Research Internship (PRE)

Field of Study: SIM
Scholar Year: 2017-2018

Numerical schemes for multidimensional SDEs with distributional coefficients

Confidentiality Notice

Non-confidential report and publishable on Internet

Author: Maximilien Germain

Promotion 2019

ENSTA ParisTech Tutor:
Francesco Russo

Host Organization Tutors:
Elena Issoglio, Tiziano de Angelis

Internship from 05-14-2018 to 08-03-2018

Host Organization: University of Leeds
Address: Leeds
LS2 9JT
United Kingdom

Confidentiality Notice

This present document is not confidential. It can be communicated outside in paper format or distributed in electronic format.

Acknowledgment

I want to express my gratitude to my internship supervisors Elena Issoglio and Tiziano de Angelis. They introduced me to research in mathematics and guided me through this project. Their availability and commitment have been really helpful during my stay in the University of Leeds.

I would also like to thank Francesco Russo for putting me in touch with Elena and Tiziano, and being my ENSTA supervisor for this research project.

Abstract

This report constructs an approximation method for multidimensional stochastic differential equations with distributional drift. It applies the Euler-Maruyama scheme to a new class of problems with irregular coefficients, after approximating the drift by Haar wavelets. We study the convergence of our algorithm, give a rate in dimension one, and apply it to the special case in which the distributional drift is obtained as the derivative of a sample path of a fractional Brownian motion.

Keywords and phrases: Stochastic differential equations; distributional drift; numerical simulation; fractional Brownian motion; Euler-Maruyama scheme; Haar wavelets.

Contents

Confidentiality Notice **3**

Acknowledgment **5**

Abstract **7**

Introduction **13**

I Context and algorithm presentation **15**

I.1 Aims and notations . . . 15
I.2 Principle of the approximation algorithm . . . 16
I.3 Virtual solutions of the original SDE and its approximation . . . 17
I.4 Approximation of the drift . . . 18

II Convergence of the algorithm **21**

II.1 Weak convergence rate of the Euler approximation . . . 21
II.2 Convergence in law of the approximated solution to the virtual solution . . . 22
II.3 Weak convergence rate of the approximated solution in dimension one . . . 23
II.3.1 Useful lemmas . . . 23
II.3.2 Main result . . . 26

III Numerical aspects **29**

III.1 Numerical simulation of fractional Brownian motion . . . 29
III.1.1 Refining a sample path of a fractional Brownian motion . . . 29
III.2 Special case of the derivative of a fractional Brownian motion . . . 31
III.2.1 Applying the framework from Flandoli, Issoglio and Russo . . . 35

IV Numerical results **37**

IV.1 Monte-Carlo method for error estimation . . . 37
IV.2 Strong convergence of the Euler approximation . . . 39
IV.3 Weak convergence of the approximated solution to the virtual solution . . . 39

Internship's organisation **41**

Conclusion **43**

Abbreviations **47**

A Technical proofs **49**
A.1 Lemma 22 49
A.2 Lemma 25 50
A.3 Proposition 26 52

List of Figures

I.1 The set $K(\beta, q)$. Figure taken from the paper [3] of Flandoli, Issoglio and Russo with the authorization of Elena Issoglio. 17
I.2 The Haar wavelet h_{jm}^k for $j \in \mathbb{N}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$. 18

II.1 The set $\widetilde{K}(\beta, q)$. Modified figure from the paper [3] of Flandoli, Issoglio and Russo. 26

III.1 Translated sample path of a fBm with 64 points 30
III.2 Refined version of the previous fBm path with 128 points 32
III.3 The set $\widetilde{K}(\beta, q)$ with $\tilde{q} = 2$. Modified figure from the paper [3] of Flandoli, Issoglio and Russo. 35
III.4 Haar approximation of the derivative of a fBm path with $N = 2$. 36
III.5 Haar approximation of the derivative of a fBm path with $N = 8$. 36

IV.1 Estimation of the L^2 error of the Euler-Maruyama scheme with a Monte-Carlo method. 38
IV.2 Estimation of the weak error of the Euler-Maruyama scheme with a Monte-Carlo method. 39

Introduction

When one wants to study a phenomenon which could be a physical, biological or economical situation, it requires to produce a model. It is a simplification of reality which allows us to (at least partially) understand and represent the behavior of the system we look at. Most of these models rely on mathematical objects such as functions, whose derivatives verify differential equations. Formally, one can look for instance at equations like

$$\mathrm{d}X_t = f(t, X_t)\ \mathrm{d}t \tag{1}$$

in which X_t represents the dynamic of the system under consideration. Equation (1) expresses a link between time, the derivative of X_t and the values taken by X_t. If f is well-behaved, one can often show that it exists one and only one solution to this equation.

Unfortunately, in most models, f is not regular enough. Thus, one may only be able to show that some solutions exist but are not unique. This is a real problem because one would like to determine only one solution, it is to say the real description of the underlying phenomenon, directly from equation (1). For example, if you want to study the dynamic effects of gravity on a system, you wish to derive from your equations only one trajectory.

Nevertheless, when one adds white noise, a random phenomenon which can be understood formally as the derivative of Brownian motion, it may be possible to regain uniqueness in the equation[1], now understood as a Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t = f(t, X_t)\ \mathrm{d}t + \mathrm{d}W_t. \tag{2}$$

You can understand this random noise as a perturbation modifying equation (1). In this project we will focus on such equations (2) for irregular f.

The motivation is the following. It is hopeless in general to search for explicit expressions for the solution to such equations. In fact, even for regular f, it is impossible in most cases[2] to do this for equation (1), which is even easier to study than equation (2). But one can produce a numerical solution which is designed to approximate the real solution as accurately as needed.

Flandoli, Issoglio and Russo in [3] study and give a meaning to equations like (2) when f is a distribution[3] from a fractional Sobolev space. I used their results during my research project in the School of Maths of the University of Leeds to construct a multidimensional simulation method, giving a numerical approximation of the solution.

[1] This phenomenon is called regularization by noise.

[2] If f is linear in space and time-homogeneous, we know the exact solution as a matrix exponential.

[3] It's a very irregular object which generalizes the notion of function.

Numerical approximations of SDEs with irregular coefficients have been studied by a lot of authors. The case of discontinuous coefficients on a set of Lebesgue measure zero is handled by Yan in [12], and the one of discontinuous monotone drift coefficient is investigated by Halidias and Kloeden in [4]. Another work from Ngo and Taguchi focuses on a drift satisfying a one-sided Lipschitz condition in [9] and more recently, Leobacher and Szölgyenyi studied in [8] a piecewise Lipschitz drift with a degenerate diffusion coefficient. Another approach from Ankirchner, Kruse and Urusov in the paper [1] requires no regularity or growth conditions but still assumes that the drift is a function in order to approximate the law of the SDE solution. We can also cite the paper [7] of Kohatsu-Higa, Lejay and Yasuda which regularizes the drift before applying the Euler-Maruyama scheme, and [13] written by Étoré and Martinez which develops an exact simulation method when the drift is discontinuous at point zero.

Nevertheless, to the best of our knowledge, no result has been proven yet concerning the approximation of SDEs with distributional drift. Contrarily to all of the previous papers on this subject which study coefficients which are functions, here we provide the first approximation algorithm which applies to time-dependent drifts with Lipschitz regularity in time but which are distributions in space.

First we recall in Part I the notations and assumptions from [3]. To study the equation (2) with a distributional drift, one needs to use the notion of virtual solution introduced in [3], constructed from mild solutions to the associated Kolmogorov equation, and a transformed SDE. In fact, the original SDE doesn't have a classical solution and we must take into account the behavior of these PDE solutions to investigate the convergence of our method. We also give in Part I the definition of Haar wavelets. These functions constitute a basis in distributional fractional Sobolev spaces and are used to construct the multidimensional approximation of the drift we are looking for.

After defining an approximated SDE with the previous Haar approximation of the drift, we use in Part II the results of Leobacher and Szölgyenyi in [8] to prove the convergence of the Euler-Maruyama scheme for this approximated equation. After that step, a result from [3] shows the convergence in law of our approximated solution to the proper solution to (1). Our main result gives a convergence rate in dimension one, by making use of Gronwall's Lemma. It also requires to conduct a careful analysis of the local time of the difference between the solution to the transformed SDE and the solution to the transformed approximated SDE, via the use of the Meyer-Tanaka formula. This step relies on the regularity properties of the PDE solutions used to construct the virtual solution.

After these theoretical aspects, we will study in Part III and IV an application of our algorithm in the one-dimensional and time-homogeneous case. We will apply our method to the special case of f as the realization of a random field, the derivative of fractional Brownian motion. We will be able to compute the coefficients of this drift on the Haar basis, thanks to a Faber representation of the fractional Brownian motion path. Finally, we will look at the numerical convergence of our method in this case.

Part I

Context and algorithm presentation

In this part, we first define the spaces used in the report, and present the ideas beyond our algorithm. After that, we recall the notion of virtual solution introduced in [3] and explain how to use Haar wavelets to approximate a distributional drift.

I.1 Aims and notations

We would like to simulate numerically sample paths of the solution to the d-dimensional stochastic differential equation

$$\mathrm{d}X_t = b(t, X_t)\ \mathrm{d}t + \mathrm{d}W_t,\ X_0 = x_0 \tag{I.1}$$

where $b : [0,T] \mapsto \mathcal{S}'(\mathbb{R}^d)$ is distribution valued, $T > 0$, $x_0 \in \mathbb{R}^d$ and W is a standard d-dimensional Brownian motion. More precisely, we will use the following function spaces:

- the tempered distributions space $\mathcal{S}'(\mathbb{R}^d)$ defined as the continuous dual of the Schwartz space
$$\mathcal{S}(\mathbb{R}^d) = \{f \in \mathcal{C}^\infty(\mathbb{R}^d) \mid \forall i, j \in \mathbb{R}^d,\ \|f\|_{i,j} < \infty\} \tag{I.2}$$
with $\|f\|_{i,j} := \sup\limits_{x\in\mathbb{R}^d} |x^i \partial^j f(x)|$ for any multi-indices $i, j \in \mathbb{R}^d$;

- the fractional Sobolev spaces $H^s_{p,q}(\mathbb{R}^d) = H^s_p(\mathbb{R}^d) \cap H^s_q(\mathbb{R}^d)$ with $s \in \mathbb{R}$, $p, q > 0$ where $H^s_p(\mathbb{R}^d) = A^{-s/2}(L^p(\mathbb{R}^d))$ and $A := I - \frac{1}{2}\Delta$. They are Banach spaces endowed with the norm $\|u\|_{H^s_p(\mathbb{R}^d)} := \left\|A^{s/2}u\right\|_{L^p(\mathbb{R}^d)}$;

- the Banach spaces $\mathcal{C}\left([0,T], B\right)$ of B-valued continuous functions, endowed with the norm $\|f\|_{\infty,B} = \sup\limits_{0\le t\le T} \|f(t)\|_B$ where B is a Banach space;

- the Banach spaces $\mathcal{C}^{1,\alpha}(\mathbb{R}^d) = \{f \in \mathcal{C}^{1,0}(\mathbb{R}^d, \mathbb{R}^d) \mid \|f\|_{\mathcal{C}^{1,\alpha}} < \infty\}$ endowed with the norm
$$\|f\|_{\mathcal{C}^{1,\alpha}} := \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty} + \sup_{x\neq y\in\mathbb{R}^d} \frac{|\nabla f(x) - \nabla f(y)|}{|x-y|^\alpha}$$
with $\mathcal{C}^{1,0}(\mathbb{R}^d, \mathbb{R}^d)$ defined as the closure of $\mathcal{S}$ with respect to the norm $\|f\|_{\mathcal{C}^{1,0}} := \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty}$ and $\alpha > 0$;

- the spaces $L_t^q(L_x^p(\mathbb{R}^d))$ of mesurable functions $b : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ verifying:

$$\int_0^T \left(\int_{\mathbb{R}^d} |b(t,x)|^p \, \mathrm{d}x \right)^{q/p} \mathrm{d}t < \infty$$

  with $d/p + 2/q < 1$.

We will often write H_q^s instead of $H_q^s(\mathbb{R}^d)$ (the same remark holds for $L_t^q(L_x^p)$ and $\mathcal{C}^{1,\alpha}$).

Assumption 1. *We will always assume that* $b \in \mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$, *with* $\beta \in \left(0, \frac{1}{2}\right)$, $\tilde{q} = \frac{d}{1-\beta}$, *and* $q \in \left(\tilde{q}, \frac{d}{\beta}\right)$. *Moreover, we assume that* $t \mapsto b(t)$ *is Lipschitz continuous.*

Equation (I.1) is studied by Flandoli, Issoglio, and Russo in [3] in which they define a concept of virtual solution. The authors prove then existence and unicity in law of this solution. Moreover, they show that this virtual solution is in fact the limit in law of classical solutions. We will develop here a numerical approximation method for this equation.

Example 2. *Let's take a look at an example of such drift b in dimension one, in the time-homogeneous case. Fix a path of a fractional Brownian motion (fBm) B_x^H with Hurst index $1/2 < H < 1$ (for $H = 1/2$, we would retrieve the usual Brownian motion). We recall that these stochastic processes are centered Gaussian processes verifying*

$$\mathbb{E}\left[B_x^H B_y^H\right] = \frac{1}{2}\left(x^{2H} + y^{2H} + |x-y|^{2H}\right). \tag{I.3}$$

Then, you smoothly cut this path to zero outside a compact and take the derivative in the distributional sense. We note $-\beta = H - 1 - \varepsilon$ with $\varepsilon > 0$ small. Given $B_x^H(\omega) \in H_{\tilde{q},q}^{1-\beta}(\mathbb{R})$ for $q > 2$ and $\tilde{q} = 2$, by taking the derivative we obtain $b = \frac{\mathrm{d}}{\mathrm{d}x} B_x^H(\omega) \in H_{\tilde{q},q}^{-\beta}(\mathbb{R})$.

I.2 Principle of the approximation algorithm

As far as the drift b is not a function but a distribution, it must be approximated if we want to evaluate it pointwise. In order to do so, we will use a series representation of b and truncate it. That is why we will consider two steps in our algorithm:

- approximate b by a function b^N meant to converge to b in $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$ as $N \to \infty$. In practise we will choose $b^N \in \mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$ piecewise constant with Supp $b^N \subset [-N,N]^d$. b^N will be defined as a truncated Haar representation. Therefore, by compact support and piecewise constant properties of b^N, and its continuity in time, we will have $b^N \in L_t^r(L_x^g(\mathbb{R}^d)) \cap \mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$ for any (r,g) satisfying $d/g + 2/r < 1$;
- approximate the solution X_t^N to the approximated SDE

$$\mathrm{d}X_t^N = b^N\left(t, X_t^N\right)\,\mathrm{d}t + \mathrm{d}W_t,\ X_0^N = x_0 \tag{I.4}$$

  by $X_t^{N,n}$ defined with the Euler-Maruyama scheme by

$$X_t^{N,n} = x_0 + \int_0^t b^N\left(\eta_n(t), X_{\eta_n(t)}^{N,n}\right)\mathrm{d}t + W_{\eta_n(t)} \tag{I.5}$$

  where $\eta_n(t) = t_k$ if and only if $t \in [t_k, t_{k+1}[$, for $t_k = \frac{k}{n}$ with $k \in [\![0, \lceil nT \rceil]\!]$.

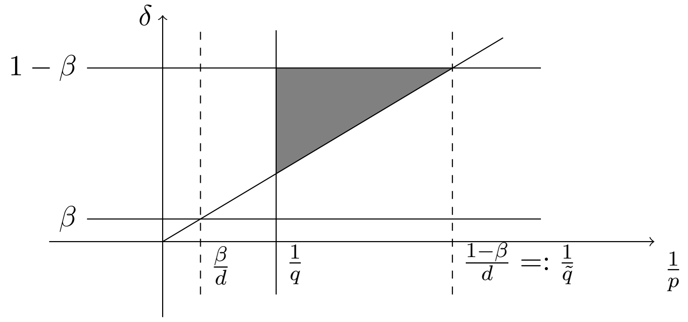

Figure I.1: The set $K(\beta, q)$. Figure taken from the paper [3] of Flandoli, Issoglio and Russo with the authorization of Elena Issoglio.

I.3 Virtual solutions of the original SDE and its approximation

In order to approximate the solution to the SDE (I.1), we must go back to the definition of its virtual solution given in [3]. Let $(\delta, p) \in K(\beta, q) := \{(\delta, p) \mid \beta < \delta < 1-\beta,\ \frac{d}{\delta} < p < q\}$ (see Figure I.1). Let's fix ρ, λ big enough to apply Theorem 14 and Lemma 22 in [3].

Definition 3 (Flandoli, Issoglio, Russo. See [3])**.** *A stochastic basis is defined as a pentuple $(\Omega, \mathcal{F}, \mathbb{F}, P, W)$ where $(\Omega, \mathcal{F}, P)$ is a complete probability space with a completed filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ and W is a d-dimensional $\mathbb{F}$-Brownian motion.*

Flandoli, Issoglio and Russo define the virtual solution to SDE (I.1) as a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P, W)$ and a continuous stochastic process $X := (X_t)_{t\in[0,T]}$ on it, $\mathbb{F}$-adapted, shortened as $(X, \mathbb{F})$ such that the integral equation

$$X_t = x_0 + u(0, x_0) - u(t, X_t) + (\lambda + 1)\int_0^t u(s, X_s)\ \mathrm{d}s + \int_0^t (\nabla u(s, X_s) + \mathrm{I}_d)\ \mathrm{d}W_s \tag{I.6}$$

holds for all $t \in [0, T]$, with probability one, where u is the mild solution in $\mathcal{C}([0, T], H_p^{1+\delta})$ of the following parabolic PDE:

$$\begin{cases} \partial_t u + \frac{1}{2}\Delta u + b\nabla u - (\lambda+1)u = -b & \text{on } [0,T]\times\mathbb{R}^d \\ u(T) = 0 & \text{on } \mathbb{R}^d \end{cases}. \tag{I.7}$$

The authors of [3] show then existence and uniqueness in law of a virtual solution to (I.1).

Moreover, $(X, \mathbb{F})$ is solution to (I.6) if and only if $(Y, \mathbb{F})$ is solution to

$$\begin{cases} Y_t = y_0 + (\lambda+1)\int_0^t u(s, \Psi(s, Y_s))\ \mathrm{d}s + \int_0^t (\nabla u(s, \Psi(s, Y_s)) + \mathrm{I}_d)\ \mathrm{d}W_s \\ X_t = \Psi(t, Y_t) \end{cases} \tag{I.8}$$

with $\varphi(t, x) = x + u(t, x)$, $y_0 = \varphi(0, x_0)$ and $\Psi(t, \cdot) = \varphi^{-1}(t, \cdot)$.

We also define another similar PDE by replacing b by $b^N \in L_t^r(L_x^g) \cap \mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}\right)$. We call u^N its unique mild solution in $\mathcal{C}([0,T], H_p^{1+\delta}(\mathbb{R}^d))$:

$$\begin{cases} \partial_t u^N + \frac{1}{2}\Delta u^N + b^N\nabla u^N - (\lambda+1)u^N = -b^N & \text{on } [0,T]\times\mathbb{R}^d \\ u^N(T) = 0 & \text{on } \mathbb{R}^d \end{cases}. \tag{I.9}$$

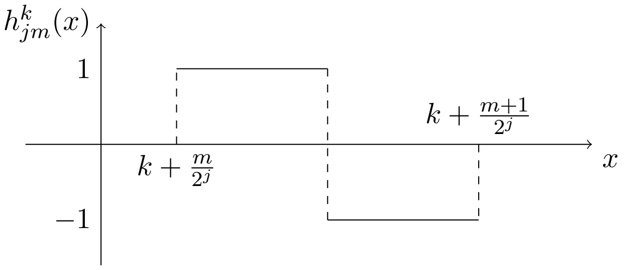

Figure I.2: The Haar wavelet h^k_{jm} for $j \in \mathbb{N}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$.

Then we consider an approximated version of (I.8):

$$\begin{cases} Y^N_t = y^N_0 + (\lambda+1)\int_0^t u^N\left(s, \Psi^N\left(s, Y^N_s\right)\right)\,\mathrm{d}s + \int_0^t \left(\nabla u^N\left(s, \Psi^N\left(s, Y^N_s\right)\right) + \mathrm{I}_d\right)\,\mathrm{d}W_s \\ X^N_t = \Psi^N(t, Y^N_t) \end{cases}. \tag{I.10}$$

with $\varphi^N(t,x) = x + u^N(t,x)$, $y^N_0 = \varphi^N(0,x_0)$ and $\Psi^N(t,\cdot) = \left(\varphi^N\right)^{-1}(t,\cdot)$.

The following result from Flandoli, Issoglio and Russo expresses the link between classical and virtual solutions and justifies our approximation algorithm.

Proposition 4 (Flandoli, Issoglio, Russo. See Proposition 26 in [3])**.** *If $b^N \in L^q_t(L^p_x(\mathbb{R}^d))$, then the classical solution $(X, \mathbb{F})$ to the SDE* (I.4) *is also a virtual solution.*

Remark 5. *Proposition 4 assures us that the virtual solution to* (I.10)*, $(X^N, \mathbb{F})$, defined above in* (I.10) *is in fact the classical solution to* (I.4)*, as far as $b^N \in L^r_t(L^g_x)$. That is why for each fixed N our Euler-Maruyamama approximation $X^{N,n}$ of the solution to* (I.4) *is meant to converge to the virtual solution X^N.*

Proposition 6. *Let u be the unique mild solution in $\mathcal{C}([0,T], H^{1+\delta}_p(\mathbb{R}^d))$ of equation $(I.7)$ for $(\delta, p) \in K(\beta, q)$. Then $u \in \mathcal{C}([0,T], \mathcal{C}^{1,\alpha}(\mathbb{R}^d))$ with $\alpha = \delta - d/p$.*

Proof. It's a direct consequence of the fractional Morrey inequality (See Theorem 16 in [3]). ☐

I.4 Approximation of the drift

We use Haar wavelets to give a series representation of b. By doing so, we will be able to approximate it numerically by truncating the series. We focus on the one dimensional case.

Definition 7 (Haar wavelets (See Figure I.2))**.** *We define the Haar wavelets $h^k_{j,m}$ on $\mathbb{R}$ with $j \in \mathbb{N}$, $m \in \mathbb{N}$ and $k \in \mathbb{Z}$ by:*

$$\begin{cases} h_M & : x \longmapsto \left(\mathbb{1}_{[0,\frac{1}{2})} - \mathbb{1}_{[\frac{1}{2},1)}\right)(x) \\ h^k_{-1,0} & : x \longmapsto \sqrt{2}\,|h_M(x-k)| \\ h_{j,m} & : x \longmapsto h_M(2^j x - m) \\ h^k_{j,m} & : x \longmapsto h_{j,m}(x-k) \end{cases}. \tag{I.11}$$

Theorem 8 (Triebel. See Theorem 2.9, Remark 2.12 in [10] and Remark 3.4 in [11]). *Let $f \in H_q^{-\beta}(\mathbb{R}) = F_{q2}^{-\beta}(\mathbb{R})$ for $2 \leq q < \infty$, and $-\beta \in \left(-\frac{1}{2}, \frac{1}{q}\right)$ (or $-\beta \in \left(\frac{1}{q} - 1, \frac{1}{2}\right)$ if $1 \leq q < 2$). Therefore, we have the unique representation*

$$f = \sum_{j=-1}^{+\infty} \sum_{k \in \mathbb{Z}} \sum_{m=0}^{2^j-1} \mu_{j,m}^k 2^{-j\left(-\beta - \frac{1}{q}\right)} h_{j,m}^k \tag{I.12}$$

with unconditional convergence in $S'(\mathbb{R})$, where $\mu_{j,m}^k = 2^{j\left(-\beta-\frac{1}{q}+1\right)} \int_{\mathbb{R}} f(x) h_{j,m}^k(x)\, \mathrm{d}x$ in the sense of dual pairing, and $\sum_{m=0}^{2^j-1}$ means $m = 0$ when $j = -1$.

Remark 9. *If $d = 1$, $\tilde{q} = \frac{1}{1-\beta} < 2$ and the Haar condition $-\beta \in \left(\frac{1}{\tilde{q}} - 1, \frac{1}{2}\right) = \left(-\beta, \frac{1}{2}\right)$ is never satisfied. Thus a Haar approximation of the drift cannot be constructed in $H_{\tilde{q}}^{-\beta}(\mathbb{R})$. However, we can adapt Assumption 1 by choosing $\tilde{q} > \frac{1}{1-\beta}$. It will only slightly modify the set of parameters $K(\beta, q)$ into $\tilde{K}(\beta, q) = \{(\delta, p) \mid \beta < \delta < 1 - \beta,\ \tilde{q} \vee \frac{d}{\delta} < p < q\}$. See figure II.1. Eventually, we solve the PDEs (I.7) and (I.9) in $H_p^{1+\delta}(\mathbb{R})$ for a restricted choice of parameters (δ, p) but the construction of the virtual solution to the SDE with distributional drift (I.1) still makes sense.*

Now we are able to define how we approximate a distribution expressed on a Haar base.

Definition 10. *With the same notation $\mu_{j,m}^k$, let $f \in H_q^{-\beta}(\mathbb{R})$ for $2 \leq q < \infty$, and $-\beta \in \left(-\frac{1}{2}, \frac{1}{q}\right)$ (or $-\beta \in \left(\frac{1}{q} - 1, \frac{1}{2}\right)$ if $1 \leq q < 2$). Given $N \in \mathbb{N}^*$ we define f^N in $H_q^{-\beta}(\mathbb{R})$ by:*

$$f^N = \sum_{j=0}^{N} \sum_{k=-N}^{N-1} \sum_{m=0}^{2^j-1} \mu_{j,m}^k 2^{-j\left(-\beta - \frac{1}{q}\right)} h_{j,m}^k. \tag{I.13}$$

Remark 11. *We can note that* Supp $f^N \subset [-N, N]$. *Moreover, we have:*

$$\left\|f - f^N\right\|_{H_q^{-\beta}(\mathbb{R})} \xrightarrow[N \to +\infty]{} 0. \tag{I.14}$$

Remark 12. *The Haar system, defined here on $\mathbb{R}$, can be extended to $\mathbb{R}^d$ and remains an unconditional basis of $H_q^{-\beta}(\mathbb{R}^d)$ for $q \geq 2$ and $-\beta \in \left(-\frac{1}{2}, \frac{1}{q}\right)$ (See Theorem 2.21 and Corollary 2.23 in [10]). Under Assumption 1 and with $d \geq 2$, it holds $q > \tilde{q} = \frac{d}{1-\beta} > 2$ and $-\beta \in \left(-\frac{1}{2}, \frac{1}{q}\right)$ so we can always construct an Haar representation in $H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)$. In this case, we can use the set $K(\beta, q)$ and fix $\tilde{q} = \frac{d}{1-\beta}$. Therefore our drift approximation with Haar wavelets is also relevant in $\mathbb{R}^d$. Nevertheless, in this case we don't know how to compute numerically the coefficients, even in the special case of the derivative of a fBm (or the derivative of any Hölder-continuous function actually). This is due to the lack of a proper Faber representation in $\mathbb{R}^d$. When $d = 1$, we are able to compute the coefficients for each given precision N. This will be done in Part III. In dimension one, Remark 9 explains why to use the modified set $\widetilde{K}(\beta, q)$.*

Remark 13. *In our case, $b \in \mathcal{C}\left([0, T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$. Thus, we have for instance in dimension 1, to simplify notations but without any loss of generality, a Haar representation:*

$$\forall t \in [0, T],\ b(t) = \sum_{j=-1}^{+\infty} \sum_{k \in \mathbb{Z}} \sum_{m=0}^{2^j-1} \mu_{j,m}^k(t) 2^{-j\left(-\beta - \frac{1}{q}\right)} h_{j,m}^k \tag{I.15}$$

where the coefficients $t \mapsto \mu_{j,m}^k(t) = 2^{j\left(-\beta-\frac{1}{q}+1\right)} \int_{\mathbb{R}} b(t,x) h_{j,m}^k(x)\,\mathrm{d}x$ *are Lipschitz continuous, because* $t \mapsto b(t)$ *is. In the proof of Lemma 5.1 in [6], Issoglio and Russo proved in this case that*

$$b^N(\cdot) = \sum_{j=-1}^{N} \sum_{k=-N}^{N-1} \sum_{m=0}^{2^j-1} \mu_{j,m}^k(\cdot) 2^{-j\left(-\beta-\frac{1}{q}\right)} h_{j,m}^k \tag{I.16}$$

belongs to $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R})\right)$ *and that* b^N *converges to* b *in* $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R})\right)$*. As they state, their argument extends to* $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$ *and the corresponding Haar basis by looking at each component of these functions in* $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R})\right)$*. Then we are able to construct with Haar wavelets a piecewise constant, with compact support and Lipschitz continuous in time bounded function* b^N *converging to* b *in* $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$*.*

Part II

Convergence of the algorithm

In this part, we present and prove convergence results for our approximation algorithm. We will show a convergence in law and give a rate in dimension one. We recall that X_t^N is defined as the solution to (I.4) (See Remark 5 and Proposition 4). Moreover, $X_t^{N,n}$ is defined as the Euler-Maruyama approximation (I.5).

II.1 Weak convergence rate of the Euler approximation

Recently, Leobacher and Szölgyenyi proved in [8] the convergence of the Euler-Maruyama scheme for multidimensional SDEs with discontinuous but piecewise Lipschitz time homogeneous drift and a degenerate diffusion coefficient. This framework applies to the scheme we use with piecewise constant in space and Lipschitz in time drift $b^N(t,x)$.

Let's define $\widetilde{X_t^N} := \begin{pmatrix} X_t^N \\ t \end{pmatrix} \in \mathbb{R}^{d+1}$, $\widetilde{b^N} : y \in \mathbb{R}^{d+1} \mapsto \begin{pmatrix} b^N(y_{d+1}, y_{|d}) \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$,

$$\widetilde{\sigma} : y \in \mathbb{R}^{d+1} \mapsto \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} \in \mathbb{R}^{(d+1)^2},$$

and $\widetilde{x_0} := \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \in \mathbb{R}^{d+1}$, with $\cdot_{|d} : y \in \mathbb{R}^{d+1} \mapsto y_{|d} = \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} \in \mathbb{R}^d$. Then $\widetilde{X_t^N}$ verifies

$$\mathrm{d}\widetilde{X_t^N} = \widetilde{b^N}\left(\widetilde{X_t^N}\right)\,\mathrm{d}t + \widetilde{\sigma}\left(\widetilde{X_t^N}\right)\,\mathrm{d}W_t,\ \widetilde{X_0^N} = \widetilde{x_0}. \tag{II.1}$$

Using notations from [8], the discontinuity set of the drift $\widetilde{b^N}$ is called Θ and $\forall \chi \in \Theta$, $n(\chi)$ is defined as the normalized normal vector to Θ at the point χ. Because [8] applies to degenerate diffusion coefficients, we are able to step up from time-homogeneous drifts to time-dependent drift. More precisely, the usual Zvonkin-Veretennikov uniform ellipticity condition on the diffusion coefficient σ is relaxed into the non-parallelity condition:

$$\forall \chi \in \Theta,\ \widetilde{\sigma}(\chi)^\top n(\chi) \neq 0.$$

As far as the discontinuity set Θ doesn't depend on time, we obtain $n(\chi) = \begin{pmatrix} n_1(\chi) \\ \vdots \\ n_d(\chi) \\ 0 \end{pmatrix}$.

Then $\forall \chi \in \Theta,\ \sigma(\chi)^\top n(\chi) = \begin{pmatrix} n_1(\chi) \\ \vdots \\ n_d(\chi) \\ 0 \end{pmatrix} = n(\chi) \neq 0$ because $n(\chi)$ is normalized. Thus the non-parallelity condition is satisfied. Moreover, Θ is the finite union of orientable compact C^3-manifolds (we recall that Supp b^N is compact, and that these manifolds are subsets of lines, by definition of Haar wavelets). $\widetilde{b^N}$ is bounded, piecewise Lipschitz continuous and $\widetilde{\sigma}$ is bounded and Lipschitz. We are therefore satisfying all the conditions to apply Example 2.6 from [8].

Theorem 14 (Leobacher, Szölgyenyi. Theorem 3.1 and Example 2.6 in [8])**.** *$\exists C_N > 0$ independent of n such that it holds $\forall \varepsilon > 0$ arbitrary small, $\exists n_0 \in \mathbb{N}$, $\forall n \geq n_0$:*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^{N,n} - X_t^N\right|^2\right]^{1/2} \leq \frac{C_N}{n^{1/4-\varepsilon}}. \tag{II.2}$$

Remark 15. *It is unclear how C_N depends on N but nevertheless, the convergence in law of X_t^N to X_t proved in Proposition 16 will allow us to choose N big enough for the weak error to be small, and then we will be able to choose n big enough to control the error between $X_t^{N,n}$ and X_t.*

II.2 Convergence in law of the approximated solution to the virtual solution

We also have a result showing that our approximated solution X^N converges to the virtual solution X.

Proposition 16 (Adapted from Flandoli, Issoglio, Russo. See Proposition 29 of [3].)**.** *Let $b^N \xrightarrow[N\to\infty]{} b$ in $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$ with $b^N \in L_t^r(L_x^g(\mathbb{R}^d)) \cap \mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R}^d)\right)$. Then the unique strong solution X^N to the associated equation* (I.4) *converges in law to the virtual solution $(X, \mathbb{F})$ of equation* (I.1)*.*

Proof. To prove the convergence in law of the solution X^N of (I.4) to the virtual solution X of (I.1), we use exactly the same arguments as in the Proposition 29 of [3] which uses the smoothness of the approximated drift only to apply Proposition 26 in [3], which extends to $L_t^q(L_x^p)$ drifts (See Proposition 4). □

Remark 17. *With Theorem 14 and Proposition 16, for each fixed $t \in [0,T]$, $X_t^{N,n} \xrightarrow[N\to\infty]{\mathcal{L}} X_t$. Therefore, $\left|\mathbb{E}\left[f(X_t) - f(X_t^{N,n})\right]\right|$, the weak error of approximation, converges to zero when f is a bounded continuous function and N, n go to $+\infty$. More precisely, if $t \in [0,T]$,*

$$\forall \varepsilon > 0,\ \exists N_0,\ \forall N \geq N_0,\ \exists n_0^N,\ \forall n \geq n_0^N,\ \left|\mathbb{E}\left[f(X_t) - f(X_t^{N,n})\right]\right| \leq \varepsilon.$$

As stated in Remark 15, we recall that n_0^N depends on N.

Eventually, we have proven that our algorithm provides a weak approximation of the virtual solution to SDE (I.1).

II.3 Weak convergence rate of the approximated solution in dimension one

The goal of this subsection, which is one of our contributions to this research project, is to give a convergence rate for the weak error in dimension one $\mathbb{E}\left[\left|X_t - X_t^{N,n}\right|\right]$. We are able to do so in the case where $\beta < 1/4$. Our main result relies on several lemmas. We will recall and use results from [3] where the stochastic differential equation (I.1) was first studied. Some of our technical lemmas proofs are developed in the appendix to simplify the reading of this report. Nevertheless, the proof of the main theorem is developped. We will assume that we have fixed a sequence b^N converging to b in $\mathcal{C}\left([0,T], H_{\tilde{q},q}^{-\beta}(\mathbb{R})\right)$. We recall that the set $\tilde{K}(\beta,q)$ (represented in figure II.1) is defined in Remark 9 and is useful only in dimension one (see Remark 12).

II.3.1 Useful lemmas

Most of the following results allow us to study the behaviour of the solutions of PDEs (I.7) and (I.9) and then of the transformations φ and Ψ. We will use these lemmas to prove our main result, Theorem 28.

We first recall a useful lemma concerning the solutions of (I.7) and (I.9). It will allow us to choose λ big enough such that u and u^N are Lipschitz.

Lemma 18 (Flandoli, Issoglio, Russo. See Lemma 20 in [3])**.** *Let $(\delta,p) \in \tilde{K}(\beta,q)$ and let u, u^N be the mild solutions to (I.7), (I.9) in $\mathcal{C}([0,T], H_p^{1+\delta}(\mathbb{R}^d))$. Fix ρ big enough such that Theorem 14 in [3] applies and let $\lambda > \lambda^*$. Then $\forall \varepsilon > 0$, $\exists \lambda_0 > 0$ such that $\forall \lambda \geq \lambda_0$*

$$\begin{cases} \sup\limits_{(t,x)\in[0,T]\times\mathbb{R}} |\nabla u(t,x)| & \leq \varepsilon \\ \sup\limits_{(t,x)\in[0,T]\times\mathbb{R}} |\nabla u^N(t,x)| & \leq \varepsilon \end{cases} \tag{II.3}$$

where the choice of λ^*, λ_0 depends only on $\delta, \beta, \|b\|_{\infty,H_{p,q}^{-\beta}}$, and $\left\|b^N\right\|_{\infty,H_{p,q}^{-\beta}}$.

Remark 19. $\left\|b^N\right\|_{\infty,H_{p,q}^{-\beta}} \in \left(1/2\,\|b\|_{\infty,H_{p,q}^{-\beta}}, 2\,\|b\|_{\infty,H_{p,q}^{-\beta}}\right)$ *for N big enough so λ^*, λ_0 can be chosen independently of N for N big enough.*

The next lemma uses the embedding between $H_p^{1+\delta}$ and $\mathcal{C}^{1,\alpha}$ to control the error of approximation of u by u^N.

Lemma 20. *Let $(\delta,p) \in \tilde{K}(\beta,q)$ and let u, u^N be the mild solutions to (I.7),(I.9) in $\mathcal{C}([0,T], H_p^{1+\delta}(\mathbb{R}^d))$, $\alpha = \delta - 1/p$. Exists $c, K > 0$ such that for both $N \in \mathbb{N}$ and ρ big enough (independently of N), $\forall t \in [0,T]$,*

$$\begin{cases} \left\|u^N(t) - u(t)\right\|_{L^\infty} & \leq \kappa \left\|b - b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} \\ \left\|\nabla u^N(t) - \nabla u(t)\right\|_{L^\infty} & \leq \kappa \left\|b - b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} \end{cases}. \tag{II.4}$$

with $\kappa = cKe^{\rho T}$.

Proof. Applying fractional Morrey inequality, $\exists c>0,\ \forall t\in[0,T]$:

$$\begin{cases} \left\|u^N(t)-u(t)\right\|_{L^\infty} & \leq \left\|u^N(t)-u(t)\right\|_{\mathcal{C}^{1,\alpha}} \leq c\left\|u^N(t)-u(t)\right\|_{H_p^{1+\delta}} \\ \left\|\nabla u^N(t)-\nabla u(t)\right\|_{L^\infty} & \leq \left\|u^N(t)-u(t)\right\|_{\mathcal{C}^{1,\alpha}} \leq c\left\|u^N(t)-u(t)\right\|_{H_p^{1+\delta}}. \end{cases}$$

Now, we can conclude with

$$\left\|u^N-u\right\|_{\infty,H_p^{1+\delta}} \leq e^{\rho T}\left\|u^N-u\right\|_{\infty,H_p^{1+\delta}}^{(\rho)} \leq Ke^{\rho T}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}$$

from Lemma 23 in [3], for both $N\in\mathbb{N}$ and ρ big enough, and where

$$\|f(t)\|_{\infty,B}^{(\rho)} := \sup_{0\leq t\leq T} e^{-\rho t}\|f(t)\|_B .$$

□

Lemma 21. *For λ big enough, independently of N,*

$$\left|\Psi^N\left(t,Y_t^N\right)-\Psi\left(t,Y_t^N\right)\right| \leq 2\kappa\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} \tag{II.5}$$

with $\kappa = cKe^{\rho T}$.

Proof. For λ big enough, by Lemma 18, $\forall t\in[0,T],\ \sup_{x\in\mathbb{R}}|\nabla u(t,x)|\leq 1/2$, so we obtain with $\varphi(t,x)=x+u(t,x)$:

$$\begin{aligned} \left|\varphi\left(t,\Psi^N\left(t,Y_t^N\right)\right)-\varphi\left(t,\Psi\left(t,Y_t^N\right)\right)\right| &\geq \inf_{x\in\mathbb{R}}|\nabla\varphi(t,x)|\left|\Psi^N\left(t,Y_t^N\right)-\Psi\left(t,Y_t^N\right)\right| \\ &\geq \frac{1}{2}\left|\Psi^N\left(t,Y_t^N\right)-\Psi\left(t,Y_t^N\right)\right| \end{aligned}$$

then

$$\begin{aligned} \left|\Psi^N\left(t,Y_t^N\right)-\Psi\left(t,Y_t^N\right)\right| &\leq 2\ \left|\varphi\left(t,\Psi^N\left(t,Y_t^N\right)\right)-\varphi\left(t,\Psi\left(t,Y_t^N\right)\right)\right| \\ &= 2\ \left|\varphi\left(t,\Psi^N\left(t,Y_t^N\right)\right)-\varphi^N\left(t,\Psi^N\left(t,Y_t^N\right)\right)\right| \\ &\leq 2\ \left\|u(t)-u^N(t)\right\|_{L^\infty} \end{aligned}$$

where we used the fact that

$$\varphi^N\left(t,\Psi^N\left(t,Y_t^N\right)\right)=\varphi\left(t,\Psi\left(t,Y_t^N\right)\right)=Y_t^N.$$

With Lemma 20 we obtain

$$\left|\Psi^N\left(t,Y_t^N\right)-\Psi\left(t,Y_t^N\right)\right| \leq 2\kappa\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}.$$

□

We will need an adapted version of a local time inequality from Yan in [12]. The method he introduced will allow us to estimate the error $\mathbb{E}[|Y-Y^N|]$ and therefore calculate the weak error of approximation our algorithm commit. In his paper he studies the case where $X_0 = 0$ whereas in our case we don't assume anything about the initial value of the semimartingale X_t. That's why the proof is mainly the same. The difference is that υ_1 is now a stopping time when its value was 0 in the proof of Yan. We also note $\tau_0 = 0$.

Lemma 22 (Adapted from Yan. See Lemma 4.2 in [12]). *Let X_t be a continuous semimartingale. For $\varepsilon > 0$ we define a double sequence of stopping times by $\tau_0 = 0$, $\upsilon_1 = \inf\{t \geq 0 | X_t = 0\}$, $\tau_1 = \inf\{t > \upsilon_1 | X_t = \varepsilon\}$, $\upsilon_n = \inf\{t > \tau_{n-1} | X_t = 0\}$, $\tau_n = \inf\{t > \upsilon_n | X_t = \varepsilon\}$. For any real function $F(\cdot) \in \mathcal{C}^2(\mathbb{R})$ with $F(0) = 0$, $F'(0) = 0$, $F(\cdot) > 0$ on $(0, \varepsilon_0)$ with some $\varepsilon_0 > 0$, for any $0 < \varepsilon < \varepsilon_0$ and for any $t \in [0,T]$we have*

$$L_t^0(X) \leq 2\varepsilon - \frac{2}{F(\varepsilon)} \int_{\upsilon_1 \wedge t}^t \theta_s(X) \left(F(\varepsilon) - \varepsilon F'\left(X_s^+\right)\right) \, \mathrm{d}X_s + \frac{\varepsilon}{F(\varepsilon)} \int_{\upsilon_1 \wedge t}^t \theta_s(X) F''(X_s^+) \, \mathrm{d}[X]_s \tag{II.6}$$

with $\theta_s(X) = \sum_{n=1}^{\infty} \mathbb{1}_{\{\upsilon_n < s \leq \tau_n, \ 0 < X_s \leq \varepsilon\}}(X)$ and $t \in [0,T]$.

Applying lemma 22 with $F : x \in \mathbb{R} \mapsto x^2$, it follows:

Corollary 23. *Let X be a continuous semimartingale. With the same notations as in Lemma 22, for any $\varepsilon > 0$ and for any $t \in [0,T]$ we have*

$$0 \leq L_t^0(X) \leq 2\varepsilon - \frac{2}{\varepsilon} \int_{\upsilon_1 \wedge t}^t \theta_s(X) \left(\varepsilon - 2X_s^+\right) \, \mathrm{d}X_s + \frac{2}{\varepsilon} \int_{\upsilon_1 \wedge t}^t \theta_s(X) \, \mathrm{d}[X]_s \tag{II.7}$$

Remark 24. *For clarity and concision purposes, we will note $\tilde{u}(s,x) = u(s, \Psi(s,x))$, $\tilde{u}^N(s,x) = u^N(s, \Psi(s,x))$, and $\overline{u}^N(s,x) = u^N\left(s, \Psi^N(s,x)\right)$. The same notations will be used for the gradient.*

Lemma 25. *Let $(\delta, p) \in \tilde{K}(\beta, q)$, $\alpha = \delta - 1/p < 1$, u, u^N be the mild solutions to (I.7), (I.9) in $H_p^{1+\delta}$, and Y, Y^N solutions of the SDEs (I.6), (I.10). Then, if $1/2 < \alpha < 1$, for λ, N big enough we have $\forall \varepsilon \in (0,1]$,*

$$0 \leq \mathbb{E}\left[L_T^0(Y - Y^N)\right] \leq g(\varepsilon). \tag{II.8}$$

where

$$\begin{aligned} g(\varepsilon) &= 4(\lambda+1)T\kappa \left\|b - b^N\right\|_{\infty, H_{\tilde{q},q}^{-\beta}} + \left(2 + 2(\lambda+1)\, T + 6\left\|u\right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T\right) \varepsilon^{2\alpha-1} \\ &\quad + 6T\left(\Lambda^2 4^\alpha \kappa^{2\alpha} \left\|b - b^N\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2 \left\|b - b^N\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^2\right) \varepsilon^{-1}. \end{aligned}$$

We have been able with Lemma 25 to bound the expectation of the local time of $Y - Y^N$ by a function of $\varepsilon > 0$. Here we choose an optimal ε such that the bound is minimal.

Proposition 26. *With assumptions and notations of Lemma 25, and $1/2 < \alpha < 1$ we have*

$$\mathbb{E}\left[L_T^0(Y - Y^N)\right] \leq \Gamma \left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1} \tag{II.9}$$

for N big enough where

$$\Gamma = 4(\lambda+1)T\kappa + 6T\kappa^{2\alpha}\left(\frac{2\alpha+1}{2\alpha-1}\Lambda^2 4^\alpha + \kappa^{2(1-\alpha)}\right)\nu_\infty^{-1}$$

and

$$\nu_\infty = \left(\frac{6T\Lambda^2 4^\alpha \kappa^{2\alpha}}{(2\alpha-1)\left(2 + 2(\lambda+1)\, T + 6\left\|u\right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T\right)}\right)^{\frac{1}{2\alpha}}.$$

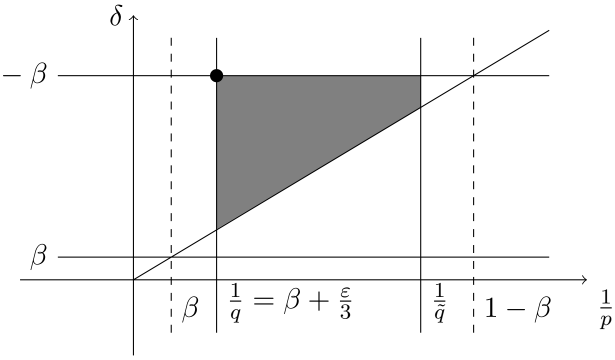

Figure II.1: The set $\widetilde{K}(\beta,q)$. Modified figure from the paper [3] of Flandoli, Issoglio and Russo.

II.3.2 Main result

Remark 27. *For the proof of our theorem to hold, we will require $\alpha=\delta-\frac{1}{p}$ to be greater than $\frac{1}{2}$. Therefore, given the constraints on the choice of $(\delta,p)\in K(\beta,q)=\{(\delta,p)\mid \beta<\delta<1-\beta,\ \frac{d}{\delta}<p<q\}$, we would like to know the greatest value α can take for a given β. First of all, $\frac{1}{p}>\frac{1}{q}>\beta$. We fix $\frac{1}{q}=\beta+\frac{\varepsilon}{3}$ with $\varepsilon>0$ arbitrarily small. Then $-\frac{1}{p}<-\beta-\frac{\varepsilon}{3}$. Moreover, $\delta<1-\beta$. Thus, $\alpha<1-2\beta-\frac{\varepsilon}{3}$, and one can take $\alpha=1-2\beta-\frac{\varepsilon}{2}$. In figure I.1 page 17, the pair (δ,p) guaranteeing the maximal value for α corresponds to the top left corner of the grey triangle, indicated by a dot.*

Theorem 28. *Let $0<\beta<1/4$. Then $\forall\varepsilon>0$ arbitrarily small, with $\frac{1}{q}=\beta+\frac{\varepsilon}{3}$, and $(\delta,p)\in K(\beta,q)$ such that $\delta-1/p=1-2\beta-\varepsilon/2$, exists ξ such that for $N\in\mathbb{N}$, ρ, λ big enough it holds:*

$$\sup_{0\le t\le T}\mathbb{E}\left[\left|X_t-X_t^N\right|\right]\le\xi\left\|b^N-b\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{1-4\beta-\varepsilon}\tag{II.10}$$

with $\xi=2\left(\zeta e^{(\lambda+1)T}+\kappa\right)$ and $\zeta=(2(\lambda+1)T+1)\kappa+\Gamma$.

Proof of Theorem 28. We note as usual $\alpha=\delta-1/p$. By Lemma 18, we choose λ big enough for ∇u and ∇u^N to be bounded by $\frac{1}{2}$. λ can be chosen independently of N as far as $\left\|b-b^N\right\|_{H_{\tilde{q}}^{s}(\mathbb{R})}\underset{N\to\infty}{\longrightarrow}0$. Therefore $u^N(t,\cdot)$ and $u(t,\cdot)$ are $\frac{1}{2}$-Lipschitz. We recall that in this case, by Lemma 22 in [3], $\Psi(t,\cdot)$ and $\Psi^N(t,\cdot)$ are 2-Lipschitz. Therefore $\widetilde{u}^N(t,\cdot)$ and $\widetilde{u}(t,\cdot)$ are 1-Lipschitz. Let $t\in[0,T]$.

$$\begin{aligned}\mathbb{E}\left[\left|X_t-X_t^N\right|\right]&=\mathbb{E}\left[\left|\Psi\left(t,Y_t\right)-\Psi^N\left(t,Y_t^N\right)\right|\right]\\&\le\mathbb{E}\left[\left|\Psi\left(t,Y_t\right)-\Psi\left(t,Y_t^N\right)\right|\right]+\mathbb{E}\left[\left|\Psi\left(t,Y_t^N\right)-\Psi^N\left(t,Y_t^N\right)\right|\right]\end{aligned}$$

And we obtain

$$\mathbb{E}\left[\left|X_t-X_t^N\right|\right]\le 2\,\mathbb{E}\left[\left|Y_t-Y_t^N\right|\right]+\mathbb{E}\left[\left|\Psi\left(t,Y_t^N\right)-\Psi^N\left(t,Y_t^N\right)\right|\right].\tag{II.11}$$

Then we study the difference between Y_t^N and Y_t:

$$\begin{aligned}Y_t-Y_t^N=y_0-y_0^N&+(\lambda+1)\int_0^t\left\{u\left(s,\Psi\left(s,Y_s\right)\right)-u^N\left(s,\Psi^N\left(s,Y_s^N\right)\right)\right\}\,\mathrm{d}s\\&+\int_0^t\left\{\nabla u\left(s,\Psi\left(s,Y_s\right)\right)-\nabla u^N\left(s,\Psi^N\left(s,Y_s^N\right)\right)\right\}\,\mathrm{d}W_s.\end{aligned}$$

We recall the notations $\tilde{u}(s,x) = u(s,\Psi(s,x))$, $\tilde{u}^N(s,x) = u^N(s,\Psi(s,x))$, and $\overline{u}^N(s,x) = u^N\left(s,\Psi^N(s,x)\right)$. The same notations are used for the gradient and the approximated mild solution. We apply Meyer-Tanaka's formula to obtain:

$$\begin{aligned}\left|Y_t - Y_t^N\right| = \left|y_0 - y_0^N\right| + (\lambda+1)\int_0^t \operatorname{sign}(Y_s - Y_s^N)\left\{\widetilde{u}(s,Y_s) - \overline{u}^N\left(s,Y_s^N\right)\right\}\,\mathrm{d}s \\ + \int_0^t \operatorname{sign}(Y_s - Y_s^N)\left\{\widetilde{\nabla u}(s,Y_s) - \overline{\nabla u}^N\left(s,Y_s^N\right)\right\}\,\mathrm{d}W_s + L_t^0(Y - Y^N).\end{aligned}$$

Taking the expectation leads to:

$$\begin{aligned}\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] &= \left|u(0,x_0) - u^N(0,x_0)\right| + \mathbb{E}\left[L_t^0(Y - Y^N)\right] \\ &\quad + (\lambda+1)\,\mathbb{E}\left[\int_0^t \operatorname{sign}(Y_s - Y_s^N)\left\{\widetilde{u}(s,Y_s) - \overline{u}^N\left(s,Y_s^N\right)\right\}\,\mathrm{d}s\right]\end{aligned}$$

because ∇u and ∇u^N are bounded so the Itô integral is a true martingale.

$$\begin{aligned}\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] &\le \left|u(0,x_0) - u^N(0,x_0)\right| + \mathbb{E}\left[L_t^0(Y - Y^N)\right] \\ &\quad + (\lambda+1)\,\mathbb{E}\left[\int_0^t \left|\widetilde{u}(s,Y_s) - \widetilde{u}\left(s,Y_s^N\right)\right|\mathrm{d}s\right] \\ &\quad + (\lambda+1)\,\mathbb{E}\left[\int_0^t \left|\widetilde{u}\left(s,Y_s^N\right) - \widetilde{u}^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right] \\ &\quad + (\lambda+1)\,\mathbb{E}\left[\int_0^t \left|\widetilde{u}^N\left(s,Y_s^N\right) - \overline{u}^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right].\end{aligned}$$

We use Lemma 20, the 1-Lipschitz property of $\widetilde{u}$, and the 1/2-Lipschitz property of u:

$$\begin{aligned}\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] &\le \kappa\left\|b^N - b\right\|_{\infty,H^{-\beta}_{\tilde{q},q}} + (\lambda+1)\,\mathbb{E}\left[\int_0^t \left|Y_s - Y_s^N\right|\mathrm{d}s\right] \\ &\quad + \frac{\lambda+1}{2}\,\mathbb{E}\left[\int_0^t \left|\Psi\left(s,Y_s^N\right) - \Psi^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right] + \mathbb{E}\left[L_t^0(Y - Y^N)\right] \\ &\le (\lambda+1)\int_0^t \mathbb{E}\left[\left|Y_s - Y_s^N\right|\right]\mathrm{d}s + (2(\lambda+1)T+1)\kappa\left\|b^N - b\right\|_{\infty,H^{-\beta}_{\tilde{q},q}} \\ &\quad + \mathbb{E}\left[L_T^0(Y - Y^N)\right] + (\lambda+1)t\kappa\left\|b^N - b\right\|_{\infty,H^{-\beta}_{\tilde{q},q}}.\end{aligned}$$

where we used Lemma 21 and the fact that $L_t^0(Y - Y^N)$ is an increasing process.

By Gronwall's Lemma, it follows:

$$\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] \le C(N)\,e^{(\lambda+1)t} \le C(N)\,e^{(\lambda+1)T} \tag{II.12}$$

with $C(N) = (2(\lambda+1)T+1)\kappa\left\|b^N - b\right\|_{\infty,H^{-\beta}_{\tilde{q},q}} + \mathbb{E}\left[L_T^0(Y - Y^N)\right]$.

With Lemma 25 and Proposition 26 we obtain

$$C(N) \le (2(\lambda+1)T+1)\kappa\left\|b^N - b\right\|_{\infty,H^{-\beta}_{\tilde{q},q}} + \Gamma\left\|b^N - b\right\|^{2\alpha-1}_{\infty,H^{-\beta}_{\tilde{q},q}} \le \zeta\left\|b^N - b\right\|^{2\alpha-1}_{\infty,H^{-\beta}_{\tilde{q},q}}.$$

for N big enough where $\zeta = (2(\lambda+1)T+1)\kappa + \Gamma$. It follows:

$$\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] \le \zeta e^{(\lambda+1)T}\left\|b^N - b\right\|^{2\alpha-1}_{\infty,H^{-\beta}_{\tilde{q},q}}. \tag{II.13}$$

Finally, combining (II.11), (II.13) and Lemma 21 we obtain:

$$\begin{aligned}\mathbb{E}\left[\left|X_t - X_t^N\right|\right] &\leq \left(2\,\mathbb{E}\left[\left|Y_t - Y_t^N\right|\right] + \mathbb{E}\left[\left|\Psi\left(t, Y_t^N\right) - \Psi^N\left(t, Y_t^N\right)\right|\right]\right) \\ &\leq \left(2\zeta e^{(\lambda+1)T}\left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1} + 2\kappa\left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}\right) \\ &\leq 2\left(\zeta e^{(\lambda+1)T} + \kappa\right)\left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}\end{aligned}$$

for N big enough, which is the expected result because every variable in the right term is independent of t. □

Combining Theorem 14 and Theorem 28, we eventually obtain a rate in dimension one for our approximation algorithm.

Corollary 29. *With the assumptions and notations of Theorem 28, for any $\varepsilon, \varepsilon' > 0$ arbitrarily small, exists ξ, C_N independent of f such that for $N \in \mathbb{N}$, ρ, λ big enough*

$$\sup_{0\leq t\leq T} \mathbb{E}\left[\left|X_t^{N,n} - X_t\right|\right] \leq \frac{C_N}{n^{1/4-\varepsilon'}} + \xi\left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{1-4\beta-\varepsilon} \tag{II.14}$$

Proof. By Jensen's inequality, we obtain, with $s \in [0, T]$

$$\begin{aligned}\mathbb{E}\left[\left|X_s^{N,n} - X_s\right|\right] &\leq \mathbb{E}\left[\left|X_s^{N,n} - X_s^N\right|\right] + \mathbb{E}\left[\left|X_s^N - X_s\right|\right] \\ &\leq \mathbb{E}\left[\left|X_s^{N,n} - X_s^N\right|^2\right]^{1/2} + \mathbb{E}\left[\left|X_s^N - X_s\right|\right] \\ &\leq \mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^{N,n} - X_t^N\right|^2\right]^{1/2} + \mathbb{E}\left[\left|X_s^N - X_s\right|\right].\end{aligned}$$

Then you take the supremum over s and apply Theorem 14 and Theorem 28. □

Corollary 30. *Let f be μ-Hölder with constant $C_f > 0$ and $\mu \in (0, 1]$. With the assumptions and notations of Theorem 28, for any $\varepsilon, \varepsilon' > 0$ arbitrarily small, exists ξ, C_N independent of f such that for $N \in \mathbb{N}$, ρ, λ big enough*

$$\sup_{0\leq t\leq T}\left|\mathbb{E}\left[f(X_t^{N,n}) - f\left(X_t\right)\right]\right| \leq C_f\left(\frac{C_N^\mu}{n^{\mu/4-\varepsilon'}} + \xi^\mu\left\|b^N - b\right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{\mu(1-4\beta-\varepsilon)}\right) \tag{II.15}$$

Part III

Numerical aspects

Our project concerns a large class of very irregular drifts. But in order to do numerical studies, we must select examples of such drifts. If we want to use a drift selected as a realization of a random process, we can think of the derivative of a sample path of a fractional Brownian motion. That's why we will explain here how we can apply our approximation algorithm in this special case.

III.1 Numerical simulation of fractional Brownian motion

To simulate a sample path of a fBm B_x^H on a finite grid $(x_k)_{k\in[\![1,N]\!]}$, we simulate n independent standard gaussian random variables $(X_k)_{k\in[\![1,N]\!]}$ and then correlate them with the definite positive correlation matrix

$$C_{k,s} = \mathbb{E}\left[B_{x_k}^H B_{x_s}^H\right] = \frac{1}{2}\left(x_k^{2H} + x_s^{2H} + |x_k - x_s|^{2H}\right).$$

To do so, we use the Cholesky decomposition method and calculate the lower triangular matrix M such that $C = MM^\top$. Therefore, defining the multidimensional random values

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \text{ and } B^H = MX,$$

B^H is a fractional Brownian motion evaluated on the grid $(x_k)_{k\in[\![1,n]\!]}$. Indeed, numerically we will only consider some realizations of these random values. Figure III.1 shows the result of such a simulation. We will always translate the path in order to have a symmetric domain. In fact, what does matter in our case is the regularity of a fixed path and not the domain of the underlying process.

III.1.1 Refining a sample path of a fractional Brownian motion

A first numerical concern was about the way to refine a path of fBm, it is to say to add new points in between the simulated path. For instance, if you have simulated a fBm on a regular grid $x_k = \frac{k}{N}$, you may wish to add new points to work on the grid 2 times more precise $\widetilde{x_k} = \frac{k}{2N}$ without changing the points already defined. Let X, Y be two vectors of N independent Gaussian random variables and C the correlation matrix associated with the grid $(x_k)_{k\in[\![1,n]\!]}$. We note $\widetilde{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$.

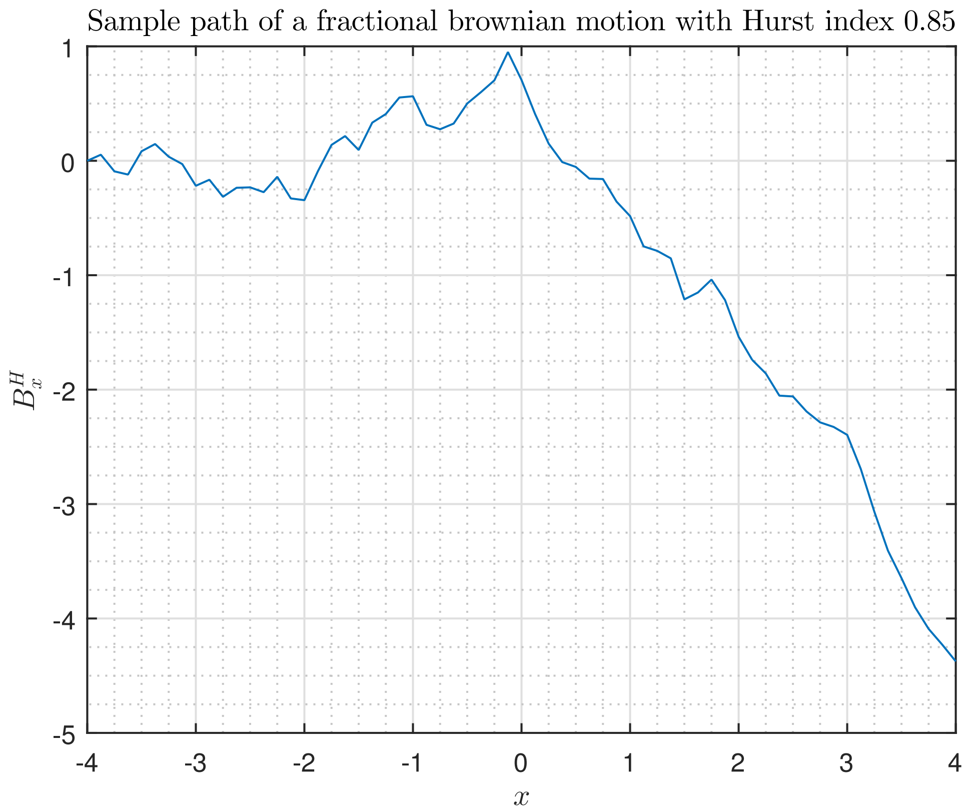

Figure III.1: Translated sample path of a fBm with 64 points

If you have already simulated B^H on the grid $(x_k)_{k\in[\![1,n]\!]}$ with the random variable X, it verifies $B^H = MX$ with $C = MM^\top$. Then, let's consider the new correlation matrix $\widetilde{C} = \begin{pmatrix} C & A^\top \\ A & B \end{pmatrix}$ with

$$A_{k,s} = \mathbb{E}\left[B^H_{\frac{x_k+x_{k-1}}{2}} B^H_{x_s}\right] = \frac{1}{2}\left(\left(\frac{x_k+x_{k-1}}{2}\right)^{2H} + x_s^{2H} + \left|\frac{x_k+x_{k-1}}{2} - x_s\right|^{2H}\right)$$

$$\begin{aligned} B_{k,s} &= \mathbb{E}\left[B^H_{\frac{x_k+x_{k-1}}{2}} B^H_{\frac{x_s+x_{s-1}}{2}}\right] \\ &= \frac{1}{2}\left(\left(\frac{x_k+x_{k-1}}{2}\right)^{2H} + \left(\frac{x_s+x_{s-1}}{2}\right)^{2H} + \left|\frac{x_k+x_{k-1}}{2} - \frac{x_s+x_{s-1}}{2}\right|^{2H}\right) \end{aligned}$$

where we chose by convention $x_0 = 0$. Looking at $\widetilde{M}$, the Cholesky root of $\widetilde{C}$, we obtain:

$$\widetilde{M} = \begin{pmatrix} D & 0 \\ E & F \end{pmatrix} \tag{III.1}$$

with D lower triangular matrix verifying $DD^\top = C$ so $D = M$. Therefore, we will have

$$\widetilde{B}^H = \widetilde{M}\widetilde{X} = \begin{pmatrix} M & 0 \\ E & F \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} MX \\ EX+FY \end{pmatrix} = \begin{pmatrix} B^H \\ EX+FY \end{pmatrix} \tag{III.2}$$

and we obtain $\widetilde{B}^H$ on the grid $(x_k, \frac{x_k+x_{k-1}}{2})$ without changing the previous values. Therefore, one just have to reorder $\widetilde{B}^H$ to obtain a realization of a fBm on the grid

$$\widetilde{x_k} = \left(\frac{x_1}{2}, x_1, \frac{x_1+x_2}{2}, \cdots, x_{k-1}, \frac{x_k+x_{k-1}}{2}, x_k, \cdots, x_N\right).$$

Numerical results are shown in figures III.1 and III.2.

III.2 Special case of the derivative of a fractional Brownian motion

This study of a special case for the approximation algorithm we developed is our other contribution to this research project. For our numerical studies, we will now fix the realization of a fractional Brownian motion B^H_x and cut its path to zero continuously. In order to apply Theorem 28, we suppose $H > \frac{3}{4}$. In this section, the drift will be time-homogeneous.

More precisely, let $K > 0$, $\gamma \in \mathcal{C}^\infty(\mathbb{R})$ such that $\forall x \in [-K, K]$, $\gamma(x) = 1$ and Supp $\gamma \subset D \subset [-K-1, K+1]$. Then for almost every $\omega \in \Omega$, $x \mapsto \gamma(x)B^H_x(\omega)$ is α-Hölder continuous for all $\alpha \in (0, H)$. We fix such an $\omega_0 \in \Omega$. To simplify notations, we still note B^H_x the cut to zero path $\gamma(x)B^H_x(\omega_0)$.

Remark 31. *It requires to define explicitly the smoothing function γ in order to simulate its values numerically. But one can also choose to restrict the domain of the drift to $[-K, K]$ and stop the process when it exits this compact set. Therefore the values of γ don't matter.*

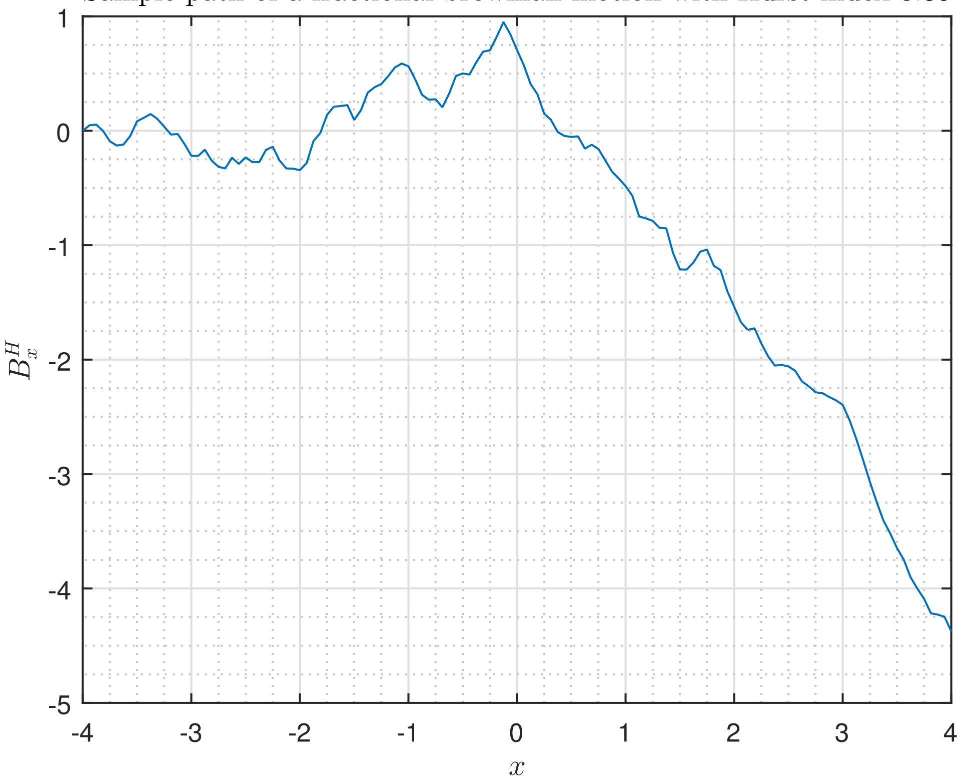

Figure III.2: Refined version of the previous fBm path with 128 points

We also recall the link between fractional Sobolev spaces and Triebel-Lizorkin spaces $F^s_{pq}(\mathbb{R})$.

Proposition 32 (Issoglio. See 4.1 in [5])**.** $\forall s \in \left(\frac{1}{2}, H\right)$, $\forall q \in [2, \infty)$

$$B^H_x \in F^s_{q2}(\mathbb{R}) = H^s_q(\mathbb{R}). \tag{III.3}$$

By Proposition 32, we fix $s = H - \varepsilon > \frac{3}{4}$ for $\varepsilon > 0$ small enough such that $B^H \in H^s_q(\mathbb{R})$ with $q \in \left(2, \frac{1}{1-s}\right)$, where we fixed $\tilde{q} = 2$.

Then, $b = \frac{\mathrm{d}}{\mathrm{d}x} B^H_x$ with support in D belongs to $H^{s-1}_q(\mathbb{R})$ so we have by Theorem 8 its Haar decomposition:

$$b = \sum_{j=-1}^{+\infty} \sum_{k \in \mathbb{Z}} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} h^k_{j,m}. \tag{III.4}$$

with $\mu^k_{j,m} = 2^{j\left(s-1-\frac{1}{q}\right)} \int_{\mathbb{R}} f(x) h^k_{j,m}(x)\, \mathrm{d}x$ in the sense of dual pairing.

Then let $Z \in \mathbb{N}$ and b_Z defined with the same coefficients (μ^k_{jm}) by

$$b_Z = \sum_{j=-1}^{+\infty} \sum_{k=-Z}^{Z-1} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} h^k_{j,m}\Big|_{(-Z,Z)}. \tag{III.5}$$

The next proposition shows that b_Z is in fact the restriction of b to $(-Z, Z)$. Therefore, by definition, it belongs to $H^{s-1}_q(-Z, Z)$.

Proposition 33. *Let $Z \in \mathbb{N}$. Then b_Z is the restriction of b to $(-Z, Z)$.*

Proof. Let $\varphi \in \mathcal{D}(\mathbb{R}) \subset S(\mathbb{R})$ with Supp $\varphi \subset (-Z, Z)$. Then

$$\begin{aligned}
\langle b, \varphi \rangle &= \left\langle \sum_{j=-1}^{\infty} \sum_{k \in \mathbb{Z}} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} h^k_{j,m}, \varphi \right\rangle \\
&= \lim_{M \to +\infty} \lim_{L \to +\infty} \sum_{j=-1}^{M} \sum_{k=-L}^{L-1} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} \left\langle h^k_{j,m}, \varphi \right\rangle
\end{aligned}$$

by unconditional convergence of the series in $S'(\mathbb{R})$ for the weak-$*$ topology (by Theorem 8). These dual pairings (which are integrals because Haar wavelets are locally integrable) are null if $k \notin [-Z, Z-1]$ so we obtain:

$$\begin{aligned}
\langle b, \varphi \rangle &= \lim_{M \to +\infty} \sum_{j=-1}^{M} \sum_{k=-Z}^{Z-1} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} \left\langle h^k_{j,m}, \varphi \right\rangle \\
&= \left\langle \sum_{j=-1}^{\infty} \sum_{k=-Z}^{Z-1} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} h^k_{j,m}, \varphi \right\rangle \\
&= \langle b^N_Z, \varphi \rangle
\end{aligned}$$

So by 1.86 in [10], b_Z is the restriction of b to $(-Z, Z)$. $\square$

Remark 34. *By definition,* $b_{|(-Z,Z)} = \left(\frac{\mathrm{d}}{\mathrm{d}x}B^H_x\right)_{|(-Z,Z)}$ *and* $\left(\frac{\mathrm{d}}{\mathrm{d}x}B^H_x\right)_{|(-Z,Z)} = \frac{\mathrm{d}}{\mathrm{d}x}B^H_{x_{|(-Z,Z)}}$ *as distributions, by definition of the derivative. That's why we are able to link* $b_Z = b_{|(-Z,Z)}$ *to* B^H*.*

When one considers the restriction $b_{|(-Z,Z)}$ of b to a bounded interval, it's possible to express it on a Faber basis. It is done in the next Theorem from Triebel. We recall that each Faber function v^k_{jm} is proportionate to the integral of the Haar function h^k_{jm}.

Definition 35 (Triebel. See 3.2.1 in [11])**.** *The Faber system on* $(-Z,Z)$ *is defined as* $\{v^k_{-1,0},\ v^k_{jm}\}$ *with* $\forall m \in \mathbb{Z}, \forall k \in (-Z, Z-1)$*:*

$$\begin{cases} v_j(x) & : x \mapsto (1-2^{j+1}|x|)_+ \ \forall j \in \mathbb{N}\cup\{-1\} \\ v_{j,m} & : x \mapsto v_j(x-2^{-j-1}-2^{-j}m) \ \forall j \in \mathbb{N} \\ v^k_{-1,0} & : x \mapsto v_{-1}(x-k) \\ v^k_{jm} & : x \mapsto v_{jm}(x-k) \ \forall j \in \mathbb{N} \end{cases} \tag{III.6}$$

Theorem 36 (Triebel. See Remark 3.4 in [11], Theorem 3.1 and Corollary 3.3 in [10])**.** *Let* $Z \in \mathbb{N}$*,* $f \in H^s_q(-Z,Z)$*,* $g = f' \in H^{s-1}_q(-Z,Z) = F^{s-1}_{q2}(-Z,Z)$ *for* $2 \le q \le \infty$*, and* $s \in \left(\frac{1}{2}, 1+\frac{1}{q}\right)$*. Therefore, we have the unique Faber and Haar representations:*

$$f = \sum_{j=-1}^{+\infty}\sum_{k=-Z}^{Z-1}\sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-\frac{1}{q}\right)-1} v^k_{j,m} \tag{III.7}$$

$$g = \sum_{j=-1}^{+\infty}\sum_{k=-Z}^{Z-1}\sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-1-\frac{1}{q}\right)} h^k_{j,m} \tag{III.8}$$

with unconditional convergence of the first series in $\mathcal{C}([-Z,Z])$ *and of the second series in* $D'(-Z,Z)$*, where*

$$\begin{cases} \mu^k_{j,m}(b) & = -2^{j\left(s-\frac{1}{q}\right)}\left(\Delta^2_{2^{-j-1}}f\right)(k+2^{-j}m) \\ \mu^k_{-1,0}(b) & = 2^{-s+\frac{1}{q}+1}(f(k+1)-f(k)) \end{cases}, \tag{III.9}$$

$-\frac{1}{2}(\Delta^2_{2^{-j-1}}f)(2^{-j}m) = f(2^{-j}m+2^{-j-1}) - \frac{1}{2}f(2^{-j}m) - \frac{1}{2}f(2^{-j}m+2^{-j})$ *and* $\sum_{m=0}^{2^j-1}$ *means* $m=0$ *when* $j=-1$*.*

Identifying the coefficients $(\mu^k_{jm})_{k\in[-N,N-1]}$ of $b_N = \frac{\mathrm{d}}{\mathrm{d}x}B^H_{x_{|(-N,N)}}$ with the ones in Theorem 36 by uniqueness, we obtain a formula to compute them numerically with the values taken by B^H.

Then, we return to the function

$$b^N = \sum_{j=0}^{N}\sum_{k=-N}^{N-1}\sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-\frac{1}{q}\right)} h^k_{j,m} \tag{III.10}$$

converging to b in $H^{s-1}_{2,q}(\mathbb{R})$ for any $q \in \left(2, \frac{1}{1-s}\right)$. For each given precision $\varepsilon > 0$, we can take $N_0 \in \mathbb{N}$ such that $\forall N \ge N_0$, $\left\|b - b^N_N\right\|_{H^{s-1}_{2,q}(\mathbb{R})} \le \varepsilon$. Then, for each fixed $N \ge N_0$, we are able to compute the coefficients $(\mu^k_{jm})_{k\in[-N,N-1]}$. Figures III.4 and III.5 show the result of such a approximation of the derivative of a fBm path by Haar wavelets.

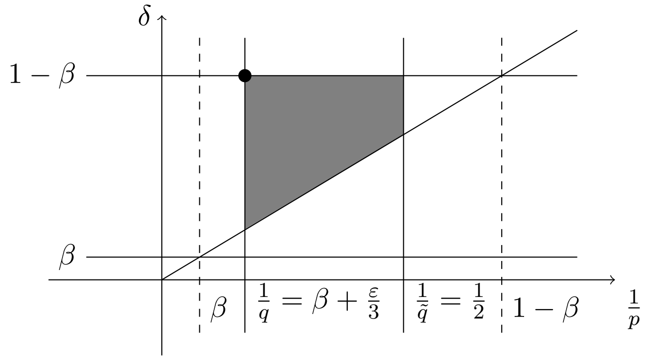

Figure III.3: The set $\widetilde{K}(\beta, q)$ with $\tilde{q} = 2$. Modified figure from the paper [3] of Flandoli, Issoglio and Russo.

Remark 37. *For a fixed N and for large values of k the μ^k_{jm} values for b^N will all be zero because of the compact support of B^H. Then we can use instead*

$$b^N = \sum_{j=0}^{N} \sum_{k=-K-1}^{K} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-\frac{1}{q}\right)} h^k_{j,m}. \tag{III.11}$$

But as explained in Remark 31, without choosing explicitly the smoothing function γ, one can use

$$b^N = \sum_{j=0}^{N} \sum_{k=-K}^{K-1} \sum_{m=0}^{2^j-1} \mu^k_{j,m} 2^{-j\left(s-\frac{1}{q}\right)} h^k_{j,m} \tag{III.12}$$

and stop the process X_t when it exits $[-K, K]$.

We are now able to numerically solve the SDE (I.4) with Euler-Maruyama scheme with drift b^N. Nevertheless, we still need to make sure that the original SDE (I.1) makes sense with respect to the framework developed in [3]. It is the object of the next subsection.

III.2.1 Applying the framework from Flandoli, Issoglio and Russo

We recall that in our special case, the drift b verifies $b = \frac{\mathrm{d}}{\mathrm{d}x} B^H_x \in H^{s-1}_{2,q}(\mathbb{R})$ for all $q \geq 2$ and a fortiori for all $q \in \left(2, \frac{1}{1-s}\right)$, with $\beta = 1 - s$. Then we can use the results from [3] and Part II by choosing $\tilde{q} = 2$ (see figure III.3, page 35 for the modified set and I.1, page 17 for the original set).

Remark 38. *Because $\beta < \frac{1}{2}$, one can still take $\alpha = 1 - 2\beta - \frac{\varepsilon}{2}$ with the new set $\widetilde{K}(\beta, q)$ (see See Remark 27), and therefore apply Theorem 28.*

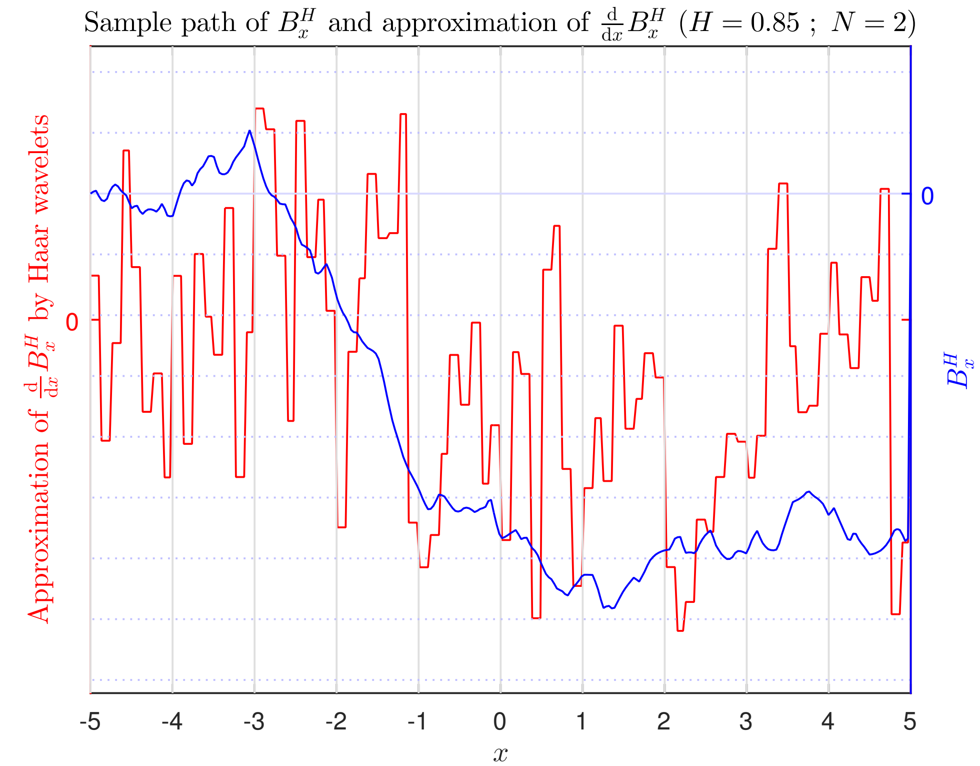

Figure III.4: Haar approximation of the derivative of a fBm path with $N = 2$.

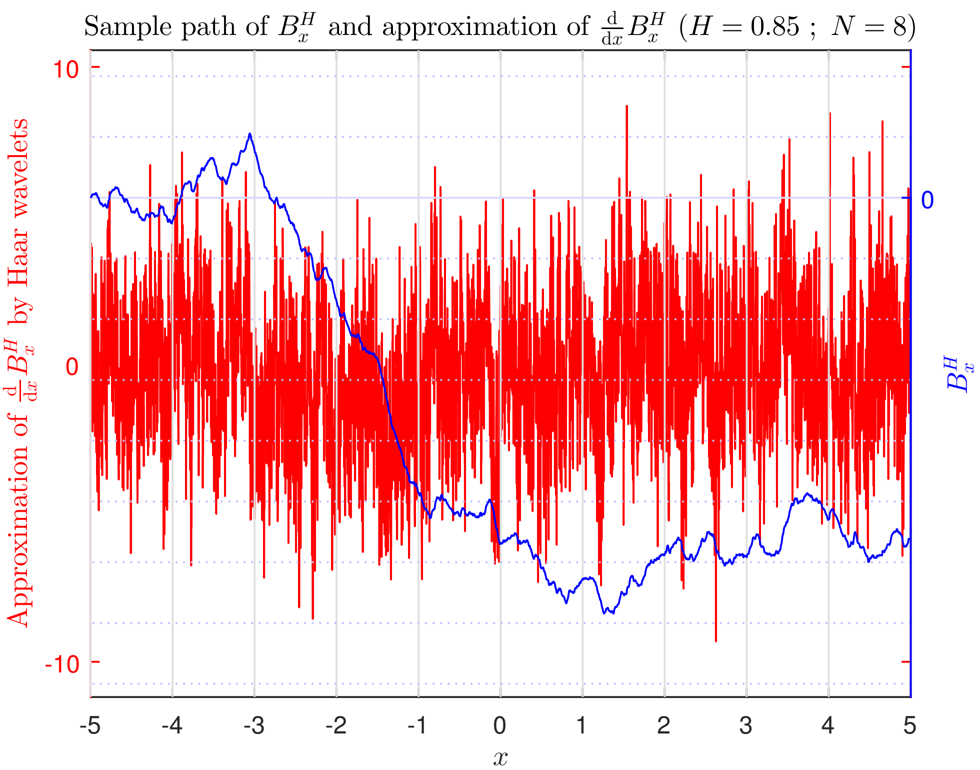

Figure III.5: Haar approximation of the derivative of a fBm path with $N = 8$.

Part IV

Numerical results

IV.1 Monte-Carlo method for error estimation

In order to study the numerical rates of convergence of our algorithm, we must numerically compute expectations. That's why we use the Monte-Carlo method which estimates expectations thanks to convergence results directly related to the Central Limit Theorem. This theorem expresses in fact the speed of convergence of the empirical mean to the expectation.

More precisely, if $X_1, \cdots, X_m$ are independent random variables with the same law as X, admitting finite moments of order up to two, and $V_m = \frac{m}{m-1}\left(\frac{1}{m}\sum_{i=1}^m X_i^2 - \frac{1}{m^2}\left(\sum_{i=1}^m X_i\right)^2\right)$ $= \frac{1}{m-1}\left(\sum_{i=1}^m X_i^2 - \frac{1}{m}\left(\sum_{i=1}^m X_i\right)^2\right)$ is the unbiased variance, the Central Limit Theorem and Slutsky's Lemma give the following convergence in law:

$$\frac{\sum_{i=1}^m (X_i - \mathbb{E}[X])}{\sqrt{V_m}\sqrt{m}} \xrightarrow[m\to\infty]{\mathcal{L}} \mathcal{N}(0,1). \tag{IV.1}$$

Then we can construct the asymptotic $100(1-\alpha)\%$ confidence interval such that:

$$\mathbb{P}\left(\mathbb{E}[X] \in \left[\frac{1}{m}\sum_{i=1}^m X_i - q\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{V_m}{m}}, \frac{1}{m}\sum_{i=1}^m X_i + q\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{V_m}{m}}\right]\right) \xrightarrow[m\to\infty]{} 1-\alpha \tag{IV.2}$$

where q is the quantile of the normal distribution. We will apply this method to construct an asymptotic 95% confidence interval for the strong error $\mathbb{E}\left[\sup_{0\le t\le T}\left|X_t^{N,n} - X_t^N\right|^2\right]$ and the weak error $\left|\mathbb{E}\left[X_T - X_T^N\right]\right|$.

A key numerical concern here is to have the smallest variance as possible, in order to reduce the length of the confidence interval for a given number m of simulation. Therefore we obtain a more precise estimation of the expectation. To do so, a lot of variance reduction methods exist. We use the antithetic variates method, which exploits the fact that the Brownian motions W and $-W$ have the same law to reduce the variance. See [2] for a more detailed presentation. Numerically, it requires to simulate at each iteration two paths of the solution $X^{N,n}$ of (I.4) with both driving Brownian motions W and $-W$.

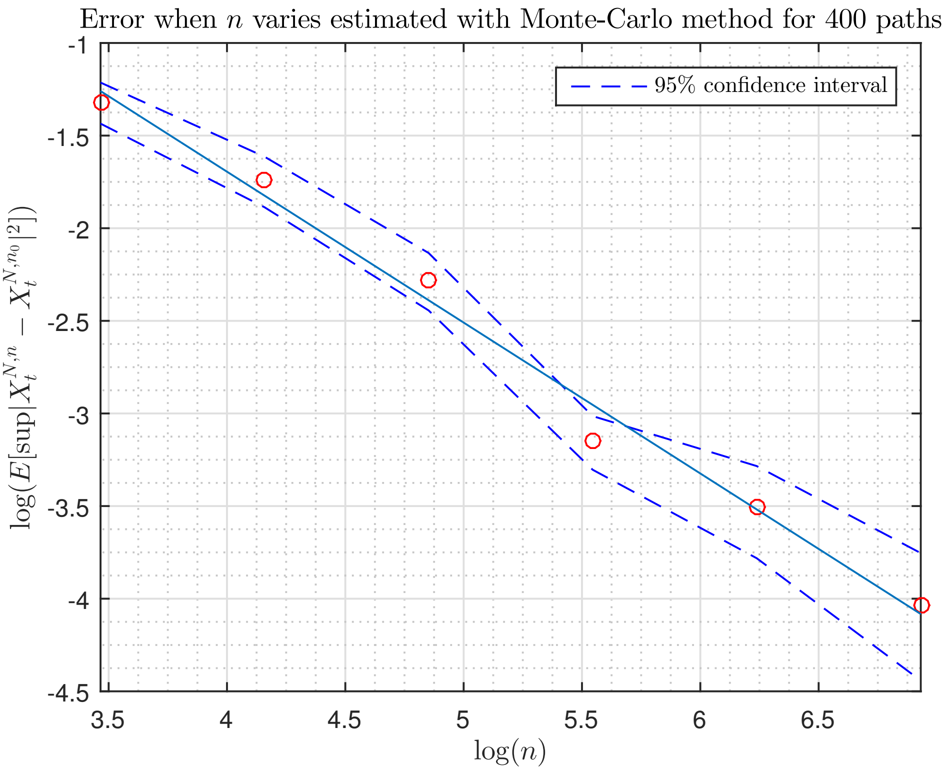

Figure IV.1: Estimation of the L^2 error of the Euler-Maruyama scheme with a Monte-Carlo method.

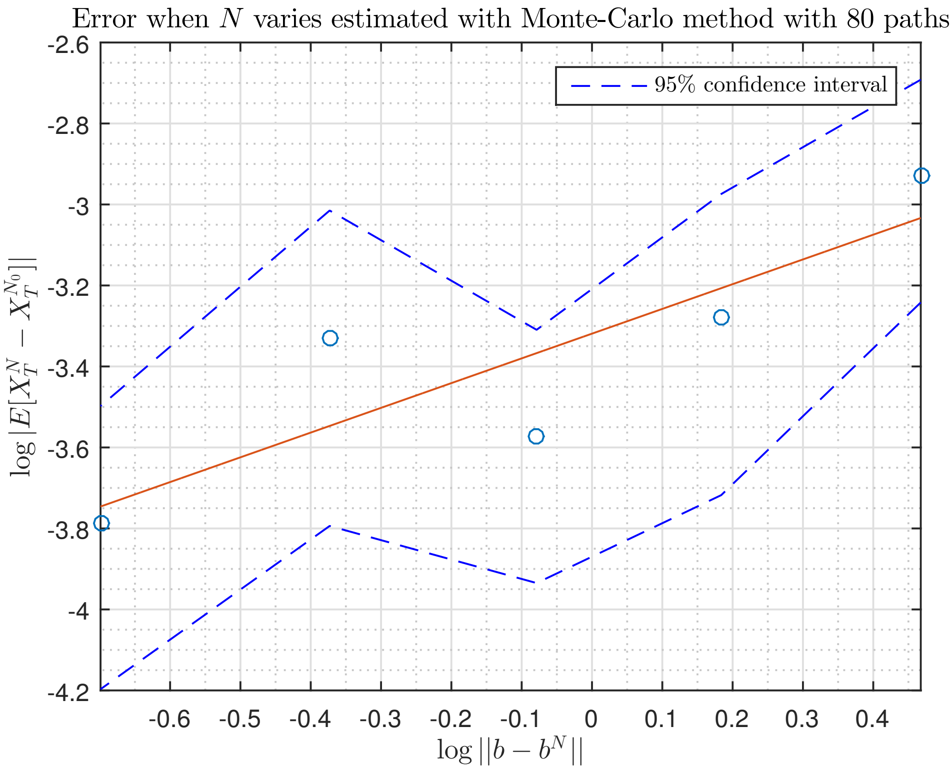

Figure IV.2: Estimation of the weak error of the Euler-Maruyama scheme with a Monte-Carlo method.

IV.2 Strong convergence of the Euler approximation

We study numerically the case $X_0 = 0$, the reference solution has $n_0 = 2^{12}$ time points, $n \in \{2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}\}$, and $N = 5$. We observe in figure IV.1 a numerical strong convergence rate of 0.81 when Theorem 14 shows a theoretical rate of $0.5 - \varepsilon$. The convergence is then better than expected with the theoretical results. This may be due to our special case, where the drift of the approximated SDE (I.4) is piecewise constant, which is more regular than a piecewise Lipschitz drift.

IV.3 Weak convergence of the approximated solution to the virtual solution

We study numerically the case $X_0 = 1$, solutions have $n_0 = 2^{12}$ time points, $N \in \{2, 3, 4, 5, 6\}$, and $N_0 = 7$. The error $\left\|b - b^N\right\|_{H^{-0.15}_{2,2}}$ is computed using the isomorphism between the coefficients space and $H^{-0.15}_{2,2}$. We observe in figure IV.2 a numerical weak convergence rate of 0.61. The simulation was done with a drift of regularity $H^{-0.15}_{2,q}$ for $q \geq 2$. Therefore, we would expect from Theorem 28 a convergence rate of $1 - 4\beta - \varepsilon = 0.4 - \varepsilon$. We obtain numerically a better rate.

First of all we need to say that the slope estimation is rather imprecise, because of the confidence interval around each point. We could increase the number of realizations taken for the Monte-Carlo estimation but these simulations take a lot of time to compute so we need to compromise between computation time and precision. Moreover, we should increase the number of time points 2^{n_0} in the Euler-Maruyama method to have the approximated solutions X^{N,n_0} and X^{N_0,n_0} closer to the proper solutions X^N and X^{N_0}. The same notion of compromise arise also in this case. Theoretically, we showed that the accuracy of the Euler approximation requires to take n bigger when N becomes bigger. It is confirmed numerically as far as the n_0 we chose here doesn't give proper results when N is bigger. It should be adapted in consequence.

Internship's organisation

Dates	Research work
May 14th–May 21st	Study of the paper [3], numerical simulation of fBm. Beginning of the implementation of the Euler-Maruyama method.
May 21st–May 28th	Haar approximation in the case of a regular drift (identity), plots of the Euler method error and reading of the paper [8].
May 28th–June 4th	Improvement of the algorithm, numerical and theoretical study of the convergence rate.
June 4th–June 11th	Study of article [12] and understanding how to use his results to find a convergence rate for our algorithm in 1D.
June 11th–June 18th	Careful writing and correction of lemmas, proof of the convergence rate result.
June 18th–June 25th	Understanding and adapting the lemma from [12] bounding the expectation of the local time.
June 25th–July 2nd	Convergence in law of the approximating process to the virtual solution with results from [3], and writing of the report
July 2nd–July 9th	Writing of the report, numerical studies of convergence rates and attempts to justify the computation of the Haar coefficients on $\mathbb{R}$.
July 9th–July 16th	Reading Triebel books in order to understand Haar and Faber representations, study of fBm's regularity with [5].
July 16th–July 23rd	Writing of the details concerning the Haar approximation in the case of the derivative of a fractional Brownian motion.
July 23rd–August 3rd	Study of Haar wavelets in $\mathbb{R}^d$ and writing of the report. Extension of the results to time-dependent drift

Conclusion

My work at the University of Leeds has been a very stimulating first research experience. During this project, I contributed to the numerical study of irregular stochastic differential equations, by designing a new algorithm which applies to distributional drifts. My main contribution was obtaining a rate in dimension one of $1 - 4\beta - \varepsilon$ for the convergence of the solution to the approximated SDE to the virtual solution, when $\beta < 1/4$.

Moreover, I was able to apply my approximation method in the case of the derivative of a fBm path. This required to produce an approximation method using Haar and Faber wavelets in order to approximate a distributional drift by a function. On this occasion, I developed various Matlab scripts and realized numerical simulations. That's why this internship allowed me to deepen my knowledge in stochastic analysis, function spaces and numerical methods. Moreover, working in a international environment was a fascinating experience.

Several points of this project could call for future work. For example, it would be great to investigate a dimension one rate when $\beta \geq 1/4$, and to find a way to obtain a convergence rate in any finite dimension. Nevertheless, no counterpart of local time exist in $\mathbb{R}^d$ thus our proof cannot extend directly to higher dimensions.

Another possible work can be dedicated to find a way to compute the Haar coefficients in $\mathbb{R}^d$ when the drift is the derivative of a Hölder-continuous function, like a fBm path. In fact, in dimension one, we were able to use a Faber representation for this path, and therefore deduce the Haar coefficients. But Triebel explains in [10] that the Faber basis cannot be extended to $\mathbb{R}^d$ as easily as the Haar basis is. It should thus be interesting to look for an alternative method to compute these coefficients.

Bibliography

[1] S. Ankirchner, T. Kruse, and M. Urusov. Numerical approximation of irregular SDEs via skorokhod embeddings. *Journal of Mathematical Analysis and Applications*, 440(2):692–715, 2016.

[2] P. Boyle, M. Broadie, and P. Glasserman. Monte carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.

[3] F. Flandoli, E. Issoglio, and F. Russo. Multidimensional stochastic differential equations with distributional drift. *Transactions of the American Mathematical Society*, 369 (3):1655–1688, 3 2017.

[4] N. Halidias and P. E. Kloeden. A note on the Euler–Maruyama scheme for stochastic differential equations with a discontinuous monotone drift coefficient. *BIT Numerical Mathematics*, 48:51–59, 2008.

[5] E. Issoglio. Transport equations with fractal noise – existence uniqueness and regularity of the solution. *The Journal of Analysis and Applications*, 32(1):37–53, 2013.

[6] E. Issoglio and F. Russo. On a class of Markov BSDEs with generalized driver. submitted. arXiv:1805.02466v1.

[7] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift. *Journal of Computational and Applied Mathematics*, 326:138–158, 2016.

[8] G. Leobacher and M. Szölgyenyi. Convergence of the Euler–Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numerische Mathematik*, 138 (1):219–239, 2018.

[9] H.-L. Ngo and D. Taguchi. Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients. *Mathematics of Computation*, 85:1793–1819, 2016.

[10] H. Triebel. *Bases in function spaces, sampling, discrepancy, numerical integration.* European Mathematical Society, 2010.

[11] H. Triebel. *Faber systems and their use in sampling, discrepancy, numerical integration.* European Mathematical Society, 2012.

[12] L. Yan. The Euler Scheme with Irregular Coefficients. *The Annals of Probability*, 30(3):1172–1194, 7 2002.

[13] P. Étoré and M. Martinez. Exact simulation for solutions of one-dimensional stochastic differential equations with discontinuous drift. *ESAIM: Probability and Statistics*, 18:686–702, 2014.

Abbreviations

fBm: fractional Brownian motion.
See the definition in Example 2.

SDE: Stochastic Differential Equation.
Differential equation involving stochastic processes.

PDE: Partial Differential Equation.
Differential equation involving partial derivatives of multivariate functions.

Appendix A

Technical proofs

A.1 Lemma 22

As explained before, this proof is an adaptation of the one of Yan in [12]. The only difference is the fact that in our case, we don't assume that the process starts at 0.

Proof of Lemma 22. If $t = 0$, the result is obvious. We now assume $t \in (0, T]$. We note $U_t(X) = \sup\{n \in \mathbb{N} | \tau_n < t\}$ and $n(t) = t \wedge \upsilon_{U_t(X)+1}$. By Meyer-Tanaka's formula, $\forall i \in \mathbb{N}^*$:

$$X^+_{\tau_i \wedge t} - X^+_{\upsilon_i \wedge t} = \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} \mathbb{1}_{\{X_s > 0\}} \, \mathrm{d}X_s + \frac{1}{2} \left\{ L^0_{\tau_i \wedge t}(X) - L^0_{\upsilon_i \wedge t}(X) \right\}.$$

Because $\forall i \in \mathbb{N}^*,\ L^0_{\tau_i \wedge t}(X) = L^0_{\upsilon_{i+1} \wedge t}(X)$ and $L^0_{\upsilon_1 \wedge t}(X) = 0$, we have

$$\sum_{i=1}^{U_t(X)+1} \left(X^+_{\tau_i \wedge t} - X^+_{\upsilon_i \wedge t} \right) = \int_{\upsilon_1 \wedge t}^{t} \theta_s(X) \, \mathrm{d}X_s + \frac{1}{2} \, L^0_t(X).$$

The left term is equal to $\varepsilon U_t(X) + X^+_t - X^+_{n(t)}$ so

$$\varepsilon U_t(X) = \int_{\upsilon_1 \wedge t}^{t} \theta_s(X) \, \mathrm{d}X_s + \frac{1}{2} \, L^0_t(X) - X^+_t + X^+_{n(t)}. \tag{A.1}$$

Now we express differently $U_t(X)$. $F \in \mathcal{C}^2(\mathbb{R})$ so by Itô's formula:

$$F\left(X^+_{\tau_i \wedge t}\right) - F\left(X^+_{\upsilon_i \wedge t}\right) = \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} F'\left(X^+_s\right) \mathrm{d}X^+_s + \frac{1}{2} \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} F''\left(X^+_s\right) \mathrm{d}[X^+]_s.$$

By (A.1), $\mathrm{d}X^+_s = \mathbb{1}_{\{X_s > 0\}} \mathrm{d}X_s + \frac{1}{2} \, \mathrm{d}L^0_t(X)$ and $\mathrm{d}[X^+]_s = \mathbb{1}_{\{X_s > 0\}} \mathrm{d}[X]_s$. It follows

$$\begin{aligned} F\left(X^+_{\tau_i \wedge t}\right) - F\left(X^+_{\upsilon_i \wedge t}\right) = & \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} F'\left(X^+_s\right) \mathbb{1}_{\{X_s > 0\}} \, \mathrm{d}X_s + \frac{1}{2} \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} F'(X^+_s) \, \mathrm{d}L^0_t(X) \\ & + \frac{1}{2} \int_{\upsilon_i \wedge t}^{\tau_i \wedge t} F''\left(X^+_s\right) \mathbb{1}_{\{X_s > 0\}} \, \mathrm{d}[X]_s. \end{aligned}$$

Adding up for i, with $F(0)=0$ we obtain

$$\begin{aligned}
&F(\varepsilon)U_t(X)+F\left(X_t^+\right)-F\left(X_{n(t)}^+\right)\\
&=\sum_{i=1}^{U_t(X)+1}\left(F\left(X_{\tau_i\wedge t}^+\right)-F\left(X_{\upsilon_i\wedge t}^+\right)\right)\\
&=\int_{\upsilon_1\wedge t}^t F'(X_s^+)\ \theta_s(X)\ \mathrm{d}X_s+\frac{1}{2}\int_{\upsilon_1\wedge t}^t F'\left(X_s^+\right)\Xi_s\ \mathrm{d}L_t^0(X)\\
&+\frac{1}{2}\int_{\upsilon_1\wedge t}^t F''\left(X_s^+\right)\theta_s(X)\ \mathrm{d}[X]_s
\end{aligned}$$

with $\Xi_s=\sum_{n=1}^{\infty}\mathbb{1}_{\{\upsilon_n<s\le\tau_n\}}$. The measure $\mathrm{d}L_t^0(X)$ is almost surely carried by $\{t|X_t=0\}$ so we can simplify $\int_{\upsilon_1\wedge t}^t F'\left(X_s^+\right)\Xi_s\ \mathrm{d}L_t^0(X)=F'(0)\int_{\upsilon_1\wedge t}^t\Xi_s\ \mathrm{d}L_t^0(X)$ in order to have, with $F'(0)=0$:

$$F(\varepsilon)U_t(X)=-F\left(X_t^+\right)+F\left(X_{n(t)}^+\right)+\int_{\upsilon_1\wedge t}^t F'\left(X_s^+\right)\theta_s(X)\ \mathrm{d}X_s \tag{A.2}$$

$$+\frac{1}{2}\int_{\upsilon_1\wedge t}^t F''\left(X_s^+\right)\theta_s(X)\ \mathrm{d}[X]_s. \tag{A.3}$$

Combining (A.1) and (A.2), it follows

$$\begin{aligned}
L_t^0(X)F(\varepsilon)&=2F(\varepsilon)(X_t^+-X_{n(t)}^+)-2\varepsilon(F\left(X_t^+\right)-F\left(X_{n(t)}^+\right)\\
&-2\int_{\upsilon_1\wedge t}^t\left(F(\varepsilon)-\varepsilon F'\left(X_s^+\right)\right)\theta_s(X)\ \mathrm{d}X_s+\varepsilon\int_{\upsilon_1\wedge t}^t F''\left(X_s^+\right)\theta_s(X)\ \mathrm{d}[X]_s.
\end{aligned}$$

Then, if $n(t)=t$, the two first right terms of the equality are equal to zero. Else, if $n(t)=\upsilon_{U_t(X)+1}$, $0\le F(\varepsilon)(X_t^+-X_{n(t)}^+)=F(\varepsilon)X_t^+\le F(\varepsilon)\varepsilon$ and $-2\varepsilon(F\left(X_t^+\right)-F\left(X_{n(t)}^+\right)$ $=-2\varepsilon F\left(X_t^+\right)\le 0$ because of the positivity of F. Finally we obtain:

$$L_t^0(X)\le 2\varepsilon-\frac{2}{F(\varepsilon)}\int_{\upsilon_1\wedge t}^t\left(F(\varepsilon)-\varepsilon F'\left(X_s^+\right)\right)\theta_s(X)\ \mathrm{d}X_s+\frac{\varepsilon}{F(\varepsilon)}\int_{\upsilon_1\wedge t}^t F''\left(X_s^+\right)\theta_s(X)\ \mathrm{d}[X]_s.$$

□

A.2 Lemma 25

Proof of Lemma 25. We will use the notations introduced in Remark 24. By Lemma 18, we choose λ big enough for ∇u and ∇u^N to be bounded by $\frac{1}{2}$. λ can be chosen independently of N as far as $\left\|b-b^N\right\|_{\infty,H^s_{\tilde{q},q}(\mathbb{R})}\underset{N\to\infty}{\longrightarrow}0$. Therefore $u^N(t,\cdot)$ and $u(t,\cdot)$ are $\frac{1}{2}$-Lipschitz. We recall that in this case, by Lemma 22 in [3], $\Psi(t,\cdot)$ and $\Psi^N(t,\cdot)$ are 2-Lipschitz. Therefore $\widetilde{u}^N(t,\cdot)$ and $\widetilde{u}(t,\cdot)$ are 1-Lipschitz.

We can notice that $\widetilde{\nabla u}$ is α-Hölder with constant $2^\alpha\left\|u\right\|_{\infty,\mathcal{C}^{1,\alpha}}$ (See Proposition 6) and ∇u^N is α-Hölder with a constant which can be bounded by Λ independently of N (see the proof of Lemma 24 in [3]). Let $\varepsilon \in (0,1]$. Corollary 23 gives us:

$$0 \le L_T^0(Y-Y^N) \le 2\varepsilon - \frac{2}{\varepsilon}\int_{\upsilon_1\wedge T}^T \theta_s(Y-Y^N)\left(\varepsilon - 2(Y_s - Y_s^N)^+\right)\,\mathrm{d}\left(Y_s - Y_s^N\right) + \frac{2}{\varepsilon}\int_{\upsilon_1\wedge T}^T \theta_s(Y-Y^N)\,\mathrm{d}\left[Y-Y^N\right]_s$$

with

$$Y_t - Y_t^N = y_0 - y_0^N + (\lambda+1)\int_0^t \left\{u\left(s,\Psi\left(s,Y_s\right)\right) - u^N\left(s,\Psi^N\left(s,Y_s^N\right)\right)\right\}\,\mathrm{d}s + \int_0^t \left\{\nabla u\left(s,\Psi\left(s,Y_s\right)\right) - \nabla u^N\left(s,\Psi^N\left(s,Y_s^N\right)\right)\right\}\,\mathrm{d}W_s.$$

Remark 39. *Note that* $\theta_s(Y-Y^N)\left|\varepsilon - 2(Y_s-Y_s^N)^+\right| \le \varepsilon\theta_s(Y-Y^N)$.

∇u and ∇u^N are bounded so the Itô integral is a true martingale. We take the expectation:

$$\mathbb{E}\left[L_T^0(Y-Y^N)\right] \le 2\varepsilon + 2(\lambda+1)\,\mathbb{E}\left[\int_{\upsilon_1\wedge T}^T \theta_s\left(Y-Y^N\right)\left|\widetilde{u}\left(s,Y_s\right) - \overline{u}^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right] + \frac{2}{\varepsilon}\,\mathbb{E}\left[\int_{\upsilon_1\wedge T}^T \theta_s\left(Y-Y^N\right)\left\{\widetilde{\nabla u}\left(s,Y_s\right) - \overline{\nabla u}\left(s,Y_s^N\right)\right\}^2\,\mathrm{d}s\right]$$

$$\begin{aligned}
&\le 2\varepsilon + 2(\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|\widetilde{u}\left(s,Y_s\right) - \widetilde{u}\left(s,Y_s^N\right)\right|\,\mathrm{d}s\right]\\
&+ 2(\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|\widetilde{u}\left(s,Y_s^N\right) - \widetilde{u}^N\left(s,Y_s^N\right)\right|\,\mathrm{d}s\right]\\
&+ 2(\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|\widetilde{u}^N\left(s,Y_s^N\right) - \overline{u}^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right]\\
&+ \frac{6}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left\{\widetilde{\nabla u}\left(s,Y_s\right) - \widetilde{\nabla u}\left(s,Y_s^N\right)\right\}^2\,\mathrm{d}s\right]\\
&+ \frac{6}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left\{\widetilde{\nabla u}\left(s,Y_s^N\right) - \widetilde{\nabla u}^N\left(s,Y_s^N\right)\right\}^2\,\mathrm{d}s\right]\\
&+ \frac{6}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left\{\widetilde{\nabla u}^N\left(s,Y_s^N\right) - \overline{\nabla u}^N\left(s,Y_s^N\right)\right\}^2\,\mathrm{d}s\right]
\end{aligned}$$

$$\begin{aligned}
&\le 2\varepsilon + 2(\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|Y_s - Y_s^N\right|\,\mathrm{d}s\right] + 2(\lambda+1)T\kappa\left\|b^N - b\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}\\
&+ (\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|\Psi\left(s,Y_s^N\right) - \Psi^N\left(s,Y_s^N\right)\right|\mathrm{d}s\right]\\
&+ \frac{6\times 4^\alpha\left\|u\right\|_{\infty,\mathcal{C}^{1,\alpha}}^2}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|Y_s - Y_s^N\right|^{2\alpha}\,\mathrm{d}s\right] + 6T\kappa^2\left\|b^N - b\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^2\,\varepsilon^{-1}\\
&+ \frac{6\Lambda^2}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|\Psi\left(s,Y_s^N\right) - \Psi^N\left(s,Y_s^N\right)\right|^{2\alpha}\,\mathrm{d}s\right]
\end{aligned}$$

where we used Lemma 20, the 1-Lipschitz property of $\widetilde{u}$, the $1/2$-Lipschitz property of u^N, the α-Hölder property of $\widetilde{\nabla u}$ (with constant $2^\alpha \|u\|_{\infty,\mathcal{C}^{1,\alpha}}$), and the α-Hölder property of ∇u^N (with constant Λ). Lemma 21 gives us:

$$
\begin{aligned}
&\mathbb{E}\left[L_T^0(Y-Y^N)\right] \\
&\leq 2\varepsilon + 2(\lambda+1)\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|Y_s-Y_s^N\right|\ \mathrm{d}s\right] + 4(\lambda+1)T\kappa\left\|b^N-b\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} \\
&+ \frac{6\times 4^\alpha \|u\|_{\infty,\mathcal{C}^{1,\alpha}}^2}{\varepsilon}\,\mathbb{E}\left[\int_0^T \theta_s\left(Y-Y^N\right)\left|Y_s-Y_s^N\right|^{2\alpha}\ \mathrm{d}s\right] + 6T\kappa^2\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^2\ \varepsilon^{-1} \\
&+ 6\Lambda^2 T 4^\alpha\kappa^{2\alpha}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{2\alpha}\ \varepsilon^{-1}.
\end{aligned}
$$

As $\theta_s(Y-Y^N)\left|Y_s-Y_s^N\right| \leq \varepsilon$ and $\theta_s(Y-Y^N)\left|Y_s-Y_s^N\right|^{2\alpha} \leq \varepsilon^{2\alpha}$, we have

$$
\begin{aligned}
&\mathbb{E}\left[L_T^0(Y-Y^N)\right] \\
&\leq 2\varepsilon + 2(\lambda+1)\ T\varepsilon + 4(\lambda+1)T\kappa\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} + 6\|u\|_{\infty,\mathcal{C}^{1,\alpha}}^2\ 4^\alpha T\varepsilon^{2\alpha-1} \\
&+ 6T\left(\Lambda^2 4^\alpha\kappa^{2\alpha}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^2\right)\varepsilon^{-1}.
\end{aligned}
$$

As $0 < 2\alpha-1 < 1$, the result follows from $\varepsilon \leq \varepsilon^{2\alpha-1}$ when $0 < \varepsilon \leq 1$.

□

A.3 Proposition 26

Proof of Proposition 26. Let $\varepsilon \in (0,1]$, by Lemma 25,

$$0 \leq \mathbb{E}\left[L_T^0(Y-Y^N)\right] \leq g(\varepsilon)$$

where

$$
\begin{aligned}
g(\varepsilon) &= 4(\lambda+1)T\kappa\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}} + \left(2+2(\lambda+1)\ T + 6\|u\|_{\infty,\mathcal{C}^{1,\alpha}}^2\ 4^\alpha T\right)\varepsilon^{2\alpha-1} \\
&+ 6T\left(\Lambda^2 4^\alpha\kappa^{2\alpha}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^2\right)\varepsilon^{-1}.
\end{aligned}
$$

With

$$
\begin{aligned}
g'(\varepsilon) &= (2\alpha-1)\left(2+2(\lambda+1)\ T + 6\|u\|_{\infty,\mathcal{C}^{1,\alpha}}^2\ 4^\alpha T\right)\varepsilon^{2\alpha-2} \\
&- 6T\left(\Lambda^2 4^\alpha\kappa^{2\alpha}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^2\right)\varepsilon^{-2},
\end{aligned}
$$

the minimum of g on $(0,1]$ is reached when N is big enough in

$$\varepsilon_N = \left(\frac{6T\left(\Lambda^2 4^\alpha\kappa^{2\alpha}\left\|b-b^N\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2\left\|b-b^N\right\|_{\infty,H_q^{-\beta}}^2\right)}{(2\alpha-1)\left(2+2(\lambda+1)\ T + 6\|u\|_{\infty,\mathcal{C}^{1,\alpha}}^2\ 4^\alpha T\right)}\right)^{\frac{1}{2\alpha}} = \nu_N\left\|b^N-b\right\|_{\infty,H_{\tilde{q},q}^{-\beta}}.$$

where

$$\nu_N = \left(\frac{6T \left(\Lambda^2 4^\alpha \kappa^{2\alpha} + \kappa^2 \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2(1-\alpha)} \right)}{(2\alpha - 1) \left(2 + 2(\lambda+1)\ T + 6 \left\| u \right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T \right)} \right)^{\frac{1}{2\alpha}}$$

$$\geq \nu_\infty = \left(\frac{6T\Lambda^2 4^\alpha \kappa^{2\alpha}}{(2\alpha - 1) \left(2 + 2(\lambda+1)\ T + 6 \left\| u \right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T \right)} \right)^{\frac{1}{2\alpha}}.$$

Therefore $\mathbb{E}\left[L_T^0 (Y - Y^N) \right] \leq g(\varepsilon_N)$

$$\leq 4(\lambda+1)T\kappa \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}} + \left(2 + 2(\lambda+1)\ T + 6 \left\| u \right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T \right) \nu_N^{2\alpha-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}$$
$$+ 6T \left(\Lambda^2 4^\alpha \kappa^{2\alpha} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha} + \kappa^2 \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2} \right) \nu_N^{-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{-1}$$
$$\leq 4(\lambda+1)T\kappa \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}} + \left(2 + 2(\lambda+1)\ T + 6 \left\| u \right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T \right) \nu_N^{2\alpha-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}$$
$$+ 6T \left(\Lambda^2 4^\alpha \kappa^{2\alpha} + \kappa^2 \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2(1-\alpha)} \right) \nu_\infty^{-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}.$$

Then we use the fact that $\nu_N^{2\alpha-1} \leq 2\nu_\infty^{2\alpha-1}$ and $\left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}} \leq 1$ when N is big enough:

$$\mathbb{E}\left[L_T^0 (Y - Y^N) \right]$$
$$\leq 4(\lambda+1)T\kappa \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1} + \left(4 + 4(\lambda+1)\ T + 12 \left\| u \right\|_{\infty, \mathcal{C}^{1,\alpha}}^2 4^\alpha T \right) \nu_\infty^{2\alpha-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}$$
$$+ 6T \left(\Lambda^2 4^\alpha \kappa^{2\alpha} + \kappa^2 \right) \nu_\infty^{-1} \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}$$
$$\leq \left(4(\lambda+1)T\kappa + \nu_\infty^{-1} \left(\frac{12}{2\alpha - 1} T\Lambda^2 4^\alpha \kappa^{2\alpha} + 6T \left(\Lambda^2 4^\alpha \kappa^{2\alpha} + \kappa^2 \right) \right) \right) \left\| b - b^N \right\|_{\infty, H_{\tilde{q},q}^{-\beta}}^{2\alpha-1}$$

The result follows. ∎