Transfer report

Luis Mario Chaparro Jaquez

November 17, 2021

Contents

1 Overview

...LET THE OVERVIEW TO THE END... Numerical schemes for Stochastic Differential Equations (SDEs) and Stochastic Partial Differential Equations (SPDEs) have been widely studied, and even for SDEs and SPDEs with low regularity coefficients. However

The works from De Angelis et al. [4], and Flandoli et al. [5] have established the framework for this project. A one dimensional SDE is considered:

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

where W is a Brownian motion, and $b(t,x)$ is a distribution taking values in a fractional Sobolev space of negative order, namely $H^{-\beta_0}_{\tilde{q}_0, q_0}$.

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 [5], and then in [4] 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 is performed a process of regularisation of the coefficient b in (1). Since b it is a distribution and cannot be computed pointwise. They use Haar systems, because those are unconditional bases in the space in which the coefficient b exists, producing a sequence $(b^N)_{N\geq 1}$, which is then submitted to a *randomisation* procedure applying the heat kernel given by
2. And finally applying the Euler-Maruyama scheme for the modified coefficient $P_{\eta_n} b^N$.

...LET THE OVERVIEW TO THE END...

2 Literature review

2.1 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 equation

the following stochastic differential equation (SDE)

$$X_t = x_0 + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s. \tag{4}$$

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 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 [7, Definition 5.2.1]) *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, we call a strong solution of equation (4) a $\mathscr{F}$-adapted, continuous, $\mathbb{R}$-valued stochastic process $(X_t)_{t\in(0,\infty)}$ such that (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 ([7, 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, $\mathscr{F}$-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 $\mathbb{P}(X_t = \tilde{X}_t; \forall 0 \leq t \leq \infty) = 1$.

2.2 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.

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 [8, Chapter 5]. In order to introduce what we need, let us set some notation.

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

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

where

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

Introduce $\mathcal{M}^1$ (call it $\mathcal{M}$) only.

no need for $\mathcal{M}$.

and $m = 1,2,3,...$ a multi-index with length $l := l(\alpha) \in \{1,2,...\}$. We denote the set of all multi-indexes as $\mathcal{M}$.

not clear to me: is this a multi-index of length zero?

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

Let us remind that m is the number of entries that the Wiener process driving our equation has. We are only interested in $\mathbb{R}$-valued solutions for the equation, and also on Wiener processes with $m = 1$. Then let us define:

Definition 3 *Let $\mathcal{M}^1 \subset \mathcal{M}$ the subset of multi-indices such that*

$$\mathcal{M}^1 = \{\alpha = (j_1,...,j_l) : j_i \in \{0,1\} \quad \forall \quad i = 1,...,l\}. \tag{10}$$

for all $l \in \mathbb{N}$.

Definition 4 (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~~ set $\mathcal{B}(\mathcal{A})$ for any hierarchical set is given by

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}. \tag{11}$$

?

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

Definition 5 (Multiple Itô Integrals [kloeden_numerical_1991]) *Let us denote by $\mathcal{H}$ the set of all càdlàg processes $f = \{f(t) \geq 0\}$. Then let us define*

where do you need this space?

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

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

$$\mathcal{H}_{(1)} = \left\{f \in \mathcal{H} : \mathbb{P}\left[\int_0^t |f(t,\omega)|^2 ds < \infty \quad \forall t \geq 0\right] = 1\right\}. \tag{14}$$

Then for a multi-index $\alpha = (j_1,...,j_l)$ and a process $f \in \mathcal{H}_\alpha$ we define the multiple Itô integral as

Circular def: you pick $f \in \mathcal{H}_\alpha$ then define I_α then define $\mathcal{H}_\alpha$ in terms of I_α. Which one comes first?

I guess here you need $\alpha \in \mathcal{M}^1$?

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

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

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

$f \in \mathcal{H}_v$, $\alpha = v$ define $I_\alpha f = f$. Then assume $I_{\alpha-}$, $\mathcal{H}_{\alpha-}$ are defined recursively. Then define I_α, $\mathcal{H}_\alpha$ with (16)(15)

Definition 6 (Itô Coefficient Functions [kloeden_numerical_1991]) *Let us define the operators*

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

$$L^1 = \sigma\frac{\partial}{\partial x}. \tag{18}$$

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

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

not really, here it looks like l must be either $l=0$ or $l=1$. ($\alpha \in \mathcal{M}^1$)

Theorem 2 ([8, Theorem 5.5.1]) *Let $t \in [0,T]$, Let $\mathscr{A} \subset \mathscr{M}^1 \subset \mathscr{M}$ be a hierarchical set, and let*

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

then the Itô-Taylor expansion

$$f(t, X_t) = \sum_{\alpha \in \mathscr{A}} I_\alpha[f_\alpha(0, X_0)]_{0,t} + \sum_{\alpha \in \mathscr{B}(\mathscr{A})} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{0,t}, \tag{21}$$

→ define this to be R

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

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 given by the remainder set $\mathscr{B}(\mathscr{A}) =: R$ (or a part of it) behind to approximate numerically the solutions to SDEs. And then it is necessary to prove that leaving that 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 [6, pp. 339–343]. Let us remind that in what follows, for any $f(t,x) \in C^2(\cdot, \mathbb{R})$ x, we will write $f' = \frac{\partial f}{\partial x}$ and $f'' = \frac{\partial^2 f}{\partial x^2}$.

Add examples of $\mathcal{A}$ (to get e.g. EM scheme).

Definition 7 *Let $[0,T]$ be a time interval, and $(t_n)_{n=0}^N$ be a sequence of elements in $[0,T]$ such that $0 = t_0 < \ldots < t_N = T$, 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

$$b := b(t, X_t), \tag{22}$$
$$\sigma := \sigma(t, X_t), \tag{23}$$

ok, but to be clear you should specify below at which time (and space) you are evaluating these, because it is not (t, X_t) but presumably (t_n, X_{t_n})?

since its dependency is clear from the context.

Definition 8 (Euler-Maruyama scheme [8, 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*

$$Y_{n+1} = Y_n + b\Delta t_n + \sigma \Delta W_n. \tag{24}$$

$\mathcal{A} = \{v, (0), (1)\} \to$ EM scheme & write explicitely $I_\alpha f = \ldots$ for $\alpha \in \mathcal{A}$.

Definition 9 (Milstein scheme [8, 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*

$\mathcal{A} = \{\ldots\} \to$ Milstein

$$\hat{Y}_{n+1} = \hat{Y}_n + b\Delta t_n + \sigma \Delta W_n + \frac{1}{2}\sigma'\sigma((\Delta W_n)^2 - \Delta t_n). \tag{25}$$

Definition 10 (Order 1.5 strong Taylor scheme [8, 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*

$\mathcal{A} = \{\ldots\} \to$

$$\begin{aligned}
\tilde{Y}_{n+1} = \tilde{Y}_n &+ b\Delta t_n + \sigma \Delta W_n + \frac{1}{2}\sigma'\sigma\left((\Delta W_n)^2 - \Delta t_n\right) \\
&+ b'\sigma \Delta Z_n + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2 \\
&+ \left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)(\Delta W_n \Delta t_n - \Delta Z_n) \\
&+ \frac{1}{2}\left(\sigma^2\sigma'' + \sigma(\sigma')^2\right)\left(\frac{1}{3}(\Delta W_n)^2 - \Delta t_n\right)\Delta W_n
\end{aligned} \tag{26}$$

where ΔZ_n is given by

$$\Delta Z_n = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{s_2} dW_{s_1}\, ds_2 \tag{27}$$

Definition 11 (Order 2.0 weak Taylor scheme [8, 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*

Do you mean $\check{Y}$?

$$\begin{aligned}\tilde{Y}_{n+1} = \tilde{Y}_n &+ b\Delta t_n + \sigma \Delta W_n + \frac{1}{2}\sigma'\sigma\left((\Delta W_n)^2 - \Delta t_n\right)\\ &+ b'\sigma\Delta Z_n + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2\\ &+ \left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)(\Delta W_n \Delta t_n - \Delta Z_n)\end{aligned} \tag{28}$$

where ΔZ_n is given by

$$\Delta Z_n = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{s_2} dW_{s_1}\, ds_2 \tag{29}$$

When it comes to numerical computations, one can make use of the fact that for any $0 \le s < t$ and $Z, \tilde{Z}$ i.i.d. as $\mathcal{N}(0,1)$ we can represent $W_t - W_s$ as $\sqrt{t-s}Z$ and $\int_s^t \int_s^z dW_u\, dz$ as $\frac{1}{2}(t-s)^{3/2}(Z + \frac{1}{\sqrt{3}}\tilde{Z})$ [8, pp. 351, 352]. Let us denote two random vectors by $Z = (Z_1, \ldots, Z_N)$ and $\tilde{Z} = (\tilde{Z}_1, \ldots, \tilde{Z}_N)$ such that for all $n \in \{1, \ldots, N\}$, $Z_n, \tilde{Z}_n \sim \mathcal{N}(0,1)$, then the n-th iteration of the schemes above are respectively

$$Y_{n+1} = Y_n + b\Delta t_n + \sigma\sqrt{\Delta t_n} Z_{n+1}, \tag{30}$$

$$\hat{Y}_{n+1} = \hat{Y}_n + b\Delta t_n + \sigma\sqrt{\Delta t_n} Z_{n+1} + \frac{1}{2}\sigma'\sigma\Delta t_n (Z_{n+1}^2 - 1), \tag{31}$$

$$\begin{aligned}\tilde{Y}_{n+1} = \tilde{Y}_n &+ b\Delta t_n + \sigma Z_{n+1}\sqrt{\Delta t_n} + \frac{1}{2}\sigma'\sigma\left(Z_{n+1}^2 - 1\right)\Delta t_n\\ &+ \frac{1}{2}b'\sigma\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2\\ &+ \frac{1}{2}\left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)\left(Z_{n+1}\left(1 - \frac{1}{2}\Delta t_n\right) - \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\Delta t_n\right)\sqrt{\Delta t_n}\\ &+ \frac{1}{2}\left(\sigma^2\sigma'' + \sigma(\sigma')^2\right)\left(\frac{1}{3}Z_{n+1}^3 - Z_{n+1}\right)(\Delta t_n)^{3/2},\end{aligned} \tag{32}$$

$$\begin{aligned}\check{Y}_{n+1} = \check{Y}_n &+ b\Delta t_n + \sigma Z_{n+1}\sqrt{\Delta t_n} + \frac{1}{2}\sigma'\sigma\left(Z_{n+1}^2 - 1\right)\Delta t_n\\ &+ \frac{1}{2}b'\sigma\left(Z_{n+1} + \frac{1}{\sqrt{3}}\check{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{2}\left(bb' + \frac{1}{2}\sigma^2 b''\right)(\Delta t_n)^2\\ &+ \frac{1}{2}\left(b\sigma' + \frac{1}{2}\sigma^2\sigma''\right)\left(Z_{n+1}\left(1 - \frac{1}{2}\Delta t_n\right) - \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\Delta t_n\right)\sqrt{\Delta t_n}.\end{aligned} \tag{33}$$

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 12 (Strong convergence [11, 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 T there is $K > 0$ such that for any discretisation $(t_n)_{n=0}^N$ of $[0, T]$ we have ~~that there exists a constant $K < \infty$~~ such that*

$$\mathbb{E}|X_T - Y_N| \leq K \left(\max_n \Delta t_n \right)^\gamma \tag{34}$$

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

repetition. (Is k indep. of the discretization $(t_n)_{n=0}^N$?)

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 [11], 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 13 (Weak convergence [11, 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 T and f there is $M_f > 0$ such that for any discretisation $(t_n)_{n=0}^N$ of $[0, T]$ we have*

$$|\mathbb{E}[f(X_T)] - \mathbb{E}[f(Y_N)]| \leq M_f \left(\max_n \Delta t_n \right)^\beta \tag{35}$$

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

For the present work we will be using equidistant time increments, of length $h \in \{2^{-3}, \dots, 2^{-7}\}$ therefore the conditions for convergence stated above can be seen simply as

$$\mathbb{E}|X_T - Y_N| \leq K h^\gamma \qquad \text{and} \qquad |\mathbb{E}[f(X_T)] - \mathbb{E}[f(Y_N)]| \leq M_f h^\beta. \tag{36}$$

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 solution as the discretisation is refined. Additionally it is also desirable to know which is the order of convergence. For that objective, Glasserman [6] and Kloeden and Platen [8] state the conditions on the coefficients and the proofs are found in the latter.

Theorem 3 (Strong convergence for E-M scheme [8, Theorem 10.2.2]) *Let the conditions from Theorem 1 hold, additionally assume that*

$$|b(s_0, x) - b(t_0, x)| + |\sigma(s_0, x) - \sigma(t_0, x)| \leq K(1 + |x|)\sqrt{|t_0 - s_0|}. \tag{37}$$

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

In the following, the condition on (37) will be called *extended linear growth*.

Theorem 4 (Strong convergence for Milstein scheme [8, 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*

$$z(t, x) = b(t, x) - \frac{1}{2}(\sigma\sigma')(t, x) \tag{38}$$

Suppose that the Lipschitz

$$\left|z(t, x_0) - z(t, y_0)\right| + \left|\sigma(t, x_0) - \sigma(t, y_0)\right| + \left|L^1\sigma(t, x_0) - L^1\sigma(t, y_0)\right| \leq K_1|x_0 - y_0|, \tag{39}$$

linear growth

$$|z(t,x_0)| + |L^1 z(t,x_0)| + |\sigma(t,x_0)| + |L^1\sigma(t,x_0)| + |L^1 L^1 \sigma(t,x_0)| \leq K_2(1+|x|), \tag{40}$$

and extended linear growth conditions

$$|z(s_0,x) - z(t_0,x)| + |\sigma(s_0,x) - \sigma(t_0,x)| \left| L^1\sigma(s_0,x) - L^1\sigma(t_0,x) \right| \leq K_3(1-|x|)|s_0 - t_0|^{1/2}, \tag{41}$$

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, namely $\gamma = 0.5, 1.0, 1.5, 2.0, \ldots$. check this result again

Theorem 5 (Strong convergence of all orders [8, 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 for any $j \geq 0$, Lipschitz continuity*

only needed for σ? Not for b?

Not clear to me: (42) (43) should hold for all $j \geq 0$???

link j with γ

$$|(L^1)^j\sigma(t,x) - (L^1)^j\sigma(t,y)| \leq K_1|x-y|, \tag{42}$$

and linear growth

ADD Assumptions on b (there must be some)

$$|(L^1)^j\sigma(t,x)| \leq K_2(1+|x|), \tag{43}$$

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

$$\mathbb{E}|X_T - Y_N|^2 \leq K_3(1+|X_0|^2)h^{2\gamma} + K_4|X_0 - Y_0|^2, \tag{44}$$

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

Theorem 6 (Weak convergence for all orders [8, 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 for all $j \geq 0$, the linear growth condition*

$$|(L^1)^j\sigma(t,x)| \leq K(1+|x|), \tag{45}$$

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

$$|\mathbb{E}[g(X_T)] - \mathbb{E}[g(Y_N)]| \leq C_g h^\beta. \tag{46}$$

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 $m \in \{1,\ldots,M\}$ then the approximation of the strong error will be then given by

$$E|X_T - Y_N| \approx \frac{1}{M}\sum_{m=1}^{M} |X_T^m - Y_N^m|, \tag{47}$$

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

$$|E[f(X_T)] - E[f(Y_N)]| \approx \left| E[f(X_T)] - \frac{1}{M}\sum_{m=1}^{M} f(Y_N^m) \right|, \tag{48}$$

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.

To verify the accuracy of the numerical schemes mentioned earlier, an SDE with a known explicit solution is used. In [8, 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)

$$\begin{cases} dX_t = \frac{1}{2}X_t dt + X_t dW_t, \\ X_0 = x_0 \end{cases} \tag{49}$$

whose solution is

$$X_t = x_0 \exp W_t. \tag{50}$$

Its mean is given by

$$\mathbb{E}[X_t] = X_0 \exp\left(\frac{1}{2}t\right). \tag{51}$$

For this SDE with the notation from Definition 8 and Definition 9, the Euler-Maruyama, Milstein, 1.5 strong and 2.0 weak approximations at time t_n are, respectively:

finish this section NOWWWWW

$$Y_{n+1} = Y_n + \frac{1}{2}Y_n\Delta t_n + Y_n\sqrt{\Delta t}Z_{n+1} \tag{52}$$

$$\hat{Y}_{n+1} = \hat{Y}_n + \frac{1}{2}\hat{Y}_n\Delta t_n + \hat{Y}_n\sqrt{\Delta t_n}Z_{n+1} + \frac{1}{2}\Delta t_n(Z_{n+1}^2 - 1) \tag{53}$$

$$\begin{aligned} \tilde{Y}_{n+1} = \tilde{Y}_n &+ \frac{1}{2}\tilde{Y}_n\Delta t_n + \tilde{Y}_n\sqrt{\Delta t_n}Z_{n+1} + \frac{1}{2}\Delta t_n(Z_{n+1}^2 - 1) \\ &+ \frac{1}{4}\tilde{Y}_n\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} \quad + \frac{1}{8}\tilde{Y}_n(\Delta t_n)^2 \\ &+ \frac{1}{2}\tilde{Y}_n\left(\frac{1}{2}Z_n + \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{2}\left(\frac{1}{3}Z_{n+1}^2 - 1\right)Z_{n+1}(\Delta t_n)^{3/2} \end{aligned} \tag{54}$$

$$\begin{aligned} \check{Y}_{n+1} = \check{Y}_n &+ \frac{1}{2}\check{Y}_n\Delta t_n + \check{Y}_n\sqrt{\Delta t_n}Z_{n+1} + \frac{1}{2}\Delta t_n(Z_{n+1}^2 - 1) \\ &+ \frac{1}{4}\check{Y}_n\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} + \frac{1}{8}\check{Y}_n(\Delta t_n)^2 \\ &+ \frac{1}{2}\check{Y}_n\left(\frac{1}{2}Z_n + \frac{1}{2\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} \end{aligned} \tag{55}$$

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 with 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.

2.3 Sobolev spaces and rough coefficients

very long sentence.

The approximation methods for SDEs mentioned above can lead us to good results if the coefficients behave in a certain way, as it was stated in Theorems 3, 4 and 6, however for the project in hand we are interested in some particular kind of drift coefficients, which are the essence of the problem and the reason it is necessary to develop new numerical schemes since those coefficients do not satisfy the necessary conditions for the schemes defined above. First let us introduce some definitions and notation.

Definition 14 (Fourier transform) *Let $f : \mathbb{R} \to \mathbb{C}$ be in $L^1(\mathbb{R})$. Then the Fourier transform of f is defined as*

$$\hat{f}(\xi) = \mathscr{F}[f](\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x} f(x) dx. \tag{56}$$

Definition 15 (Inverse Fourier transform) *Let $f : \mathbb{R} \to \mathbb{R}$ be in $L^1(\mathbb{R})$. Then the inverse Fourier transform of $\hat{f}$ is defined as*

$$f(\xi) = \mathscr{F}^{-1}[\hat{f}](\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x} \hat{f}(x) dx. \tag{57}$$

Definition 16 (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*

$$\int_{\mathbb{R}} u\phi^{(\alpha)} dx = (-1)^\alpha \int_{\mathbb{R}} v\phi dx \tag{58}$$

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

Definition 17 (Schwartz functions) *We denote the space of Schwartz functions as*

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty : \sup_{x\in\mathbb{R}} |x^\alpha f^{(\beta)}(x)| < \infty, \forall \alpha, \beta \in \mathbb{N}\}, \tag{59}$$

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

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 18 (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 $\mathcal{S}'(\mathbb{R})$.*

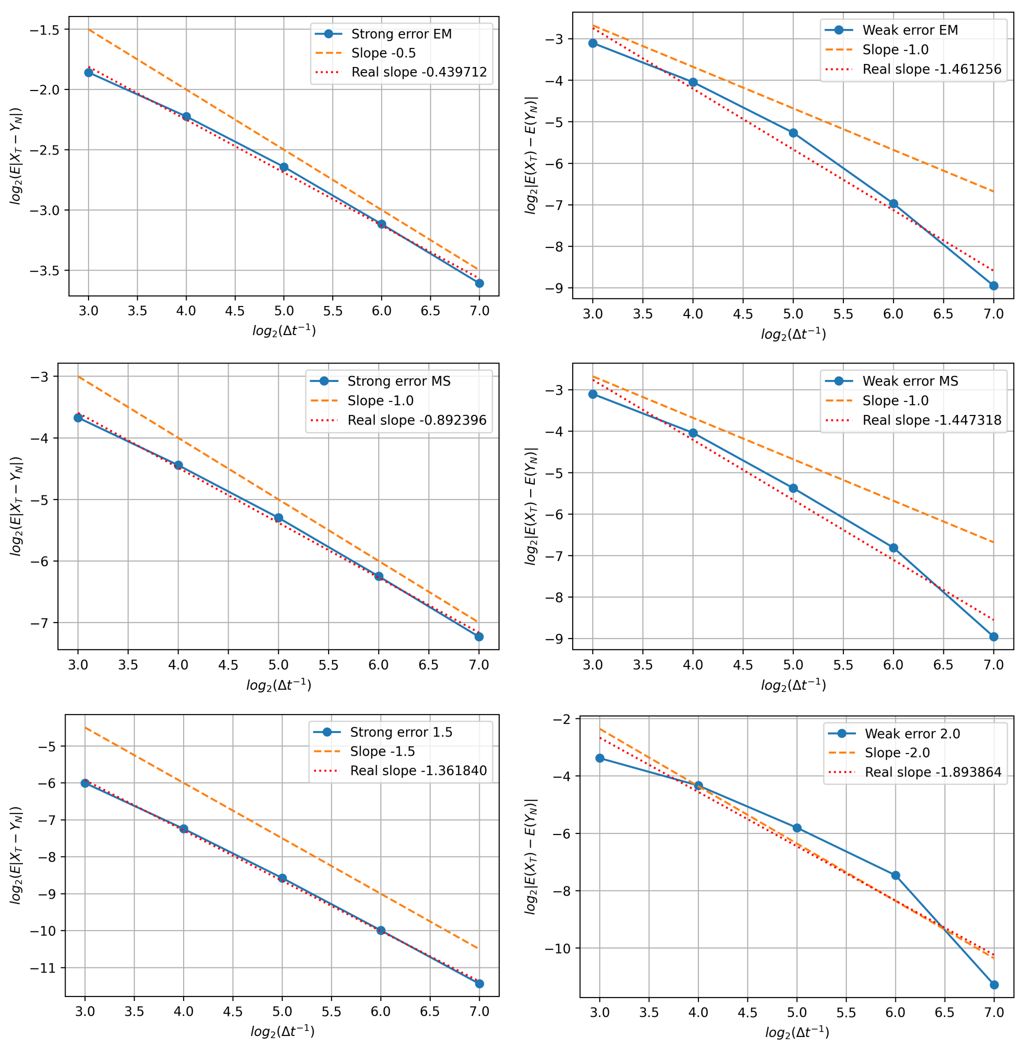

Figure 1: Log-log graphs of the error of approximation. Left column from top to bottom: Strong error for Euler scheme, strong error for Milstein scheme and strong error for 1.5 strong scheme. Right column from top to bottom: Weak error for Euler scheme, weak error for Milstein scheme and weak error for 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.

did you check on Triebel if this is true for all $s \in \mathbb{R}$, as you claim? I don't think it can hold for $s<0$...

$\mathcal{S}'(\mathbb{R})$

Definition 19 *[12, Remark 1.2] Let $f \in L^p(\mathbb{R})$ for $0 < p < \infty$ and let $\xi, s \in \mathbb{R}$. The map*

$$I^s : f \mapsto \mathscr{F}^{-1}\left[\left(1+|\xi|^2\right)^{\frac{s}{2}} \mathscr{F}[f]\right] \tag{60}$$

defines a $\boxed{\text{bijection on } L^p(\mathbb{R})}$. No! it is bijection on $\mathcal{S}$ and $\mathcal{S}'$.

The definition of fractional Sobolev spaces can be found in different formats by different authors, we have chosen the one in [12, Remark 1.2], and it is presented below.

Definition 20 (Fractional Sobolev spaces) *Let $s \in \mathbb{R}$, $0 < p < \infty$, and I^s as in (60) we call the space*

$$H_p^s := I^{-s} L^p(\mathbb{R}) \tag{61}$$

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

$$\|f\|_{H_p^s} := \|I^{-s} f\|_{L^p(\mathbb{R})}. \tag{62}$$

include this.

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 $W^{k,p}(\mathbb{R})$.

Add that if $s<0$ then H^s_p includes distributions

Definition 21 (Sobolev spaces) *Let $1 < p < \infty$ and $k \in \mathbb{N}$. Then the space $W^{k,p}(\mathbb{R})$ is the collection of all $f \in L^p(\mathbb{R})$ such that*

$$\|f\|_{W^{k,p}} := \left(\sum_{\alpha=0}^{k} \|f^{(\alpha)}\|_{L^p}^p\right)^{\frac{1}{p}} < \infty, \tag{63}$$

where $f^{(\alpha)}$ denotes the weak derivative of order $\alpha \in \mathbb{N}$.

If $s=0$ then $H^s_p = L^p$.

Add that there are equivalent definitions of these spaces, see Triebel, Def 1.1., where the spaces are defined as subspaces of $\mathcal{S}'$.

The reason Definition 20 and Definition 21 are equivalent is that under the Fourier transform, the derivative is a multiplier operator, that is

$$\mathscr{F}\left[\frac{d^k}{dx^k} f(x)\right](\xi) = (i\xi)^k [\mathscr{F} f](\xi). \tag{64}$$

So, for an example, let us set $p = 2$ and $k = 1$. For a function $f : \mathbb{R} \to \mathbb{R}$ to belong to the Sobolev space $W^{1,2}$ as in Definition 21 we need that $f, \frac{df}{dx} \in L^2$. And in terms of the Fourier transform we need that $\mathscr{F} f, i\xi \mathscr{F} f \in L^2$. Those last two conditions can be combined as

$$(1+|\xi|)\mathscr{F} f \in L^2. \tag{65}$$

But we can note at least two problems in that condition: first, $(1+|\xi|)$ is not smooth, and second, it does not account for the *"derivatives of order $s < k$"*. But the the term $(1+|\xi|^2)^{1/2}$ solves both issues. So for a general $k = s \in \mathbb{R}$, Definition 20 is better since the Fourier transform is well defined in Sobolev spaces.

Sobolev spaces provide us with an appropriate framework to find solutions to Partial Differential Equations (PDEs), relaxing the conditions of differentiability. One simple example of this feature can be seen in the equation

$$\begin{cases} -u''(x) + u(x) = f(x) & \text{on } [a,b], \\ u(a) = u(b) = 0. \end{cases} \tag{66}$$

In the classical sense a solution to this equation is a function in $C^2([a,b])$ that satisfies the equation. However, if we take a function $\phi \in C^1([a,b])$ to multiply Eq. (66), and then integrate by parts we obtain

$$\int_a^b u'\phi' dx + \int_a^b u\phi dx = \int_a^b f\phi dx, \tag{67}$$

we can now see that the solution u to the equation only needs to belong to $C^1([a,b])$, and its derivative u' has to be one time weak differentiable as the previous equation indicates. Further to solve equation (67), it is sufficient if $u, u' \in L^1(a,b)$, because in Sobolev spaces we are concerned with functions only in L^p spaces, therefore we have that the solution to equation (66) belongs to the space $W^{2,1}(a,b)$. include motivation to use fractional and negative sobolev spaces

$H^s_p := I^{-s} L_p$

what do you mean?

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). On that direction, note that the definition of fractional Sobolev spaces presented above not only makes sense for functions in $L^p(\mathbb{R})$ but also for the space of Schwartz functions $\mathcal{S}(\mathbb{R})$ and its dual, the space of tempered distributions $\mathcal{S}'(\mathbb{R})$. For that reason in what follows we will talk about objects belonging to the last two spaces. 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. We must add that the definition of the bases for different spaces can be written for a generic interval I, or even on $\mathbb{R}$, ~~however~~ we will only define the approximation of the distributional coefficient on $I = (0,1)$, and then with a rescaling will compute the coefficient at any point of an interval $[-K,K]$, more on this will be discussed on the implementation section.

$H^p_s = I^{-s} S$ no!

You have $S = I^s S$

$S' = I^s S'$

in fact, it is not just the def of the basis, but of the space itself, which is $H^s_p(I)$ rather than $H^s_p(\mathbb{R})$ as above. This space should be defined too. Use (87) from [De Angelis et al]

Definition 22 *Let $h_M : \mathbb{R} \to \mathbb{R}$*

$$h_M := \mathbb{1}_{[0,\frac{1}{2})} - \mathbb{1}_{[\frac{1}{2},1)} \tag{68}$$

be the mother Haar wavelet. Then the Haar wavelet system on $(0,1)$ is given by

$$\{h_0, h_{j,m} : \cancel{j \in \mathbb{N} \cup \{-1\}, m \in \mathbb{Z}}\} \tag{69}$$

No

YES

where

$$h_0(x) := \mathbb{1}_{(0,1)}(x) \tag{70}$$

and

$$h_{j,m}(x) := h_M(2^j x - m) \tag{71}$$

for $j \in \mathbb{N}$ and $m \in \{0, \dots, 2^j - 1\}$. Alternatively the Haar wavelets can be expressed as

$$h_{j,m}(x) = \mathbb{1}_{\left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right)} - \mathbb{1}_{\left[\frac{m+1/2}{2^j}, \frac{m+1}{2^j}\right)}. \tag{72}$$

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 ([12, Theorem 2.13]) *Let $0 \le r < \infty$, $-\frac{1}{2} < s < \frac{1}{r}$, and let $f \in \mathcal{S}'(0,1)$. Then $f \in H^s_r(0,1)$ if and only if*

$$f = \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} \mu_{j,m} 2^{-j(s-\frac{1}{r})} h_{j,m}, \tag{73}$$

with unconditional convergence in any space $H^\sigma_r(0,1)$ with $\sigma < s$. The representation is unique with the coefficients given by

$$\mu_0 := \int_0^1 f(x) h_0(x) dx \tag{74}$$

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

$$\mu_{j,m} := 2^{j(s-\frac{1}{r}+1)} \int_0^1 f(x) h_{j,m}(x) dx, \tag{75}$$

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

$$\left\{h_0, 2^{-j(s-\frac{1}{r})}h_{j,m} : j \in \mathbb{N}, m = 0, \dots, 2^j - 1\right\} \tag{76}$$

is an unconditional normalised basis of $H^s_r(0,1)$.

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, just define the coefficient in terms of dual pairing and explain why is not possible to compute 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 we can see that for $g \in H^s_r$ with $2 \le r < \infty$ and $1/2 < s < 1 + 1/r$ have basis with coefficients that are more easily computable and are also related to the coefficients of Haar basis, after all, in a way we are adding 1 to the order of derivation of the space in which the Haar functions are basis to get this new space. The systems for the new spaces are called Faber bases.

too informal

• $\forall s \in \mathbb{R}$ $f \in H^s \Rightarrow f' \in H^{s-1}$

• Formally we also have $v'_{jm} = h_{jm} 2^{j+1}$

Definition 23 *The Faber system in $(0,1)$ is given by*

$$\{v_0, v_1, v_{j,m} : j \in \mathbb{N}, m = 0, \dots, 2^j - 1\} \tag{77}$$

where

$$v_0(x) := \begin{cases} 1 - x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad v_1(x) := \begin{cases} x & 0 \le x \le 1 \\ 0 & \text{otherwise,} \end{cases} \tag{78}$$

and the hat functions are defined by

$$v_{j,m} := 2^{j+1} \int_0^x h_{j,m}(y)dy = \begin{cases} 2^{j+1}(x - 2^{-j}m) & x \in \left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right) \\ 2^{j+1}(2^{-j}(m-1) - x) & x \in \left[\frac{m}{2^j}, \frac{m+1/2}{2^j}\right) \\ 0 & \text{otherwise.} \end{cases} \tag{79}$$

• Hence we can think that having f expanded in Faber series can be formally derived term by term to get an expansion of f' in Haar series

This systems allows us to represent elements in fractional Sobolev spaces with positive order, this spaces are comprised by measurable functions and is definitely not where the drift $b(t,\cdot)$ lives. Nevertheless, the following theorem establishes the series representation of elements in the appropriate space with the Faber system.

Theorem 8 ([12, Theorem 3.1]) *Let $g \in H^s_r(0,1)$ for $2 \le r < \infty$ and $1/2 < s < 1 + 1/r$. Then we have the unique Faber representation for g*

$$g = \hat{\mu}_0 v_0 + \hat{\mu}_1 v_1 + \sum_{j=0}^{+\infty} \sum_{m=0}^{2^j - 1} \hat{\mu}_{j,m} v_{j,m} \tag{80}$$

with unconditional convergence in $C(0,1)$ and in $H^\sigma_r(0,1)$ with $\sigma < s$. Here *~~Where~~ the coefficients $\hat{\mu}$ are given by*

$$\begin{cases} \hat{\mu}_{j,m} & = -\frac{1}{2}(\Delta^2_{2^{-j-1}} g)(2^{-j}m) \\ \hat{\mu}_0 & = g(0) \\ \hat{\mu}_1 & = g(1), \end{cases} \tag{81}$$

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

Content of Remark A.10 in [4]. Highlight more.

After setting the above theorem it remains to see how those coefficients, that are clearly easier to compute, are related to the coefficients for the Haar representation. As pointed out in [4], if $s < 0$, the space $H^s_r(\mathbb{R})$ contains distributions. We can then define $H^s_{p,q}(\mathbb{R}) := H^s_p(\mathbb{R}) \cap H^s_q(\mathbb{R})$, and we have that if $f \in H^s_{p,q}(\mathbb{R})$ then $f \in H^s_r(\mathbb{R})$ for all $p \wedge q < r < p \vee q$. We can see, as is discussed

out of topic here. Move elsewhere.

if $g \in H^s_r(I)$ with $\frac{1}{2} < s < 1+\frac{1}{r}$ and $2 \le r < \infty$ then

in [4, Remark A.10] and proved in [12, Theorem 3.1] that we have $g' \in H_r^{s-1}(I)$ ~~with $2 \le r < \infty$; $1/2 < s < 1 + 1/r$, which~~ can be written as follows

and it.

$$g' = (\hat{\mu}_1 - \hat{\mu}_0)h_0 + \sum_{j=0}^{+\infty} \sum_{m=0}^{2^j-1} 2^{j+1} \hat{\mu}_{j,m} h_{j,m}. \tag{82}$$

Comparing this expansion with (73) from Thm 7 we see that the coeff μ_0, μ_{jm} in (74)(75) can be equivalently written as

~~the result above is a representation of elements in the space where the distributional coefficient lives as seen in Theorem 7 setting the coefficients to be for the series expansion as~~

$$\mu_0 = \hat{\mu}_1 - \hat{\mu}_0 \qquad and \qquad \mu_{j,m} = 2^{j+1} \hat{\mu}_{j,m}. \tag{83}$$

2.4 Numerical schemes for SDEs with distributional drift

The foundations of this research project are in [4], in which it is devised and 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.

why two spaces?

Definition 24 *We call $C^k([0,T];B)$, to the space of B-valued functions that are k-Hölder continuous in time, where B is some Banach space. If $\alpha \in (0,1)$ the space $C^\alpha([0,T];B)$ will be the set of functions in $C([0,T];B)$ such that*

what norm here?

$$[f]_{\alpha,B} := \sup_{s \neq t \in [0,T]} \frac{\|f(t) - f(s)\|_B}{|t-s|^\alpha} < \infty. \tag{84}$$

In fact, in C^α you want to have also the sup $\|f(t)\|_B$ bounded.

This depends on the space B, but so far, B is too generic to talk about distributions. You should say that we will pick $B = H^{-\beta_0}_{\tilde{q}_0 q_0}$

review this paragraph Let us note that when we have a map $g(t,\cdot) \in C^\alpha([0,T];B)$ it is potentially a distribution and we cannot say much about its behaviour other than in the time variable is well behaved and in the case of our problem in which it is time homogeneous, it is constant in time.

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 ([4]) *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 $b \in C^\kappa\left([0,T]; H^{-\beta_0}_{\tilde{q}_0,q_0}\right)$.*

Assumption 2 ([4]) *Let $(b^N)_{N \ge 1} \subset C^{\frac{1}{2}}\left([0,T]; H^0_{\tilde{q}_0,q_0}\right)$ be such that*

$$\lim_{N\to\infty} b^N = b \qquad in \qquad \mathscr{C}^{\frac{1}{2}}\left([0,T]; H^{-\beta_0}_{\tilde{q}_0,q_0}\right). \tag{85}$$

$H^s_p \subset H^r_p \quad s > r$

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 $\boxed{\text{inclusion property}}$ that those spaces have, in particular we could have $H^0_? = L^2$. Because of that inclusion property the assumption on the limit of b^N makes sense, as the sequence will belong also to the rough space in which the coefficient lives originally and therefore the convergence in said space will make sense. The coefficient b^N will be defined in Theorem 10, and although this is already a good enough object for the numerical schemes to work, the approximation will still be rough, in particular it will not be smooth. For this coefficient to be better behaved in that area, we can apply a heat kernel to it.

repetition

just add a proper reference, mention convolution

not really... it's hard (or not possible) to show that the num. scheme converges in this case (although it is possible to define the approximation)

Definition 25 (Heat kernel) *The heat kernel is a function $\Phi(t,y) : [0,T] \times \mathbb{R} \to \mathbb{R}$ defined as*

$$\Phi(t,y) := \frac{1}{\sqrt{4\pi t}} \tag{86}$$

here you mean the heat semigroup, which still needs to be defined.

define heat kernel here, probably some properties might be useful to include as well what is this eqn doing here!!!!????

$$P_{\eta_N}\mathbb{1}_{[x_1,x_2)} = \exp(-\eta_N)\left(\Phi\left(\frac{x_2-x}{\sqrt{\eta_N}}\right)-\Phi\left(\frac{x_1-x}{\eta_N}\right)\right) \tag{87}$$

rewrite It is worth to notice as in [4, Remark 3.2] move this properties above and also add the discussion by Jan to rescale and shift the coefficient b that the heat kernel has the following properties:

$$\begin{aligned}\|P_t f\|_{\infty,H_q^{-s}} &\le \|f\|_{\infty,H_q^{-s}}\\ \|P_t f - f\|_{\infty,H_q^{-s-\epsilon}} &\le ct^{\epsilon/2}\|f\|_{\infty,H_q^{-s}},\end{aligned} \tag{88}$$

Note that when we apply the semigroup P_{η_N} to the Haar functions in the Haar series approximation b^N the following form for the numerical computations is obtained:

$$\begin{aligned}P_{\eta_N}h_{j,m} &= P_{\eta_N}\mathbb{1}_{\left[\frac{m}{2^j},\frac{m+1/2}{2^j}\right)} - P_{\eta_N}\mathbb{1}_{\left[\frac{m+1/2}{2^j},\frac{m+1}{2^j}\right)}\\ &= \exp(-\eta_N)\left(-\Phi\left(\frac{\frac{m+1}{2^j}-x}{\sqrt{\eta_N}}\right)+2\Phi\left(\frac{\frac{m+1/2}{2^j}-x}{\sqrt{\eta_N}}\right)-\Phi\left(\frac{\frac{m}{2^j}-x}{\sqrt{\eta_N}}\right)\right),\end{aligned} \tag{89}$$

where N is fixed, η_N is a constant that depends on N, $j=1,2,\dots,N$ and $m=1,2,\dots,2^j$.

For ease of notation let us denote

$$b_\eta^N := P_{\eta_N}b^N. \tag{90}$$

Theorem 9 ([4, Proposition 3.1]) *Let Assumption 1 and 2 hold. Take any (β,q) such that $\beta\in(\beta_0,\frac{1}{2})$ and $q_0\ge q>\tilde{q}\ge\tilde{q}_0$, where $\tilde{q}:=(1-\beta)^{-1}$. Then for any $\frac{1}{2}<\gamma<\gamma_0$ there is a constant $C_\gamma>0$ such that*

$$\sup_{0\le t\le T}\mathbb{E}|X_t^N - X_t| \le C_\gamma\|b_\eta^N - b\|_{\infty,H_q^{-\beta}}^{2\gamma-1} \tag{91}$$

The previous theorem shows the rate of convergence of the approximation X^N to the solution X by finding a bound for the error, and also that 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.

Remind properties (88) so that we have the inequality

very brief. Can you elaborate?

$$\|b_\eta^N - b\|_{\infty,H_q^{-\beta}} \le \|b^N - b\|_{\infty,H_q^{-\beta}} + c\eta_N^{\frac{\beta-\beta_0}{2}}\|b\|_{\infty,H_{q_0}^{-\beta_0}}. \tag{92}$$

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

Theorem 10 ([4, Proposition 3.3]) *Let Assumption 1 hold and let the sequence $(b^N)_{N\ge1}$ be defined as*

$$b^N(t) := \sum_{j=-1}^{N}\sum_{j=-2^j}^{2^j}\mu_{j,m}(t)2^{-j(-\beta-\frac{1}{q})}h_{j,m}, \tag{93}$$

m

which by construction $b^N(t)\in H^0_{\tilde{q}_0,q_0}\subset H^{-\beta_0}_{\tilde{q}_0,q_0}$. Then $(b^N)_{N\ge1}$ satisfies assumption Assumption 2 and for any $\beta\in(\beta_0,\frac{1}{2})$ it holds that

$$\|b^N - b\|_{\infty,H_2^{-\beta}} \le c2^{-(N+1)(\beta-\beta_0)}\|b\|_{\infty,H_2^{-\beta_0}}. \tag{94}$$

zero, not σ.

You need to reverse the order in which you introduce things. The reason why you can apply Theorem 8 is because we define $b = \frac{\partial}{\partial x}(B^H(\omega) \mathbb{1}_D)$ So first you should introduce fBM, then b, then say that $b \in H^{-\beta_0}_{q_0,q}$ then use Thm 8 to calculate the coeff. of the expansion.

For the computation of the coefficients $\mu_{j,m}$ of equation Eq. (94) we use the relation between the coefficients of Haar and Faber expansions in the discussion of Theorem 8 in particular with the function g being a fractional Brownian motion (fBm), which is defined as follows.

Definition 26 (Fractional Brownian motion) *A fractional Brownian motion is a Gaussian process $\{B^H(x), x \in \mathbb{R}\}$ with covariance given by*

$$\mathbb{E}[B^H(x)B^H(y)] = \frac{1}{2}(x^{2H} + y^{2H} + |x-y|^{2H}). \tag{95}$$

We can see that $B^H \in H_{q_0}^{-\beta_0}(\mathbb{R})$ and as $supp(B^H) \subset I$, $B^H \in H_{q_0}^{-\beta_0}(I)$.

Hence the coefficients have the form

$$\begin{cases} \mu_0 & = B^H(1) - B^H(0) \\ \mu_{j,m} & = -2^j \left(B^H\left(\frac{m+1}{2^j}\right) - 2B^H\left(\frac{m+1/2}{2^j}\right) + B^H\left(\frac{m}{2^j}\right)\right) \end{cases} \tag{96}$$

The fBm is particularly useful because is easy to simulate numerically.

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^N and with that coefficient we have a new SDE → not exactly - indeed you then take a further smoothing of b^N, namely $P_\eta b^N$.

$$\begin{cases} dX_t^N = P_\eta b^N(X_t^N)dt + dW_t, & t \in [0,T] \\ X_0^N = x_0. \end{cases} \tag{97}$$

It tells you about convergence

We can find a numerical solution X_t^N for the system above, and the following theorem tells us that the numerical solution coincides with the solution for the new system. NO

Theorem 11 ([4, Proposition 3.4]) *Let Assumption 1 hold and let $b^N \in C^{\frac{1}{2}}\left([0,T]; H^0_{\tilde{q}_0,q_0}\right)$ for fixed N. Then as $m \to \infty$*

$$\sup_{0\le t\le T} \mathbb{E}\left|X_t^{N,m} - X_t^N\right| \le C_2(N)m^{-1} + C_3(N)m^{-\frac{1}{2}} \tag{98}$$

where

$$C_2 := c\|b_\eta^N\|_{\infty,L^\infty}\left(1 + \|\nabla b_\eta^N\|_{\infty,L^\infty}\right), \tag{99}$$

$$C_3 := c'\left(\|\nabla b_\eta^N\|_{\infty,L^\infty} + \left[b_\eta^N\right]_{\frac{1}{2},L^\infty}\right), \tag{100}$$

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

this is what is done in Thm 11 above!

No

The second step of the numerical scheme, i.e. approximate the solution after we approximate ~~the drift,~~ is addressed in the Theorem below. This happens because the numerical solution $X_t^{N,m}$ No coincides with the approximated solution X_t^N and the latter, with the solution X_t. Being the following result essentially a combination of Theorem 9 and 11.

Theorem 12 *Let Assumption 1 hold and also b^N defined as in (93) 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*

$$\sup_{0\le t\le T} \mathbb{E}\left|X_t^{N,m} - X_t\right| \le c_\epsilon \left(m^{-\Theta_*\left(\frac{1}{2}-\beta_0\right)\left(\gamma_0-\frac{1}{2}\right)-\epsilon}\right) \tag{101}$$

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 [4] actually proved a stronger statement, because the result is strong convergence not only for the terminal time T, but for all times $t \in [0,T]$.

What do you mean by "the converse"?

2.5 BSDEs and FBSDEs

Given $\sigma_s \in L^2(\mathbb{R})$ we know that $M_t := \int_0^t \sigma_s dW_s$ is a square integrable martingale. However, the converse is not always true. Consider the notation $\mathscr{L}^p(\mathbb{R})$ to represent the set of $\{F_t\}_{t\geq 0}$ measurable processes ξ such that $\mathbb{E}[|\xi|^p] < \infty$. Let us consider $\xi \in \mathscr{L}^2(\mathbb{R})$, and $\{\mathscr{F}_t\}_{t\geq 0}$ be the filtration generated by the Wiener process W. Let us define the martingale $Y_t := \mathbb{E}[\xi|\mathscr{F}_t]$. Then we can see that there exists a unique $Z \in \mathscr{L}^2(\mathbb{R})$ such that

$$Y_t = \xi - \int_t^T Z_s dW_s. \tag{102}$$

this is consequence of Thm 13

Theorem 13 (Martingale representation theorem [13, Theorem 2.5.2]) *Let M be a martingale such that $\mathbb{E}[|M_T|^2] < \infty$, then there exists a unique $\sigma \in \mathscr{L}^2(\mathbb{R})$ such that*

$$M_t = M_0 + \int_0^t \sigma_s dW_s \tag{103}$$

The martingale representation theorem shows that the simplest case of a BSDE has indeed a solution, however we need to define more general statements to study BSDEs in general.

Definition 27 (Backward SDE) *Let $f : [0,T] \times \mathbb{R}^2 \to \mathbb{R}$ be measurable in all variables, $f_t(x,y) := f(t,x,y)$ uniformly Lipschitz continuous in (x,y), and denote $f_t^0 = f(t,0,0) \in L^2(\mathbb{R})$, finally let $Y_T = \xi \in L^2(\mathbb{R})$. Then the following is called a Backward SDE:*

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s) ds - \int_t^T Z_s dW_s, \tag{104}$$

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

In contrast with an SDE, which is a completely different object, we have here The conditions established in the definition above are enough to have a unique solution of a BSDE.

Theorem 14 ([13, Theorem 4.3.1]) *Under the assumptions of Definition 27, the BSDE has a unique solution $(Y,Z) \in L^2(\mathbb{R}) \times L^2(\mathbb{R})$.*

One can go one step further and take the terminal condition ξ to be defined by an SDE. This generalisation is called Forward-Backward SDE (FBSDE).

Definition 28 (Forward-Backward SDE) *Let $f : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ satisfy analogous conditions as in Definition 27 and X defined by (4), the following system is called a Forward-Backward SDE:*

$$\begin{cases} X_t &= X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \\ Y_t &= g(X_T) - \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \end{cases} \tag{105}$$

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

Theorem 15 *[10, Theorem 5.1] Under the assumptions of Definition 28 there exists a T_0, such that for any $T \in [0, T_0]$ a FBSDE (105) admits a unique solution (X,Y,Z).*

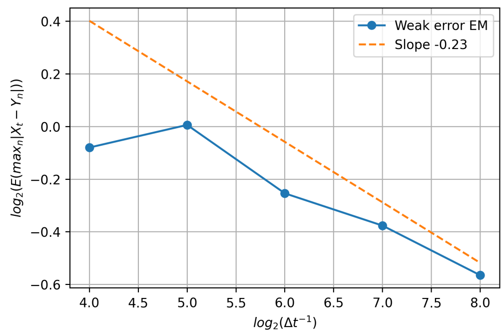

Figure 2: Log-log graphs of the estimation error of the Euler-Maruyama scheme with $N = 9$, $m \in \{2^4, ..., 2^8\}$.

3 Contributions

At the moment the same algorithm in the article by De Angelis et al. [4] has been implemented with satisfactory results, confirming the rate of convergence established in said article. Figure 2 shows the error for the approximation **...elaborate in the estimation of error as in the article...**

It is clear that Milstein scheme is not useful for the problem we have, since we require of the derivative of the diffusion coefficient, but since said coefficient is equals one, the derivative vanishes. That is why in Definition 10 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. Fortunately we have a integral representation of the drift

The 1.5 strong scheme for this problem, i.e: b a distribution and $\sigma = 1$ reads

$$\begin{aligned}\tilde{Y}_{n+1} &= \tilde{Y}_n + b(t_n, \tilde{Y}_n)\Delta t_n + Z_{n+1}\sqrt{\Delta t_n} \\ &+ \frac{1}{2}(b')(t_n, \tilde{Y}_{t_n})\left(Z_{n+1} + \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2} \\ &+ \frac{1}{2}\left((bb')(t_n, \tilde{Y}_{t_n}) + \frac{1}{2}(b'')(t_n, \tilde{Y}_{t_n})\right)(\Delta t_n)^2 \\ &+ \frac{1}{4}\left(Z_{n+1} - \frac{1}{\sqrt{3}}\tilde{Z}_{n+1}\right)(\Delta t_n)^{3/2}\end{aligned} \tag{106}$$

As we require the derivative of the coefficient

$$b^N(x) = \mu_0 h_0(x) + \sum_{j=0}^{N}\sum_{m=1}^{2^j-1}\mu_{j,m}h_{j,m}(x). \tag{107}$$

Applying the heat kernel P_{η_N} to the object above gives

$$
\begin{aligned}
P_{\eta_N} b^N(x) &= \mu_0 P_{\eta_N} h_0(x) + \sum_{j=0}^{N} \sum_{m=1}^{2^j-1} \mu_{j,m} P_{\eta_N} h_{j,m}(x) \\
&= \mu_0 \frac{1}{\sqrt{2\pi\eta_N}} \int_{-\infty}^{\infty} h_0(y) e^{-\frac{(x-y)^2}{2\eta_N}} dy \\
&\quad + \sum_{j=0}^{N} \sum_{m=1}^{2^j-1} \mu_{j,m} \frac{1}{\sqrt{2\pi\eta_N}} \int_{-\infty}^{\infty} h_{j,m}(y) e^{-\frac{(x-y)^2}{2\eta_N}} dy \\
&= \mu_0 \frac{1}{\sqrt{2\pi\eta_N}} \int_{0}^{1} e^{-\frac{(x-y)^2}{2\eta_N}} dy \\
&\quad + \sum_{j=0}^{N} \sum_{m=1}^{2^j-1} \mu_{j,m} \frac{1}{\sqrt{2\pi\eta_N}} \left(\int_{\frac{m}{2^j}}^{\frac{m+1/2}{2^j}} e^{-\frac{(x-y)^2}{2\eta_N}} dy - \int_{\frac{m+1/2}{2^j}}^{\frac{m+1}{2^j}} e^{-\frac{(x-y)^2}{2\eta_N}} dy \right).
\end{aligned}
\tag{108}
$$

The previous notation 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 $P_{\eta_N} b^N$ is a finite sum of Haar functions multiplied by some coefficients, we can rewrite $\frac{\partial}{\partial x} P_{\eta_N} b^N(x)$ and $\frac{\partial^2}{\partial x^2} P_{\eta_N} b^N(x)$, 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. Then those building blocks are as follows:

$$
\frac{\partial}{\partial x}(P_{\eta_N} \mathbb{1}_{(a,b)})(x) = \frac{1}{\sqrt{2\pi\eta_N}} \left(\exp\left(-\frac{(x-a)^2}{2\eta_N} \right) - \exp\left(-\frac{(x-b)^2}{2\eta_N} \right) \right), \tag{109}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2}(P_{\eta_N} \mathbb{1}_{(a,b)})(x) = \frac{1}{\sqrt{2\pi\eta_N}} \Bigg(&(x-a) \exp\left(-\frac{(x-a)^2}{2\eta_N} \right) - (x-b) \exp\left(-\frac{(x-b)^2}{2\eta_N} \right) \\
&+ \Phi\left(\frac{x-a}{\sqrt{\eta_N}} \right) - \Phi\left(\frac{x-b}{\sqrt{\eta_N}} \right) \Bigg).
\end{aligned}
\tag{110}
$$

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 $P_{\eta_N} b^N$ by means of multiplying by the appropriate coefficients $\mu_{j,m}$, and the Haar functions $h_{j,m}$.

4 Research plan

Following the foundations that have been established at the moment, the following steps are 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

$$
\begin{cases}
X_t &= X_0 + \int_0^t b(s, X_s) ds + \int_0^t dW_s, \\
Y_t &= g(X_T) - \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s.
\end{cases}
\tag{111}
$$

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 [3, Section 3]. It is possible to study first the special case for $f = 0$ which leaves us with an equation with the same form as (102) 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 [4], 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

$$\begin{cases} X_t &= X_0 + \int_0^t dW_s, \\ Y_t, &= g(X_T) + \int_t^T b(s, X_s) Z_s ds + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s. \end{cases} \tag{112}$$

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 [4] 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

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t dW_s^H. \tag{113}$$

Results in this direction can be seen in [2, 1] among others. comment about those two articles

5 Training record

- **Measure and Integration Theory**. Collegio Carlo Alberto. Turin, Italy *(Online)*. (Pass 65%).
- **MATH5734M: Advanced Stochastic Calculus and Applications to Finance**. University of Leeds. Spring Semester, 20 credits. (Pass 72%).
- Reading of [7, Sections 3.2, 3.3, 5.2] as complement for MATH5734M.
- **Seminar Series on Probability and Financial Mathematics**. University of Leeds. Leeds, UK *(Online)*. 2020/2021.
- **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] R. Catellier and M. Gubinelli. “Averaging along irregular curves and regularisation of ODEs”. In: *Stochastic Processes and their Applications* 126.8 (Aug. 2016), pp. 2323–2366.

[3] 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.

[4] 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.

[5] 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.

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

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

[8] 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.

[9] Giovanni Leoni. *A first course in Sobolev spaces*. Second edition. Graduate studies in mathematics volume 181. Providence, Rhode Island: American Mathematical Society, 2017. 734 pp.

[10] 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.

[11] Eckhard Platen. “An introduction to numerical methods for stochastic differential equations”. In: *Acta Numerica* 8 (Jan. 1999). Publisher: Cambridge University Press, pp. 197–246.

[12] Hans Triebel. *Bases in function spaces, sampling, discrepancy, numerical integration*. EMS tracts in mathematics 11. OCLC: ocn640095745. Zürich: European Mathematical Society, 2010. 296 pp.

[13] 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.