{2}}. \tag{4.18}$$

Proof. We have $B+\phi = K_H(W + K_H^{-1}\phi)$. The condition (4.17) corresponds to the strong Novikov's condition, and the first claim follows from Girsanov's theorem of the Brownian motion. To obtain the last inequality, by the Cauchy–Schwarz inequality,

$$\mathbb{E}[\xi_T(\phi)^\lambda] \le \mathbb{E}[\xi_T(-2\lambda\phi)]^{\frac{1}{2}}\mathbb{E}[\exp(\lambda^2\|\phi\|^2_{\mathcal{H}^H})]^{\frac{1}{2}}.$$

Note that $t \mapsto \xi_t(-2\lambda\phi)$ is a martingale, hence $\mathbb{E}[\xi_T(-2\lambda\phi)] = 1$. □

VMO estimates

In order to prove weak well-posedness of (4.1), we rely on Girsanov's theorem (Lemma 4.2.5). To this end (and also for later purposes), we need the notion of processes of vanishing mean oscillation (VMO processes) and sharp inequalities for those processes.

Definition 4.2.6. We say that a map $(s,t) \mapsto w(s,t)$ $(s \leq t)$ is a *control* if it is continuous with $w(s,s) = 0$ and

$$w(s,u) + w(u,t) \leq w(s,t) \quad \text{for all } s < u < t.$$

Note that the map $(s,t) \mapsto \|b\|^q_{L^q([s,t];L^p_x)}$ is a control. Given two controls w_1 and w_2 and two nonnegative numbers ν_1, ν_2 with $\nu_1 + \nu_2 \geq 1$, the map

$$(s,t) \mapsto w_1(s,t)^{\nu_1} w_2(s,t)^{\nu_2}$$

is again a control, see [FV10, Exercise 1.9]. In particular, if $\frac{dH}{p} + \frac{1}{q} < 1$, the map

$$(s,t) \mapsto \left(\|b\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}}\right)^{(1-\frac{dH}{p})^{-1}}$$

is a control. These facts will be used throughout the chapter.

Definition 4.2.7 ([Lê22, Definitions 1.1 and 3.2])**.** Recall the definition of the conditional moment from (4.15). A $(\mathcal{F}_t)$-adapted continuous process $(Z_t)_{t\in[0,T]}$ is called a *process of vanishing mean oscillation*, which will be called a VMO process afterwards, if

$$\lim_{h\downarrow 0} \sup_{0\leq s<t\leq T, t-s\leq h} \|\|Z_{s,t}|\mathcal{F}_s\|_1\|_\infty = 0.$$

For a VMO process Z and $r \in [1,\infty)$, we write $Z \in \mathrm{VMO}^{r\text{-var}}$ if there exists a control w such that

$$\|\|Z_{s,t}|\mathcal{F}_s\|_1\|_\infty \leq w(s,t)^{\frac{1}{r}}, \quad \forall s \leq t. \tag{4.19}$$

The following result is called the quantitative John–Nirenberg inequality for VMO processes.

Proposition 4.2.8 ([Lê22, Corollary 3.5])**.** *Let $Z \in \mathrm{VMO}^{r\text{-var}}$ with $r \in (1,\infty)$ and (4.19) satisfied. Then, there exists a constant $c = c_r$, depending only on r, such that provided $\lambda w(0,T)^{\frac{1}{r-1}} \leq c$ we have*

$$\mathbb{E}[e^{\lambda \sup_{t\leq T}|Z_t - Z_0|^{\frac{r}{r-1}}}] \lesssim_{r,\lambda,w(0,T)} 1. \tag{4.20}$$

The proportional constant depends increasingly on λ and $w(0,T)$. Furthermore, for every $m \in [1,\infty)$ we have

$$\|\|Z_{s,t}|\mathcal{F}_s\|^m_m\|_\infty \lesssim_r \Gamma(1 + m(1-1/r))w(s,t)^{\frac{m}{r}}. \tag{4.21}$$

4.2.2 Weak solutions and stochastic estimates

To construct a weak solution, we apply Girsanov's theorem. To this end, the following quantitative bounds will be important.

Lemma 4.2.9. *Under $\frac{dH}{p}+\frac{1}{q}<1$, let $f\in L_t^qL_x^p$ and set*

$$\phi_t:=\int_0^t f(r,B_r)\mathrm{d}r.$$

Then, for any $\varepsilon\in(0,1-\frac{dH}{p}-\frac{1}{q})$, there exists a constant Λ depending only on d,H,p,q,ε such that

$$\mathbb{E}\Big[\exp\Big\{\Lambda\Big(\sup_{0\le s<t\le T}\frac{|\phi_{s,t}|}{\|b\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}-\varepsilon}}\Big)^{\frac{p}{dH}}\Big\}\Big]\lesssim 1. \tag{4.22}$$

Proof. In view of Kolmogorov-type estimate (e.g. [GG23, Lemma A.2][1]), it suffices to show

$$\sup_{0\le s<t\le T}\mathbb{E}\Big[\exp\Big\{\lambda\Big(\frac{|\phi_{s,t}|}{\|b\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}}}\Big)^{\frac{p}{dH}}\Big\}\Big]\lesssim 1 \tag{4.23}$$

for some small λ. To this end, we will apply the quantitative John–Nirenberg inequality (4.21). By [Lê20, Lemma 6.4], we have

$$\mathbb{E}_s\int_s^t|f(r,B_r)|\mathrm{d}r\lesssim\|f\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}}.$$

That is, ϕ belongs to $\mathrm{VMO}^{(1-dH/p)^{-1}\text{-var}}$. Hence, the estimate (4.21) yields

$$\|\phi_{s,t}\|_m\le C\Gamma\Big(1+\frac{dHm}{p}\Big)^{\frac{1}{m}}\|f\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}} \tag{4.24}$$

for every $m\in[1,\infty)$ and for some C depending only on H,p,q,d. For simplicity, we set $\alpha:=\frac{p}{dH}$ and

$$\eta(s,t):=\|f\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}}.$$

Applying (4.24), we obtain

$$\mathbb{E}\Big[\exp\Big\{\lambda\Big(\frac{|\phi_{s,t}|}{\eta(s,t)}\Big)^{\alpha}\Big\}\Big]=\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\frac{\lambda}{\eta(s,t)}\Big)^{n\alpha}\mathbb{E}[|\phi_{s,t}|^{n\alpha}]\le\sum_{n=0}^{\infty}(C\lambda)^{n\alpha},$$

which gives the desired estimate provided $\lambda\le(2C)^{-1}$. $\square$

[1] In the cited reference, the estimate is given only for $dH/p=2$, but it is evident that the same argument yields the estimate for the general case.

Lemma 4.2.10. *In the setting of Lemma 4.2.9, there exists a constant $\mathfrak{C}$, depending only on d, H, p, q, with the following property: for any $\lambda, M \in [0,\infty)$ with*

$$\lambda(\|f\mathbf{1}_{\{|f|\geq M\}}\|_{L_t^q L_x^p} T^{1-\frac{1}{q}-\frac{dH}{p}})^{\frac{p}{dH}-1} \leq \mathfrak{C},$$

we have

$$\mathbb{E}[e^{\lambda \sup_{t\leq T}|\phi_t|^{\frac{p}{dH}}}] \lesssim_{T,\lambda} e^{\lambda(2TM)^{\frac{p}{dH}}}.$$

Proof. It is shown that ϕ belongs to $\mathrm{VMO}^{(1-dH/p)^{-1}\text{-var}}$ in the proof of Lemma 4.2.9. Therefore, the estimate (4.20) yields

$$\mathbb{E}[e^{\lambda \sup_{t\leq T}|\phi_t|^{\frac{p}{dH}}}] \lesssim_{T,\lambda,\|f\|_{L_t^q L_x^p}} 1$$

provided that

$$\lambda\|f\|_{L_t^q L_x^p}^{\frac{p}{dH}-1} T^{(1-\frac{1}{q}-\frac{dH}{p})(\frac{p}{dH}-1)} \leq c,$$

where c is a constant depending only on d, H, p, q. To deal with large λ, as in [GG23, Lemma C.3], we decompose b by $f = f^{(1)} + f^{(2)}$, where $f^{(1)}(t,x) := f(t,x)\mathbf{1}_{\{|f(t,x)|\geq M\}}$. We observe, with $\alpha := \frac{p}{dH}$, that

$$\begin{aligned} |\phi_t|^\alpha &\leq \Big(MT + \int_0^T |f_r^{(1)}(B_r)|\mathrm{d}r\Big)^\alpha \\ &\leq 2^{\alpha-1}(MT)^\alpha + 2^{\alpha-1}\Big(\int_0^T |f_r^{(1)}(B_r)|\mathrm{d}r\Big)^\alpha \end{aligned}$$

Hence,

$$\mathbb{E}[e^{\lambda \sup_{t\leq T}|\phi_t|^\alpha}] \leq e^{2^{\alpha-1}\lambda(MT)^\alpha}\mathbb{E}\exp\Big\{2^{\alpha-1}\lambda\Big(\int_0^T |f_r^{(1)}(B_r)|\mathrm{d}r\Big)^\alpha\Big\}.$$

We choose M so that

$$2^\alpha\lambda\|f^{(1)}\|_{L_t^q L_x^p}^{\frac{p}{dH}-1} T^{(1-\frac{1}{q}-\frac{dH}{p})(\frac{p}{dH}-1)} \leq c,$$

and we complete the proof. $\square$

Now we can prove weak existence.

Proposition 4.2.11. *Under (4.2), there exists a weak solution to (4.1) such that the law of the solution is equivalent to that of B.*

Proof. We set

$$\phi_t := \int_0^t b_r(B_r)\mathrm{d}r.$$

As in [Lê20, Theorem 6.1], it suffices to check (strong) Novikov's condition:

$$\mathbb{E}\exp(\lambda\|\phi\|^2_{\mathcal{H}^H}) < \infty, \quad \forall \lambda \geq 0.$$

To estimate it, as used in [GG23, Appendix C], we introduce the Besov–Nikolskii norm:

$$\|f\|_{N^{\beta,r}} := \sup_{h\in[0,T]} h^{-\beta}\Big(\int_0^{T-h} |f_{s,s+h}|^r \mathrm{d}s\Big)^{\frac{1}{r}}.$$

Fix a small $\varepsilon \in (0, H)$. According to [GG23, Proposition C.1], we have

$$\|\phi\|_{\mathcal{H}^H} \lesssim_{T,\varepsilon} \|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}}.$$

We will estimate $\|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}}$. Decomposing b by

$$b = b\mathbf{1}_{\{|b|<M\}} + b\mathbf{1}_{\{|b|\geq M\}},$$

we can assume that $\|b\|_{L_t^q L_x^p} \leq 1$, which eases the task of tracking constants.

Case 1: $p \geq (1-H)^{-1}$. By Hölder's inequality, we obtain

$$|\phi_{s,t}| \leq (t-s)^H \Big(\int_s^t |b_r(B_r)|^{\frac{1}{1-H}}\mathrm{d}r\Big)^{1-H}.$$

Therefore, by setting

$$\psi_t := \int_0^t |b|^p(r, B_r)\mathrm{d}r,$$

we obtain

$$\int_0^{T-h} |\phi_{s,s+h}|^2 \mathrm{d}s \leq h^{2H}\int_0^{T-h} \psi_{s,s+h}^{2(1-H)}\mathrm{d}s.$$

Note that $|b|^{\frac{1}{1-H}} \in L_t^{q(1-H)} L_x^{p(1-H)}$ and, due to (4.2),

$$p(1-H) \geq 1, \quad q(1-H) \geq 1, \quad \frac{dH}{p(1-H)} + \frac{1}{q(1-H)} < 1.$$

By Lemma 4.2.9, there exists a constant Λ, depending only on d, H, p, q, ε, such that

$$\mathbb{E}[\exp(\Lambda\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{\frac{p(1-H)}{dH}})] \lesssim 1. \tag{4.25}$$

We deonote by $M(b,\lambda)$ the smallest M such that

$$\lambda(2\||b|^{\frac{1}{1-H}}\mathbf{1}_{\{|b|^{\frac{1}{1-H}}\geq M\}}\|_{L_t^{q(1-H)}L_x^{p(1-H)}}T^{1-\frac{1}{q(1-H)}-\frac{dH}{p(1-H)}})^{\frac{p(1-H)}{dH}-1} \leq \mathfrak{C},$$

where $\mathfrak{C}$ is the constant from Lemma 4.2.10. Then, by Lemma 4.2.10,

$$\mathbb{E}[e^{\lambda\psi_T^{\frac{p(1-H)}{dH}}}] \lesssim_{T,\lambda,M(b,\lambda)} 1, \quad \forall\lambda. \tag{4.26}$$

We thus estimate

$$\begin{aligned}
&\int_0^{T-h}|\psi_{s,s+h}|^{2(1-H)}\mathrm{d}s\\
&\qquad\leq\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{1-2H}h^{(1-2H)(1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon)}\int_0^{T-h}|\psi_{s,s+h}|\mathrm{d}s\\
&\qquad\leq\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{1-2H}\psi_T h^{(1-2H)(1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon)+1}
\end{aligned}$$

and

$$\int_0^{T-h}|\phi_{s,s+h}|^2\mathrm{d}s \leq h^{2H+1+(1-2H)(1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon)}\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{1-2H}\psi_T.$$

Choosing ε so small that $1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon>\varepsilon$, we obtain

$$\begin{aligned}
\|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}}^2 &\lesssim_T \|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{1-2H}\psi_T\\
&\qquad\lesssim \delta(2\lambda)^{-1}\Lambda\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{2(1-H)}+(2\delta^{-1}\Lambda^{-1}\lambda)^{\frac{1-2H}{2(1-H)}}\psi_T^{2(1-H)}.
\end{aligned}$$

Hence, for an appropriately small $\delta=\delta(d,H,p,q,T)$ we have

$$2\|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}}^2 \leq \lambda^{-1}\Lambda\|\psi\|_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}^{2(1-H)}+C\lambda^{\frac{1-2H}{2(1-H)}}\psi_T^{2(1-H)}$$

for some $C=C(d,H,p,q,T,\varepsilon)$. In view of (4.25) and (4.26), setting

$$\tilde{M} := M(b, 2(2\Lambda^{-1}\lambda)^{\frac{1-2H}{2(1-H)}})$$

we have

$$\begin{aligned}\mathbb{E}[e^{\lambda\|\phi\|^2_{N^{H+\frac{1}{2}+\varepsilon,2}}}] &\leq \mathbb{E}[e^{\Lambda\|\psi\|^{2(1-H)}_{C^{1-\frac{dH}{p(1-H)}-\frac{1}{q(1-H)}-\varepsilon}}}]^{\frac{1}{2}}\mathbb{E}[e^{C\lambda^{\frac{1-2H}{2(1-H)}}\psi_T^{2(1-H)}}]^{\frac{1}{2}} \\ &\lesssim_{T,\lambda,\tilde{M}} 1\end{aligned}$$

provided that $2(1-H) \leq \frac{p(1-H)}{dH}$, but this is the condition (4.2c).

Case 2: $q \leq 2$**.** Set

$$[\![\phi]\!] := \sup_{0\leq s<t\leq T} \frac{|\phi_{s,t}|}{\|b\|_{L^q([s,t];L^p_x)}(t-s)^{1-\frac{dH}{p}-\frac{1}{q}-\varepsilon}}.$$

We estimate

$$\begin{aligned}\int_0^{T-h}|\phi_{s,s+h}|^2\mathrm{d}s &\leq [\![\phi]\!]^2 h^{2(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon)}\int_0^{T-h}\|b\|^2_{L^q([s,s+h];L^p_x)}\mathrm{d}s \\ &\lesssim [\![\phi]\!]^2 h^{2(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon)+1}\|b\|^2_{L^q_tL^p_x}.\end{aligned}$$

Therefore, if ε is small,

$$\|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}} \lesssim \|b\|_{L^q_tL^p_x}[\![\phi]\!],$$

and, noting that $\frac{p}{dH} > 2$ due to (4.2b), it remains to apply Lemma 4.2.9.

Case 3: $p < (1-H)^{-1}$ **and** $q > 2$**.** We define $\nu \in [0,1]$ by

$$\nu := \frac{2^{-1}-q^{-1}}{p^{-1}-q^{-1}}$$

so that the relation

$$\frac{2\nu}{p} + \frac{2(1-\nu)}{q} = 1$$

holds. We then bound

$$|\phi_{s,t}| = |\phi_{s,t}|^{1-\nu}|\phi_{s,t}|^{\nu} \leq [\![\phi]\!]^{1-\nu}h^{(1-\nu)(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon)}\|b\|^{1-\nu}_{L^q([s,t];L^p_x)}|\phi_{s,t}|^{\nu}.$$

By Hölder's inequality,

$$|\phi|_{s,t} \leq (t-s)^{1-\frac{1}{p}}\Big(\int_s^t |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{1}{p}}.$$

Hence,

$$\int_0^{T-h} |\phi_{s,s+h}|^2 \mathrm{d}s \leq [\![\phi]\!]^{2(1-\nu)} h^{2(1-\nu)(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon)+2\nu(1-\frac{1}{p})} \times \int_0^{T-h} \|b\|^{2(1-\nu)}_{L^q([s,s+h];L^p_x)} \Big(\int_s^{s+h} |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}} \mathrm{d}s.$$

Since

$$(s,t) \mapsto \|b\|^{2(1-\nu)}_{L^q([s,t];L^p_x)} \Big(\int_s^t |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}}$$

is a control, we have

$$\int_0^{T-h} \|b\|^{2(1-\nu)}_{L^q([s,s+h];L^p_x)} \Big(\int_s^{s+h} |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}} \mathrm{d}s \lesssim h\|b\|^{2(1-\nu)}_{L^q_t L^q_x} \Big(\int_0^T |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}}.$$

We therefore arrive at the estimate

$$\int_0^{T-h} |\phi_{s,s+h}|^2 \mathrm{d}s \lesssim [\![\phi]\!]^{2(1-\nu)} h^{2(1-\nu)(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon)+2\nu(1-\frac{1}{p})+1} \times \|b\|^{2(1-\nu)}_{L^q_t L^q_x} \Big(\int_0^T |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}}.$$

The condition (4.2d) exactly ensures that the exponent

$$2(1-\nu)\Big(1-\frac{dH}{p}-\frac{1}{q}-\varepsilon\Big)+2\nu\Big(1-\frac{1}{p}\Big)+1$$

is greater than $2H+1+\varepsilon$ for small ε. Thus,

$$\begin{aligned}\|\phi\|_{N^{H+\frac{1}{2}+\varepsilon,2}} &\lesssim [\![\phi]\!]^{2(1-\nu)} \Big(\int_0^T |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2\nu}{p}} \\ &\leq \lambda^{-1}\Lambda[\![\phi]\!]^2 + (\Lambda^{-1}\lambda)^{1-\nu}\Big(\int_0^T |b|^p(r,B_r)\mathrm{d}r\Big)^{\frac{2}{p}}.\end{aligned}$$

It remains to apply Lemma 4.2.9 and Lemma 4.2.10. □

Remark 4.2.12. The above proof shows that there exist $\mathfrak{d} = \mathfrak{d}(d, H, p)$ and $\mathfrak{D} = \mathfrak{D}(d, H, p, q, T)$ such that provided that

$$\|b\mathbf{1}_{\{|b|\geq M\}}\|_{L_t^q L_x^p} \leq \mathfrak{D}\lambda^{-\mathfrak{d}}$$

we have

$$\mathbb{E}\exp(\lambda\|\phi\|_{\mathcal{H}^H}^2) \lesssim_{T,\lambda,M} 1, \tag{4.27}$$

where $\phi_t := \int_0^t b_r(B_r)\mathrm{d}r$. This is carefully demonstrated in Case 1 of the proof.

Proposition 4.2.13. *Weak uniqueness holds for* (4.1).

Proof. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a bounded continuous function and

$$t_1, \dots, t_n \in [0, T].$$

Let X be a solution to (4.1). We need to show that

$$\mathbb{E}[f(X_{t_1}, \dots, X_{t_n})]$$

is uniquely characterized. For $M \geq 1$, we set

$$\tau_M(w) := \inf\Big\{t \leq T : \int_0^t \Big|K_H^{-1}\Big(\int_0^\cdot b_r(w_r)\mathrm{d}r\Big)(s)\Big|^2 \mathrm{d}s \geq M\Big\}$$

with $\inf\varnothing := T$. By the condition (iii) of Definition 4.2.1, it suffices to show that

$$\mathbb{E}[F(X)], \quad F(X) := f(X_{t_1\wedge\tau_M(X)}, \dots, X_{t_n\wedge\tau_M(X)}),$$

is uniquely characterized. We set $\xi := \xi_\tau(\int_0^\cdot b_r(X_r)\mathrm{d}r)$ and $\mathrm{d}\tilde{\mathbb{P}}(\omega) := \xi(\omega)\mathrm{d}\mathbb{P}(\omega)$. By Girsanov's theorem, the law of $(X_t - x)_{t\leq\tau(X)}$ under $\tilde{\mathbb{P}}$ is equal to that of $(B_t)_{t\leq\tau(B+x)}$ under $\mathbb{P}$. Furthermore, since

$$W = K_H^{-1}\Big(X - x - \int_0^\cdot b_r(X_r)\mathrm{d}r\Big),$$

with some measurable map $G\colon C([0,T];\mathbb{R}^d) \to \mathbb{R}$, we can write

$$\xi^{-1}F(X) = G(X).$$

Therefore,

$$\mathbb{E}[F(X)] = \mathbb{E}[G(B)],$$

and the right-hand side depends only on the law of B. $\square$

As the final result of this section, we prove a quantitative version of Girsanov's theorem.

Lemma 4.2.14. *Let X be a pathwise solution to* (4.1) *and set*

$$\phi_t := \int_0^t b_r(X_r)\mathrm{d}r.$$

Under (4.2)*, there exist $\mathfrak{e} = \mathfrak{e}(d, H, p)$ and $\mathfrak{E} = \mathfrak{E}(d, H, p, q, T)$ such that for any $\lambda \geq 2$ provided that*

$$\|b\mathbf{1}_{\{|b|\geq M\}}\|_{L_t^q L_x^p} \leq \mathfrak{E}\lambda^{-\mathfrak{e}}$$

we have

$$\mathbb{E}\exp(\lambda\|\phi\|_{\mathcal{H}^H}) \lesssim_{T,\lambda,M} 1. \tag{4.28}$$

In particular, the law of $X - x$ is equivalent to that of B.

Proof. In view of Lemma 4.2.5, it suffices to show (4.28). By Proposition 4.2.13, weak uniqueness is established. Therefore, we can prove the estimate for a particular solution, which will be the one constructed in Proposition 4.2.11. Namely, we will show that for

$$\tilde{\phi}_t := \int_0^t b_r(B_r)\mathrm{d}r,$$

we have

$$\mathbb{E}[\xi_T(\tilde{\phi})e^{\lambda\|\tilde{\phi}\|^2_{\mathcal{H}^H}}] \lesssim_{T,\lambda,M} 1$$

for all $\lambda \geq 0$. To this end, let $\mathfrak{d}$ and $\mathfrak{D}$ be the constants in Remark 4.2.12. By the Cauchy–Schwarz inequality,

$$\mathbb{E}[\xi_T(\tilde{\phi})e^{\lambda\|\tilde{\phi}\|^2_{\mathcal{H}^H}}] \leq \mathbb{E}\Big[\exp\Big(2\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}K_H^{-1}\tilde{\phi}(t)\mathrm{d}W_t\Big)\Big]^{\frac{1}{2}}\mathbb{E}[e^{2\lambda\|\tilde{\phi}\|^2_{\mathcal{H}^H}}]^{\frac{1}{2}}.$$

By Lemma 4.2.5,

$$\mathbb{E}\Big[\exp\Big(2\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}K_H^{-1}\tilde{\phi}(t)\mathrm{d}W_t\Big)\Big]^{\frac{1}{2}} \leq \mathbb{E}[e^{4\|\tilde{\phi}\|^2_{\mathcal{H}^H}}]^{\frac{1}{4}}.$$

Hence,

$$\mathbb{E}[\xi_T(\tilde{\phi})e^{\lambda\|\tilde{\phi}\|^2_{\mathcal{H}^H}}] \leq \mathbb{E}[e^{2\lambda\|\tilde{\phi}\|^2_{\mathcal{H}^H}}]^{\frac{3}{4}}.$$

We can set $\mathfrak{e} := \mathfrak{d}$ and $\mathfrak{E} := 2^{-\mathfrak{d}}\mathfrak{D}$. ◻

The following observation will be used in Section 4.4.1. Let $\tilde{B}$ be the Riemann–Liouville process

$$\tilde{B}_t := c_H \int_0^t (t-r)^{H-\frac{1}{2}} \mathrm{d}W_r, \tag{4.29}$$

for some constant $c_H \in (0,\infty)$. Then, we have analogous results of this section for $\tilde{X}$. In fact, we replace the operator K_H by the Riemann–Liouville operator

$$I_H w(t) := c_H \int_0^t (t-r)^{H-\frac{1}{2}} \mathrm{d}w_r,$$

and repeat the argument.

The operator I_H, $H < \frac{1}{2}$, corresponds to the fractional derivative $\partial^{\frac{1}{2}-H}$, and it is invertible with

$$I_H^{-1} w(t) = \bar{c}_H \int_0^t (t-r)^{-H-\frac{1}{2}} w_r \mathrm{d}r$$

for some constant $\bar{c}_H$. The operator I_H^{-1} corresponds to the fractional integral $J_{\frac{1}{2}-H}$ of order $\frac{1}{2}-H$. The Cameron–Martin space $\tilde{\mathcal{H}}^H$ of $\tilde{B}$ is

$$\{\, w : w_0 = 0,\ I_H^{-1} w \in H^1([0,T]) \,\} = \{\, w : w_0 = 0,\ \partial^{H+\frac{1}{2}} w \in L^2([0,T]) \,\}.$$

To see an analogue of Proposition 4.2.11, first observe that the local nondeterminism estimate

$$\mathrm{Var}(\tilde{B}_t - \mathbb{E}_s[\tilde{B}_t]) \gtrsim (t-s)^{2H}$$

holds. Therefore, we have the VMO estimate

$$\mathbb{E}_s \int_s^t |f(r,\tilde{B}_r)| \mathrm{d}r \lesssim \|f\|_{L^q([s,t];L^p_x)} (t-s)^{1-\frac{dH}{p}-\frac{1}{q}}.$$

To estimate the Cameron–Martin norm of $\tilde{\phi} := \int_0^\cdot b_r(\tilde{B}_r)\mathrm{d}r$, by [GLY15, Theorem 3.1], we have

$$\|\tilde{\phi}\|_{\tilde{\mathcal{H}}^H} = \|\partial^{H+\frac{1}{2}} \tilde{\phi}\|_{L^2([0,T])} \lesssim_T \|\tilde{\phi}\|_{W^{H+\frac{1}{2},2}([0,T])}.$$

Thus, by [GG23, Proposition C.1],

$$\|\tilde{\phi}\|_{\tilde{\mathcal{H}}^H} \lesssim_{T,\varepsilon} \|\tilde{\phi}\|_{N^{H+\frac{1}{2}+\varepsilon,2}}.$$

Hence, repeating the arguments of Proposition 4.2.11 and Lemma 4.2.14, we obtain the following.

Lemma 4.2.15. *Let η be a deterministic continuous path in $\mathbb{R}^d$ with $\eta_0 = 0$, and let $\tilde{X}$ be a pathwise solution to the SDE*

$$\mathrm{d}\tilde{X}_t = \tilde{b}_t(X_t)\mathrm{d}t + \mathrm{d}\eta_t + \mathrm{d}\tilde{B}_t, \quad \tilde{X}_0 = \tilde{x}$$

with $\tilde{b} \in L_t^q L_x^p$, and set

$$\tilde{\phi}_t := \int_0^t \tilde{b}_r(\tilde{X}_r)\mathrm{d}r.$$

Under (4.2)*, there exist $\tilde{\mathfrak{e}} = \tilde{\mathfrak{e}}(d, H, p)$ and $\tilde{\mathfrak{E}} = \tilde{\mathfrak{E}}(d, H, p, q, T)$ such that for any $\lambda \geq 2$ provided that*

$$\|\tilde{b}\mathbf{1}_{\{|\tilde{b}| \geq M\}}\|_{L_t^q L_x^p} \leq \tilde{\mathfrak{E}}\lambda^{-\tilde{\mathfrak{e}}}$$

we have

$$\mathbb{E}\exp(\lambda \|\tilde{\phi}\|_{\tilde{\mathcal{H}}^H}) \lesssim_{T,\lambda,M} 1.$$

In particular, the law of $\tilde{X} - \tilde{x} - \eta$ is equivalent to that of $\tilde{B}$.

Proof. Setting $\bar{X}_t := \tilde{X}_t - \eta_t$ and $\bar{b}_t(x) := b_t(x + \eta_t)$, we have

$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \mathrm{d}\tilde{B}_t.$$

Since $\|\bar{b}\mathbf{1}_{\{|\bar{b}| \geq M\}}\|_{L_t^q L_x^p} = \|b\mathbf{1}_{\{|b| \geq M\}}\|_{L_t^q L_x^p}$, we can assume that $\eta = 0$. Then, as the preceeding discussion shows, the claim follows as in Lemma 4.2.14. □

4.3 Strong well-posedness

As explained in Remark Remark 4.2.4, in order to prove Theorem 4.1.1, it remains to prove pathwise uniqueness. We essentially follow the idea from [Lê20] which deduces pathwise uniqueness from pathwise regularity of the stochastic integral

$$\int_0^t \nabla b(r, B_r)\mathrm{d}r. \tag{4.30}$$

The main difference in our approach is that, instead of using Hölder regularity, we use variational regularity, which only requires estimations of moments up to order p-th. This allows us to obtain a desirable estimate on (4.30) beyond the regime of the previous work [Lê20]. Before estimating (4.30), we prepare some technical lemmas.

4.3.1 Some technical estimates

We will need two technical lemmas: taming singularity (Lemma 4.3.1) and Kolmogorov's continuity theorem for ρ-variation (Lemma 4.3.2).

Lemma 4.3.1. *Let $(\mathcal{Y}, \|\cdot\|)$ be a normed vector space and let $Y \colon [0,T] \to \mathcal{Y}$ be continuous. Suppose that for $0 < \eta \le \alpha < \infty$ and $C \in [0,\infty)$ we have*

$$\|Y_{s,t}\| \le C s^{\eta-\alpha}(t-s)^{\alpha}, \quad \textit{for } 0 < s < t \le T.$$

Then

$$\|Y_{s,t}\| \lesssim_{\alpha,\eta} C(t^{\eta/\alpha} - s^{\eta/\alpha})^{\alpha}, \quad \textit{for } 0 \le s < t \le T.$$

Proof. To get rid of the singularity, as in [BFG21], we set $\tau(t) := t^{\beta}$ for some $\beta \ge 1$. Then

$$\|Y_{\tau(s),\tau(t)}\| \le C s^{\beta(\eta-\alpha)}|t^{\beta} - s^{\beta}|^{\alpha}.$$

If $t - s \le s$, we have

$$t^{\beta} - s^{\beta} = \int_s^t \beta r^{\beta-1}\mathrm{d}r \le \beta t^{\beta-1}(t-s) \lesssim_{\beta} s^{\beta-1}|t-s|.$$

Therefore,

$$\|Y_{\tau(s),\tau(t)}\| \lesssim_{\beta} s^{\beta(\eta-\alpha)+(\beta-1)\alpha}(t-s)^{\alpha}.$$

We choose $\beta = \alpha/\eta$ so that $\beta(\eta-\alpha) + (\beta-1)\alpha = 0$. Hence, if $(t-s) \le s$, we have

$$\|Y_{\tau(s),\tau(t)}\| \lesssim_{\alpha,\eta} C(t-s)^{\alpha}. \tag{4.31}$$

To remove the condition $(t-s) \le s$, we use an idea from [LL22, Lemma 3.4]. For $s < t$, without assuming $(t-s) \le s$, we set

$$t_n := s + (t-s)2^{-n}.$$

By the continuity of Y,

$$Y_{\tau(s),\tau(t)} = \lim_{n\to\infty} \sum_{i=1}^{n} Y_{\tau(t_i),\tau(t_{i-1})}$$

and

$$\|Y_{\tau(s),\tau(t)}\| \le \sum_{n=1}^{\infty} \|Y_{\tau(t_n),\tau(t_{n-1})}\|.$$

By our choice of t_n, we have $|t_n - t_{n-1}| = (t-s)2^{-n} \le t_n$. Therefore,

$$\|Y_{\tau(s),\tau(t)}\| \lesssim_{\alpha,\eta} C \sum_{n=1}^{\infty} 2^{-n\alpha}(t-s)^{\alpha} \lesssim_{\alpha} C(t-s)^{\alpha}.$$

The estimate therefore (4.31) holds without the condition $|t-s| \le s$. We obtain the claimed estimate by change of variables $(s,t) \mapsto (s^{1/\beta}, t^{1/\beta})$. $\square$

Lemma 4.3.2. *Let $(V_t)_{t\in[0,T]}$ be a stochastic process. Suppose that*

$$\|V_{s,t}\|_m \le w(s,t)^{\alpha}$$

for some control w, $m \in (1,\infty)$ and $\alpha \in (\frac{1}{m}, 1]$. Then for every $\alpha' < \alpha$ with $m\alpha' > 1$

$$\Big\|\|V\|_{(1/\alpha')\text{-var}}\Big\|_m \lesssim_{m,\alpha,\alpha'} w(0,T)^{\alpha}.$$

Proof. Replacing $w(s,t)$ by $w(s,t) + \varepsilon(t-s)$, we may assume that $s \mapsto w(0,s)$ is strictly increasing. If we set

$$\tau(t) := (w(0,\cdot))^{-1}(t), \quad \tilde{V}_t := V_{\tau(t)}, \quad \tilde{T} := \tau^{-1}(T) = w(0,T),$$

then for $0 \le s < t \le \tilde{T}$

$$\|\tilde{V}_{s,t}\|_m \le (t-s)^{\alpha}.$$

Thus, for any $\alpha' < \alpha$ we have

$$\mathbb{E}\Big[\int_0^{\tilde{T}}\int_0^{\tilde{T}} \frac{|\tilde{V}_{s,t}|^m}{|t-s|^{1+m\alpha'}} \mathrm{d}t\mathrm{d}s\Big] \lesssim_{m,\alpha,\alpha'} \tilde{T}^{1+m(\alpha-\alpha')}.$$

By [FV10, Corollary A.3], if $m\alpha' > 1$,

$$\|\tilde{V}\|_{(1/\alpha')\text{-var}} \lesssim_{m,\alpha'} \tilde{T}^{\alpha'-\frac{1}{m}} \int_0^{\tilde{T}}\int_0^{\tilde{T}} \frac{|\tilde{V}_{s,t}|^m}{|t-s|^{1+m\alpha'}} \mathrm{d}t\mathrm{d}s.$$

Therefore, since the variation is invariant under reparametrization, we obtain

$$\|\|V\|_{(1/\alpha')\text{-var}}\|_m \lesssim_{m,\alpha,\alpha'} \tilde{T}^{\alpha} = w(0,T)^{\alpha}.$$ $\square$

4.3.2 Pathwise uniqueness

Now we prove the most important estimate of Section 4.3.

Lemma 4.3.3. *We consider*

$$U_t := \int_0^t \nabla f_r(B_r)\mathrm{d}r, \quad \textit{for } f \in C([0,T], C^2(\mathbb{R}^d)).$$

Let p, q satisfy

$$dH\Big(\frac{1}{p} - \frac{1}{p \vee 2}\Big) < \min\Big\{1 - H - \frac{1}{q}, \frac{1}{2} - H\Big\}. \tag{4.32}$$

For $0 < s < t \leq T$, we have the estimate

$$\|U_{s,t}\|_{p\vee 2} \lesssim s^{-\frac{dH}{p\vee 2}} \|f\|_{L^q([s,t];L^p_x)} (t-s)^{1-H-\frac{1}{q}-dH(\frac{1}{p}-\frac{1}{p\vee 2})}. \tag{4.33}$$

Furthermore, under (4.2a), (4.2b) and (4.5), in which case (4.32) is satisfied, for any ρ with

$$\rho^{-1} < 1 - H - dH\Big(\frac{1}{p} - \frac{1}{p \vee 2}\Big), \tag{4.34}$$

we have

$$\|\|U\|_{\rho\text{-var}}\|_{p\vee 2} \lesssim_\rho \|f\|_{L^q_t L^p_x} T^{1-H-\frac{dH}{p}-\frac{1}{q}}. \tag{4.35}$$

Remark 4.3.4. It is straightforward to see that (4.5) follows from (4.2d). Hence, the estimates of Lemma 4.3.3 hold under (4.2). Note also that the condition (4.32) enforces that the right-hand side of (4.34) is greather than $\frac{1}{2}$.

Remark 4.3.5. In fact, the estimate (4.35) holds in higher moments, see Lemma 4.4.3.

Proof. It is convenient to use the Mandelbrot–van Ness representation (1.25). We denote by $(\mathcal{F}_t)_{t\in\mathbb{R}}$ the filtration generated by W. Since U is measurable with respect to B, assuming such representation does not lose generality. To apply the stochastic sewing with control [Lê23, Theorem 3.1], we consider the germ

$$A_{s,t} := \int_s^t \mathbb{E}_s[\nabla f(B_r)]\mathrm{d}r = \int_s^t \nabla P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\mathrm{d}r, \tag{4.36}$$

where

$$P_\sigma f(x) := (2\pi\sigma)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{2\sigma}} f(y)\mathrm{d}y,$$

$$\sigma(s,r)^2 := \mathbb{E}\Big[\Big(\int_s^r \mathcal{K}(r,u)\mathrm{d}W_u^i\Big)^2\Big] = \frac{1}{2H}|r-s|^{2H}.$$

Using the estimate

$$|A_{s,t} - \nabla f(B_s)(t-s)| \le \|f\|_{C^2}\|B\|_{C^{H/2}}(t-s)^{1+\frac{H}{2}},$$

we easily see that a.s.

$$U_{s,t} = \lim_{\substack{\pi \text{ is a partition of } [s,t],\\ |\pi|\to 0}} \sum_{[u,v]\in\pi} A_{u,v}.$$

Obviously we have $\mathbb{E}[\delta A_{s,u,t}|\mathcal{F}_s] = 0$. We observe that

$$(s,t) \mapsto \|f\|^q_{L^q([s,t];L^p_x)}$$

is a control, and by [FV10, Exercise 1.9]

$$(s,t) \mapsto \|f\|^{\frac{1}{1-H}}_{L^q([s,t];L^p_x)}(t-s)^{\frac{1-H-1/q}{1-H}}$$

is a control. By the stochastic sewing with control [Lê23, Theorem 3.1], in order to obtain the estimate (4.33), it suffices to show

$$\|A_{s,t}\|_{p_\star} \lesssim s^{-\frac{dH}{p_\star}}\|f\|_{L^q([s,t];L^p_x)}(t-s)^{1-H-\frac{1}{q}-dH(\frac{1}{p}-\frac{1}{p_\star})} \tag{4.37}$$

with $p_\star := p \vee 2$. Indeed, the condition (4.32) ensures that the sum of the exponents

$$\frac{1}{q} + \Big(1 - H - \frac{1}{q} - dH\Big(\frac{1}{p} - \frac{1}{p_\star}\Big)\Big)$$

is greater than $\frac{1}{2}$.

Recalling (4.36) and applying Minkowski's inequality, we get

$$\|A_{s,t}\|_{p_\star} \le \int_s^t \|\nabla P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_{p_\star}\mathrm{d}r.$$

To compute further, we set $g_\sigma(x) := (2\pi\sigma)^{-d/2}e^{-\frac{|x|^2}{2\sigma}}$ and

$$\rho^2(s,t) := \mathbb{E}\Big[\Big(\int_{-\infty}^s \mathcal{K}(t,r)\mathrm{d}W_r^i\Big)^2\Big],$$

which is the variance of the Gaussian $\mathbb{E}_s[B_t^i]$. We denote by $*$ the convolution in space. Then

$$\begin{aligned}
&\|\nabla P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_{p_\star} \\
&\qquad = \Big(\int_{\mathbb{R}^d} |\nabla(g_{\sigma(s,r)^2} * f_r)(x)|^{p_\star} g_{\rho(s,r)^2}(x)\mathrm{d}x\Big)^{1/p_\star} \\
&\qquad \leq \|\nabla g_{\sigma(s,r)^2} * f_r\|_{L^{p_\star}(\mathbb{R}^d)} \|g_{\rho(s,r)^2}\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{p_\star}} \\
&\qquad \leq \|\nabla g_{\sigma(s,r)^2}\|_{L^{p'}(\mathbb{R}^d)} \|f_r\|_{L^p(\mathbb{R}^d)} \|g_{\rho(s,r)^2}\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{p_\star}}, \qquad (4.38)
\end{aligned}$$

where we applied Young's convolution inequality in the last step with

$$1 + \frac{1}{p_\star} = \frac{1}{p} + \frac{1}{p'}. \qquad (4.39)$$

We have

$$\|g_{\rho(s,r)^2}\|_{L^\infty(\mathbb{R}^d)} = (2\pi)^{-d/2} \rho(s,r)^{-d}$$

and

$$\begin{aligned}
\|\nabla g_{\sigma(s,r)^2}\|_{L^{p'}(\mathbb{R}^d)}^{p'} &= \int_{\mathbb{R}^d} \frac{1}{(2\pi\sigma(s,r)^2)^{dp'/2}} \Big|\frac{x}{\sigma(s,r)^2}\Big|^{p'} e^{-\frac{p'|x|^2}{2\sigma(s,r)^2}} \mathrm{d}x \\
&\lesssim \sigma(s,r)^{-(d+1)p'+d}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\nabla P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_{p_\star} &\lesssim \|f_r\|_{L^p(\mathbb{R}^d)} \sigma(s,r)^{-(d+1)+d/p'} \rho(s,r)^{-\frac{d}{p_\star}} \\
&= \|f_r\|_{L^p(\mathbb{R}^d)} \sigma(s,r)^{-1-d(\frac{1}{p}-\frac{1}{p_\star})} \rho(s,r)^{-\frac{d}{p_\star}}.
\end{aligned}$$

Since

$$\sigma(s,r) \gtrsim (r-s)^H \qquad (4.40)$$

and

$$\rho(s,r)^2 \geq \int_0^s (r-u)^{2H-1} \mathrm{d}r \gtrsim r^{2H} - (r-s)^{2H} \geq s^{2H}, \qquad (4.41)$$

we obtain

$$\|\nabla P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_p \lesssim s^{-\frac{dH}{p_\star}} (r-s)^{-H-dH(\frac{1}{p}-\frac{1}{p_\star})} \|f_r\|_{L^p(\mathbb{R}^d)}. \qquad (4.42)$$

Then,

$$
\begin{aligned}
\|A_{s,t}\|_p &\lesssim s^{-\frac{dH}{p_\star}} \int_s^t (r-s)^{-H-dH(\frac{1}{p}-\frac{1}{p_\star})} \|f_r\|_{L^p} \mathrm{d}r \\
&\le s^{-\frac{dH}{p_\star}} \Big(\int_s^t (r-s)^{-(H+dH(\frac{1}{p}-\frac{1}{p_\star}))\frac{q}{q-1}} \Big)^{1-\frac{1}{q}} \|f\|_{L^q([s,t];L^p_x)} \\
&\lesssim s^{-\frac{dH}{p_\star}} \|f\|_{L^q([s,t];L^p_x)} (t-s)^{1-H-\frac{1}{q}-dH(\frac{1}{p}-\frac{1}{p_\star})},
\end{aligned}
$$

where (4.32) is used in the second inequality to have the integral finite. The estimate (4.37) is now proven, hence so is the estimate (4.33).

We notice that the conditions (4.2a), (4.2b) and (4.5) imply (4.32). Indeed, the only non-trivial part is to see

$$
\frac{dH}{p} + H < \frac{1}{2} + \frac{dH}{2} \tag{4.43}
$$

for $p \in [(1-H)^{-1}, 2)$. If $Hd > 1 - \frac{2}{q}$, then $\frac{1}{2} + \frac{Hd}{2} > 1 - \frac{1}{q}$ and the (4.43) follows from (4.2b). If $Hd < 1 - \frac{2}{q}$, then

$$
Hd\Big(\frac{1}{p} - \frac{1}{2}\Big) < \frac{1}{p} - \frac{1}{2} \le \frac{1}{2} - H
$$

thanks to the condition $p > (1-H)^{-1}$.

To show (4.35), we thus apply Lemma 4.3.1 to obtain that

$$
\|U_{s,t}\|_p \lesssim \|f\|_{L^q([s,t];L^p_x)} (t^\beta - s^\beta)^{1-H-\frac{1}{q}-dH(\frac{1}{p}-\frac{1}{p_\star})},
$$

where $\beta := (1 - H - \frac{dH}{p} - \frac{1}{q})/(1 - H - \frac{1}{q} - dH(\frac{1}{p} - \frac{1}{p_\star}))$. We set

$$
w(s,t) := (\|f\|^q_{L^q([s,t];L^p_x)})^{\frac{1/q}{1-H-dH(\frac{1}{p}-\frac{1}{p_\star})}} (t^\beta - s^\beta)^{\frac{1-H-1/q-dH(\frac{1}{p}-\frac{1}{p_\star})}{1-H-dH(\frac{1}{p}-\frac{1}{p_\star})}}.
$$

Note that the function $(s,t) \mapsto w(s,t)$ is a control by [FV10, Exercise 1.9]. Applying Lemma 4.3.2, for any ρ satisfying (4.34), we have

$$
\|\|U\|_{\rho\text{-var}}\|_p \lesssim w(0,T)^{1-H-dH(\frac{1}{p}-\frac{1}{p_\star})},
$$

which deduces the estimate (4.35). □

Remark 4.3.6. The only places where the specific property of B is used are (4.40) and (4.41). Hence, the above proof works if B is replaced by a Gaussian process G such that

$$\|G_{s,t} - \mathbb{E}[G_{s,t}|\mathcal{F}_s]\|_2 \gtrsim (t-s)^H, \quad \|\mathbb{E}[G_{s,t}|\mathcal{F}_s]\|_2 \gtrsim s^H.$$

In particular, it works for the Riemann–Liouville process (4.29).

Remark 4.3.7. Since $W^{s,p}$ can be identified with the Triebel–Lizorkin space $F^s_{p,2}$, Young's convolution inequality for the Triebel–Lizorkin space [KS22, Theorem 2.2] yields

$$\|g * f\|_{L^{p_\star}} \lesssim \|g\|_{W^{1,p'}} \|f\|_{W^{-1,p}}, \tag{4.44}$$

where $p_\star = p \vee 2$ and p' is defined by (4.39). Using this, under (4.32) we have

$$\Big\|\int_s^t f(B_r)\mathrm{d}r\Big\|_{p_\star} \lesssim s^{-\frac{dH}{p_\star}} \|f\|_{L^q([s,t];W_x^{-1,p})} (t-s)^{1-H-dH(\frac{1}{p}-\frac{1}{p_\star})}$$

if $|t-s| \le 1$. Indeed, by the reasoning of the above proof, it suffices to show that

$$\Big\|\mathbb{E}_s \int_s^t f(B_r)\mathrm{d}r\Big\|_{p_\star} \lesssim s^{-\frac{dH}{p_\star}} \|f\|_{L^q([s,t];W_x^{-1,p})} (t-s)^{1-H-dH(\frac{1}{p}-\frac{1}{p_\star})},$$

which, as demonstrated above, follows from the estimate

$$\|P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_{p_\star} \lesssim s^{-\frac{dH}{p_\star}} \|f_r\|_{W_x^{-1,p}} (r-s)^{-H-dH(\frac{1}{p}-\frac{1}{p_\star})}, \tag{4.45}$$

corresponding to (4.42). To prove (4.45), as in the computation (4.38),

$$\begin{aligned}\|P_{\sigma(s,r)^2} f_r(\mathbb{E}_s[B_r])\|_{p_\star} &\le \|g_{\sigma(s,r)^2} * f_r\|_{L^{p_\star}(\mathbb{R}^d)} \|g_{\rho(s,r)^2}\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{p_\star}} \\ &\lesssim \|g_{\sigma(s,r)^2}\|_{W^{1,p'}(\mathbb{R}^d)} \|f_r\|_{W^{-1,p}(\mathbb{R}^d)} \|g_{\rho(s,r)^2}\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{p_\star}},\end{aligned}$$

where the last inequality is a consequence of (4.44). Since

$$\begin{aligned}\|g_{\sigma(s,r)^2}\|_{W^{1,p'}(\mathbb{R}^d)} &= \|g_{\sigma(s,r)^2}\|_{L^{p'}(\mathbb{R}^d)} + \|\nabla g_{\sigma(s,r)^2}\|_{L^{p'}(\mathbb{R}^d)} \\ &\lesssim 1 + \sigma(s,r)^{-1-d(\frac{1}{p}-\frac{1}{p_\star})},\end{aligned}$$

we obtain the estimate (4.45).

Proof of Theorem Theorem 4.1.1. As noted in Remark Remark 4.2.4, it suffices to prove pathwise uniqueness. Let $X^{(i)}$ ($i = 1, 2$) be two pathwise solutions to (4.1). Set

$$\phi^{(i)} := X^{(i)} - B - x.$$

Then, we have
$$\phi_t^{(i)} = \int_0^t b_r(x + B_r + \phi_r^{(i)})\mathrm{d}r.$$
Replacing b_r by $b_r(x + \cdot)$, we can assume $x = 0$. If (b^n) is a smooth approximation of b (see Remark Remark 4.1.5), then by [Lê20, Proposition 6.8],
$$\lim_{n\to\infty} \Big\| \int_0^\cdot b_r^{(n)}(B_r + \phi_r^{(i)})\mathrm{d}r - \int_0^\cdot b_r(B_r + \phi_r^{(i)})\mathrm{d}r \Big\|_{L^\infty([0,T])} = 0$$
in $L^m(\mathbb{P})$ for every $m \in [2, \infty)$. As in [Lê20, Lemma 6.12], we write
$$\int_0^t b^{(n)}(r, B_r + \phi_r^{(1)})\mathrm{d}r - \int_0^t b^{(n)}(r, B_r + \phi_r^{(2)})\mathrm{d}r = \int_0^t (\phi_r^{(1)} - \phi_r^{(2)}) \cdot \mathrm{d}V_r^{(n)} \quad (4.46)$$
where
$$V_r^{(n)} := \int_0^r \int_0^1 \nabla b^{(n)}(B_u + \theta\phi_u^{(1)} + (1-\theta)\phi_u^{(2)})\mathrm{d}\theta\mathrm{d}u.$$
Thus,
$$\phi_t^{(1)} - \phi_t^{(2)} = \lim_{n\to\infty} \int_0^t (\phi_r^{(1)} - \phi_r^{(2)}) \cdot \mathrm{d}V_r^{(n)}. \quad (4.47)$$
Our goal is to show that $V^{(n)}$ has some limit V.

Let $p_1 \in (1, 2)$ be sufficiently close to 1 and set $m := p/p_1$. Let ρ satisfy (4.34), and we can suppose that $\rho < 2$. We see that
$$\|\|V^{(n_1)} - V^{(n_2)}\|_{\rho\text{-var}}\|_m \le \int_0^1 \Big\|\Big\|\int_0^\cdot \nabla(b^{(n_1)} - b^{(n_2)})_u(B_u + \theta\phi_u^{(1)} + (1-\theta)\phi_u^{(2)})\mathrm{d}u\Big\|_{\rho\text{-var}}\Big\|_m \mathrm{d}\theta.$$
By Lemma 4.2.14, we have
$$\mathbb{E}e^{\lambda\|\theta\phi^{(1)} + (1-\theta)\phi^{(2)}\|^2_{\mathcal{H}^H}} \lesssim_{\lambda,b} 1, \quad \forall \lambda \ge 0.$$
Therefore, by Lemma 4.2.5 and Hölder's inequality,
$$\Big\|\Big\|\int_0^\cdot \nabla(b^{(n_1)} - b^{(n_2)})_u(B_u + \theta\phi_u^{(1)} + (1-\theta)\phi_u^{(2)})\mathrm{d}u\Big\|_{\rho\text{-var}}\Big\|_m \lesssim_{p_1,b} \Big\|\Big\|\int_0^\cdot \nabla(b^{(n_1)} - b^{(n_2)})_u(B_u)\mathrm{d}u\Big\|_{\rho\text{-var}}\Big\|_p. \quad (4.48)$$

Applying the estimate (4.35) of Lemma Lemma 4.3.3, we get

$$\|\|V^{(n_1)} - V^{(n_2)}\|_{\rho\text{-var}}\|_m \lesssim_{p_1,b,T} \|b^{(n_1)} - b^{(n_2)}\|_{L_t^q L_x^p}.$$

Therefore, there exists a stochastic process $(V_t)_{t\in[0,T]}$ such that

$$\lim_{n\to\infty} \mathbb{E}[\|V - V^{(n)}\|_{\rho\text{-var}}^m] = 0.$$

Now observe that $\phi^{(i)}$ has a finite 1-variation, as dominated by the increasing process

$$t \mapsto \int_0^t |b|_r(X_r^{(i)})\mathrm{d}r.$$

Therefore, by (4.47) a.s.

$$\phi_t^{(1)} - \phi_t^{(2)} = \int_0^t \{\phi_r^{(1)} - \phi_r^{(2)}\} \cdot \mathrm{d}V_r, \quad \phi_0^{(1)} - \phi_0^{(2)} = 0. \tag{4.49}$$

The uniqueness of Young's differential equation implies $\phi^{(1)} = \phi^{(2)}$ or $X^{(1)} = X^{(2)}$ a.s. $\square$

4.4 Stability

We have shown that under (4.2), for given x and b, there exists a unique strong solution to the SDE (4.1). In this section we are interested in the stability of the solution with respect to the input data (x, b). **Throughout this section, we assume (4.2).**

With the result of Section 4.3, a standard argument easily enables us to prove the pathwise stability. For $i \in \{1, 2\}$, let X^i be the solution to

$$\mathrm{d}X_t^i = b_t^i(X_t^i)\mathrm{d}t + \mathrm{d}B_t, \quad X_0^i = x^i. \tag{4.50}$$

We define

$$V_t := \int_0^t \int_0^1 \nabla b_r^1(\theta X_r^1 + (1-\theta)X_r^2)\mathrm{d}\theta\mathrm{d}r, \tag{4.51}$$

whose construction was discussed in the proof of Theorem 4.1.1. Therein we showed that $\|V\|_{\rho\text{-var}} < \infty$ a.s. for any ρ satisfying (4.34), and in particular we can suppose that $\rho < 2$. We also set

$$R_t := \int_0^t (b^1 - b^2)(r, X_r^2)\mathrm{d}r. \tag{4.52}$$

Let $X^{\Delta} := X^1 - X^2$. We observe

$$X^{\Delta}_{s,t} = \int_s^t \{b^1_r(X^1_r) - b^1_r(X^2_r)\}\mathrm{d}r + \int_s^t (b^1 - b^2)(r, X^2_r)\mathrm{d}r.$$

Then, as in [Lê20, Lemma 6.12] (or repeating the argument leading to (4.49)), we have

$$\int_s^t \{b^1_r(X^1_r) - b^1_r(X^2_r)\}\mathrm{d}r = \int_s^t X^{\Delta}_r \mathrm{d}V_r,$$

and

$$X^{\Delta}_{s,t} = \int_s^t X^{\Delta}_r \mathrm{d}V_r + R_{s,t}, \tag{4.53}$$

where the integral $\int_s^t X^{\Delta}_r \mathrm{d}V_r$ is understood in pathwise Young sense.

Proposition 4.4.1. *In the above setting, let $\rho < 2$ satisfy (4.34), and let $T \in (0, \infty)$. Then there exists a positive constant $c = c(\rho)$ such that a.s.*

$$\sup_{t \in [0,T]} |X^1_t - X^2_t| \leq e^{c(1 + \|V\|^{\rho}_{\rho\text{-var};[0,T]})}\big(|x^1 - x^2| + \|R\|_{\rho\text{-var};[0,T]}\big). \tag{4.54}$$

Consequently, if X^n is the solution to (4.1) with drift b^n and initial condition x^n such that $\lim_n b^n = b^{\infty}$ in $L^q_t L^p_x$ and $\lim_n x^n = x^{\infty}$, then

$$\lim_n \sup_{t \in [0,T]} |X^n_t - X^{\infty}_t| = 0 \quad \textit{in probability.}$$

Proof. From (4.53), we apply the Gronwall's estimate in ρ-variation (e.g. [CDH18, Lemma 3.3], [GG23, Lemma B.1]) to obtain the estimate (4.54).

To prove the claim on the convergence in probability, we estimate as in (4.48), but this time we pay more attention to proportional constants so that estimates are uniform with respect to n. We set

$$\begin{aligned} V^{\infty,n}_t &:= \int_0^t \int_0^1 \nabla b^{\infty}_r(\theta X^{\infty}_r + (1-\theta)X^n_r)\mathrm{d}\theta\mathrm{d}r, \\ R^{\infty,n}_t &:= \int_0^t (b^{\infty} - b^n)(r, X^n_r)\mathrm{d}r. \end{aligned}$$

We set

$$\phi^n_t := \int_0^t \nabla b^n_r(X^n_r)\mathrm{d}r$$

and $\phi^{n,\theta} := \theta\phi^\infty + (1-\theta)\phi^n$. Let $m < p \vee 2$ and m' be such that

$$\frac{1}{m} = \frac{1}{p \vee 2} + \frac{1}{m'}. \tag{4.55}$$

By Lemma 4.2.5 and Hölder's inequality we have

$$\Big\|\Big\|\int_0^\cdot \nabla b_r^\infty(B_r + \phi_r^{n,\theta})\mathrm{d}r\Big\|_{\rho\text{-var}}\Big\|_m \le \mathbb{E}[\xi_T(\phi^{n,\theta})^{-m'+1}]^{\frac{1}{m'}}\Big\|\Big\|\int_0^\cdot \nabla b_r^\infty(B_r)\mathrm{d}r\Big\|_{\rho\text{-var}}\Big\|_{p\vee 2}$$

By (4.18),

$$\begin{aligned}\mathbb{E}[\xi_T(\phi^{n,\theta})^{-m'+1}] &\le \mathbb{E}[e^{(m'-1)^2\|\phi^{n,\theta}\|^2_{\mathcal{H}^H}}]^{\frac{1}{2}} \\ &\le \mathbb{E}[e^{4(m'-1)^2\|\phi^n\|^2_{\mathcal{H}^H}}]^{\frac{1}{4}}\mathbb{E}[e^{4(m'-1)^2\|\phi^\infty\|^2_{\mathcal{H}^H}}]^{\frac{1}{4}}.\end{aligned} \tag{4.56}$$

Let $\mathfrak{e}$ and $\mathfrak{E}$ be the constants of Remark 4.2.12. We denote by $M(\lambda)$ the smallest M such that

$$\sup_n \|b^n \mathbf{1}_{\{|b^n|\ge M\}}\|_{L_t^q L_x^p} \le \mathfrak{E}\lambda^{-\mathfrak{e}}.$$

By (4.28),

$$\mathbb{E}[e^{4(m'-1)^2\|\phi^n\|^2_{\mathcal{H}^H}}] + \mathbb{E}[e^{4(m'-1)^2\|\phi^\infty\|^2_{\mathcal{H}^H}}] \lesssim_{T,m',M(4(m'-1)^2)} 1.$$

On the other hand, Lemma 4.3.3 yields

$$\Big\|\Big\|\int_0^\cdot \nabla b_r^\infty(B_r)\mathrm{d}r\Big\|_{\rho\text{-var}}\Big\|_{p\vee 2} \lesssim_{\rho,T} \|b^\infty\|_{L_t^q L_x^p}.$$

Hence, we get

$$\|\|V^{\infty,n}\|_{\rho\text{-var}}\|_m \lesssim_{\rho,m,M(4(m'-1)^2),T} \|b^\infty\|_{L_t^q L_x^p}. \tag{4.57}$$

In view of Remark 4.3.7, we similarly obtain

$$\|\|R^{\infty,n}\|_{\rho\text{-var}}\|_m \lesssim_{\rho,m,M(4(m'-1)^2),T} \|b^\infty - b^n\|_{L_t^q W_x^{-1,p}}. \tag{4.58}$$

It is important that the proportional constants do not depend on n.

Let $\delta > 0$ and $N \ge 1$ be arbitrary. By (4.54),

$$\begin{aligned}&\mathbb{P}(\sup_{t\in[0,T]} |X_t^\infty - X_t^n| \ge \delta) \\ &\qquad \le \mathbb{P}(\|V^{\infty,n}\|_{\rho\text{-var}} \ge N) + \mathbb{P}(e^{c(1+N)}(|x^\infty - x^n| + \|R^{\infty,n}\|_{\rho\text{-var}}) \ge \delta).\end{aligned}$$

By (4.57), we have

$$\lim_{N\to\infty}\sup_n \mathbb{P}(\|V^{\infty,n}\|_{\rho\text{-var}} \geq N) = 0,$$

and by (4.58), for each N we have

$$\lim_{\delta\to 0}\mathbb{P}(e^{c(1+N)}(|x^\infty - x^n| + \|R^{\infty,n}\|_{\rho\text{-var}}) \geq \delta) = 0.$$

The convergence in probability now readily follows. $\square$

Remark 4.4.2. Proposition 4.4.1 shows that the solution to (4.1) is strong, without resorting to the Yamada–Watanabe theorem.

Proposition 4.4.1 proves the first part of Theorem 4.1.2. The rest of the section is devoted to the second part (stability in moment norms).

4.4.1 VMO estimates

The easiest way to deduce the stability estimate in $L^m(\mathbb{P})$ is via the pathwise stability estimate (4.54). However, this is only possible provided that

$$\mathbb{E}[e^{\lambda\|V\|^{\rho}_{\rho\text{-var}}}] < \infty \quad \text{for any } \lambda > 0. \tag{4.59}$$

As shown in Remark 4.4.6 below, the VMO technique allows us to obtain

$$\mathbb{E}[e^{\lambda\|V\|^{(H+\frac{dH}{p})^{-1}}_{\rho\text{-var}}}] < \infty \quad \text{for any } \lambda > 0.$$

For (4.59) to hold, in view of (4.34) we must have

$$\frac{dH}{p} + H \leq 1 - H - dH\Big(\frac{1}{p} - \frac{1}{p\vee 2}\Big),$$

which does not necessarily hold under (4.2).

Hence, in the full regime of (4.2) the pathwise stability estimate in Proposition 4.4.1 does not imply stability in moment norms, and we propose a new method. Our strategy is to view (4.53) as a rough differential equation driven by a rough path lifted from V and R, and to employ Gronwall's argument in the probabilistic setting. To this end, we need to control VMO norms of the rough path lifted from V and R. In this regard, the key is to obtain a VMO-type estimate on ρ-variations of V and R, which is the goal of this section. Recall the notation of the conditional moment from (4.15). Let $\tilde{\mathfrak{e}}$ and $\tilde{\mathfrak{E}}$ be the constants of Lemma 4.2.15, and set

$$\mathcal{B}_M := \{b \in L^q_t L^p_x : \|b\mathbf{1}_{\{|b|\geq M\}}\|_{L^q_t L^p_x} \leq \tilde{\mathfrak{E}} 4^{-\tilde{\mathfrak{e}}}\}. \tag{4.60}$$

The constant $\mathfrak{c}$ in (4.3) is given by $\tilde{\mathfrak{E}}4^{-\tilde{\mathfrak{e}}}$.

Lemma 4.4.3. *Let b^1, b^2 be smooth, let M be large enough to have $b^1, b^2 \in \mathcal{B}_M$, and let $\rho < 2$ satisfy (4.34). We define V and R by (4.51) and (4.52) respectively. For any $m \in (0, \infty)$ and $s < t$, we have*

$$\|\|\|V\|_{\rho\text{-var};[s,t]}|\mathcal{F}_s\|_m\|_\infty \lesssim_{T,\rho,m,M} \|b^1\|_{L^q([s,t];L^p_x)}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}},$$
$$\|\|\|R\|_{\rho\text{-var};[s,t]}|\mathcal{F}_s\|_m\|_\infty \lesssim_{T,\rho,m,M} \|b^1 - b^2\|_{L^q([s,t];W_x^{-1,p})}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}.$$

Remark 4.4.4. To proceed further, we make the following simple observation. Our goal is to estimate

$$\|\|X^1 - X^2\|_{C^\gamma}\|_m. \tag{4.61}$$

As we know from Theorem 4.1.1 that X^1, X^2 are strong solutions, (4.61) is determined by the law of the driver B. In particular, for the sake of estimating (4.61), we can assume that B has the Mandelbrot–van Ness representation (1.25).

As validated in the above remark, without loss of generality, we assume the representation (1.25). This will simplify our arguments. We denote by $(\mathcal{F}_t)_{t\in\mathbb{R}}$ the filtration generated by W in the representation (1.25).

To prove Lemma 4.4.3, our strategy is to redo the argument of Lemma 4.3.3 under conditioning. For a while, as assumed in Lemma 4.4.3, we suppose that b^1 and b^2 are smooth. Furthermore, we fix a $v \in [0, T)$. For any continuous function η, the differential equation

$$\mathrm{d}x_t = b^i_{v+t}(x_t)\mathrm{d}t + \mathrm{d}\eta_t, \quad x_0 = y$$

has a unique solution. Recall that we write $\tilde{B}$ for the Riemann–Liouville process defined by (4.29). We denote by $\tilde{X}^i[y,\eta]$ the unique solution to the SDE

$$\mathrm{d}\tilde{X}^i_t[y,\eta] = b^i_{v+t}(\tilde{X}^i_t[y,\eta])\mathrm{d}t + \mathrm{d}\eta_t + \mathrm{d}\tilde{B}_t, \quad \tilde{X}^i_0[y,\eta] = y.$$

Note that $\tilde{X}^i[y,\eta]$ is adapted to the filtration generated by $\tilde{B}$.

We then have

$$\begin{aligned}&\mathbb{E}[F((X^1_r, X^2_r)_{v\le r\le r+t})|\mathcal{F}_v]\\ &\qquad = \mathbb{E}[F((\tilde{X}^1_r[y^1,\eta], \tilde{X}^2_r[y^2,\eta]))]|_{y^i = X^i_v, \eta = \int_{-\infty}^v \mathcal{K}(v+\cdot,r)\mathrm{d}W_r}.\end{aligned} \tag{4.62}$$

Due to this relation, we will work on $\tilde{X}^i[y,\eta]$. We set

$$\tilde{V}_t[y^1,y^2,\eta] := \int_0^t \int_0^1 \nabla b^1_{v+r}(\theta \tilde{X}^1_r[y^1,\eta] + (1-\theta)\tilde{X}^2_r[y^2,\eta])\mathrm{d}\theta \mathrm{d}r, \tag{4.63}$$

$$\tilde{U}_t := \int_0^t \nabla b^1_{v+r}(\tilde{B}_r)\mathrm{d}r. \tag{4.64}$$

As in the proof of Theorem 4.1.1, we deduce the estimate of $\tilde{V}$ from $\tilde{U}$ by Girsanov's theorem (Lemma 4.2.15). The following is an improvement of the estimate (4.35), in that we can remove the restriction on the moment.

Lemma 4.4.5. *Let $\rho < 2$ satisfy (4.34) and $m \in (0,\infty)$. We define $\tilde{U}$ by (4.64) with a smooth b^1. We then have*

$$\| \|\tilde{U}\|_{\rho\text{-var};[0,t]}\|_m \lesssim_{\rho,m} \|b^1\|_{L^q([v,v+t];L^p_x)} t^{1-H-\frac{dH}{p}-\frac{1}{q}}.$$

Proof. Let $(\mathcal{G}_t)$ be the filtration generated by $\tilde{B}$. We first prove

$$\| \| \|\tilde{U}\|_{\rho\text{-var};[s,t]}|\mathcal{G}_s\|_{p\vee 2}\|_\infty \lesssim_\rho \|b^1\|_{L^q([v+s,v+t];L^p_x)}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}. \tag{4.65}$$

Setting

$$\bar{U}_t[\eta] := \int_0^t \nabla b^1_{v+r}(\eta_r + \tilde{B}_r)\mathrm{d}r,$$

we have

$$\| \|\tilde{U}\|_{\rho\text{-var};[s,t]}|\mathcal{G}_s\|_{p\vee 2} = \| \|\bar{U}[\eta]\|_{\rho\text{-var};[0,t-s]}\|_{p\vee 2}|_{\eta=\int_0^s \mathcal{K}(\cdot,r)\mathrm{d}W_r}.$$

Since $\bar{b}^1_{v+t}(x) := b^1_{v+t}(x+\eta_t)$ satisfies $\|\bar{b}^1\|_{L^q_tL^p_x} = \|b^1\|_{L^q_tL^p_x}$, it suffices to prove (4.65) with $s=0$, i.e.

$$\| \|\tilde{U}\|_{\rho\text{-var};[0,t]}\|_{p\vee 2} \lesssim_\rho \|b^1\|_{L^q([v,v+t];L^p_x)} t^{1-H-\frac{dH}{p}-\frac{1}{q}},$$

which is essentially proven in Lemma 4.3.3 as noted in Remark 4.3.6.

Now we set $\mathfrak{U}_t := \|\tilde{U}\|_{\rho\text{-var};[0,t]}$. The process $\mathfrak{U}$ is continuous (recall that b^1 is smooth), and due to the second inequality of (4.14), we have $\mathfrak{U}_{s,t} \leq \|\tilde{U}\|_{\rho\text{-var};[s,t]}$. Therefore, the estimate (4.65) implies that

$$\| \|\mathfrak{U}_{s,t}|\mathcal{F}_s\|_1\|_\infty \lesssim_\rho \|b^1\|_{L^q([v+s,v+t];L^p_x)}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}.$$

The John–Nirenberg inequality (4.21) thus completes the proof. $\square$

Remark 4.4.6. Setting $U_t := \int_0^t \nabla b^1_r(B_r)\mathrm{d}r$, we similarly get

$$\|\|\|\|U\|_{\rho\text{-var};[s,t]}|\mathcal{F}_s\|_{p\vee 2}\|_\infty \lesssim_\rho \|b^1\|_{L^q([s,t];L^p_x)}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}.$$

Hence, the John–Nirenberg inequality (4.20) yields

$$\mathbb{E}[e^{\lambda\|U\|_{\rho\text{-var}}^{(H+\frac{dH}{p})^{-1}}}] < \infty$$

for small λ. As in Lemma 4.2.10, we can remove the smallness condition of λ. Furthermore, by Girsanov's theorem we can replace U by V defined by (4.51).

Corollary 4.4.7. *Let $\rho < 2$ satisfy (4.34) and $m \in (0,\infty)$. If M is large enough to have $b^1, b^2 \in \mathcal{B}_M$ and if $\tilde{V}[y^1,y^2,\eta]$ is defined by (4.63), then for any $t \le T$ we have*

$$\|\|\tilde{V}[y^1,y^2,\eta]\|_{\rho\text{-var};[0,t]}\|_m \lesssim_{T,\rho,m,M} \|b^1\|_{L^q([v,v+t];L^p_x)} t^{1-H-\frac{dH}{p}-\frac{1}{q}}. \tag{4.66}$$

Proof. We drop dependence on y^1, y^2, η. Setting

$$\tilde{\phi}^i_t := \int_0^t b^i_{v+r}(X^1_r[y^i,\eta_r])\mathrm{d}r,$$

then

$$\tilde{V}_{s,t} = \int_s^t \int_0^1 \nabla \tilde{b}^1_r(\theta \tilde{X}^1_r + (1-\theta)\tilde{X}^2_r)\mathrm{d}\theta\mathrm{d}r.$$

The Girsanov theorem (Lemma 4.2.15) and the bound similar to (4.56) yield

$$\begin{aligned}\|\|V\|_{\rho\text{-var};[0,t]}\|_m \lesssim_{T,m} \mathbb{E}\Big[e^{4\|\tilde{\phi}^1\|^2_{\tilde{\mathcal{H}}^H}}\Big]^{\frac{1}{4}} \mathbb{E}\Big[e^{4\|\tilde{\phi}^2\|^2_{\tilde{\mathcal{H}}^H}}\Big]^{\frac{1}{4}} \\ \times \int_0^1 \Big\|\Big\|\int_0^\cdot \nabla b^1_{v+r}(\tilde{B}_r + \theta y^1 + (1-\theta)y^2)\mathrm{d}r\Big\|_{\rho\text{-var};[0,t]}\Big\|_{2m}\mathrm{d}\theta,\end{aligned}$$

where m' is defined by (4.55). By Lemma 4.4.5,

$$\begin{aligned}\Big\|\Big\|\int_0^\cdot \nabla b^1_{v+r}(\tilde{B}_r + \theta y^1 + (1-\theta)y^2)\mathrm{d}r\Big\|_{\rho\text{-var};[0,t]}\Big\|_{2m} \\ \lesssim_{\rho,m} \|b^1\|_{L^q([v,v+t];L^p_x)} t^{1-H-\frac{dH}{p}-\frac{1}{q}}.\end{aligned}$$

By Lemma 4.2.15, if $b^i \in \mathfrak{B}_M$, we get

$$\mathbb{E}\Big[e^{4\|\tilde{\phi}^i\|^2_{\tilde{\mathcal{H}}^H}}\Big] \lesssim_{T,M} 1,$$

and the proof is complete. □

Proof of Lemma 4.4.3. In view of (4.62) and Corollary 4.4.7, the estimate for V follows. As for R, the proof is similar in view of Remark 4.3.7. □

4.4.2 Lifting paths

As in the previous section, we assume that b^1, b^2 are smooth, and we define V and R by (4.51) and (4.52) respectively. Lemma 4.4.3 shows that

$$\|\|V_{s,t}|\mathcal{F}_s\|_m\|_\infty \lesssim_m \|b^1\|_{L^q([s,t];L^p_x)}(t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}},$$

that is the VMO regularity of V, and similarly that of R, is $1-H-\frac{dH}{p}$. This exponent can be arbitrarily close to 0 under (4.2). To get a closed estimate of X^Δ from the Young–Lyons affine equation (4.53), we have to lift V and R as a rough path of sufficient order. Here we recall notation and very basic properties of rough paths.

Let I and Θ be the index sets defined by

$$I := \{1,2,\dots,d\}^2 \quad \text{and} \quad \Theta := \bigcup_{n=1}^{\infty} I^{\{1,2,\dots,n\}}. \tag{4.67}$$

For $w \in \Theta$, we write $|w| = n$ if $w \in I^{\{1,2,\dots,n\}}$. For $w \in \Theta$ and $k < |w|$ we set

$$w_k^+ := w|_{\{1,2,\dots,k\}}, \quad w_k^- := w|_{\{|w|-k+1,|w|-k+2,\dots,|w|\}}.$$

We define the lift $\mathbb{V} = (\mathbb{V}^w)_{w\in\Theta}$ from V as follows. If $|w| = 1$ then we set $\mathbb{V}^w_{s,t} := V_t^{w(1)} - V_s^{w(1)}$. For $|w| = n$, we define inductively

$$\mathbb{V}^w_{s,t} := \int_s^t \mathbb{V}^{w_{n-1}^+}_{s,r} \mathrm{d}V_r^{w(|w|)}$$

as Young's integral [You36]. The lift $\mathbb{V}$ satisfies the algebraic condition called Chen's relation:

$$\mathbb{V}^w_{s,t} - \mathbb{V}^w_{s,u} - \mathbb{V}^w_{u,t} = \sum_{k=1}^{n-1} \mathbb{V}^{w_{n-k}^+}_{s,u} \mathbb{V}^{w_k^-}_{u,t}, \quad \forall s \le u \le t. \tag{4.68}$$

It also follows from [Lyo98] that for each $w \in \Theta$, there exists a (deterministic) constant $C(w)$ such that

$$|\mathbb{V}^w_{s,t}| \le C(w)\|V\|^{|w|}_{\rho\text{-var};[s,t]}, \quad \forall s \le t. \tag{4.69}$$

In the same way, we can construct a joint lift

$$(\mathbb{U}^{jw})_{j\in\{1,\dots,d\};w\in\Theta}$$

of R and V as follows. For $|w|=1$ we set

$$\mathbb{U}^{jw}_{s,t} := \int_s^t R^j_{s,r} \mathrm{d}V^{w(1)}_r$$

and inductively for $|w|>1$,

$$\mathbb{U}^{jw}_{s,t} := \int_s^t \mathbb{U}^{jw^+_{|w|-1}}_{s,r} \mathrm{d}V^{w(|w|)}_r,$$

where integrals are defined in Young sense. We similarly have Chen's identity and

$$|\mathbb{U}^{jw}_{s,t}| \lesssim_w \|R\|_{\rho\text{-var};[s,t]} \|V\|^{|w|}_{\rho\text{-var};[s,t]}. \tag{4.70}$$

Lemma 4.4.8. *Let M be large enough to have $b^1, b^2 \in \mathcal{B}_M$. For every $m \in (0,\infty)$, we have*

$$\begin{aligned}
\|\|\|\mathbb{V}^w_{s,t}|\mathcal{F}_s\|_m\|_\infty &\lesssim_{m,M,w} \|b^1\|_{L^q([s,t];L^p_x)} (t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}, \\
\|\|\|\mathbb{U}^{jw}_{s,t}|\mathcal{F}_s\|_m\|_\infty &\lesssim_{m,M,w} \|b^1-b^2\|_{L^q([s,t];W^{-1,p}_x)} (t-s)^{1-H-\frac{dH}{p}-\frac{1}{q}}.
\end{aligned}$$

Proof. It easily follows from Lemma 4.4.3, (4.69) and (4.70). □

4.4.3 Gronwall's argument in the probabilistic setting

The goal of this section is to prove the second part of Theorem 4.1.2, see the end of this section. As in the previous sections, we define V and R by (4.51) and (4.52). The signatures $(\mathbb{V}^w)_w$ and $(\mathbb{U}^{jw})_{j,w}$ are discussed in Section 4.4.2. We assume that b^1 and b^2 are smooth, until we come to the proof of Theorem 4.1.2.

In the rest of this section, we fix $m \in [2,\infty)$ and set

$$\gamma := 1 - H - \frac{dH}{p} - \frac{1}{q}.$$

Recalling (4.53), set $Z := X^\Delta$ and $z := x^1 - x^2$ so that

$$Z^i_t = z^i + \sum_{j=1}^d \int_0^t Z^j_r \mathrm{d}V^{ij}_r + R^i_t, \quad i = 1,2,\dots,d. \tag{4.71}$$

Our strategy is to employ Gronwall's arguments directly in $L^m(\mathbb{P})$, using VMO estimates from Lemma 4.4.8.

For $w \in \Theta$ with $w(i) = (a_1, a_2)$ we write $w(i)(j) := a_j$ ($j = 1, 2$). We denote by Ξ_n^{ji} the subset of Θ such that $w \in \Xi_n^{ji}$ if and only if

$$|w| \le n, \quad w(1)(1) = j, \quad w(|w|)(2) = i,$$
$$w(i)(2) = w(i+1)(1) \quad \forall i \le |w| - 1.$$

We then set

$$\llbracket Z \rrbracket_{s,t} := \sup_{s \le u < v \le t} \frac{\|Z_{u,v}\|_m}{(v-u)^\gamma} + \sum_{i=1}^{d} \sum_{k=1}^{\lfloor \gamma^{-1} \rfloor - 1} \sup_{s \le u < v \le t} (v-u)^{-(k+1)\gamma} \times \Big\| Z_{u,v}^i - R_{u,v}^i - \sum_{j=1}^{d} \sum_{w \in \Xi_k^{ji}} Z_u^j \mathbb{V}_{u,v}^w - \sum_{j=1}^{d} \sum_{w \in \Xi_{k-1}^{ji}} \mathbb{U}_{u,v}^{jw} \Big\|_m.$$

Since b^1, b^2 are smooth, we have $\llbracket Z \rrbracket \lesssim_{\|b^1\|_{C^1}, \|b^2\|_{C^1}} 1$.

Lemma 4.4.9. *Let $b^1, b^2 \in \mathcal{B}_M$. If we set*

$$A_{s,t}^i := \sum_{j=1}^{d} \Big(\sum_{w \in \Xi_n^{ji}} Z_s^j \mathbb{V}_{s,t}^w + \sum_{w \in \Xi_{n-1}^{ji}} \mathbb{U}_{s,t}^{jw} \Big),$$

then with $n = \lfloor \gamma^{-1} \rfloor$

$$\Big\| \sum_{j=1}^{d} \int_s^t Z_r^j \mathrm{d} V_r^{ji} - A_{s,t}^i \Big\|_m \lesssim_{T,m,M} (\|b^1\|_{L^q([s,t];L_x^p)} \vee \|b^1\|_{L^q([s,t];L_x^p)}^n) \llbracket Z \rrbracket_{s,t} (t-s)^{(n+1)\gamma}.$$

Proof. Since

$$|\mathbb{V}_{s,t}^w| \lesssim_{\|b^1\|_{C^1}} (t-s)^{|w|}, \quad |\mathbb{U}_{s,t}^{jw}| \lesssim_{\|b^1 - b^2\|_{L^\infty}} (t-s)^{1+|w|},$$

we have

$$\lim_{\substack{\pi \text{ is a partition of } [s,t], \\ |\pi| \to 0}} \sum_{[u,v] \in \pi} A_{u,v}^i = \lim_{\substack{\pi \text{ is a partition of } [s,t], \\ |\pi| \to 0}} \sum_{[u,v] \in \pi} \sum_{j=1}^{d} Z_u^j V_{u,v}^{ji} = \sum_{j=1}^{d} \int_s^t Z_r^j \mathrm{d} V_r^{ji}.$$

Hence, to obtain the claimed estimate, we apply the sewing lemma in $L^m(\mathbb{P})$. We have

$$\delta A^i_{s,u,t} = \sum_{j=1}^d \Big\{ -\sum_{w\in\Xi^{ji}_n} Z^j_{s,u}\mathbb{V}^w_{u,t} \\ + \sum_{w\in\Xi^{ji}_n} Z^j_s(\mathbb{V}^w_{s,t}-\mathbb{V}^w_{s,u}-\mathbb{V}^w_{u,t}) + \sum_{w\in\Xi^{ji}_{n-1}} (\mathbb{U}^{jw}_{s,t}-\mathbb{U}^{jw}_{s,u}-\mathbb{U}^{jw}_{u,t})\Big\}.$$

By Chen's identity (4.68),

$$\mathbb{V}^w_{s,t}-\mathbb{V}^w_{s,u}-\mathbb{V}^w_{u,t} = \sum_{k=1}^{|w|-1} \mathbb{V}^{w^+_{|w|-k}}_{s,u}\mathbb{V}^{w^-_k}_{u,t},$$

$$\mathbb{U}^{jw}_{s,t}-\mathbb{U}^{jw}_{s,u}-\mathbb{U}^{jw}_{u,t} = R^j_{s,u}\mathbb{V}^w_{u,t} + \sum_{k=1}^{|w|-1} \mathbb{U}^{jw^+_{|w|-k}}_{s,u}\mathbb{V}^{w^-_k}_{u,t}.$$

Therefore,

$$\delta A^i_{s,u,t} = -\sum_{j=1}^d \sum_{w\in\Xi^{ji}_{n-1}} \Big(Z^j_{s,u}-R^j_{s,u} \\ -\sum_{l=1}^d \sum_{w_1\in\Xi^{lj}_{n-|w|}} Z^l_s\mathbb{V}^{w_1}_{s,u} - \sum_{l=1}^d \sum_{w_2\in\Xi^{lj}_{n-1-|w|}} \mathbb{U}^{lw_2}_{s,u}\Big)\mathbb{V}^w_{u,t} \\ -\sum_{j=1}^d \sum_{w\in\Xi^{ji}_n\setminus\Xi^{ji}_{n-1}} Z^j_{s,u}\mathbb{V}^w_{u,t}.$$

Using the estimate $\|\|\mathbb{V}^w_{u,t}|\mathcal{F}_u\|_m\|_\infty \lesssim_{T,m,M,w} \|b\|^{|w|}_{L^q([u,t];L^p_x)}(t-s)^{|w|\gamma}$ from Lemma 4.4.8, we obtain

$$\|\delta A^i_{s,u,t}\|_m \lesssim_m (\|b\|_{L^q([u,t];L^p_x)} \vee \|b\|^n_{L^q([u,t];L^p_x)})[\![Z]\!]_{s,t}(t-s)^{(n+1)\gamma}.$$

By our choice of n, the exponent is greater than 1. The claimed estimate follows from the sewing lemma. □

Lemma 4.4.10. *Let $b^1, b^2 \in \mathcal{B}_M$. With some positive constant*

$$C = C(d,H,p,q,T,m,M)$$

we have

$$\|Z_{s,t}\|_m \le C(1\vee\Gamma^n)e^{C(1\vee\Gamma)^{\frac{n}{\gamma}}}\left\{|z| + \|b^1-b^2\|_{L_t^q W_x^{-1,p}}\right\}(t-s)^\gamma,$$

where $n=\lfloor\gamma^{-1}\rfloor$, $\Gamma := \|b^1\|_{L_t^q L_x^p}$ *and* $z = Z_0 = x^1-x^2$.

Proof. By (4.71) we have

$$\begin{aligned} Z^i_{s,t} - R^i_{s,t} - \sum_{j=1}^d \Big(\sum_{w\in\Xi_n^{ji}} Z_s^j \mathbb{V}^w_{s,t} + \sum_{w\in\Xi^{ji}_{n-1}} \mathbb{U}^{jw}_{s,t}\Big) \\ = \sum_{j=1}^d\int_s^t Z_r^j \mathrm{d}V_r^{ji} - \sum_{j=1}^d \Big(\sum_{w\in\Xi_n^{ji}} Z_s^j \mathbb{V}^w_{s,t} + \sum_{w\in\Xi^{ji}_{n-1}} \mathbb{U}^{jw}_{s,t}\Big).\end{aligned}$$

Note by Lemma 4.4.8 that

$$\begin{aligned}\|Z_s^j\mathbb{V}^w_{s,t}\|_m &\lesssim_{T,m,M} \|b^1\|_{L_t^qL_x^p}\|Z_s^j\|_m(t-s)^{|w|\gamma},\\ \|\mathbb{U}^{jw}_{s,t}\|_m &\lesssim_{T,m,M} \|b^1\|^{|w|}_{L_t^qL_x^p}\|b^1-b^2\|_{L_t^qW_x^{-1,p}}(t-s)^{(|w|+1)\gamma}.\end{aligned}$$

Therefore, if $t-s\le 1$, Lemma 4.4.9 yields (with $\Gamma = \|b^1\|_{L_t^qL_x^p}$)

$$\begin{aligned}\Big\| Z^i_{s,t} - R^i_{s,t} - \sum_{j=1}^d \Big(\sum_{w\in\Xi_k^{ji}} Z_s^j \mathbb{V}^w_{s,t} + \sum_{w\in\Xi^{ji}_{k-1}} \mathbb{U}^{jw}_{s,t}\Big)\Big\|_m \\ \lesssim_{T,m,M} (\Gamma\vee\Gamma^n)\Big\{[\![Z]\!]_{s,t}(t-s)^{(n+1)\gamma} + \|Z_s\|_m(t-s)^{(k+1)\gamma} \\ + \|b^1-b^2\|_{L_t^qW_x^{-1,p}}(t-s)^{(k+1)\gamma}\Big\}.\end{aligned}$$

This implies

$$[\![Z]\!]_{s,t} \le C(1\vee\Gamma^n)\big\{[\![Z]\!]_{s,t}(t-s)^\gamma + \sup_{r\in[s,t]}\|Z_r\|_m + \|b^1-b^2\|_{L_t^qW_x^{-1,p}}\big\}$$

for some $C = C(d,H,p,q,T,m,M)$. Hence, if

$$(t-s)^\gamma \le \frac{1}{2}(C(1\vee\Gamma^n))^{-1},$$

we have

$$[\![Z]\!]_{s,t} \le 2C(1\vee\Gamma^n)\big\{\sup_{r\in[s,t]}\|Z_r\|_m + \|b^1-b^2\|_{L_t^qW_x^{-1,p}}\big\}.$$

In particular,

$$\|Z_{s,t}\|_m \lesssim_{T,m,M} (1 \vee \Gamma^n)\big\{ \sup_{r\in[s,t]} \|Z_r\|_m + \|b^1 - b^2\|_{L_t^q W_x^{-1,p}} \big\}(t-s)^{\gamma}. \tag{4.72}$$

Now set $G_t := \|Z_t\|_m + \|b^1 - b^2\|_{L_t^q W_x^{-1,p}}$. By (4.72)

$$|G_{s,t}| \lesssim_{T,m,M} (1 \vee \Gamma^n) G_s (t-s)^{\gamma} \quad \text{if } t-s \lesssim_{T,m,M} (1 \vee \Gamma^n)^{-\frac{1}{\gamma}}.$$

By the Gronwall lemma (e.g. [Dey+19, Lemma 2.12]), we obtain

$$\sup_{t\in[0,T]} G_t \lesssim e^{CT(\Gamma\vee 1)^{\frac{n}{\gamma}}} G_0 = e^{CT(\Gamma\vee 1)^{\frac{n}{\gamma}}} \big\{ |z| + \|b^1 - b^2\|_{L_t^q W_x^{-1,p}} \big\}.$$

Plugging this estimate into (4.72), we obtain the claimed estimate. □

Proof of Theorem 4.1.2. The pathwise stability is proven by Proposition 4.4.1, with $A = c(1 + \|V\|_{\rho\text{-var}}^{\rho})$ and $D = \|R\|_{\rho\text{-var}}$. To prove the stability in moment norms, thanks to the pathwise stability, we may suppose that b^1, b^2 are smooth. Then the claim follows from Lemma 4.4.10. □

Bibliography

[ALL23] L. Anzeletti, K. Lê, and C. Ling. *Path-by-path uniqueness for stochastic differential equations under Krylov-Röckner condition*. 2023. arXiv: 2304.06802 [math.PR].

[Ath+21] S. Athreya, O. Butkovsky, K. Lê, and L. Mytnik. *Well-posedness of stochastic heat equation with distributional drift and skew stochastic heat equation*. 2021. arXiv: 2011.13498 [math.PR].

[BNP20] D. Baños, T. Nilssen, and F. Proske. "Strong Existence and Higher Order Fréchet Differentiability of Stochastic Flows of Fractional Brownian Motion Driven SDEs with Singular Drift". *Journal of Dynamics and Differential Equations* 32.4 (2020), pp. 1819–1866.

[BFG21] C. Bellingeri, P. K. Friz, and M. Gerencsér. "Singular paths spaces and applications". *Stochastic Analysis and Applications* (Oct. 2021), pp. 1–24.

[BM19] O. Butkovsky and L. Mytnik. "Regularization by noise and flows of solutions for a stochastic heat equation". *The Annals of Probability* 47.1 (Jan. 2019), pp. 165–212.

[CLL13] T. Cass, C. Litterer, and T. Lyons. "Integrability and tail estimates for Gaussian rough differential equations". *The Annals of Probability* 41.4 (July 2013), pp. 3026–3050.

[CG16] R. Catellier and M. Gubinelli. "Averaging along irregular curves and regularisation of ODEs". *Stochastic Processes and their Applications* 126.8 (2016), pp. 2323–2366.

[CD22] R. Catellier and R. Duboscq. *Regularization by noise for rough differential equations driven by Gaussian rough paths*. 2022. arXiv: 2207.04251 [math.PR].

[CDH18] N. D. Cong, L. H. Duc, and P. T. Hong. "Nonautonomous Young Differential Equations Revisited". *Journal of Dynamics and Differential Equations* 30.4 (2018), pp. 1921–1943.

[DG22] K. Dareiotis and M. Gerencsér. *Path-by-path regularisation through multiplicative noise in rough, Young, and ordinary differential equations*. 2022. arXiv: 2207.03476 [math.PR].

[DGL21] K. Dareiotis, M. Gerencsér, and K. Lê. *Quantifying a convergence theorem of Gyöngy and Krylov*. 2021. arXiv: 2101.12185 [math.PR].

[Dav07] A. M. Davie. "Uniqueness of Solutions of Stochastic Differential Equations". *International Mathematics Research Notices* 2007 (Dec. 2007), rnm124.

[Dey+19] A. Deya, M. Gubinelli, M. Hofmanová, and S. Tindel. "A priori estimates for rough PDEs with application to rough conservation laws". *Journal of Functional Analysis* 276.12 (2019), pp. 3577–3645.

[FHL23] P. K. Friz, A. Hocquet, and K. Lê. *Rough stochastic differential equations*. 2023. arXiv: 2106.10340 [math.PR].

[FV10] P. K. Friz and N. B. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[GG23] L. Galeati and M. Gerencsér. *Solution theory of fractional SDEs in complete subcritical regimes*. 2023. arXiv: 2207.03475 [math.PR].

[GH22] L. Galeati and F. A. Harang. "Regularization of multiplicative SDEs through additive noise". *The Annals of Applied Probability* 32.5 (Oct. 2022), pp. 3930–3963.

[GL23] L. Galeati and C. Ling. "Stability estimates for singular SDEs and applications". *Electronic Journal of Probability* 28.none (2023), pp. 1–31.

[GM23] P. Gassiat and Ł. Mądry. "Perturbations of singular fractional SDEs". *Stochastic Processes and their Applications* 161 (2023), pp. 137–172.

[GLY15] R. Gorenflo, Y. Luchko, and M. Yamamoto. "Time-fractional diffusion equation in the fractional Sobolev spaces". *Fractional Calculus and Applied Analysis* 18.3 (2015), pp. 799–820.

[Kry21] N. V. Krylov. "On strong solutions of Itô's equations with σ in W^1_d and b in L_d". *The Annals of Probability* 49.6 (Nov. 2021), pp. 3142–3167.

[KR05] N. V. Krylov and M. Röckner. "Strong solutions of stochastic equations with singular time dependent drift". *Probability Theory and Related Fields* 131.2 (2005), pp. 154–196.

[KS22] F. Kühn and R. L. Schilling. "Convolution inequalities for Besov and Triebel–Lizorkin spaces, and applications to convolution semigroups". *Studia Mathematica* 262 (2022). Published online: 12 July 2021, pp. 93–119.

[Kur07] T. G. Kurtz. "The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities". *Electron. J. Probab* 12 (2007), pp. 951–965.

[Lê20] K. Lê. "A stochastic sewing lemma and applications". *Electron. J. Probab.* 25 (2020), Paper No. 38, 55.

[Lê22] K. Lê. *Quantitative John–Nirenberg inequality for stochastic processes of bounded mean oscillation*. 2022. arXiv: 2210.15736 [math.PR].

[Lê23] K. Lê. "Stochastic sewing in Banach spaces". *Electronic Journal of Probability* 28.none (Jan. 2023), pp. 1–22.

[LL22] K. Lê and C. Ling. *Taming singular stochastic differential equations: A numerical method*. 2022. arXiv: 2110.01343 [math.PR].

[Lyo94] T. Lyons. "Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young". *Math. Res. Lett.* 1.4 (1994), pp. 451–464.

[Lyo98] T. J. Lyons. "Differential equations driven by rough signals". *Rev. Mat. Iberoamericana* 14.2 (1998), pp. 215–310.

[NO02] D. Nualart and Y. Ouknine. "Regularization of differential equations by fractional noise". *Stochastic Processes and their Applications* 102.1 (2002), pp. 103–116.

[NO03] D. Nualart and Y. Ouknine. "Stochastic Differential Equations with Additive Fractional Noise and Locally Unbounded Drift". In: *Stochastic Inequalities and Applications*. Ed. by E. Giné, C. Houdré, and D. Nualart. Basel: Birkhäuser Basel, 2003, pp. 353–365.

[RZ21] M. Röckner and G. Zhao. *SDEs with critical time dependent drifts: strong solutions*. 2021. arXiv: 2103.05803 [math.PR].

[RZ23] M. Röckner and G. Zhao. "SDEs with critical time dependent drifts: Weak solutions". *Bernoulli* 29.1 (Feb. 2023), pp. 757–784.

[Ver81] A. J. Veretennikov. "On strong solutions and explicit formulas for solutions of stochastic integral equations". *Mathematics of the USSR-Sbornik* 39.3 (1981), p. 387.

[You36] L. C. Young. "An inequality of the Hölder type, connected with Stieltjes integration". *Acta Math.* 67.1 (1936), pp. 251–282.

[Zha10] X. Zhang. "Stochastic flows of SDEs with irregular coefficients and stochastic transport equations". *Bulletin des Sciences Mathématiques* 134.4 (2010), pp. 340–378.

[Zvo74] A. K. Zvonkin. "A transformation of the phase space of a diffusion process that removes the drift". *Mathematics of the USSR-Sbornik* 22.1 (1974), p. 129.