Transfer report

Author:
Luis Mario Chaparro Jáquez
201481193

Supervisors:
Dr. Elena Issoglio
Dr. Jan Palczewski

November 2021

Contents

1 Overview

Stochastic Differential Equations (SDEs) are widely used tools in fields such as finances, neuroscience and engineering. They help to model behaviours in systems with certain level of noise, or randomness, that deterministic methods cannot model.

Numerical schemes strong convergence for SDEs with low regularity coefficients have been widely studied, examples of it are [1, 3, 4, 5, 6].

The works from De Angelis et al. [5], and Flandoli et al. [6] have established the framework for this project. A one dimensional SDE is considered:

$$\begin{cases} dX_t = b(t, X_t)dt + dW_t, & t \in [0, T], \\ X_0 = x_0, \end{cases} \tag{1}$$

where W is a Brownian motion, and $b(t, \cdot)$ is a distribution taking values in a fractional Sobolev space of negative order, namely $H_q^{-s}(\mathbb{R})$ as defined in Definition 18.

This type of equations immediately introduce a challenge because the coefficient b can not be evaluated pointwise and it is necessary to give a meaning to the term $\int_0^t b(s, X_s)ds$. This problem is solved in [6], and then in [5] an algorithm for the one dimensional version of the problem is described. The algorithm proposed has two steps for it to produce the numerical solutions:

1. First the coefficient b in (1) is approximated to be able to work numerically with it using a sum of Haar functions and then submitted to a smoothing procedure applying the heat kernel as in Definition 22 via convolution. This is described in Section 4.

2. And finally applying the Euler-Maruyama scheme for the modified coefficient $P_{\eta_n} b^N$ as discussed in Section 5.

We implemented the same process with good results for the Euler scheme, and attempted to use a scheme with higher order of convergence, namely the 1.5 strong scheme. In the latter we did not achieve good results, however this is an scheme that has not been studied theoretically

for this type of equations and even with desirable results the link between the original problem and the numerical scheme would be unclear.

Further we have a brief introduction to the theory of Backward SDEs (BSDEs) and Forward-Backward SDEs (SDEs), since the following steps are to develop the theory for numerical schemes of BSDEs and FBSDEs.

The report is organised as follows: In Section 2 we have the definition of SDEs and conditions for uniqueness of solutions. Section 3 is the introduction of the classical theory on numerical schemes for SDEs, introducing several different schemes, modes of convergence and an example for an SDE with analytical solution, in order to try the numerical schemes discussed. Section 4 presents the theory necessary to understand the approximation of the distributional coefficient b, and in Section 5 we have a review of the theoretical results from [5] and the numerical implementations concerning those results. Finally the short review about BSDEs and FBSDEs in Section 6.

2 Background material on SDEs

Let

$$b(t,x) : [0,\infty) \times \mathbb{R} \to \mathbb{R}, \tag{2}$$

$$\sigma(t,x) : [0,\infty) \times \mathbb{R} \to \mathbb{R}, \tag{3}$$

be Borel measurable functions and $W = \{W_t; 0 < t < \infty\}$ be a Brownian motion. For $T > 0$ consider the following Stochastic Differential Equation (SDE)

$$X_t = x_0 + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s, \quad t \in [0,T], \quad x_0 \in \mathbb{R}. \tag{4}$$

Often for ease of notation we can refer to (4) as

$$\begin{cases} dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t, & t \in [0,T] \\ X_0 = x_0, & x_0 \in \mathbb{R}. \end{cases} \tag{5}$$

However this last equation is just notation and it is always intended to represent (4). Let us note that usually the coefficients $b(t,x)$ and $\sigma(t,x)$ are called drift and diffusion coefficients, respectively.

Definition 1 (Strong solution [9, Definition 5.2.1]) *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, we call a strong solution of equation (4) a $\mathcal{F}$-adapted, continuous, $\mathbb{R}$-valued stochastic process $(X_t)_{t \in [0,\infty)}$ such that $\mathbb{P}[X_0 = x_0] = 1$, and (4) holds almost surely for every $t \geq 0$.*

The following theorem states sufficient conditions for the existence and uniqueness for a solution of the SDE.

Theorem 1 ([9, Theorem 5.2.9]) *Suppose that there exists a constant $K \geq 0$ such that for all $x,y \in \mathbb{R}$ and $t \geq 0$,*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K|x-y|, \tag{6}$$

$$|b(t,x)| + |\sigma(t,x)| \leq K(1 + |x|). \tag{7}$$

Then there exist a continuous, $\mathcal{F}$-adapted process $X = (X_t)_{t \in [0,\infty)}$ which is a strong solution to equation (5). This solution is unique up to indistinguishability, i.e: if $\tilde{X}$ is also a strong solution then $\mathbb{P}(X_t = \tilde{X}_t; \forall 0 \leq t \leq \infty) = 1$.

3 Numerical schemes for SDEs

As in the deterministic theory of Differential Equations, most of the equations have no closed form solution, so that it becomes natural to develop numerical schemes to treat such objects which arise in a variety of problems.

3.1 Itô-Taylor expansions

In order to construct numerical schemes to solve SDEs, one can use a procedure that is an analogous to the Taylor expansion used in ODEs. Said procedure is called *Itô-Taylor expansion* and is presented in the following theorem whose further study can be found in [10, Chapter 5]. In order to introduce the theorem, let us set some notation.

Definition 2 (Multi-indices) *We call a row vector*

$$\alpha = (j_1, ..., j_l) \tag{8}$$

where

$$j_i \in \{0, 1\} \tag{9}$$

a multi-index with length $l := l(\alpha) \in \{1, 2, ...\}$. We denote the set of all multi-indices as $\mathcal{M}$.

We call also v to the multi-index with length zero, i.e: $l(v) = 0$, and $-\alpha$ and $\alpha-$ the multi-indices obtained by removing the first and last element, respectively, of α. Also $n := n(\alpha) \in \{0, 1, 2, ...\}$ is the amount of elements equal to zero in the multi-index.

Definition 3 (Hierarchical sets) *A subset $\mathcal{A} \subset \mathcal{M}$ is called a hierarchical set if:*

- *$\mathcal{A} \neq \emptyset$,*
- *$\sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$, and*
- *$-\alpha \in \mathcal{A}$ for each $\alpha \in \mathcal{A} \setminus \{v\}$.*

The remainder set $\mathcal{B}(\mathcal{A})$ for any hierarchical set is given by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{10}$$

The definition of multiple Itô integrals can be obtained in a recursive way, as follows:

Definition 4 (Multiple Itô Integrals [10, pp. 168, 169]) *Let us denote by $\mathcal{H}$ the set of all càdlàg processes $f = \{f(t) \geq 0\}$. Then let us define*

$$\mathcal{H}_v = \{f \in \mathcal{H} : \mathbb{P}[|f(t, \omega)| < \infty \quad \forall t \geq 0] = 1\}, \tag{11}$$

$$\mathcal{H}_{(0)} = \left\{f \in \mathcal{H} : \mathbb{P}\left[\int_0^t |f(t, \omega)| ds < \infty \quad \forall t \geq 0\right] = 1\right\}, \tag{12}$$

$$\mathcal{H}_{(1)} = \left\{f \in \mathcal{H} : \mathbb{P}\left[\int_0^t |f(t, \omega)|^2 ds < \infty \quad \forall t \geq 0\right] = 1\right\}. \tag{13}$$

Then for a multi-index $\alpha = (j_1, ..., j_l) \in \mathcal{M}$, let $f \in \mathcal{H}_v$ and $\alpha = v$ we define

$$I_\alpha f = f. \tag{14}$$

Assume $I_{\alpha-}$ and $\mathcal{H}_{(\alpha-)-}$ are defined. Then

$$I_\alpha[f(\cdot)]_{0,t} := \begin{cases} \int_0^t I_{\alpha-}[f(\cdot)]_{0,t} ds & , \quad l \geq 1, j_l = 0 \\ \int_0^t I_{\alpha-}[f(\cdot)]_{0,t} dW_s & , \quad l \geq 1, j_l = 1 \end{cases}, \tag{15}$$

where the set $\mathcal{H}_\alpha$ is defined as

$$\mathcal{H}_\alpha = \{f \in \mathcal{H} : I_{\alpha-}[f(\cdot)]_{0,t} \in \mathcal{H}_{(j_l)} \quad \forall t \geq 0\}. \tag{16}$$

Definition 5 (Itô Coefficient Functions [10, p. 177]) *Let us define the operators*

$$L^0 = \frac{\partial}{\partial t} + b(t, X_t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2}{\partial x^2} \tag{17}$$

$$L^1 = \sigma(t, X_t)\frac{\partial}{\partial x}. \tag{18}$$

Then for each multi-index $\alpha = (j_1, ..., j_l) \in \mathcal{M}$ and a function $f \in C^{l(\alpha)+n(\alpha)}([0, T] \times \mathbb{R}, \mathbb{R})$ we define the Itô coefficient function

$$f_\alpha = \begin{cases} f & , \quad l = 0 \\ L^{j_1} f_{-\alpha} & , \quad l \geq 1 \end{cases}. \tag{19}$$

Theorem 2 (Itô-Taylor expansion [10, Theorem 5.5.1]) *Let $t \in [t_0, T]$, for any $t_0 \geq 0$. Let $\mathcal{A} \subset \mathcal{M}$ be a hierarchical set, and let*

$$f : [0, T] \times \mathbb{R} \to \mathbb{R}, \tag{20}$$

then the Itô-Taylor expansion

$$f(t, X_t) = \sum_{\alpha \in \mathcal{A}} I_\alpha[f_\alpha(t_0, X_{t_0})]_{0,t} + R, \tag{21}$$

holds provided all the derivatives of f, b and σ, and all Itô multiple integrals in (21) exist, and

$$R = \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{0,t}. \tag{22}$$

R is called the reminder of the expansion.

The result above is the more general setting for the Itô-Taylor expansion, however, since the length of the multi-indices can go up to infinity, we have to leave the term R behind to approximate numerically the solutions to SDEs. This is done by selecting an appropriate set $\mathcal{A}$ which will define the number of terms that the scheme will have.

As an example, let us have $f(t, x) = x$ and $\mathcal{A} = \{v, (0), (1), (1, 1)\}$, then we will have the following coefficient functions according to (19). Let us also recall that in what follows, for any function $f(t, x)$ such that $x \mapsto f(t, x) \in C^2(\mathbb{R})$ we will write $f' = \frac{\partial f}{\partial x}$ and $f'' = \frac{\partial^2 f}{\partial x^2}$.

$$f_v = x,$$

$$f_{(0)} = L^0 f_v = \left[\frac{\partial f}{\partial t} + bf' + \frac{1}{2}\sigma^2 f''\right](t, X_t) = b(t, X_t),$$

$$f_{(1)} = L^1 f_v = [\sigma f'](t, X_t) = \sigma(t, X_t),$$

$$f_{(1,1)} = L^1(L^1 f_v) = \left[L^1(\sigma f')\right](t, X_t) = \sigma\sigma'(t, X_t)$$

and then the integrals

$$I_v[f_v(t_0, X_{t_0})]_{t_0,t} = X_{t_0}$$

$$I_{(0)}[f_{(0)}(t_0, X_{t_0})]_{t_0,t} = b(t_0, X_{t_0})\int_{t_0}^t ds$$

$$I_{(1)}[f_{(1)}(t_0, X_{t_0})]_{t_0,t} = \sigma(t_0, X_{t_0})\int_{t_0}^t dW_s$$

$$I_{(1,1)}[f_{(1,1)}(t_0, X_{t_0})]_{t_0,t} = \sigma\sigma'(t_0, X_{t_0})\int_{t_0}^t \int_{t_0}^s dW_r dW_s$$

3.2 Four numerical schemes

Given the Itô-Taylor expansion from Theorem 2 then it is necessary to prove that leaving the remainder R behind allows a scheme to produce an acceptable approximation to the real solution. The derivation of numerical methods such as Euler-Maruyama and Milstein schemes, can be found in [7, pp. 339–343], and here we present the definitions as in [10].

Definition 6 *Let* $[0, T]$ *be a time interval, and* $(t_n)_{n=0}^N$ *be a sequence of elements in* $[0, T]$ *such that* $0 = t_0 < ... < t_N = T$, *that is called a discretisation of the interval* $[0, T]$. *Let also denote* $\Delta t_n = t_{n+1} - t_n$ *and for a Brownian motion* $(W_t)_{t \geq 0}$, $\Delta W_n = W_{t_{n+1}} - W_{t_n}$. *We will denote by* $Y_n := Y_{t_n}$ *a discrete time approximation of* X_{t_n} *for any* n.

In the following definitions, for ease and clarity of notation we will write

$$b := b(t, X_t), \tag{23}$$
$$\sigma := \sigma(t, X_t), \tag{24}$$

since its dependency is clear from the context. Let us also note that in the numerical schemes we are using the function $f(t, x) = x$.

Definition 7 (Euler-Maruyama scheme [10, Section 10.2]) *Let* $(t_n)_{n=0}^N$ *be a discretisation of the interval* $[0, T]$. *Then an Euler-Maruyama approximation for the solution of* (5) *at time* t_{n+1} *is given by plugging the hierarchical set*

$$\mathcal{A} = \{v, (0), (1)\} \tag{25}$$

into (21), *i.e.*

$$Y_{n+1} = Y_n + b\Delta t_n + \sigma \Delta W_n. \tag{26}$$

Definition 8 (Milstein scheme [10, Section 10.3]) *Let* $(t_n)_{n=0}^N$ *be a discretisation of the interval* $[0, T]$. *Then a Milstein approximation for the solution of* (5) *at time* t_{n+1} *is given by plugging the hierarchical set*

$$\mathcal{A} = \{v, (0), (1), (1, 1)\} \tag{27}$$

into (21), *i.e.*

$$\hat{Y}_{n+1} = \hat{Y}_n + b\Delta t_n + \sigma \Delta W_n + \frac{1}{2}\sigma'\sigma((\Delta W_n)^2 - \Delta t_n). \tag{28}$$

Definition 9 (Order 1.5 strong Taylor scheme [10, Section 10.4]) *Let* $(t_n)_{n=0}^N$ *be a discretisation of the interval* $[0, T]$. *Then an order 1.5 strong approximation for the solution of* (5) *at time* t_{n+1} *is given by plugging the hierarchical set*

$$\mathcal{A} = \{v, (0), (1), (1, 1), (0, 1), (1, 0), (0, 0), (1, 1, 1)\} \tag{29}$$

into (21), *i.e.*

$$\begin{aligned}
\tilde{Y}_{n+1} = \tilde{Y}_n &+ b\Delta t_n + \sigma \Delta W_n + \frac{1}{2}\sigma'\sigma\left((\Delta W_n)^2 - \Delta t_n\right) \\
&+ b'\sigma \Delta Z_n + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2 \\
&+ \left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)(\Delta W_n \Delta t_n - \Delta Z_n) \\
&+ \frac{1}{2}\left(\sigma^2\sigma'' + \sigma(\sigma')^2\right)\left(\frac{1}{3}(\Delta W_n)^2 - \Delta t_n\right)\Delta W_n
\end{aligned} \tag{30}$$

where ΔZ_n *is given by*

$$\Delta Z_n = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} dW_{s_1}\, ds_2. \tag{31}$$

Definition 10 (Order 2.0 weak Taylor scheme [10, Section 14.2]) *Let $(t_n)_{n=0}^N$ be a discretisation of the interval $[0,T]$. Then an order 2.0 weak approximation for the solution of (5) at time t_{n+1} is given by plugging the hierarchical set*

$$\mathscr{A} = \{v, (0), (1), (1,1), (0,1), (1,0), (0,0)\} \tag{32}$$

into (21), i.e.

$$
\begin{aligned}
\check{Y}_{n+1} = \check{Y}_n &+ b\Delta t_n + \sigma\Delta W_n + \frac{1}{2}\sigma'\sigma\left((\Delta W_n)^2 - \Delta t_n\right) \\
&+ b'\sigma\Delta Z_n + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2 \\
&+ \left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)(\Delta W_n\Delta t_n - \Delta Z_n)
\end{aligned}
\tag{33}
$$

where ΔZ_n is given by

$$\Delta Z_n = \int_{t_n}^{t_{n+1}}\int_{t_n}^{s_2} dW_{s_1}\, ds_2 \tag{34}$$

When it comes to numerical computations, one can make use of the fact that for any $0 \leq s < t$ and $Z, \tilde{Z}$ i.i.d. as $\mathscr{N}(0,1)$ we can represent $W_t - W_s$ as $\sqrt{t-s}Z$ and $\int_s^t\int_s^z dW_u\, dz$ as $\frac{1}{2}(t-s)^{3/2}(Z + \frac{1}{\sqrt{3}}\tilde{Z})$ [10, pp. 351, 352]. Let us denote two random vectors by $Z = (Z_1, ..., Z_N)$ and $\tilde{Z} = (\tilde{Z}_1, ..., \tilde{Z}_N)$ such that for all $n \in \{1, ..., N\}$, $Z_n, \tilde{Z}_n \sim \mathscr{N}(0,1)$, then the n-th iteration of the schemes above are respectively

$$Y_{n+1} = Y_n + b\Delta t_n + \sigma\sqrt{\Delta t_n}Z_{n+1}, \tag{35}$$

$$\hat{Y}_{n+1} = \hat{Y}_n + b\Delta t_n + \sigma\sqrt{\Delta t_n}Z_{n+1} + \frac{1}{2}\sigma'\sigma\Delta t_n(Z_{n+1}^2 - 1), \tag{36}$$

$$
\begin{aligned}
\tilde{Y}_{n+1} = \tilde{Y}_n &+ b\Delta t_n + \sigma Z_{n+1}\sqrt{\Delta t_n} + \frac{1}{2}\sigma'\sigma\left(Z_{n+1}^2 - 1\right)\Delta t_n \\
&+ \frac{1}{2}b'\sigma\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2 \\
&+ \frac{1}{2}\left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)\left(Z_{n+1}\left(1 - \frac{1}{2}\Delta t_n\right) - \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\Delta t_n\right)\sqrt{\Delta t_n} \\
&+ \frac{1}{2}\left(\sigma^2\sigma'' + \sigma(\sigma')^2\right)\left(\frac{1}{3}Z_{n+1}^3 - Z_{n+1}\right)(\Delta t_n)^{3/2},
\end{aligned}
\tag{37}
$$

$$
\begin{aligned}
\check{Y}_{n+1} = \check{Y}_n &+ b\Delta t_n + \sigma Z_{n+1}\sqrt{\Delta t_n} + \frac{1}{2}\sigma'\sigma\left(Z_{n+1}^2 - 1\right)\Delta t_n \\
&+ \frac{1}{2}b'\sigma\left(Z_{n+1} + \frac{1}{\sqrt{3}}\check{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2 \\
&+ \frac{1}{2}\left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)\left(Z_{n+1}\left(1 - \frac{1}{2}\Delta t_n\right) - \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\Delta t_n\right)\sqrt{\Delta t_n}.
\end{aligned}
\tag{38}
$$

3.3 Modes of convergence

Now that reasonable numerical approximations to the solutions of SDEs are given, it is necessary to check whether said approximations will converge to the real solution of a given SDE. For this, we will discuss two forms of convergence that are usually of interest for numerical analysis of SDEs: strong and weak convergence.

First, strong convergence will tell us how close the sample paths of the approximations are to the sample paths of the solution. Naturally the numerical verification of this mode of convergence requires the simulations of all the paths we want to compare, both from the solution and the approximation.

Definition 11 (Strong convergence [12, Section 4]) *Let (X_t) be a strong solution of equation* (5). *We call $\gamma \in (0, \infty)$ a strong convergence order of the discrete time approximation (Y_n) if for any discretisation $(t_n)_{n=0}^N$ of $[0, T]$ there exists a constant $K < \infty$ such that*

$$\mathbb{E}|X_T - Y_N| \leq K \left(\max_n \Delta t_n \right)^{\gamma} \tag{39}$$

where (Y_n) is computed over the discretisation (t_n).

On the other hand, there are problems in which the condition for convergence can be relaxed because we might only be interested in the approximation of a function f applied on the process X_t. Examples of such functions f are the moments of the stochastic process that solves the equation considered. As mentioned by [12], for such situations it is possible to save computation time by defining another type of convergence in which we can compute only the approximations of the numerical solution and then compare it to a known value that has to be computed only once. For this purpose we define weak convergence.

Definition 12 (Weak convergence [12, Section 4]) *Let (X_t) be a strong solution of equation* (5). *We call $\beta \in (0, \infty)$ a weak convergence order of the discrete time approximation (Y_n) if for any T and f there is $M_f > 0$ such that for any discretisation $(t_n)_{n=0}^N$ of $[0, T]$ we have*

$$|\mathbb{E}[f(X_T)] - \mathbb{E}[f(Y_N)]| \leq M_f \left(\max_n \Delta t_n \right)^{\beta} \tag{40}$$

where (Y_n) is computed over the discretisation (t_n).

For the present work we will be using equidistant time increments, therefore the conditions for convergence stated above can be seen simply as

$$\mathbb{E}|X_T - Y_N| \leq K h^{\gamma} \qquad \text{and} \qquad |\mathbb{E}[f(X_T)] - \mathbb{E}[f(Y_N)]| \leq M_f h^{\beta}, \tag{41}$$

where $h = \Delta t_n$ for all $n = 0, ..., N$.

Once we have schemes as above to approximate numerically the solution of an SDE it is necessary to prove that said schemes will provide an appropriate output, meaning that one must know if those schemes converge (in either of the two senses that have been discussed earlier) to the real solution, as the discretisation is refined. Additionally it is also desirable to know which is the order of convergence. For that objective, [7] and [10] state the conditions on the coefficients and the proofs are found in the latter.

Theorem 3 (Strong convergence for E-M scheme [10, Theorem 10.2.2]) *Let the conditions from Theorem 1 hold for the coefficients b and σ, additionally assume that*

$$|b(s_0, x) - b(t_0, x)| + |\sigma(s_0, x) - \sigma(t_0, x)| \leq K(1 + |x|)\sqrt{|t_0 - s_0|}. \tag{42}$$

Then the Euler-Maruyama scheme has a strong order of convergence $\gamma = 1/2$.

In the following, the condition on (42) will be called *extended linear growth*.

Theorem 4 (Strong convergence for Milstein scheme [10, Theorem 10.3.5]) *Let the conditions from Theorem 1 and Theorem 3 hold for the coefficients b and σ. Recall that the operators L^0 and L^1 have been defined in* (17) *and* (18), *and let us consider the following notation*

$$z(t, x) = b(t, x) - \frac{1}{2}(\sigma \sigma')(t, x) \tag{43}$$

Suppose that the Lipschitz

$$\left| z(t, x_0) - z(t, y_0) \right| + \left| \sigma(t, x_0) - \sigma(t, y_0) \right| + \left| L^1 \sigma(t, x_0) - L^1 \sigma(t, y_0) \right| \leq K_1 |x_0 - y_0|, \tag{44}$$

linear growth

$$|z(t, x_0)| + |L^1 z(t, x_0)| + |\sigma(t, x_0)| + |L^1 \sigma(t, x_0)| + |L^1 L^1 \sigma(t, x_0)| \le K_2(1 + |x|), \tag{45}$$

and extended linear growth conditions

$$|z(s_0, x) - z(t_0, x)| + |\sigma(s_0, x) - \sigma(t_0, x)| \left| L^1 \sigma(s_0, x) - L^1 \sigma(t_0, x) \right| \le K_3(1 - |x|)|s_0 - t_0|^{1/2}, \tag{46}$$

hold for all $x, y, x_0, y_0 \in \mathbb{R}$, $t, s_0, t_0 \in [0, T]$, and constants $K_1, K_2, K_3 > 0$. Then the Milstein scheme has a strong rate of convergence $\gamma = 1$.

In general one can have a more robust strong convergence criterion statement that accounts for the strong approximations of all orders.

Theorem 5 (Strong convergence of all orders [10, Theorem 10.6.3]) *Let $N \in \mathbb{N}$, and $Y = (Y_n)_{n=0}^N$ be an approximation to the solution of the SDE (4) with order γ of strong convergence for some $\gamma = 0.5, 1.0, 1.5, 2.0, ...$, corresponding to an equidistant time discretisation $(t_n)_{n=0}^N$ with $\Delta t_n = h$. Suppose that the coefficient functions f_α satisfy Lipschitz continuity*

$$|f_\alpha(t, x) - f_\alpha(t, y)| \le K_1 |x - y|, \tag{47}$$

and linear growth

$$|f_\alpha(t, x)| \le K_2(1 + |x|), \tag{48}$$

hold for all $t \in [0, T]$ and $x, y \in \mathbb{R}$. Then

$$\mathbb{E}|X_T - Y_N|^2 \le K_3(1 + |X_0|^2)h^{2\gamma} + K_4 |X_0 - Y_0|^2, \tag{49}$$

holds for $K_1, K_2, K_3, K_4 < \infty$ constants independent of the time grid.

Theorem 6 (Weak convergence for all orders [10, Theorem 14.5.1]) *Let $N \in \mathbb{N}$, and $Y = (Y_n)_{n=0}^N$ be a approximation to the solution of the SDE (4) with order β of weak convergence for some $\beta = 1, 2, 3, ...$ corresponding to the time discretisation $(t_n)_{n=0}^N$ with $\Delta t_n = h$. Suppose that the coefficients of the equation are $b, \sigma \in C^{2(\beta+1)}(\mathbb{R})$ and Lipschitz continuous, and that f_α for $f(t, x) := x$ satisfy*

$$|f_\alpha(t, x)| \le K(1 + |x|), \tag{50}$$

holds for $K < \infty$. Then for each $g \in C^{2(\beta+1)}(\mathbb{R})$ there exists a constant C_g, which does not depend on the time discretisation, such that

$$|\mathbb{E}[g(X_T)] - \mathbb{E}[g(Y_N)]| \le C_g h^\beta. \tag{51}$$

Note that for strong convergence we present results for Euler scheme and Milstein scheme separately, plus the general statement for all orders, but for weak convergence only the general statement is mentioned. We do this because for weak convergence, Euler and Milstein schemes are both of convergence order 1 and in order to achieve weak convergence of a greater order, it is necessary to include more terms than those the Milstein scheme has.

Just as seen above, from the Euler scheme the following weak convergence order is 2.0. Said scheme includes two more terms than the Milstein scheme, but one less than the 1.5 strong scheme, so one should be aware that not all strong schemes will have a counterpart in the weak convergence side. Although for sure, even if the 1.5 scheme is not necessarily of interest in terms of weak convergence, just by having more terms from the Taylor expansion, it has to have at least order 1.0 of weak convergence, in fact it must have at least order 2.0 since it contains more elements from the Taylor approximation than the 2.0 weak scheme.

Once the numerical approximations are found, there are different methods to confirm convergence of numerical methods, in particular we decided to use Monte Carlo methods to compute the error of approximation, i.e: we produce a large number or simulations and then computing a deterministic operation over them. We denote by X_t^m and Y_t^m the m-th sample path of the solution and the numerical approximation respectively.

- For strong convergence to be confirmed numerically is necessary to compare a considerable number of paths from both the exact solution of the equation and the approximation with the strong scheme that has been used, and using the law of large numbers to approximate the mean. Then if M sample paths of each are computed and $m \in \{1,...,M\}$ then the approximation of the strong error will be then given by

$$E|X_T - Y_N| \approx \frac{1}{M} \sum_{m=1}^{M} |X_T^m - Y_N^m|, \tag{52}$$

which by the law of large numbers converges to $E|X_T - Y_N|$ as $m \to \infty$.

- Similarly, weak convergence is numerically verified using the following approximation for the weak error

$$|E[f(X_T)] - E[f(Y_N)]| \approx \left| E[f(X_T)] - \frac{1}{M} \sum_{m=1}^{M} f(Y_N^m) \right|, \tag{53}$$

where we already know $E[f(X_T)]$, therefore it is only necessary to compute the M sample paths from the numerical approximation and not from the real solution.

3.4 A numerical example: Geometric Brownian Motion

To verify the accuracy of the numerical schemes mentioned earlier, an SDE with a known explicit solution is used. In [10, pp. 117–126] a comprehensive list of such equations is given, and we will use the following particular case of a *geometric Brownian motion* (GBM)

$$\begin{cases} dX_t = \frac{1}{2} X_t dt + X_t dW_t, \\ X_0 = x_0 \end{cases} \tag{54}$$

whose solution is

$$X_t = x_0 \exp W_t. \tag{55}$$

Its mean is given by

$$\mathbb{E}[X_t] = x_0 \exp\left(\frac{1}{2}t\right). \tag{56}$$

For this SDE with the notation from Definition 7 and Definition 8, the Euler-Maruyama, Milstein, 1.5 strong and 2.0 weak approximations at time t_n are, respectively:

$$Y_{n+1} = Y_n + \frac{1}{2} Y_n \Delta t_n + Y_n \sqrt{\Delta t} Z_{n+1} \tag{57}$$

$$\hat{Y}_{n+1} = \hat{Y}_n + \frac{1}{2} \hat{Y}_n \Delta t_n + \hat{Y}_n \sqrt{\Delta t_n} Z_{n+1} + \frac{1}{2} \Delta t_n (Z_{n+1}^2 - 1) \tag{58}$$

$$\begin{aligned} \tilde{Y}_{n+1} = {} & \tilde{Y}_n + \frac{1}{2} \tilde{Y}_n \Delta t_n + \tilde{Y}_n \sqrt{\Delta t_n} Z_{n+1} + \frac{1}{2} \Delta t_n (Z_{n+1}^2 - 1) \\ & + \frac{1}{4} \tilde{Y}_n \left(Z_{n+1} + \frac{1}{\sqrt{3}} \tilde{Z}_{n+1} \right) (\Delta t_n)^{3/2} + \frac{1}{8} \tilde{Y}_n (\Delta t_n)^2 \\ & + \frac{1}{2} \tilde{Y}_n \left(\frac{1}{2} Z_n + \frac{1}{2\sqrt{3}} \tilde{Z}_{n+1} \right) (\Delta t_n)^{3/2} + \frac{1}{2} \left(\frac{1}{3} Z_{n+1}^2 - 1 \right) Z_{n+1} (\Delta t_n)^{3/2} \end{aligned} \tag{59}$$

$$\begin{aligned} \check{Y}_{n+1} = {} & \check{Y}_n + \frac{1}{2} \check{Y}_n \Delta t_n + \check{Y}_n \sqrt{\Delta t_n} Z_{n+1} + \frac{1}{2} \Delta t_n (Z_{n+1}^2 - 1) \\ & + \frac{1}{4} \check{Y}_n \left(Z_{n+1} + \frac{1}{\sqrt{3}} \tilde{Z}_{n+1} \right) (\Delta t_n)^{3/2} + \frac{1}{8} \check{Y}_n (\Delta t_n)^2 \\ & + \frac{1}{2} \check{Y}_n \left(\frac{1}{2} Z_n + \frac{1}{2\sqrt{3}} \tilde{Z}_{n+1} \right) (\Delta t_n)^{3/2} \end{aligned} \tag{60}$$

We ran batches of approximations for the gBm and its approximations with the numerical schemes mentioned above, each with step sizes $\Delta t \in \{2^{-3}, ..., 2^{-7}\}$, 10^5 sample paths on each batch. In Fig. 1 we can see how increasing the amount of points in the approximation to the solution reduces the error. Also one might notice that strong weak error behave in different ways, this is because strong error is computed comparing paths of the real solution with paths of the approximation, making the error smaller in each step, thus, probably we do not need as many sample paths to get a well behaved error. Meanwhile for weak error we compare approximations of the solution against a fixed quantity $\mathbb{E}[f(X_t)]$, in this case when we are using Monte Carlo methods to compute the error the amount of paths that we compare is important, and the more we use the better behaved will be the error.
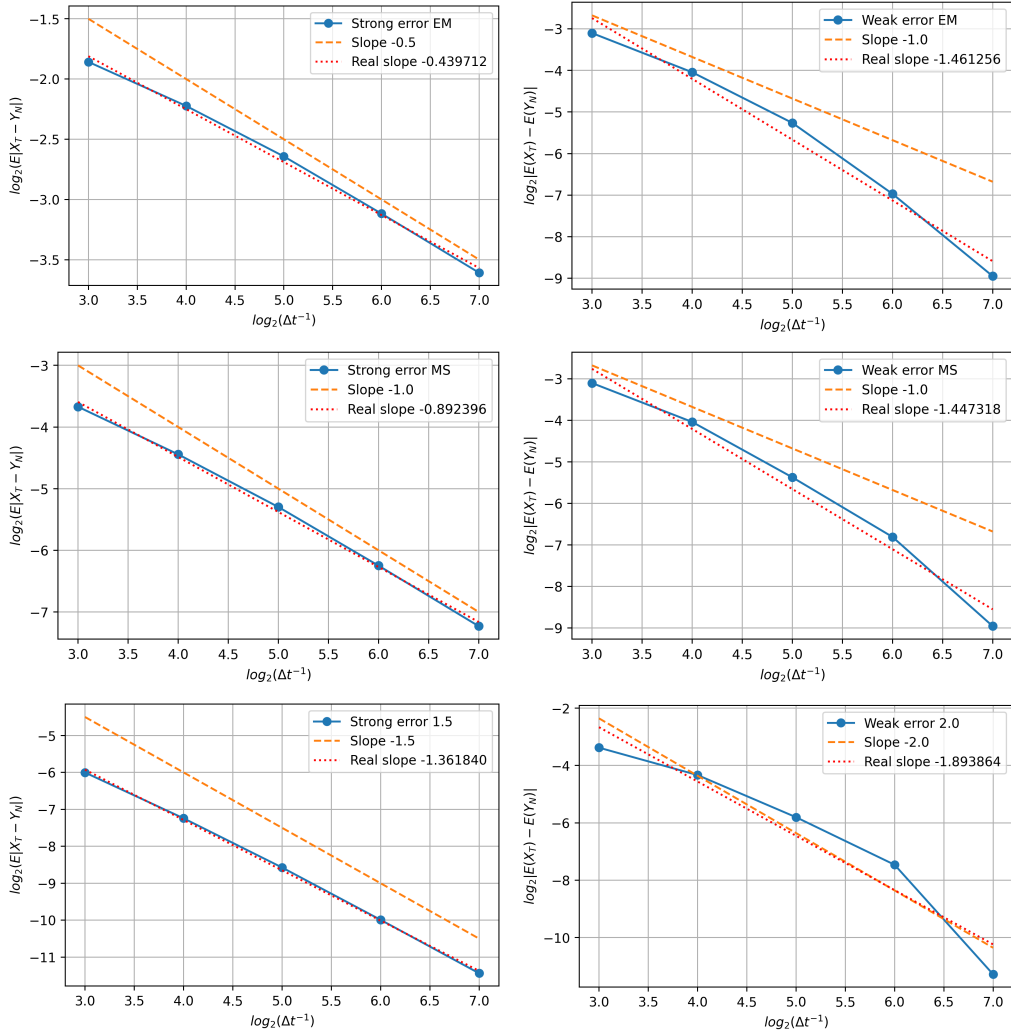

Figure 1: Log-log graphs of the error of approximation. Left column from top to bottom: Strong error for Euler scheme, strong error for Milstein scheme and strong error for 1.5 strong scheme. Right column from top to bottom: Weak error for Euler scheme, weak error for Milstein scheme and weak error for 2.0 weak scheme. The blue line with markers represents the real values of the errors, the orange dashed line is the theoretical slope that the error for each scheme should have in each context and the red dotted line is the real slope of the error computed via linear regression.

4 Sobolev spaces and rough coefficients

The approximation methods for SDEs mentioned above can lead us to good results if the coefficients behave as stated in Theorems 3 to 6. However we are interested in some particular kind of drift coefficients that do not satisfy the necessary conditions for the schemes defined above, hence we need to find a way to approximate those coefficients so that we can use them numerically.

4.1 Fractional Sobolev spaces

Here we will define the spaces to which the drift of the equations we want to study lives. Let us introduce some definitions and notation.

Definition 13 (Fourier transform) *Let $f : \mathbb{R} \to \mathbb{C}$ be in $L^1(\mathbb{R})$. Then the Fourier transform of f is defined as*

$$\hat{f}(\xi) = \mathscr{F}[f](\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x} f(x) dx. \tag{61}$$

Definition 14 (Weak derivative) *Suppose $u, v \in L^1(\mathbb{R})$, and $\alpha \in \mathbb{N}$, we say that $u^{(\alpha)} = v$ is the α-th weak derivative of u, provided*

$$\int_{\mathbb{R}} u \phi^{(\alpha)} dx = (-1)^\alpha \int_{\mathbb{R}} v \phi dx \tag{62}$$

for all test functions $\phi \in C_c^\infty(\mathbb{R})$.

Definition 15 (Schwartz functions) *We denote the space of Schwartz functions as*

$$\mathcal{S}(\mathbb{R}) = \{ f \in C^\infty : \sup_{x \in \mathbb{R}} |x^\alpha f^{(\beta)}(x)| < \infty, \forall \alpha, \beta \in \mathbb{N} \}, \tag{63}$$

where $f^{(\beta)} = \frac{d^\beta f}{dx^\beta}$.

As the definition points out, Schwartz functions are those whose derivatives are rapidly decreasing. This means that any derivative decays faster than any power of x, and thus the product of those is bounded.

Definition 16 (Tempered distributions) *A tempered distribution on $\mathbb{R}$ is a linear mapping $\phi \to (f, \phi)$ from $\mathcal{S}(\mathbb{R})$ to $\mathbb{C}$. Therefore the set of all tempered distributions is the dual of the space of Schwartz functions, and is denoted by $\mathcal{S}'(\mathbb{R})$.*

Definition 17 *[13, Remark 1.2] Let $f \in \mathcal{S}'(\mathbb{R})$ for $0 < p < \infty$ and let $\xi, s \in \mathbb{R}$. We define the map $I^s : \mathcal{S}'(\mathbb{R}) \to \mathcal{S}'(\mathbb{R})$ as*

$$I^s : f \mapsto \mathscr{F}^{-1} \left[\left(1 + |\xi|^2 \right)^{\frac{s}{2}} \mathscr{F}[f] \right]. \tag{64}$$

Said map is a bijection in $\mathcal{S}'(\mathbb{R})$.

The definition of fractional Sobolev spaces can be found in different formats by different authors, we have chosen the one in [13, Remark 1.2].

Definition 18 (Fractional Sobolev spaces) *Let $s \in \mathbb{R}$, $0 < p < \infty$, and I^s as in (64) we call the space*

$$H_p^s := I^{-s} L^p(\mathbb{R}) \tag{65}$$

a fractional Sobolev space. Further, this is a Banach space if it is equipped with the norm

$$\|f\|_{H_p^s} := \|I^{-s} f\|_{L^p(\mathbb{R})}. \tag{66}$$

We must add that in the case that $s = k \in \mathbb{N}$ we have the special case of classical Sobolev spaces, denoted usually by $W^{k,p}(\mathbb{R})$ [14, Section 2.3.8]. Also notice that if $s < 0$, the space $H_p^s(\mathbb{R})$ will contain distributions and if $s = 0$, $H_p^s(\mathbb{R}) = L^p(\mathbb{R})$.

4.2 Bases in Sobolev spaces

We are interested in the study of a coefficient b which belongs to an appropriate fractional Sobolev space of negative order, acting as the drift for an SDE such as (4). We want to know how to compute the distributional coefficient numerically for the simulations, however as the coefficient is a distribution, its pointwise evaluation does not make sense and we need some way to approximate it numerically. One way to achieve this is to find bases for those spaces, that are comprised of measurable functions.

Let us denote the interval $I = (0,1)$, we can define as in [5]

$$H_p^s(I) := \{f \in (C_c^\infty)' : f = h|_I \text{ for some } h \in H_p^s(\mathbb{R})\}, \tag{67}$$

with the norm

$$\|f\|_{H_p^s(I)} = \inf\{\|h\|_{H_p^s}(\mathbb{R}) : h \in H_p^s(\mathbb{R}) \text{ with } f = h|_I\}. \tag{68}$$

Definition 19 (Haar wavelets on $I = (0,1)$) *The Haar wavelet system on $I = (0,1)$ is given by*

$$\{h_0, h_{j,m} : j \in \mathbb{N}, m = 0,...,2^j - 1\} \tag{69}$$

where

$$h_0(x) := \mathbb{1}_I(x) \tag{70}$$

and

$$h_{j,m}(x) = \mathbb{1}_{\left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right)} - \mathbb{1}_{\left[\frac{m+1/2}{2^j}, \frac{m+1}{2^j}\right)}. \tag{71}$$

Haar systems are bases for the fractional Sobolev spaces to which the drift coefficient lives and are particularly useful for numerical schemes because they are piecewise constants and therefore its values can be stored in computer without any loss of information. The following theorem tells us how to represent elements in said space as infinite sums of Haar functions.

Theorem 7 ([13, Theorem 2.13]) *Let $0 \le p < \infty$, $-\frac{1}{2} < s < \frac{1}{p}$, and let $f \in (C_c^\infty)'(I)$. Then $f \in H_p^s(I)$ if and only if*

$$f = \sum_{j=0}^\infty \sum_{m=0}^{2^j - 1} \mu_{j,m} 2^{-j\left(s - \frac{1}{r}\right)} h_{j,m}, \tag{72}$$

with unconditional convergence in any space $H_p^\sigma(I)$ with $\sigma < s$. The representation is unique with the coefficients given by

$$\mu_0 := \int_0^1 f(x) h_0(x) dx \tag{73}$$

and for $j \in \mathbb{N}$ and $m = 0,...,2^j - 1$ by

$$\mu_{j,m} := 2^{j\left(s - \frac{1}{p} + 1\right)} \int_0^1 f(x) h_{j,m}(x) dx, \tag{74}$$

where the integral above is in the sense of dual pairing. Moreover the system

$$\left\{h_0, 2^{-j\left(s - \frac{1}{p}\right)} h_{j,m} : j \in \mathbb{N}, m = 0,...,2^j - 1\right\} \tag{75}$$

is an unconditional normalised basis of $H_p^s(I)$.

There is one more difficulty with that theorem, and is that the coefficients $\mu_{j,m}$ are defined in the form of an integral in the sense of dual pairing, which even though it is a well defined mathematical object, it is not possible to make a proper numerical computation of it. To overcome this problem let us introduce Faber systems, which are basis for the spaces $H_p^s(I)$, for $2 \ge r < \infty$ and $1/2 < s < 1 + 1/p$. It is important to note that Faber series representation will only hold on bounded intervals, which for numerical schemes is perfectly acceptable.

Definition 20 (Faber wavelets in $I = (0,1)$) *The Faber system in I is given by*

$$\{v_0, v_1, v_{j,m} : j \in \mathbb{N}, m = 0, ..., 2^j - 1\} \tag{76}$$

where

$$v_0(x) := \begin{cases} 1 - x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad v_1(x) := \begin{cases} x & 0 \le x \le 1 \\ 0 & \text{otherwise,} \end{cases} \tag{77}$$

and the hat functions are defined by

$$v_{j,m} := 2^{j+1} \int_0^x h_{j,m}(y)dy = \begin{cases} 2^{j+1}(x - 2^{-j}m) & x \in \left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right) \\ 2^{j+1}(2^{-j}(m-1) - x) & x \in \left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right) \\ 0 & \text{otherwise.} \end{cases} \tag{78}$$

The following theorem establishes the series representation of elements in the appropriate space with the Faber system.

Theorem 8 ([13, Theorem 3.1]) *Let $g \in H_r^s(I)$ for $2 \le r < \infty$ and $1/2 < s < 1 + 1/r$. Then we have the unique Faber representation for g*

$$g = \hat{\mu}_0 v_0 + \hat{\mu}_1 v_1 + \sum_{j=0}^{+\infty} \sum_{m=0}^{2^j-1} \hat{\mu}_{j,m} v_{j,m} \tag{79}$$

with unconditional convergence in $C(I)$ and in $H_r^\sigma(I)$ with $\sigma < s$. Where the coefficients $\hat{\mu}$ are given by

$$\begin{cases} \hat{\mu}_{j,m} & = -\frac{1}{2}(\Delta^2_{2^{-j-1}}g)(2^{-j}m) \\ \hat{\mu}_0 & = g(0) \\ \hat{\mu}_1 & = g(1), \end{cases} \tag{80}$$

and where $(\Delta_h^2 g)(x) := g(x + 2h) - 2g(x + h) + g(x)$.

Notice how in Definition 20 the Faber functions are defined as the normalised integrals of Haar functions, and also from Theorem 8 we have that the Faber functions are basis in a space $H_p^s(I)$ for $1/2 < s < 1 + 1/p$ and $2 \le p < \infty$ meanwhile Haar functions are basis for spaces $H_p^s(I)$ for $-1/2 < s < 1/p$ and $0 \le p < \infty$. It is true that

$$f \in H_p^{s+1} \implies f' \in H_p^s, \tag{81}$$

whenever the derivative exists. We also have that

$$v'_{j,m} = 2^{j+1} h_{j,m}, \tag{82}$$

whereby, we can think that having f expanded in Faber series we can be formally differentiate term by term to get the Haar representation of f'.

Theorem 9 ([13, Theorem 3.1]) *Let $g' \in H_r^s(I)$, then $g' \in H_r^{s-1}(I)$ with $2 \le r < \infty$, $1/2 < s < 1 + 1/r$ which can be written as*

$$g' = (\hat{\mu}_1 - \hat{\mu}_0)h_0 + \sum_{j=0}^{+\infty} \sum_{m=0}^{2^j-1} 2^{j+1} \hat{\mu}_{j,m} h_{j,m}, \tag{83}$$

where

$$\mu_0 = \hat{\mu}_1 - \hat{\mu}_0 \quad \text{and} \quad \mu_{j,m} = 2^{j+1} \hat{\mu}_{j,m}. \tag{84}$$

5 Numerical schemes for SDEs with distributional drift

5.1 Theoretical results

The foundations of this research project are in [5], in which it is devised an algorithm to work with drifts that belong to an appropriate fractional Sobolev space. Below the main results of the paper are mentioned, and for this the following notation and assumptions are in order.

Let us have the SDE

$$\begin{cases} dX_t = b(t, X_t)dt + dW_t, \\ X_0 = x_0, \quad t \in [0, T], \end{cases} \tag{85}$$

where $b \in H_p^{-s}(\mathbb{R})$.

Definition 21 *Let B be a Banach space, then let us denote by $C([0, T]; B)$ the spaces of B-valued functions of time, this is a Banach space with the norm*

$$\|f\|_{\infty, B} = \sup_{0 \geq t \geq T} \|f(t)\|_B. \tag{86}$$

Let $\kappa \in (0, 1)$, the space $C^\kappa([0, T]; B)$ will denote the set of functions in $C([0, T]; B)$ that are κ-Hölder continuous in time i.e:

$$[f]_{\kappa, B} := \sup_{s \neq t \in [0, T]} \frac{\|f(t) - f(s)\|_B}{|t - s|^\kappa} < \infty. \tag{87}$$

This is a Banach space when endowed with the norm

$$\|f\|_{\kappa, B} = \|f\|_{\infty, B} + [f]_{\kappa, B}. \tag{88}$$

In our case, we will select $B = H_p^s$, therefore a map $g(t, \cdot) \in C^\alpha([0, T]; B)$ it is potentially a distribution depending in the sign of s. Let us denote $H_{p,q}^s = H_p^s \cap H_q^s$. Also note that if $f \in H_{p,q}^s(\mathbb{R})$ for $p \wedge q < r < p \vee q$, it will also be true that $f \in H_r^s(\mathbb{R})$. Now consider the equation (4), the aim is to first make an approximation of the coefficient b with a better behaved coefficient b^N, for this purpose let the following assumptions for b and b^N hold:

Assumption 1 ([5]) *Let $\beta_0 \in \left(0, \frac{1}{4}\right)$ and $q_0 \in \left(4, \frac{1}{\beta_0}\right)$, fix $\tilde{q}_0 := (1 - \beta_0)^{-1}$. Then for some $\kappa \in \left(\frac{1}{2}, 1\right)$* ==take== $b \in C^\kappa\left([0, T]; H_{\tilde{q}_0, q_0}^{-\beta_0}\right)$.

Assumption 2 ([5]) *Let $(b^N)_{N \geq 1} \subset C^{\frac{1}{2}}\left([0, T]; H_{\tilde{q}_0, q_0}^0\right)$ be such that*

$$\lim_{N \to \infty} b^N = b \qquad in \qquad C^{\frac{1}{2}}\left([0, T]; H_{\tilde{q}_0, q_0}^{-\beta_0}\right). \tag{89}$$

The assumptions above allow us to select the elements of the sequence b^N living in a space of coefficients that is not as rough as the space in which the real coefficient lives, but will still belong to the fractional negative Sobolev space because of the inclusion property $H_p^s \in H_p^r$ if $s > r$. Because of that inclusion property, the convergence of $\lim_{N \to \infty} b^N$ makes sense with the norm defined for $C^{\frac{1}{2}}([0, T]; H_{\tilde{q}_0, q_0}^{-\beta_0})$.

Definition 22 (Heat kernel and heat semigroup) *The heat kernel is a function $\phi(t, x) : [0, T] \times \mathbb{R} \to \mathbb{R}$ defined as*

$$\phi(t, x) := \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right). \tag{90}$$

Let $f \in \mathcal{S}'(\mathbb{R})$ then we define the heat semigroup $P_t f$ via convolution of the heat kernel with f,

$$[P_t f](x) = \int_{-\infty}^{\infty} \phi(t, x - y) f(y) dy. \tag{91}$$

The heat kernel as introduced here is the fundamental solution of the heat equation

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) - \frac{\partial^2 u}{\partial x^2}(t,x) = 0 & (t,x) \in [0,T] \times \mathbb{R} \\ u(0,x) = g(x), \end{cases} \tag{92}$$

and the general solution would be $u(t,x) = [P_t g](x)$. Note that the heat kernel is the probability distribution function of a random variable distributed as $\mathcal{N}(0,t)$.

Let us now remind some properties of the heat semigroup, mentioned in [5, Remark 3.2]:

$$\|P_t f\|_{\infty, H_q^{-s}} \leq \|f\|_{\infty, H_q^{-s}} \tag{93}$$

$$\|P_t f - f\|_{\infty, H_q^{-s-\epsilon}} \leq c t^{\epsilon/2} \|f\|_{\infty, H_q^{-s}}, \tag{94}$$

The distributional coefficient b is going to be approximated by a coefficient b^N defined as the sum of Haar functions. When the heat kernel is applied to Haar functions it has the form

$$
\begin{aligned}
[P_{\eta_N} h_{j,m}](x) &= \left[P_{\eta_N} \mathbb{1}_{\left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right)} \right](x) - \left[P_{\eta_N} \mathbb{1}_{\left[\frac{m+1/2}{2^j}, \frac{m+1}{2^j}\right)} \right](x) \\
&= \exp(-\eta_N)\left(-\Phi\left(\frac{x - \frac{m+1}{2^j}}{\sqrt{\eta_N}}\right) + 2\Phi\left(\frac{x - \frac{m+1/2}{2^j}}{\sqrt{\eta_N}}\right) - \Phi\left(\frac{x - \frac{m}{2^j}}{\sqrt{\eta_N}}\right) \right),
\end{aligned}
\tag{95}
$$

where $j = 1, 2, ..., N$, $m = 1, 2, ..., 2^j$, N is fixed, η_N is a constant that depends on N, and Φ is the cumulative distribution function of a standard normal random variable.

Let us note that the heat kernel applied to the approximation b_η^N will be denoted by

$$b_\eta^N := P_{\eta_N} b^N. \tag{96}$$

Theorem 10 ([5, Proposition 3.1]) *Let Assumption 1 and 2 hold. Take any (β, q) such that $\beta \in (\beta_0, \frac{1}{2})$ and $q_0 \geq q > \tilde{q} \geq \tilde{q}_0$, where $\tilde{q} := (1-\beta)^{-1}$. Then for any $\frac{1}{2} < \gamma < \gamma_0$ there is a constant $C_\gamma > 0$ such that*

$$\sup_{0 \leq t \leq T} \mathbb{E}|X_t^N - X_t| \leq C_\gamma \|b_\eta^N - b\|_{\infty, H_q^{-\beta}}^{2\gamma-1}. \tag{97}$$

The previous theorem shows the rate of convergence of the approximation X^N to the solution X by finding a bound for the error, said bound is related to the rate of convergence of the approximation for the distributional drift itself.

Since the convergence of the approximation to the solution is dependant on the convergence of the approximation to the distributional drift, it is necessary to see that the approximation of the drift also has a sensible bound. Using properties Eqs. (93) and (94) and setting $\beta = \beta_0 + \epsilon$ we can get a bound for the right hand term in Eq. (97):

$$
\begin{aligned}
\|b_\eta^N - b\|_{\infty, H_q^{-\beta}} &= \|b_\eta^N - P_\eta b + P_\eta b - b\|_{\infty, H_q^{-\beta}} \\
&\leq \|P_\eta(b^N - b)\|_{\infty, H_q^{-\beta}} - \|P_\eta b - b\|_{\infty, H_q^{-\beta}} \\
&\leq \|b^N - b\|_{\infty, H_q^{-\beta}} + c\eta^{\frac{\beta-\beta_0}{2}} \|b\|_{\infty, H_{q_0}^{-\beta_0}}.
\end{aligned}
$$

Then as $\eta \to \infty$, as $\beta > \beta_0$ the second term in the inequality goes to zero and then it is necessary to find a bound for the first term, as is made in the following theorem.

Theorem 11 ([5, Proposition 3.3]) *Let Assumption 1 hold and let the sequence $(b^N)_{N \geq 1}$ be defined as*

$$b^N(t) := \sum_{j=-1}^{N} \sum_{m=-2^j}^{2^j} \mu_{j,m}(t) 2^{-j(-\beta-\frac{1}{q})} h_{j,m}, \tag{98}$$

which by construction $b^N(t) \in H^0_{\tilde{q}_0, q_0} \subset H^{-\beta_0}_{\tilde{q}_0, q_0}$. Then $(b^N)_{N \geq 1}$ satisfies assumption Assumption 2 and for any $\beta \in (\beta_0, \frac{1}{2})$ it holds that

$$\|b^N - b\|_{\infty, H_2^{-\beta}} \leq c 2^{-(N+1)(\beta - \beta_0)} \|b\|_{\infty, H_2^{-\beta_0}}. \tag{99}$$

We mentioned that since Haar functions $h_{j,m}$ are the derivatives, up to a constant, of Faber functions $v_{j,m}$, it is natural to think that we can differentiate elements in a space $H_p^{s+1}(I)$ and get elements in $H_p^s(I)$. Let us define

$$b(t, x) := \frac{\partial g}{\partial x}(t, x) \tag{100}$$

for $g \in H_p^s(I)$. Now let us introduce a definition to get a particular function g.

Definition 23 (Fractional Brownian motion) *A Fractional Brownian Motion (FBM) is a Gaussian process $\{\hat{B}^H(x), x \in \mathbb{R}\}$ with covariance given by*

$$\mathbb{E}[\hat{B}^H(x)\hat{B}^H(y)] = \frac{1}{2}(x^{2H} + y^{2H} + |x - y|^{2H}). \tag{101}$$

Let us denote $B^H := \psi \hat{B}^H$, for some $\psi \in C_c^\infty$ supported in the interval I, it can be proven that $\hat{B}^H \in H^{-\beta_0+1}_{q_0}(\mathbb{R})$, $\mathbb{P}$-a.s. [8, Proposition 4.1]. Nonetheless, Faber functions are only basis of Sobolev spaces $H^{\beta_0+1}_{q_0}(I)$ where I is a bounded interval, but as $supp(B^H(\omega)) \subset I$, then $B^H(\omega) \in H^{-\beta_0+1}_{q_0}(I)$, $\mathbb{P}$-a.s.

FBM is particularly useful because is easy to simulate numerically, we can see a procedure for it in [5, Appendix B]. We define g in Eq. (100) to be a single path of the FBM B^H, and thanks to Theorem 9 we can compute the coefficients for the series expansion of b as

$$\begin{cases} \mu_0 &= B^H(1) - B^H(0) \\ \mu_{j,m} &= -2^j \left(B^H \left(\frac{m+1}{2^j} \right) - 2B^H \left(\frac{m+1/2}{2^j} \right) + B^H \left(\frac{m}{2^j} \right) \right). \end{cases} \tag{102}$$

In order to abide by the definition of Haar and Faber wavelets in the interval $I = (0,1)$, the coefficients μ_0 and $\mu_{j,m}$ for the series expansion will be computed using a FBM defined over the same interval. However, the coefficient $b(t,x)$ has to be applied to a stochastic process X, therefore if we restrict the coefficient to $x \in [0,1]$ we will not be able to compute it over a large portion of the domain of X, thus we need to rescale the coefficient to apply it over a large interval.

For the sake of the numerical implementations, since we cannot compute anything over the whole real line, let us say that we need to compute it over an interval $[-K, K]$ for a large K, which essentially will be the same interval as $[0, 2K]$ after a switch which can be performed thanks to the self-similarity property of FBM. So let us define FBMs $\tilde{B}^H$ and B^H over the interval $[0, 2K]$ and $[0,1]$ respectively, then we have

$$\tilde{B}^H(x) = (2K)^H B^H \left(\frac{x}{2K} \right), \tag{103}$$

and then

$$\frac{\partial}{\partial x} \tilde{B}^H(x) = (2K)^H \frac{\partial}{\partial x} B^H \left(\frac{x}{2K} \right) \frac{1}{2K}$$

$$= (2K)^{H-1} \frac{\partial}{\partial x} B^H \left(\frac{x}{2K} \right).$$

We then need to apply the heat kernel to the derivative above in order to obtain our approximation to the real distributional coefficient, which is given by

$$P_\eta \left[\frac{\partial}{\partial x} \tilde{B}^H \right](x) = \frac{1}{\sqrt{2\pi\eta}} \int_{\mathbb{R}} \exp \left(-\frac{(\xi - x)^2}{2\eta} \right) \frac{\partial}{\partial x} \tilde{B}^H(\xi) d\xi.$$

Using the change of variables $y = \xi/2K$, $dy = d\xi/2K$ we can change $\tilde{B}^H$ by B^H according to (103) and have

$$P_\eta\left[\frac{\partial}{\partial x}\tilde{B}^H\right](x) = \frac{1}{\sqrt{2\pi\eta}}\int_{\mathbb{R}}\exp\left(-\frac{(2K)^2(y - x/2K)^2}{2\eta}\right)(2K)^{H-1}\frac{\partial}{\partial x}B^H(y)(2K)dy$$

$$= \frac{1}{\sqrt{2\pi\eta/(2K)^2}}\int_{\mathbb{R}}\exp\left(-\frac{(y - x/2K)^2}{2\eta/(2K)^2}\right)(2K)^{H-1}\frac{\partial}{\partial x}B^H(y)(2K)dy$$

$$= (2K)^{H-1}P_{\frac{\eta}{(2K)^2}}\left[\frac{\partial}{\partial x}B^H\right]\left(\frac{x}{2K}\right).$$

From there, we can compute the approximation of the coefficient on $[0, 2K]$, but then for a FBM $\check{B}^H$ over $[-K, K]$ we compute

$$P_\eta\left[\frac{\partial}{\partial x}\check{B}^H\right](x) = (2K)^{H-1}P_{\frac{\eta}{(2K)^2}}\left[\frac{\partial}{\partial x}B^H\right]\left(\frac{x+K}{2K}\right). \tag{104}$$

Hence in the numerical schemes we end up using this form above given that we compute the coefficients μ_0 and $\mu_{j,m}$ with a FBM over $I = (0,1)$.

The core of the present problem is the approximation of the distributional coefficient in a way that is manageable numerically. Note that it is first found an approximation to the coefficient b, which we call b_η^N and with that coefficient we have a new SDE with a time homogeneous drift

$$\begin{cases} dX_t^N = b_\eta^N(X_t^N)dt + dW_t, & t \in [0, T] \\ X_0^N = x_0. \end{cases} \tag{105}$$

We can find a numerical solution $X_t^{N,m}$ using an Euler scheme with m time steps, for the system above and the following theorem tells us that the numerical solution $X_t^{N,m}$ converges to the solution X_t^N.

Theorem 12 ([5, Proposition 3.4]) *Let Assumption 1 hold and let $b^N \in C^{\frac{1}{2}}\left([0, T]; H_{\bar{q}_0, q_0}^0\right)$ for fixed N. Then as $m \to \infty$*

$$\sup_{0 \leq t \leq T}\mathbb{E}\left|X_t^{N,m} - X_t^N\right| \leq C_2(N)m^{-1} + C_3(N)m^{-\frac{1}{2}} \tag{106}$$

where

$$C_2 := c\|b_\eta^N\|_{\infty, L^\infty}\left(1 + \|\nabla b_\eta^N\|_{\infty, L^\infty}\right), \tag{107}$$

$$C_3 := c'\left(\|\nabla b_\eta^N\|_{\infty, L^\infty} + \left[b_\eta^N\right]_{\frac{1}{2}, L^\infty}\right), \tag{108}$$

and $c, c' > 0$ are constants independent of (N, m).

The following result essentially is a combination of Theorem 10 and 12.

Theorem 13 *Let Assumption 1 and also b^N defined as in (98) so that Assumption 2 holds too, and let $\Theta_* := \frac{1}{2}\left[\frac{3}{4} - \beta_0\left(\gamma_0 - \frac{1}{2}\right)\right]^{-1}$. Then as $m \to \infty$, let $\eta_N = m^{-\Theta_*}$ and $N = 2\Theta_* \log_2 m$ it holds that*

$$\sup_{0 \leq t \leq T}\mathbb{E}\left|X_t^{N,m} - X_t\right| \leq c_\epsilon\left(m^{-\Theta_*\left(\frac{1}{2} - \beta_0\right)\left(\gamma_0 - \frac{1}{2}\right) - \epsilon}\right) \tag{109}$$

where $\epsilon > 0$ is arbitrarily small and $c_\epsilon > 0$ is a constant depending on ϵ.

If we recall the definition of strong convergence established before, we can see that the authors in [5] actually proved a stronger statement, because the result is strong convergence not only for the terminal time T, but for all times $t \in [0, T]$.

5.2 Numerical implementations

In [5] it is proven that the Euler scheme modified to work with the approximation of the distributional coefficient has a certain order of convergence, which is not the same as in an Euler scheme, this is because in the Euler scheme (or any other scheme that we discussed earlier), the only parameter that changes is size of the time grid that we use, meanwhile here we have that for size of the time grid, the parameter η of the semigroup will vary, so there are two parameters changing, assuming that N for the series representation of the coefficient is fixed.

We have explored, only numerically, the possibilities to improve the Euler scheme by adding terms from the Taylor expansion, and also looked at weak convergence of the schemes mentioned. It is clear that Milstein scheme is not useful for an equation such as (85), since we require of the derivative of the diffusion coefficient, but as said coefficient is equals one the derivative vanishes. That is why in Definition 9 we introduced a different numerical scheme that is useful for this particular problem, some of the terms will vanish again as they are multiplying the derivative of the diffusion term, and also we need the derivative and second derivative of the drift coefficient.

The 1.5 strong scheme for this problem reads

$$
\begin{aligned}
\tilde{Y}_{n+1} = {}& \tilde{Y}_n + b(t_n, \tilde{Y}_n)\Delta t_n + Z_{n+1}\sqrt{\Delta t_n} \\
&+ \frac{1}{2}(b')(t_n, \tilde{Y}_{t_n})\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} \\
&+ \frac{1}{2}\left((bb')(t_n, \tilde{Y}_{t_n}) + \frac{1}{2}(b'')(t_n, \tilde{Y}_{t_n})\right)(\Delta t_n)^2 \\
&+ \frac{1}{4}\left(Z_{n+1} - \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2}
\end{aligned}
\tag{110}
$$

Of course in numerical implementations we cannot use the distributional coefficient, as we discussed earlier and we have the approximation of the drift as

$$
b_\eta^N(x) = \mu_0 P_\eta h_0(x) + \sum_{j=0}^{N}\sum_{m=1}^{2^j-1} \mu_{j,m} P_\eta h_{j,m}(x).
\tag{111}
$$

The equation above with the heat semigroup applied to the Haar functions reads

$$
\begin{aligned}
b_\eta^N(x) = {}& \mu_0 \frac{1}{\sqrt{2\pi\eta_N}} \int_{-\infty}^{\infty} h_0(y) e^{-\frac{(x-y)^2}{2\eta_N}}\, dy \\
&+ \sum_{j=0}^{N}\sum_{m=1}^{2^j-1} \mu_{j,m} \frac{1}{\sqrt{2\pi\eta_N}} \int_{-\infty}^{\infty} h_{j,m}(y) e^{-\frac{(x-y)^2}{2\eta_N}}\, dy \\
= {}& \mu_0 \frac{1}{\sqrt{2\pi\eta_N}} \int_{0}^{1} e^{-\frac{(x-y)^2}{2\eta_N}}\, dy \\
&+ \sum_{j=0}^{N}\sum_{m=1}^{2^j-1} \mu_{j,m} \frac{1}{\sqrt{2\pi\eta_N}} \left(\int_{\frac{m}{2^j}}^{\frac{m+1/2}{2^j}} e^{-\frac{(x-y)^2}{2\eta_N}}\, dy - \int_{\frac{m+1/2}{2^j}}^{\frac{m+1}{2^j}} e^{-\frac{(x-y)^2}{2\eta_N}}\, dy \right).
\end{aligned}
\tag{112}
$$

The equation above is useful when we want to compute the derivatives of $P_{\eta_N} b^N$, because we can differentiate inside the integral sign, and then as b_η^N is a finite sum of Haar functions multiplied by some coefficients, we can rewrite $\frac{\partial}{\partial x} b_\eta^N(x)$ and $\frac{\partial^2}{\partial x^2} b_\eta^N(x)$, as the sum of the building blocks of the sum, i.e. terms of the form $\frac{\partial}{\partial x} P_{\eta_N} \mathbb{1}_{(a,b)}(x)$ and $\frac{\partial^2}{\partial x^2} P_{\eta_N} \mathbb{1}_{(a,b)}(x)$, where a,b are the points defining the corresponding Haar functions.

$$
\frac{\partial}{\partial x}[P_{\eta_N}\mathbb{1}_{(a,b)}](x) = \frac{1}{\sqrt{2\pi\eta_N}}\left(\exp\left(-\frac{(x-a)^2}{2\eta_N}\right) - \exp\left(-\frac{(x-b)^2}{2\eta_N}\right)\right),
\tag{113}
$$

$$\frac{\partial^2}{\partial x^2}[P_{\eta_N}\mathbb{1}_{(a,b)}](x) = \frac{1}{\sqrt{2\pi\eta_N}}\left((x-a)\exp\left(-\frac{(x-a)^2}{2\eta_N}\right) - (x-b)\exp\left(-\frac{(x-b)^2}{2\eta_N}\right)\right.$$
$$\left. + (e^{\eta_N}-1)\left(\Phi\left(\frac{x-a}{\sqrt{\eta_N}}\right) - \Phi\left(\frac{x-b}{\sqrt{\eta_N}}\right)\right)\right). \tag{114}$$

The derivatives of the heat kernel applied to the indicator functions are then arranged in the form of the numerical schemes that we want to apply, and considering of course that we have to recover the coefficient b_η^N by means of multiplying by the appropriate coefficients $\mu_{j,m}$, and the Haar functions $h_{j,m}$.

We must insist that the schemes that involve the derivative of b_η^N have not been studied theoretically, and therefore the link between those and the derivatives of a distributional coefficient b are not clear.

First we implement the Euler scheme, and since this equation does not have a closed form solution, we decide in a finest grid to compute what we will use as the real solution with the Euler scheme. As an example we use the finest time grid $(t_m)_{m=0}^M$ where $M = 2^9$ and $t_0 = 0 < ... < 1 = t_M$. We generate a FBM with Hurst coefficient $H = 0.85$, so that $\beta_0 = 0.15$, and according to Assumption 1 we select $q_0 = 1/\beta_0$. Now, for b_η^N we have $N = 9$ and let us remind that in [5], after Assumption 1, γ_0 is defined as

$$\gamma_0 = 1 - \beta_0 - \frac{1}{q_0}, \tag{115}$$

and our selection of q_0 leaves us with $\gamma_0 = 1 - 2\beta_0$. For the approximations we use grids $(t_{m'})_{m'=0}^{M'}$ where $t_0 = 0 < ... < 1 = t_{M'}$, for $M' = 2^2,...,2^6$, and we compute again the Euler scheme. For both the proxy of the real solution and the coarser approximations we compute 100 paths and compute strong and weak error, and we can see both of them in Fig. 2.
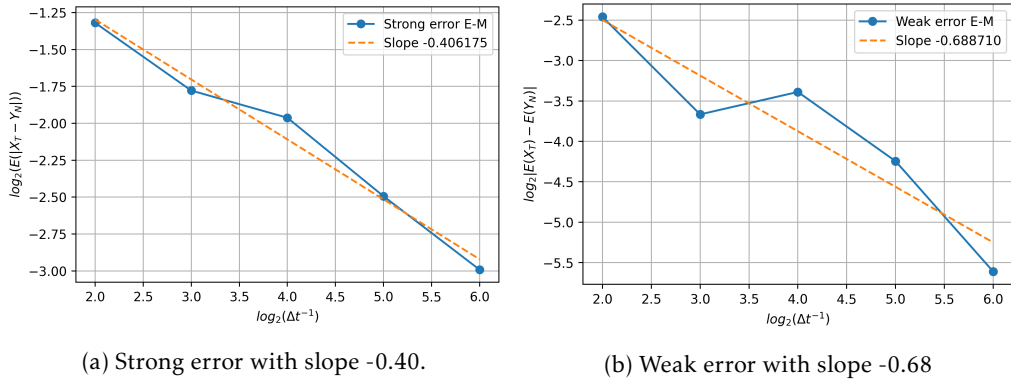

(a) Strong error with slope -0.40. (b) Weak error with slope -0.68

Figure 2: Log-log graphs of the error of approximation with $N = 9$, and time steps $M' = 2^2,...,2^6$ for $b \in H_{6.66}^{-0.15}$ for the Euler scheme.

In Fig. 2 we can observe how the error does not behave as in a regular Euler scheme. Also, although changing the regularity of the coefficient, i.e: changing the Hurst coefficient of FBM, should change the rate of convergence, it does not show a noticeable difference, which agrees with the results in [5, Table 1]. in contrast with said table we are getting slightly better results for the empirical rate of convergence, having a slope in the neighbourhood of -0.45 in most of the runs.

Additionally, we compared the standard deviation of the differences at terminal time between the sample paths of the real solution and the approximations at different refinements of the time grid, and we obtain a similar behaviour between that and the error, as we can see in Fig. 3.

20

In our case, we see that the standard deviation also decreases with the increase in step sizes.
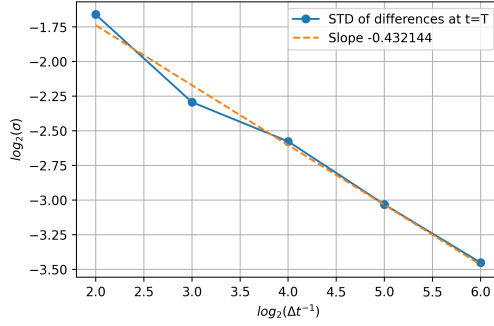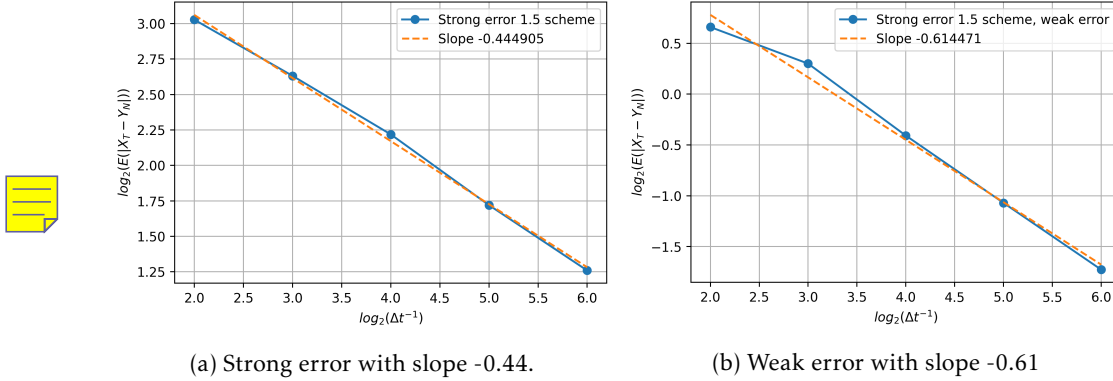

Figure 3: Log-log graphs standard deviation of approximations for the Euler Scheme.

On the other hand, when we implement the 1.5 strong scheme we do not have encouraging results, although the slope obtained for the scheme is similar to the Euler scheme as we can see in Fig. 4a, the error is way larger, of the order of 2 to 2^3. The weak error in contrast, behaves better in the sense that is of the order of $2^{0.5}$ to $2^{-1.5}$.

(a) Strong error with slope -0.44.   (b) Weak error with slope -0.61

Figure 4: Log-log graphs of the error of approximation with $N = 9$, and time steps $M' = 2^2, ..., 2^6$ for $b \in H_{6.66}^{-0.15}$ ==for the Euler scheme==.

With the standard deviation of the differences, we have a similar behaviour as in the strong error, a reasonable slope, but a large deviation between the observations as we can see in Fig. 5.

Although the convergence of the 1.5 strong scheme, or any other scheme adding more terms from the Taylor expansion, has not been studied theoretically, one could expect the error to behave in a better way since the Euler scheme is being applied to a function b_η^N that has properties more suited for the convergence of the Euler scheme. This is an important limitation of the lack of theoretical study in the subject, although we would expect a certain behaviour we can not be certain that it is the behaviour that we are actually supposed to have for a given choice of parameters for the approximation.

Most of efforts on the numerical schemes for SDEs are made around the Euler scheme, and this partly is because we only need the drift and diffusion coefficients properties to be known, and not those of their derivatives, which adds a further level of complexity.

In this case, even if the 1.5 strong scheme (or any other) behaves in a more desirable way, we do not really know the relationships between the derivatives of b_η^N and the ones of b, which if it proves relevance for the project could be studied in the future.
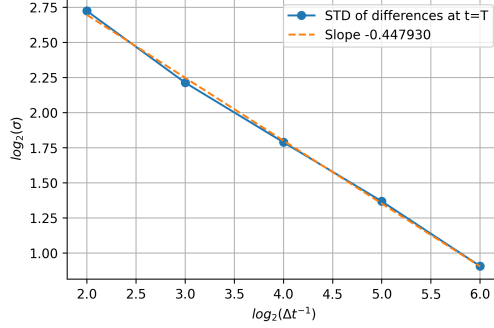
Figure 5: Log-log graphs standard deviation of approximations for the Euler scheme.

6 BSDEs and FBSDEs

In what follows let $\mathbb{F} := (\mathscr{F}_t)_{t \geq 0}$ be a filtration, $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$ a complete filtered probability space, and consider the notation $\mathscr{L}^p(\mathbb{R})$ to represent the set

$$\mathscr{L}^p(\mathbb{R}) = \{\xi : \mathbb{F}\text{-adapted and } \mathbb{E}[|\xi|^p] < \infty\}. \tag{116}$$

Definition 24 (Backward SDE) *Let $f : [0, T] \times \mathbb{R}^2 \to \mathbb{R}$ be measurable in all variables, $f(t, y, z)$ uniformly Lipschitz continuous in (y, z), and let $\xi \in \mathscr{L}^2(\mathbb{R})$ be $\mathscr{F}_T$-measurable. Then the following is called a Backward SDE (BSDE):*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \tag{117}$$

for $0 \leq t \leq T$, $\mathbb{P}$-a.s. where $Y, Z \in \mathscr{L}^2(\mathbb{R})$. f and ξ are called, respectively, the driver and the terminal condition of the BSDE.

A particular case of a BSDE when the driver $f = 0$ is the well known martingale representation theorem, in which we see that under certain conditions we can write a given square integrable martingale as a stochastic integral.

Theorem 14 (Martingale representation theorem [15, Theorem 2.5.2]) *For any $\xi \in \mathscr{L}^2(\mathbb{R})$, there exists a unique $\sigma \in \mathscr{L}^2(\mathbb{R})$ such that*

$$\xi = \mathbb{E}[\xi] + \int_0^T \sigma_t dW_t. \tag{118}$$

Consequently if M is a martingale such that $\mathbb{E}[|M_T|^2] < \infty$, then there exists a unique $\sigma \in \mathscr{L}^2(\mathbb{R})$ such that

$$M_t = M_0 + \int_0^t \sigma_s dW_s \tag{119}$$

As an example let us consider $\xi \in \mathscr{L}^2(\mathbb{R})$, and $\mathbb{F} = (\mathscr{F}_t)_{t \geq 0}$ be the filtration generated by the Wiener process W. It induces the martingale $Y_t := \mathbb{E}[\xi | \mathscr{F}_t]$, and by the martingale representation theorem, there exists a unique $Z \in \mathscr{L}^2(\mathbb{R})$ such that

$$Y_t = \xi - \int_t^T Z_s dW_s. \tag{120}$$

Theorem 15 ([15, Theorem 4.3.1]) *Under the assumptions of Definition 24, the BSDE* (117) *has a unique solution $(Y, Z) \in \mathscr{L}^2(\mathbb{R}) \times \mathscr{L}^2(\mathbb{R})$.*

One can go one step further and take the terminal condition ξ to be defined by an SDE, moreover we can choose the driver f to depend on the solution of the SDE. This generalisation is called Forward-Backward SDE (FBSDE).

Definition 25 (Forward-Backward SDE) *Let $f : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ be measurable in all variables, uniformly Lipschitz continuous in (x,y,z), and let $\xi \in \mathscr{L}^2(\mathbb{R})$ be the following system is called a Forward-Backward SDE:*

$$\begin{cases} X_t & = X_0 + \displaystyle\int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s, \\ Y_t & = g(X_T) - \displaystyle\int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dW_s, \end{cases} \tag{121}$$

for $0 \le t \le T$, $\mathbb{P}$-a.s.

Theorem 16 ([11, Theorem 5.1]) *Under the assumptions of Definition 25, the FBSDE (120) admits a unique solution $(X,Y,Z) \in \mathscr{L}^2(\mathbb{R}) \times \mathscr{L}^2(\mathbb{R}) \times \mathscr{L}^2(\mathbb{R})$.*

Following the foundations that have been established at the moment, the plan for the research project is to develop numerical schemes to solve FBSDEs with a rough drift in the SDE associated. The equations that we are going to attack are of the form

$$\begin{cases} X_t & = X_0 + \displaystyle\int_0^t b(s,X_s)ds + \int_0^t dW_s, \\ Y_t & = g(X_T) - \displaystyle\int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dW_s. \end{cases} \tag{122}$$

We can see that the forward element of the equation above is exactly the same equation with distributional for which the numerical schemes have already been implemented. With the first equation solved we can proceed to solve the BSDE by plugging X into f and in the terminal condition. Since the distributional coefficient only appears in the forward equation, we can use known numerical methods for BSDEs such as the ones discussed in [2, Section 3]. In the future we will study first the special case for $f = 0$ which leaves us with an equation with the same form as (119) and then the component Y will be exactly the conditional expectation of the terminal condition ξ.

Once the numerical methods for such equations are studied, the following task is to prove the convergence of the scheme. Let us remember that the convergence for numerical schemes for the forward equation is dependant on the convergence for the approximation of the distributional coefficient as proved in [5], then we expect to see a dependency in a similar fashion for the convergence rate of the solution for the BSDE.

Further, we can have a system in which the BSDE contains a rough driver, the system would be

$$\begin{cases} X_t & = X_0 + \displaystyle\int_0^t dW_s, \\ Y_t, & = g(X_T) + \displaystyle\int_t^T b(s,X_s)Z_s ds + \int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dW_s. \end{cases} \tag{123}$$

This is a completely new problem, and we aim to come up with schemes to solve these kind of BSDEs. The first approach could be to use the techniques from [5] to deal with the coefficient b and then use known methods for BSDEs. Numerically we can always try to adapt the schemes that are already known, however the theoretical results of convergence might be more challenging, and that should be the next big step for the project.

Another problem that is of our interest is the case of an SDE with a rough coefficient just as the one for which the numerical methods have been implemented, but with the difference of if being driven by a different process than a regular Brownian motion. Namely, we can have an

equation driven by a fractional Brownian motion, that is different to the one used to approximate the distributional coefficient. Such equation reads

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t dW_s^H. \tag{124}$$

Results in this direction can be seen in [1, 3] where theoretical rates of convergence for the equation above have been given, in particular in [3] it is discussed how regardless of the regularity of the drift b, it is possible to obtain a rate of convergence $1/2 + \epsilon$ provided that b is bounded, not necessarily Lipschitz.

7 Training record

- **Taught work and readings**

  - **Measure and Integration Theory**. Collegio Carlo Alberto. Turin, Italy. January-February 2021. *(Online)*. (Pass 65%).

  - **MATH5734M: Advanced Stochastic Calculus and Applications to Finance**. University of Leeds. Spring Semester. 20 credits. (Pass 72%).

  - Reading of [9, Sections 3.2, 3.3, 5.2] as complement for MATH5734M.

- **Seminar series**

  - **Seminar Series on Probability and Financial Mathematics**. University of Leeds. Leeds, UK. *(Online)*. 2020/2021.

  - **Seminar Series on Probability and Financial Mathematics**. University of Leeds. Leeds, UK. *(Online)*. 2021/2022.

- **Conferences**

  - **Stochastic Processes and Their Friends**. University of Leeds. Leeds, UK. *(Online)*. 18-19 March 2021.

  - **Conference Beyond the Boundaries**. University of Leeds. Leeds, UK. *(Online)*. 4-7 May 2021.

  - **British Early Career Mathematicians' Colloquium**. University of Birmingham. Birmingham, UK. *(Online)*. 15-16 July 2021.

  - **6th Berlin Workshop for Young Researchers on Mathematical Finance**. Humboldt-Universität zu Berlin. Berlin, Germany. *(Online)*. 23-25 August 2021.

  - **Bath Mathematical Symposium on PDE and Randomness: Summer School**. University of Bath. Bath, UK. *(Online)*. 1-3 September 2021.

References

[1] Oleg Butkovsky, Konstantinos Dareiotis, and Máté Gerencsér. "Approximation of SDEs: a stochastic sewing approach". In: *Probability Theory and Related Fields* (July 30, 2021).

[2] Jared Chessari and Reiichiro Kawai. "Numerical Methods for Backward Stochastic Differential Equations: A Survey". In: *arXiv:2101.08936 [cs, math]* (Jan. 21, 2021). arXiv: 2101.08936.

[3] Konstantinos Dareiotis and Máté Gerencsér. "On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift". In: *Electronic Journal of Probability* 25 (none Jan. 1, 2020). arXiv: 1812.04583.

[4]   Konstantinos Dareiotis, Máté Gerencsér, and Khoa Lê. "Quantifying a convergence theorem of Gy\"ongy and Krylov". In: *arXiv:2101.12185 [math]* (Jan. 28, 2021). arXiv: 2101. 12185.

[5]   Tiziano De Angelis, Maximilien Germain, and Elena Issoglio. "A Numerical Scheme for Stochastic Differential Equations with Distributional Drift". In: *arXiv:1906.11026 [cs, math]* (Oct. 22, 2020). arXiv: 1906. 11026.

[6]   Franco Flandoli, Elena Issoglio, and Francesco Russo. "Multidimensional stochastic differential equations with distributional drift". In: *Transactions of the American Mathematical Society* 369.3 (June 20, 2016), pp. 1665–1688.

[7]   Paul Glasserman. *Monte Carlo methods in financial engineering*. Applications of mathematics 53. New York: Springer, 2004. 596 pp.

[8]   Elena Issoglio. "Transport Equations with Fractal Noise - Existence, Uniqueness and Regularity of the Solution". In: *Zeitschrift für Analysis und ihre Anwendungen* 32.1 (2013), pp. 37–53.

[9]   Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Second edition. Graduate texts in mathematics 113. New York ; Springer-Verlag, 1991.

[10]   Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. Corr. 3rd print. Applications of mathematics 23. Berlin ; New York: Springer, 1999. 636 pp.

[11]   Jin Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*. Corr. 3rd print. Lecture notes in mathematics 1702. Berlin ; New York: Springer, 2007. 270 pp.

[12]   Eckhard Platen. "An introduction to numerical methods for stochastic differential equations". In: *Acta Numerica* 8 (Jan. 1999), pp. 197–246.

[13]   Hans Triebel. *Bases in function spaces, sampling, discrepancy, numerical integration*. EMS tracts in mathematics 11. Zürich: European Mathematical Society, 2010. 296 pp.

[14]   Hans Triebel. *Theory of function spaces*. Modern Birkhäuser classics. Basel ; Boston: Birkhäuser Verlag, 2010. 285 pp.

[15]   Jianfeng Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*. 1st ed. 2017. Probability Theory and Stochastic Modelling 86. New York, NY: Springer New York : Imprint: Springer, 2017. 1 p.