

CONVERGENCE RATE OF NUMERICAL SCHEME FOR SDES WITH A DISTRIBUTIONAL DRIFT IN BESOV SPACE

LUIS MARIO CHAPARRO JÁQUEZ¹, ELENA ISSOGLIO^{2,*} AND JAN PALCZEWSKI¹

Abstract. This paper is concerned with numerical solutions of one-dimensional SDEs with the drift being a generalised function in the spatial variable, in particular being a $\frac{1}{2}$ -Hölder continuous function of time taking values in a Hölder-Zygmund space $\mathcal{C}^{-\gamma}$ of negative order $-\gamma < 0$. We design an Euler–Maruyama numerical scheme and prove its convergence, obtaining an upper bound for the strong L^1 convergence rate. We finally implement the scheme and discuss the results obtained.

Mathematics Subject Classification. Primary 65C30, Secondary 60H35, 65C20, 46F99.

Received October 13, 2023. Accepted July 10, 2025.

1. INTRODUCTION

This paper is concerned with the Euler–Maruyama scheme and its rate of convergence for a one-dimensional stochastic differential equation (SDE) of the form

$$X_t = x + \int_0^t b(s, X_s) ds + W_t, \quad (1)$$

where $(W_t)_{t \geq 0}$ is a Brownian motion and the drift $b(t, \cdot)$ belongs to the space of Schwartz distributions $\mathcal{S}'(\mathbb{R})$ for all $t \in [0, T]$. More precisely, the map $t \mapsto b(t)$ is $\frac{1}{2}$ -Hölder continuous for any $t \in [0, T]$ with values in the Hölder-Zygmund space $\mathcal{C}^{-\gamma}$ of negative order $-\gamma < 0$, which we denote by $b \in C_T^{1/2} \mathcal{C}^{-\gamma}$. For a precise definition of these spaces see Section 2.1 below.

SDEs with distributional coefficients have been studied by several authors in different settings and with different noises, starting from the early 2000s with [2, 11, 12] and then in recent years by Flandoli *et al.* [13], Delarue and Diel [10], Cannizzaro and Chouk [5], Chaudru de Raynal and Menozzi [7] and Issoglio and Russo [17]. In all these works the drift is a distribution and the authors investigate theoretical questions of existence and uniqueness of solution, without exploring numerical aspects. The specific setting we consider here is the one studied in [17], where the authors formulate the notion of solution to (1) as a suitable martingale problem (c.f. Sect. 2.4). SDE (1) is solved in [17] in any general dimension d , and the notion of solution is intrinsically a weak solution, since it is formulated as a martingale problem. Here we restrict ourselves to dimension 1 in

Keywords and phrases. Distributional drift, Euler–Maruyama scheme, rate of convergence, Besov space, stochastic differential equation, numerical scheme.

¹ School of Mathematics, University of Leeds, Leeds LS2 9JT, UK.

² Department of Mathematics “G. Peano”, University of Turin, Via Carlo Alberto 10, 10123 Torino, Italy.

*Corresponding author elena.issoglio@unito.it

which case there is a strong solution X (see Lem. 2.4); the strong convergence studied in this paper can only be defined for strong solutions. The research of weak convergence for multi-dimensional SDEs of the above type is left for the future.

The first results on numerical schemes for SDEs date back to the 1980s, see the book by Kloeden and Platen [26] for the case of smooth coefficients. On the other hand, numerical schemes for SDEs with low-regularity coefficients is an active area of research, but almost all contributions deal with SDEs with coefficients that are at least functions. We refer to the introduction of [20] and of [8] for a list of other relevant papers and a short summary of techniques used for numerical schemes with low-regularity coefficients. In [24] the authors study SDEs with additive noise in dimension 1 and with the drift in Sobolev spaces $W_2^\gamma \cap L^1 \cap L^\infty$ for $\gamma \in (0, 1)$ and prove a strong rate of convergence of $\frac{1+\gamma}{2} \wedge \frac{3}{4}$ for the Euler–Maruyama scheme. Notice that the drift has a positive Sobolev regularity of γ , hence it is a possibly discontinuous function. Paper [8] extends this result to any dimension d and removes the bound $\frac{3}{4}$. Indeed, they prove a strong L^p -rate of $\frac{1+\gamma}{2}$ when the drift b is time-homogeneous and an element of the homogeneous Sobolev space $\dot{W}_{d\nu}^\gamma$ for $\gamma \in (0, 1)$ and the diffusion coefficient is the identity. In [8] they also consider multiplicative noise and prove a strong L^p -rate of $\frac{1}{2}$ for non-unitary C^2 diffusion coefficient with a bounded measurable drift. Another relevant paper is [19], where the authors deal with the case of L^q – L^p drift and unitary diffusion. They prove that the weak error of the Euler–Maruyama scheme is $\frac{\alpha}{2}$, where α is the distance from the singularity $\alpha := 1 - \left(\frac{d}{p} + \frac{2}{q}\right)$, where d is the dimension of the problem. They also conjecture that their methods of proof should produce a rate of $\frac{1+\gamma}{2}$ for a time-inhomogeneous drift that belong to $C_T^{\frac{\gamma}{2}} C^\gamma$. A different line of research investigates discontinuous drifts and possibly degenerate diffusion coefficients, where the discontinuities lie on finitely many points or hypersurfaces, see [22, 23, 25, 27] for more details, and [29] for a review.

Only a few works deal with numerical schemes for SDEs with distributional coefficients. In [9], the SDE is like (1) but the drift b belongs to a different distribution space, namely to the fractional Sobolev space of negative order $b \in C_T^\kappa H_{(1-\gamma)^{-1}, q}^{-\gamma}$ for $\kappa \in (\frac{1}{2}, 1)$, $\gamma \in (0, \frac{1}{4})$ and $q \in (4, \frac{1}{\gamma})$. The authors obtain a strong L^1 -rate of convergence depending on γ which vanishes as $-\gamma$ approaches the boundary $-\frac{1}{4}$, and tends to $\frac{1}{6}$ for when $-\gamma$ approaches 0 (*i.e.*, when it approaches measurable drifts). In [14], the authors study SDEs in d -dimensions with drifts in negative Besov spaces $B_p^{-\gamma}$, and the noise is a fractional Brownian motion with Hurst index $H \in (0, \frac{1}{2})$. They require that $1 - \frac{1}{2H} < -\gamma - \frac{d}{p} < 0$, *i.e.*, the roughness of the drift is compensated by the roughness of the noise. The case $p = \infty$, $d = 1$ and $H = \frac{1}{2}$ would correspond to our case, but this combination of parameters violates the above condition as the left-hand side becomes 0. Techniques used in [14] cannot be easily extended to our case. Indeed, for their proofs of convergence the authors rely on the well-known fact that a rougher noise gives more regularity to the solution, hence allowing for a rougher drift coefficient b (or a higher dimension).

In this paper, we set up a two-step numerical scheme. The first step is to regularise the distributional drift with the action of the heat semigroup, which gives a smooth function and allows us to use Schauder estimates to control the approximation error bounds for the solution of the SDE with smoothed drift. Proving this step is the bulk of the paper. The second step is to bound the error of the Euler–Maruyama scheme, which requires ad-hoc estimates (rather than standard EM estimates that can be found in most of the literature) to be able to control the constant in front of the rate in terms of the properties of the smoothed drift. To do so, we borrow ideas and results from [9], but we still have to prove a delicate L^1 -bound of the local time of the error process (see Lem. 4.6). Notice that we consider the L^1 strong error, and not the more common L^2 error, because the diffusion coefficient of an auxiliary process (Y_t) which is used to define a virtual solution to SDE (1) is only Hölder continuous; hence, the quadratic variation of the square of the difference between the solution and its approximation is of a lower order than the square of the approximation error, so a Gronwall lemma argument cannot be applied, see Remark 4.5 for more details. Finally, we link the smoothing parameter in Step 1 and the time step in Step 2 to obtain a one-step scheme and its convergence rate.

This paper is organised as follows. In Section 2 we set up the problem and the notation, recalling all relevant theoretical tools such as Hölder–Zygmund spaces, the heat kernel and semigroup, Schauder estimates and the

notion of virtual solution to SDE (1). In Section 3 we describe the numerical scheme and state the main result (Thm. 3.4) which provides a convergence rate in terms of the regularity parameter γ (Cor. 3.6). Section 4 contains the proof of the building bloc of the main theoretical result (Prop. 3.3), which is a bound on the difference between the solution to the original SDE (1) and its approximation after smoothing the drift b . Here is where we make use of the bound on the local time borrowed from [9]. Finally in Section 5, we describe a numerical implementation of the scheme and analyse numerical results. It is striking to see that the empirical convergence rate seems to be $\frac{1}{2} - \frac{\gamma}{2}$, which would be the extension of the rate found in [8] if they allowed for negative regularity index (and hence for distributional drifts). A straightforward application of their techniques is not possible and further investigations in this direction, for example using the stochastic sewing lemma introduced in [21], are left for future research.

2. PRELIMINARIES

2.1. Notation

For a function $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ that is sufficiently smooth, we denote by f_t the partial derivative with respect to t , by f_x the partial derivative with respect to x , and by f_{xx} the second partial derivative with respect to x . For a function $g : \mathbb{R} \rightarrow \mathbb{R}$ sufficiently smooth we denote its derivative by g' .

We now recall some useful definitions and facts from the literature. First of all, let $\mathcal{S}(\mathbb{R})$ be the space of Schwartz functions on \mathbb{R} and $\mathcal{S}'(\mathbb{R})$ the space of tempered distributions. We denote $(\cdot)^\wedge$ and $(\cdot)^\vee$ the Fourier transform and inverse Fourier transform on \mathcal{S} respectively, extended to \mathcal{S}' in the standard way. For $\gamma \in \mathbb{R}$, the Hölder-Zygmund space is defined by

$$\mathcal{C}^\gamma(\mathbb{R}) = \left\{ f \in \mathcal{S}' : \|f\|_\gamma := \sup_{j \geq -1} 2^{j\gamma} \left\| (\phi_j \hat{f})^\vee \right\|_{L^\infty} < \infty \right\}, \tag{2}$$

where $(\phi_j)_j$ is any partition of unity. The Hölder-Zygmund space $\mathcal{C}^\gamma(\mathbb{R})$ is also known as Besov space $B_{\infty, \infty}^\gamma(\mathbb{R})$. For more details see [1, 30]. To shorten notation we write \mathcal{C}^γ instead of $\mathcal{C}^\gamma(\mathbb{R})$. Note that if $\gamma \in \mathbb{R}^+ \setminus \mathbb{N}$ the space above coincides with the classical Hölder space. These spaces will be used widely in the paper, so we recall the norms that we will use in the paper. If $\gamma \in (0, 1)$, the classical γ -Hölder norm

$$\|f\|_{L^\infty} + \sup_{\substack{x \neq y \\ |x-y| < 1}} \frac{|f(x) - f(y)|}{|x - y|^\gamma} \tag{3}$$

is an equivalent norm in \mathcal{C}^γ . If $\gamma \in (1, 2)$ an equivalent norm is

$$\|f\|_{L^\infty} + \|f'\|_{L^\infty} + \sup_{\substack{x \neq y \\ |x-y| < 1}} \frac{|f'(x) - f'(y)|}{|x - y|^{\gamma-1}}. \tag{4}$$

We will write $C_T \mathcal{C}^\gamma := C([0, T]; \mathcal{C}^\gamma)$ with the norm

$$\|f\|_{C_T \mathcal{C}^\gamma} := \sup_{t \in [0, T]} \|f(t)\|_{\mathcal{C}^\gamma}.$$

We will also use a family of equivalent norms $\|\cdot\|_{C_T \mathcal{C}^\gamma}^{(\rho)}$, for $\rho \geq 0$, given by

$$\|f\|_{C_T \mathcal{C}^\gamma}^{(\rho)} := \sup_{t \in [0, T]} e^{-\rho(T-t)} \|f(t)\|_{\mathcal{C}^\gamma}.$$

Indeed, it is easy to see that

$$\|f\|_{C_T \mathcal{C}^\gamma} \leq e^{\rho T} \|f\|_{C_T \mathcal{C}^\gamma}^{(\rho)}. \tag{5}$$

For any given $\gamma \in \mathbb{R}$ we denote by $\mathcal{C}^{\gamma+}$ and $\mathcal{C}^{\gamma-}$ the following spaces

$$\mathcal{C}^{\gamma+} := \cup_{\alpha > \gamma} \mathcal{C}^\alpha, \tag{6}$$

$$\mathcal{C}^{\gamma-} := \cap_{\alpha < \gamma} \mathcal{C}^\alpha. \tag{7}$$

Similarly, we also write $C_T \mathcal{C}^{\gamma+} := \cup_{\alpha > \gamma} C_T \mathcal{C}^\alpha$.

The following bound in Hölder-Zygmund spaces will be useful later, and it is known as the Bernstein inequality.

Lemma 2.1 (Bernstein inequality). *For any $\gamma \in \mathbb{R}$, there is $c > 0$ such that*

$$\|f'\|_\gamma \leq c \|f\|_{\gamma+1}, \quad f \in \mathcal{C}^{\gamma+1}. \tag{8}$$

Let $\kappa \in (0, 1)$. We denote by $C_T^\kappa L^\infty$ the space of κ -Hölder continuous functions with values in L^∞ (the space of bounded functions), and by $C_T^\kappa C_b^1$ the space of κ -Hölder continuous functions from $[0, T]$ with values in the space of C^1 functions which are bounded and have a bounded derivative. We define the following norms and seminorms for $g \in C_T^\kappa L^\infty$:

$$\|g\|_{\infty, L^\infty} := \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}} |g(t, x)| \tag{9}$$

and

$$[g]_{\kappa, L^\infty} := \sup_{t, s \in [0, T], t \neq s} \frac{\|g(t) - g(s)\|_{L^\infty}}{|t - s|^\kappa}. \tag{10}$$

Notice that if $g \in C_T^\kappa C_b^1$ then $g_x \in C_T^\kappa L^\infty$.

We finish this section by introducing an asymptotic relation between functions. For functions f, g defined on an unbounded subset of \mathbb{R}^+ , we write $f(x) = o(g(x))$ if $\lim_{x \rightarrow \infty} |f(x)|/|g(x)| = 0$.

2.2. Standing assumption

The following assumption will hold throughout the paper.

Assumption 2.2. *We have $b \in C_T^{1/2} \mathcal{C}^{(-\hat{\beta})+}$ for some $0 < \hat{\beta} < 1/2$.*

We will derive our numerical scheme and prove its convergence under the above assumption, but solutions to the SDE (1) exist under a weaker assumption $b \in C_T \mathcal{C}^{-\hat{\beta}}$ as explained in detail in Section 2.4 below, and this fact will be exploited in the derivation of the rate of convergence. Although solutions to SDE (1) exist in higher dimensions, we work in dimension 1 because in this case one can construct a strong solution, see Lemma 2.4, which is fundamental to the definition of strong convergence error.

2.3. Heat kernel and heat semigroup

We will use heat kernel smoothing to derive a sequence of approximating SDEs. Here we introduce notation and provide background information about the action of the heat semigroup on elements of \mathcal{C}^γ .

The function

$$p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \tag{11}$$

is called the heat kernel and it is the fundamental solution to the heat equation. The operator acting as a convolution of the heat kernel with a (generalised) function is called heat semigroup, and it is denoted by P_t : for any $g \in \mathcal{S}$ we have

$$(P_t g)(y) = (p_t * g)(y) = \int_{\mathbb{R}} p_t(x) g(y - x) dx. \tag{12}$$

The semigroup P_t can be extended to \mathcal{S}' by duality. For a distribution $\varrho \in \mathcal{S}'$, the convolution and derivative commute as mentioned in Section 5.3 from [28] and Remark 2.5 from [16], that is,

$$(p_t * \varrho)' = p_t' * \varrho = p_t * \varrho'. \tag{13}$$

This fact is useful for efficient construction of regularised SDEs when $b = B_x$ for some function $B \in C_T^{1/2} \mathcal{C}^{(1-\hat{\beta})+}$ as we do in the numerical example studied in Section 5.

We recall the so-called Schauder estimates which quantify the effect of heat semigroup smoothing.

Lemma 2.3 (Schauder estimates). *For any $\gamma \in \mathbb{R}$, and $\theta \geq 0$ there is a positive constant c such that for any $f \in \mathcal{C}^\gamma$*

$$\|P_t f\|_{\gamma+2\theta} \leq ct^{-\theta} \|f\|_\gamma. \tag{14}$$

Also, for any $\gamma \in \mathbb{R}$, and $0 < \theta < 1$ there is a positive constant c such that for any $f \in \mathcal{C}^{\gamma+2\theta}$ then

$$\|P_t f - f\|_\gamma \leq ct^\theta \|f\|_{\gamma+2\theta}. \tag{15}$$

For a proof of the result above see Theorem 2.10 from [18] with $d = 0$ and $N = 1$.

2.4. Existence of solutions to the SDE

From [16, 17] we recall the main results on construction, existence and uniqueness of solutions to SDE (1) under the following assumption:

$$b \in C_T \mathcal{C}^{(-\beta)+}, \quad \beta \in (0, 1/2). \tag{16}$$

There exist two equivalent notions of solution to SDE (1): virtual solutions and solutions *via* martingale problem. The formulation *via* martingale problem is omitted and the reader is referred to [17]. We describe below the formulation *via* virtual solutions, because it is particularly suited to our analysis of the numerical scheme.

For $\lambda > 0$, let us consider a Kolmogorov-type PDE

$$\begin{cases} u_t + \frac{1}{2}u_{xx} + bu_x = \lambda u - b \\ u(T) = 0 \end{cases} \tag{17}$$

with the solution understood in a mild sense, *i.e.*, as a solution to the integral equation

$$u(t) = \int_t^T P_{s-t} (u_x(s)b(s)) \, ds - \int_t^T P_{s-t} (\lambda u(s) - b(s)) \, ds, \quad \forall t \in [0, T]. \tag{18}$$

It is shown in Theorem 4.7 from [16] that a solution u exists in $C_T \mathcal{C}^{(2-\beta)-}$ and is unique in $C_T \mathcal{C}^{(1+\beta)+}$. Hence, u_x is α -Hölder continuous for any $\alpha < 1 - \beta$. By Proposition 4.13 of [16], we have $\|u_x\|_\infty < 1/2$ for λ large enough.

From now on we fix λ large enough. By Proposition 4.13 of [16], the mapping

$$\phi(t, x) := x + u(t, x) \tag{19}$$

is invertible in the space dimension, and we denote this space-inverse by $\psi(t, \cdot)$. Consider now a weak solution Y to SDE

$$Y_t = y_0 + \lambda \int_0^t u(s, \psi(s, Y_s)) \, ds + \int_0^t (u_x(s, \psi(s, Y_s)) + 1) \, dW_s. \tag{20}$$

We say that $X_t := \psi(t, Y_t)$ is a *virtual solution*¹ to SDE (1). Clearly, $Y_t = \phi(t, X_t) = X_t + u(t, X_t)$ solves (20) when X is a weak solution to the equation

$$X_t = x + u(0, x) - u(t, X_t) + \lambda \int_0^t u(s, X_s) ds + \int_0^t (u_x(s, X_s) + 1) dW_s, \quad (21)$$

so any solution (X_t) of (21) is also a virtual solution to SDE (1). We also note that the virtual solution does not depend on λ thanks to the links between virtual solutions and martingale problem developed in [17]; the reader is referred to the aforementioned paper for a complete presentation of those links.

We complete this section by showing that the virtual solution of SDE (1) is a strong solution, in the sense that the solution Y to SDE (20) is a strong solution.

Lemma 2.4. *Under condition (16), there exists process (Y_t) on the original probability space with the Brownian motion (W_t) that satisfies (20), and this process is unique up to modifications. Hence, there is a strong solution (X_t) of SDE (1).*

Proof. The drift is Lipschitz in y uniformly in t and the volatility is α -Hölder in y uniformly in t , for some $\alpha > \frac{1}{2}$. This allows us to use arguments in the proof of Chapter IV, Theorem 3.2 from [15] (see also the corollary following it) to claim the existence and uniqueness of a strong solution Y to SDE (20). \square

3. THE NUMERICAL SCHEME AND MAIN RESULTS

Our numerical scheme for SDE (1) is based on two approximations. The first one replaces the distributional drift with a sequence of functional drifts so that the solutions of the respective SDEs converge to the solution of the original SDE. Subsequently, the approximating SDEs are simulated with an Euler–Maruyama scheme. We will then balance errors coming from the approximation of the drift and the discretisation of time to maximise the rate of convergence.

We regularise the drift by applying the heat semigroup. For a real number $N > 0$, we define

$$b^N = P_{\frac{1}{N}} b. \quad (22)$$

By an application of Lemma 2.3 with parameter $t = 1/N$ we know that $b^N(t, \cdot) \in \mathcal{C}^\gamma$ for any $\gamma > 0$ and for all $t \in [0, T]$, thus we also have $b^N(t, \cdot) \in C_b^1$, $t \in [0, T]$, hence, it is Lipschitz continuous in the spatial variable. Hence, the SDE

$$dX_t^N = b^N(t, X_t^N) dt + dW_t \quad (23)$$

has a unique strong solution. Now we write the standard EM scheme for the above SDE. For $m \in \mathbb{N}$ we take an equally spaced partition $t_k = t_k^m = kT/m$, $k = 0, \dots, m$, of the interval $[0, T]$. We define

$$k(t) = k^m(t) = \max \{k : t_k \leq t\}, \quad t \in [0, T]. \quad (24)$$

Consider an Euler–Maruyama approximation of (X^N) with m time steps

$$X_t^{Nm} = x + \int_0^t b^N(t_{k(s)}, X_{t_{k(s)}}^{Nm}) ds + W_t, \quad t \in [0, T]. \quad (25)$$

We first obtain a bound on the strong error between the approximation (X^{Nm}) and the process (X^N) with explicit dependence of constants on the properties of the drift b^N . This is needed in order to balance the smoothing *via* choice of N and the number of time steps m to optimise the convergence rate of the Euler–Maruyama approximation to the true solution (X) . Following arguments in the proof of Proposition 3.4 from [9] we obtain the following result.

¹We borrow here the term virtual solution from [13], where the authors use an analogous equation for Y to define the solution of an SDE with distributional drift b in a fractional Sobolev space. In [17] the authors instead define the solution *via* martingale problem, but the equivalence with the notion of virtual solutions follows from their Theorem 3.9.

Proposition 3.1. *Assume that $b^N \in C_T^{1/2}L^\infty \cap L_T^\infty C_b^1$. Then*

$$\sup_{0 \leq t \leq T} \mathbb{E} [|X_t^{Nm} - X_t^N|] \leq A^N m^{-1} + B^N m^{-1/2},$$

where

$$\begin{aligned} A^N &= \|b^N\|_{\infty, L^\infty} \left(1 + \|b_x^N\|_{\infty, L^\infty}\right), \\ B^N &= \|b_x^N\|_{\infty, L^\infty} + [b^N]_{\frac{1}{2}, L^\infty}. \end{aligned}$$

Corollary 3.2. *Under Assumption 2.2, the condition of Proposition 3.1 is satisfied and for any $\epsilon > 0$ there is a constant $c > 0$ such that*

$$A^N \leq cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b\|_{C_T C^{-\hat{\beta}}} \left(1 + N^{\frac{\epsilon+\hat{\beta}+1}{2}} \|b\|_{C_T C^{-\hat{\beta}}}\right)$$

and

$$B^N \leq cN^{\frac{\epsilon+\hat{\beta}+1}{2}} \|b\|_{C_T C^{-\hat{\beta}}} + cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b\|_{C_T^{\frac{1}{2}} C^{-\hat{\beta}}}.$$

Proof. We first show that $b^N \in C_T^{\frac{1}{2}}L^\infty \cap L_T^\infty C_b^1$. In the proof, a constant c may change from line to line.

We apply Lemma 2.3 with ϵ from the statement of the corollary, $\theta = \frac{\epsilon+\hat{\beta}}{2}$ and $\gamma = -\hat{\beta}$ to get

$$\begin{aligned} \|b^N(t, \cdot) - b^N(s, \cdot)\|_{L^\infty} &= \|P_{\frac{1}{N}}(b(t, \cdot) - b(s, \cdot))\|_{L^\infty} \leq c \|P_{\frac{1}{N}}(b(t, \cdot) - b(s, \cdot))\|_{C^\epsilon} \\ &\leq cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b(t, \cdot) - b(s, \cdot)\|_{C^{-\hat{\beta}}}, \end{aligned}$$

where the first inequality is by (3). Hence,

$$[b^N]_{\frac{1}{2}, L^\infty} \leq cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b\|_{C_T^{\frac{1}{2}} C^{-\hat{\beta}}}. \quad (26)$$

By the same arguments applied to $b^N(t, \cdot)$, we obtain $\|b^N(t, \cdot)\|_{L^\infty} \leq cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b(t, \cdot)\|_{C^{-\hat{\beta}}}$, which yields

$$\|b^N\|_{\infty, L^\infty} \leq \|b^N\|_{C_T C^\epsilon} \leq cN^{\frac{\epsilon+\hat{\beta}}{2}} \|b\|_{C_T C^{-\hat{\beta}}}. \quad (27)$$

Bounds (26) and (27) allow us to conclude that $b^N \in C_T^{\frac{1}{2}}L^\infty$.

It remains to show that $b_x^N(t, \cdot)$ exists and is bounded uniformly in $t \in [0, T]$. The derivative $b_x^N(t, \cdot)$ is well defined for all $t \in [0, T]$ because $b^N \in C_T C^\gamma$ for all $\gamma > 0$. Using the equivalent norm (3) and the Bernstein's inequality (8) we have

$$\|b_x^N\|_{\infty, L^\infty} \leq \|b_x^N\|_{C_T C^\epsilon} \leq c \|b^N\|_{C_T C^{\epsilon+1}} \leq cN^{\frac{\epsilon+\hat{\beta}+1}{2}} \|b\|_{C_T C^{-\hat{\beta}}}, \quad (28)$$

where the last inequality holds by Lemma 2.3. Inequalities (27) and (28) show that $b^N \in L_T^\infty C_b^1$. It remains to insert the bounds derived above into formulas for A^N and B^N from Proposition 3.1. \square

Before stating the main result, we state another auxiliary result whose proof is the main content of Section 4 below.

Proposition 3.3. *Under Assumption 2.2, for any $\hat{\alpha} \in (1/2, 1 - \hat{\beta})$ and any $\beta \in (\hat{\beta}, 1/2)$, there is a constant c such that*

$$\sup_{t \in [0, T]} \mathbb{E}[|X_t^N - X_t|] \leq c \|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1}$$

for all N sufficiently large so that $\|b^N - b\|_{C_T C^{-\beta}} < 1$.

The approximation error of our numerical scheme comes from two sources: the time discretisation error of the Euler–Maruyama scheme, which depends on m and N , and the smoothing error coming from replacing b with b^N in the SDE, which depends on N . We will now show how to balance those two sources of errors and bound the resulting convergence rate. To this end, we parametrise N in terms of m and choose this parametrisation to be of the form $N(m) = m^\eta$ for some $\eta > 0$. We denote

$$\hat{X}^m := X^{N(m)m}$$

and consider the strong error

$$\Upsilon(m) = \sup_{0 \leq t \leq T} \mathbb{E} \left[\left\| \hat{X}_t^m - X_t \right\| \right].$$

Take any $\epsilon \in (0, 1/2 - \hat{\beta})$, $\beta \in (\hat{\beta}, 1/2)$ and $\hat{\alpha} \in (1/2, 1 - \hat{\beta})$. By triangle inequality, we have

$$\begin{aligned} \Upsilon(m) &\leq \sup_{0 \leq t \leq T} \mathbb{E} \left[\left\| \hat{X}_t^m - X_t^{N(m)} \right\| \right] + \sup_{0 \leq t \leq T} \mathbb{E} \left[\left\| X_t^{N(m)} - X_t \right\| \right] \\ &\leq A^{N(m)} m^{-1} + B^{N(m)} m^{-1/2} + c \left\| b^{N(m)} - b \right\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1} \\ &\leq c \left[m^{\eta \frac{\epsilon + \hat{\beta}}{2}} \|b\|_{C_T C^{-\hat{\beta}}} \left(1 + m^{\eta \frac{\epsilon + \hat{\beta} + 1}{2}} \|b\|_{C_T C^{-\hat{\beta}}} \right) m^{-1} \right. \\ &\quad \left. + \left(m^{\eta \frac{\epsilon + \hat{\beta} + 1}{2}} \|b\|_{C_T C^{-\hat{\beta}}} + m^{\eta \frac{\epsilon + \hat{\beta}}{2}} \|b\|_{C_T^{\frac{1}{2}} C^{-\hat{\beta}}} \right) m^{-\frac{1}{2}} \right. \\ &\quad \left. + m^{-\eta \frac{\beta - \hat{\beta}}{2} (2\hat{\alpha} - 1)} \|b\|_{C_T C^{-\hat{\beta}}}^{2\hat{\alpha} - 1} \right], \end{aligned} \tag{29}$$

where in the second inequality we used Propositions 3.1, 3.3, and in the last inequality Corollary 3.2, and the following estimate arising from Lemma 2.3:

$$\left\| b^{N(m)}(t, \cdot) - b(t, \cdot) \right\|_{C^{-\beta}} \leq c N(m)^{-\frac{\beta - \hat{\beta}}{2}} \|b(t, \cdot)\|_{C^{-\hat{\beta}}}, \quad t \in [0, T].$$

Since all the norms appearing on the right hand side are finite and independent of m , they can be absorbed by a constant and we have

$$\Upsilon \leq c \left[m^{\eta \frac{\epsilon + \hat{\beta}}{2} - 1} + m^{\eta \frac{2\epsilon + 2\hat{\beta} + 1}{2} - 1} + m^{\eta \frac{\epsilon + \hat{\beta} + 1}{2} - \frac{1}{2}} + m^{\eta \frac{\epsilon + \hat{\beta}}{2} - \frac{1}{2}} + m^{-\eta \frac{(\beta - \hat{\beta})(2\hat{\alpha} - 1)}{2}} \right]. \tag{30}$$

Before proceeding further, we optimise the last term in β and $\hat{\alpha}$ to maximise its rate of decrease of the last term in m . Recalling the constraints for β and $\hat{\alpha}$, the product $(\beta - \hat{\beta})(2\hat{\alpha} - 1)$ is maximised for $\hat{\alpha} \approx 1 - \hat{\beta}$ and $\beta \approx 1/2$. We take $\hat{\alpha} = 1 - \hat{\beta} - \epsilon$ and $\beta = 1/2 - \epsilon$ which yields the value $2(1/2 - \hat{\beta} - \epsilon)^2$. The last term of (30) takes the form

$$m^{-\eta(\frac{1}{2} - \hat{\beta} - \epsilon)^2}.$$

The monotonicity of the remaining four terms in (30) depends on η . For the scheme to converge, we need to make sure that they all decrease, which is guaranteed if $\eta \frac{2\epsilon + 2\hat{\beta} + 1}{2} - 1 < 0$ and $\eta \frac{\epsilon + \hat{\beta} + 1}{2} - \frac{1}{2} < 0$. This leads to the constraint

$$0 < \eta < \frac{1}{\epsilon + \hat{\beta} + 1}. \tag{31}$$

At this point we have to find the optimal value of η . It is easy to see that the slowest decreasing term within the first four terms of (30) is $m^{\eta \frac{\epsilon + \hat{\beta} + 1}{2} - \frac{1}{2}}$. To balance Euler–Maruyama error measured by the first four terms and the error of approximating b with $b^{N(m)}$ in the last term we equate the rates

$$\eta \frac{\epsilon + \hat{\beta} + 1}{2} - \frac{1}{2} = -\eta \left(\frac{1}{2} - \hat{\beta} - \epsilon \right)^2. \tag{32}$$

This leads to

$$\eta = \frac{1}{\epsilon + \hat{\beta} + 1 + 2 \left(\frac{1}{2} - \hat{\beta} - \epsilon\right)^2}. \tag{33}$$

Inserting this expression for η into the right-hand side of (32) we obtain the rate of convergence of our scheme as

$$\left(\frac{\epsilon + \hat{\beta} + 1}{\left(\frac{1}{2} - \hat{\beta} - \epsilon\right)^2} + 2 \right)^{-1}.$$

We summarise the above derivation in the following theorem.

Theorem 3.4. *Let Assumption 2.2 hold and fix $\epsilon \in (0, 1/2 - \hat{\beta})$. By taking*

$$N(m) = m^{\frac{1}{\epsilon + \hat{\beta} + 1 + 2\left(\frac{1}{2} - \hat{\beta} - \epsilon\right)^2}},$$

the strong error of our scheme is bounded as follows:

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[\left\| X_t^{N(m)m} - X_t \right\| \right] \leq cm^{-\left(\frac{\epsilon + \hat{\beta} + 1}{\left(\frac{1}{2} - \hat{\beta} - \epsilon\right)^2} + 2\right)^{-1}}, \tag{34}$$

where constant c depends on ϵ and the drift b .

Remark 3.5. Note that the bound stated in Proposition 3.3 holds for N large enough so that $\|b^{N(m)} - b\|_{C_T C^{-\beta}} < 1$. Formally, the error for small $N(m)$, *i.e.*, small m , could be incorporated into the constant c in (34) when the condition $\|b^{N(m)} - b\|_{C_T C^{-\beta}} < 1$ is not satisfied. However, when doing numerical estimation of the convergence rate (see Sect. 5) we have to pick m large enough so that $\|b^{N(m)} - b\|_{C_T C^{-\beta}} < 1$, otherwise the numerical estimate of the rate could not be reliable.

Corollary 3.6. *Our construction allows us to achieve any strong convergence rate strictly smaller than*

$$\lim_{\epsilon \downarrow 0} \left(\frac{\epsilon + \hat{\beta} + 1}{\left(\frac{1}{2} - \hat{\beta} - \epsilon\right)^2} + 2 \right)^{-1} = \left(\frac{\hat{\beta} + 1}{\left(\frac{1}{2} - \hat{\beta}\right)^2} + 2 \right)^{-1} =: r(\hat{\beta}).$$

Remark 3.7. We rewrite $r(\hat{\beta}) = \frac{(\frac{1}{2} - \hat{\beta})^2}{2(\frac{1}{2} - \hat{\beta})^2 + \hat{\beta} + 1}$ from Corollary 3.6. Figure 3 in Section 5 displays this function over the range of $\hat{\beta} \in (0, \frac{1}{2})$. We take a closer look at the limits as $\hat{\beta}$ approaches 0 and $\frac{1}{2}$:

- $\lim_{\hat{\beta} \downarrow 0} r(\hat{\beta}) = \frac{1}{6}$. This corresponds to $b \in C_T C^{0+}$, which is comparable to the case of a measurable function $b \in C_T C^0$.
- $\lim_{\hat{\beta} \uparrow \frac{1}{2}} r(\hat{\beta}) = 0$, so when the roughness of the drift approaches the boundary $\frac{1}{2}$, the scheme deteriorates. This is expected due to the nature of the estimate in Proposition 3.3 that we use.

A direct comparison of our rate with other rates found in the literature is only possible in the case of measurable drifts with Brownian noise, which has been treated in [4]. The rate obtained there is $\frac{1}{2}$, so our estimate is clearly not optimal because we obtain $\frac{1}{6}$. However, the technique they use is different from ours, and in particular they use the stochastic sewing lemma from [21] to drive up the rate, but it seems it is not straightforward to apply techniques used in [4] to our setting.

Paper [14] considers time-homogeneous distributional drifts and, between others, covers drifts $b \in C^{-\beta}$ for some $\beta > 0$, but the driving noise is a fractional Brownian motion with Hurst index $H \in (0, \frac{1}{2})$. For $\beta \in (0, \frac{1}{2H} - 1)$, the rate of convergence is $\frac{1}{2(1+\beta)} - \epsilon$ for any $\epsilon > 0$, which excludes the case $H = \frac{1}{2}$ of Brownian noise which would lead to a rate of $1/2$ for measurable functions (since β would also approach 0).

In [9], the authors consider SDE (1) with the drift $b \in C_T^\kappa H_{(1-\beta)^{-1},q}^{-\beta}$ for $\kappa \in (\frac{1}{2}, 1)$, $\beta \in (0, \frac{1}{4})$ and $q \in (4, \frac{1}{\beta})$. The space $H_{(1-\beta)^{-1},q}^{-\beta}$ is a fractional Sobolev space of negative regularity $-\beta$. Their β is closely related to our $\hat{\beta}$ as fractional Sobolev spaces $H_{p,q}^s$ are related, albeit different, to Besov spaces $B_{p,q}^s$, see e.g., [30]. In both cases the index s is a measure of smoothness of the elements of the space. When $\beta \downarrow 0$, the convergence rate of the numerical scheme in [9] tends to $\frac{1}{6}$, as in our case. However, when $\beta \rightarrow \frac{1}{4}$ the convergence rate in [9] vanishes, while we get $r(\frac{1}{4}) = \frac{1}{22}$.

Finally, we would like to mention that in our numerical study in Section 5 we obtain a numerical estimate of the convergence rate equal to $\frac{1+(-\hat{\beta})}{2}$, which is the equivalent of the rate $\frac{1+\gamma}{2}$ found in [4] for positive γ , i.e., for γ -Hölder continuous drifts. This suggests that the rate from the latter paper could apply in our setting, but it would require a different approach and is left as future work.

4. PROOF OF PROPOSITION 3.3

Let us introduce an auxiliary process Y^N , which is the (weak) solution of the SDE

$$Y_t^N = y_0^N + \lambda \int_0^t u^N(s, \psi^N(s, Y_s^N)) ds + \int_0^t (u_x^N(s, \psi^N(s, Y_s^N)) + 1) dW_s, \tag{35}$$

where u^N is the unique solution to the regularised Kolmogorov equation

$$\begin{cases} u_t^N + \frac{1}{2}u_{xx}^N + b^N u_x^N = \lambda u^N - b^N \\ u^N(T) = 0, \end{cases} \tag{36}$$

and $\psi^N(t, x)$ is the space-inverse of

$$\phi^N(t, x) := x + u^N(t, x), \tag{37}$$

which exists by Proposition 4.16 of [16] for λ large enough. Since $b^N \rightarrow b$ in $C_T C^{-\hat{\beta}}$ by Lemma 2.3 and Assumption 2.2, we can apply Lemma 4.19 of [16]. This lemma and its proof provide key properties of the above system and its relation to (17) and (20). The solution of the regularised Kolmogorov equation (36) enjoys a bound $\|u_x^N\|_\infty < 1/2$ for λ large enough; the constant λ can be chosen independently of N (but it depends on the drift b of the original SDE), and from now on we fix such λ . We also have $u^N \rightarrow u$ and $u_x^N \rightarrow u_x$ uniformly on $[0, T] \times \mathbb{R}^d$.

This section is devoted to the proof of Proposition 3.3 with the overall structure inspired by De Angelis *et al.* [9]. The main idea of the proof is to rewrite the solutions X^N and X in their equivalent virtual formulations and thus study the error between the auxiliary processes Y^N defined in (35) and Y defined in (20). This transformation has the advantage that the SDEs for Y^N and Y are classical SDEs with strong solutions, see Lemma 2.4. After writing the difference $X^N - X$ in terms of $Y^N - Y$ we apply Itô's formula to $|Y^N - Y|$. We will control the stochastic and Lebesgue integrals by $u^N - u$ and $\psi^N - \psi$, which in turn are bounded by some function of $b^N - b$. The term involving the local time at 0 of $Y^N - Y$ is handled in Lemma 4.6.

In the remainder of this section, we fix $\beta \in (\hat{\beta}, 1/2)$ as in the statement of Proposition 3.3. Since $\beta \in (\hat{\beta}, 1/2)$, we have $b \in C_T C^{(-\beta)^+}$ by Assumption 2.2. Notice that the solutions u and u^N obtained when taking b as an element of $C_T^{1/2} C^{(-\hat{\beta})^+}$ are the same as when viewing the same b as an element of $C_T C^{(-\beta)^+}$. We obviously have $u, u^N \in C_T C^{1+\alpha}$ for any $\alpha < 1 - \beta$ but also for any $\alpha < 1 - \hat{\beta}$.

Lemma 4.1. *There are ρ_0 and c such that for any $\rho \geq \rho_0$ and for any $\alpha < 1 - \beta$ we have*

$$\|u - u^N\|_{C_T C^{1+\alpha}}^{(\rho)} \leq 2c \|b - b^N\|_{C_T C^{-\beta}} \left(\|u\|_{C_T C^{1+\alpha}}^{(\rho)} - 1 \right) \rho^{\frac{\alpha+\beta-1}{2}}.$$

Proof. The bound in the statement of the lemma forms the main part of the proof of Lemma 4.17 from [16] in the special case when the terminal condition is zero and $g^N = b^N$, $g = b$. □

The parameter ρ and the exact dependence of the bound on it is not important for our arguments, so we may fix ρ that satisfies the conditions of the above lemma. However, for completeness of the presentation, we will mention the dependence on ρ_0 and ρ in results below.

Using Lemma 4.1 we can derive a uniform bound on the L^∞ -norm of the difference $u(t) - u^N(t)$ and $u_x(t) - u_x^N(t)$.

Lemma 4.2. *For any $\rho \geq \rho_0$, where ρ_0 is from Lemma 4.1, there is $\kappa > 0$ such that for all $t \in [0, T]$*

$$\|u(t) - u^N(t)\|_{L^\infty} + \|u_x(t) - u_x^N(t)\|_{L^\infty} \leq \kappa \|b - b^N\|_{C_T C^{-\beta}}.$$

The constant κ depends also on c from Lemma 4.1 and on T and u .

Proof. We recall that $u, u^N \in C_T \mathcal{C}^{1+\alpha}$ and $u_x, u_x^N \in C_T \mathcal{C}^\alpha$ for any $\alpha < 1 - \beta$. Fix such α . Recall also that the norm in \mathcal{C}^γ for $\gamma \in (0, 1)$ is given by (3) and in \mathcal{C}^γ for $\gamma \in (1, 2)$ is given by (4). Using this together with the Bernstein inequality (8) we get for all $t \in [0, T]$ that

$$\|u(t) - u^N(t)\|_{L^\infty} + \|u_x(t) - u_x^N(t)\|_{L^\infty} \leq c' \|u(t) - u^N(t)\|_{\mathcal{C}^{1+\alpha}},$$

for some $c' > 0$. We conclude using (5) and Lemma 4.1. The inequality in the statement of the lemma holds with constant κ given by

$$\kappa = c' e^{\rho T} 2c \left(\|u\|_{C_T \mathcal{C}^{1+\alpha}}^{(\rho)} - 1 \right) \rho^{\frac{\alpha+\beta-1}{2}}.$$

□

Next we derive a bound for the difference $\psi - \psi^N$, where we recall that the two functions are the space-inverses of ϕ and ϕ^N defined in (19) and (37).

Lemma 4.3. *Take κ from Lemma 4.2. We have*

$$\sup_{(t,y) \in [0,T] \times \mathbb{R}} |\psi(t, y) - \psi^N(t, y)| \leq 2\kappa \|b - b^N\|_{C_T C^{-\beta}}. \tag{38}$$

Proof. Let us first recall that $\|u_x\|_\infty \leq \frac{1}{2}$, see the discussion at the beginning of this section. Hence u is $\frac{1}{2}$ -Lipschitz. Using this we have for any $x, x' \in \mathbb{R}$

$$\begin{aligned} |\phi(t, x) - \phi(t, x')| &= |(x + u(t, x)) - (x' + u(t, x'))| \\ &\geq |x - x'| - |u(t, x) - u(t, x')| \\ &\geq |x - x'| - \frac{1}{2}|x - x'| \\ &\geq \frac{1}{2}|x - x'|. \end{aligned} \tag{39}$$

Insert $x = \psi(t, y)$ and $x' = \psi^N(t, y)$ for some $y \in \mathbb{R}$, to get the bound

$$|\phi(t, \psi(t, y)) - \phi(t, \psi^N(t, y))| \geq \frac{1}{2} |\psi(t, y) - \psi^N(t, y)|.$$

This implies the first inequality below

$$\begin{aligned} |\psi(t, y) - \psi^N(t, y)| &\leq 2 |\phi(t, \psi(t, y)) - \phi(t, \psi^N(t, y))| \\ &= 2 |\phi^N(t, \psi^N(t, y)) - \phi(t, \psi^N(t, y))| \\ &= 2 |u^N(t, \psi^N(t, y)) - u(t, \psi^N(t, y))|, \end{aligned} \tag{40}$$

where we used $\phi^N(t, \psi^N(t, y)) = y = \phi(t, \psi(t, y))$ and the definition of ϕ and ϕ^N . Thus

$$|\psi(t, y) - \psi^N(t, y)| \leq 2 \|u(t) - u^N(t)\|_{L^\infty} \leq 2\kappa \|b - b^N\|_{C_T C^{-\beta}},$$

having used Lemma 4.2 in the last inequality.

□

Finally we derive a bound for $|u^N(s, \psi^N(s, y)) - u(s, \psi(s, y'))|$ and for $|u_x^N(s, \psi^N(s, y)) - u_x(s, \psi(s, y'))|$.

Lemma 4.4. *For any $\alpha < 1 - \hat{\beta}$ and any $y, y' \in \mathbb{R}$, the following bounds are satisfied*

$$|u^N(s, \psi^N(s, y')) - u(s, \psi(s, y))| \leq 2\kappa \|b^N - b\|_{C_T C^{-\beta}} + |y - y'|, \quad (41)$$

and

$$\begin{aligned} |u_x^N(s, \psi^N(s, y')) - u_x(s, \psi(s, y))| &\leq \kappa \|b - b^N\|_{C_T C^{-\beta}} + 2^\alpha \kappa^\alpha \|u\|_{C_T C^{1+\alpha}} \|b^N - b\|_{C_T C^{-\beta}}^\alpha \\ &\quad + |u_x(s, \psi(s, y)) - u_x(s, \psi(s, y'))|, \end{aligned} \quad (42)$$

where κ is from Lemma 4.2.

Proof. We start with (41). We rewrite the left-hand side as

$$\begin{aligned} |u^N(s, \psi^N(s, y')) - u(s, \psi(s, y))| &\leq |u^N(s, \psi^N(s, y')) - u(s, \psi^N(s, y'))| \\ &\quad + |u(s, \psi^N(s, y')) - u(s, \psi(s, y'))| \\ &\quad + |u(s, \psi(s, y')) - u(s, \psi(s, y))|. \end{aligned} \quad (43)$$

Using Lemma 4.2, we bound the first term above by

$$|u^N(s, \psi^N(s, y')) - u(s, \psi^N(s, y'))| \leq \|u^N(s) - u(s)\|_{L^\infty} \leq \kappa \|b - b^N\|_{C_T C^{-\beta}}.$$

The second term in (43) is bounded using the fact that $\|u_x\|_\infty \leq \frac{1}{2}$ (see the discussion at the beginning of this section) and Lemma 4.3:

$$|u(s, \psi^N(s, y')) - u(s, \psi(s, y'))| \leq \frac{1}{2} |\psi^N(s, y') - \psi(s, y')| \leq \kappa \|b - b^N\|_{C_T C^{-\beta}}.$$

For the third term in (43), we use again that $u(t, \cdot)$ is $\frac{1}{2}$ -Lipschitz

$$|u(s, \psi(s, y')) - u(s, \psi(s, y))| \leq \frac{1}{2} |\psi(s, y') - \psi(s, y)| \leq |y - y'|,$$

where the last inequality follows from the fact that $\psi(t, \cdot)$ is 2-Lipschitz, which we argue as follows. By the definition of ψ as the space inverse of ϕ , we have $z = \psi(t, z) + u(t, \psi(t, z))$. Hence

$$|\psi(t, y) - \psi(t, y')| \leq |u(t, \psi(t, y)) - u(t, \psi(t, y'))| + |y - y'| \leq \frac{1}{2} |\psi(t, y) - \psi(t, y')| + |y - y'|,$$

where the last inequality uses again that $u(t, \cdot)$ is $\frac{1}{2}$ -Lipschitz. Combining the above estimates proves (41).

Let us now prove (42). Similarly as above we write

$$\begin{aligned} |u_x^N(s, \psi^N(s, y')) - u_x(s, \psi(s, y))| &\leq |u_x^N(s, \psi^N(s, y')) - u_x(s, \psi^N(s, y'))| \\ &\quad + |u_x(s, \psi^N(s, y')) - u_x(s, \psi(s, y'))| \\ &\quad + |u_x(s, \psi(s, y')) - u_x(s, \psi(s, y))|. \end{aligned} \quad (44)$$

The first term is bounded using Lemma 4.2 as follows

$$|u_x^N(s, \psi^N(s, y')) - u_x(s, \psi^N(s, y'))| \leq \|u_x^N(s) - u_x(s)\|_{L^\infty} \leq \kappa \|b - b^N\|_{C_T C^{-\beta}}.$$

Since $b \in C_T C^{-\hat{\beta}}$ and $\alpha \in (1/2, 1 - \hat{\beta})$ then $u \in C_T C^{1+\alpha}$, so by the Bernstein inequality (Lem. 2.1) we have $u_x(t) \in C_T C^\alpha$ with $\|u_x\|_{C_T C^\alpha} \leq \|u\|_{C_T C^{1+\alpha}}$. Recalling the norm (3) in C^α , we conclude that $u_x(t)$ is α -Hölder continuous with constant $\|u\|_{C_T C^{1+\alpha}}$. This allows us to bound the second term in (44) as

$$\begin{aligned} |u_x(s, \psi^N(s, y')) - u_x(s, \psi(s, y'))| &\leq \|u\|_{C_T C^{1+\alpha}} |\psi^N(s, y') - \psi(s, y')|^\alpha \\ &\leq \|u\|_{C_T C^{1+\alpha}} 2^\alpha \kappa^\alpha \|b - b^N\|_{C_T C^{-\beta}}^\alpha, \end{aligned}$$

where the last inequality uses Lemma 4.3. This concludes the proof. \square

Remark 4.5. Note that in our main result we find the strong error in L^1 instead of the most commonly used estimates in L^2 , the reason for this is illustrated here. Let us apply Itô’s formula to $(Y_t^N - Y_t)^2$, integrate from 0 to t and take expectation. Apply further Lemma 4.4 and recall that $\|u_x\|_\infty < 1/2$, ψ is 2-Lipschitz, and u_x is α -Hölder continuous for any $\alpha < 1 - \beta$. We get

$$\begin{aligned} \mathbb{E} (Y_t^N - Y_t)^2 - (Y_0^N - Y_0)^2 &\leq 2\mathbb{E} \int_0^t |Y_s^N - Y_s|^2 + 4\kappa \|b^N - b\|_{C_T C^{-\beta}} \mathbb{E} \int_0^t |Y_s^N - Y_s| \, ds \\ &\quad + t \left(\kappa \|b^N - b\|_{C_T C^{-\beta}} + 2^\alpha \kappa^\alpha \|u\|_{C_T C^{-\beta}} \|b^N - b\|_{C_T C^{-\beta}}^\alpha \right)^2 \\ &\quad + \mathbb{E} \int_0^t c^2 2^{2\alpha} |Y_s^N - Y_s|^{2\alpha} \, ds. \end{aligned}$$

Note that the second and last terms in the inequality have the difference $|Y_s^N - Y_s|$ in smaller powers than 2 (the latter since $\alpha < 1$), so they decrease slower than the term on the left-hand side and Gronwall inequality cannot be applied. The power 2α in the last term follows from the fact that u_x is α -Hölder continuous which is an inherent feature of the construction of the virtual solution to (1).

In order to bound the L^1 norm of the difference $X^N - X$, we need to bound the local time of $Y^N - Y$ at zero. To this end, we recall the definition of a local time of a continuous semimartingale Z and a bound for this local time established in [9]. We define the local time of Z at 0 by

$$L_t^0(Z) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^t \mathbb{1}_{\{|Z| \leq \epsilon\}} \, d\langle Z \rangle_s, \quad t \geq 0. \tag{45}$$

Lemma 4.6 ([9], Lem. 5.1). *For any $\epsilon \in (0, 1)$ and any real-valued, continuous semi-martingale Z we have*

$$\mathbb{E}[L_t^0(Z)] \leq 4\epsilon - 2\mathbb{E} \left[\int_0^t \left(\mathbb{1}_{\{Z_s \in (0, \epsilon)\}} + \mathbb{1}_{\{Z_s \geq \epsilon\}} e^{1-Z_s/\epsilon} \right) \, dZ_s \right] + \frac{1}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{Z_s > \epsilon\}} e^{1-Z_s/\epsilon} \, d\langle Z \rangle_s \right]. \tag{46}$$

The next proposition is a bound for the local time of $Y^N - Y$, solutions to the SDEs (20) and (35). Due to the application of Itô’s formula to $|Y^N - Y|$, this is a key bound in the proof of the L^1 -convergence of X^N to X .

Proposition 4.7. *For any $\alpha \in (1/2, 1 - \hat{\beta})$ and N such that $\|b^N - b\|_{C_T C^{-\beta}} < 1$ we have*

$$\mathbb{E} [L_t^0(Y^N - Y)] \leq A \mathbb{E} \left[\int_0^t |Y_s^N - Y_s| \, ds \right] + B \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha-1} + o\left(\|b^N - b\|_{C_T C^{-\beta}}^{2\alpha-1}\right), \tag{47}$$

for some constants $A, B > 0$.

Proof. This proof follows the same steps as the proof of Proposition 5.4 from [9] with differences coming from the spaces that b and b^N belong to. It is provided for the reader’s convenience.

Recall that Y_t, Y_t^N are strong solutions of (20) and (35) respectively, see Lemma 2.4. Their difference satisfies

$$\begin{aligned} Y_t^N - Y_t &= (y_0^N - y_0) + \lambda \int_0^t (u^N(s, \psi^N(s, Y_s^N)) - u(s, \psi(s, Y_s))) \, ds \\ &\quad + \int_0^t (u_x^N(s, \psi^N(s, Y_s^N)) - u_x(s, \psi(s, Y_s))) \, dW_s. \end{aligned} \tag{48}$$

We apply Lemma 4.6 to $Y_t^N - Y_t$ for $\epsilon \in (0, 1)$:

$$\mathbb{E} [L_t^0 (Y^N - Y)] \leq 4\epsilon - 2\lambda \mathbb{E} \left[\int_0^t \left(\mathbb{1}_{\{Y_s^N - Y_s \in (0, \epsilon)\}} + \mathbb{1}_{\{Y_s^N - Y_s \geq \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} \right) (u^N (s, \psi^N (s, Y_s^N)) - u (s, \psi (s, Y_s))) \, ds \right] \tag{49}$$

$$+ \frac{1}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} (u_x^N (s, \psi^N (s, Y_s^N)) - u_x (s, \psi (s, Y_s)))^2 \, ds \right], \tag{50}$$

where the expectation of the integral with respect to the Brownian motion (W_t) is zero thanks to the fact that u_x and u_x^N are bounded uniformly in N .

Notice that if $Y_s^N - Y_s \geq \epsilon$, then $e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} \leq 1$. Hence, (49) is bounded from above by

$$4\lambda \mathbb{E} \left[\int_0^t (u (s, \psi (s, Y_s)) - u^N (s, \psi^N (s, Y_s^N))) \, ds \right] \leq 4\lambda \left(2\kappa \|b^N - b\|_{C_T C^{-\beta}} t + \mathbb{E} \left[\int_0^t |Y_s^N - Y_s| \, ds \right] \right),$$

where the last inequality is by Lemma 4.4.

Now for (50), we use again the observation that $\mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} \leq 1$, the estimate (42) from Lemma 4.4 choosing $\alpha' \in (\alpha, 1 - \hat{\beta})$, and the inequality $(x_1 + x_2 + x_3)^2 \leq 3(x_1^2 + x_2^2 + x_3^2)$ to get the bound

$$\begin{aligned} & \frac{1}{\epsilon} \mathbb{E} \int_0^t \left(3\kappa^2 \|b - b^N\|_{C_T C^{-\beta}}^2 + 3 \cdot 2^{2\alpha'} \kappa^{2\alpha'} \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'} \|u\|_{C_T C^{1+\alpha'}}^2 \right) \, ds \\ & + \frac{1}{\epsilon} \mathbb{E} \int_0^t 3 \mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x (s, \psi (s, Y_s^N)) - u_x (s, \psi (s, Y_s))|^2 \, ds \\ & \leq \frac{3t}{\epsilon} \|b^N - b\|_{C_T C^{-\beta}} \left(\kappa^2 \|b^N - b\|_{C_T C^{-\beta}} + (2\kappa)^{2\alpha'} \|u\|_{C_T C^{1+\alpha'}}^2 \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'-1} \right) \\ & + \frac{3}{\epsilon} \mathbb{E} \left(\int_0^t \mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x (s, \psi (s, Y_s^N)) - u_x (s, \psi (s, Y_s))|^2 \, ds \right). \end{aligned}$$

Putting the above estimates together yields the following bound for the expectation of the local time:

$$\begin{aligned} \mathbb{E} [L_t^0 (Y^N - Y)] & \leq 4\epsilon + 4\lambda \left(2\kappa \|b^N - b\|_{C_T C^{-\beta}} t + \mathbb{E} \left[\int_0^t |Y_s^N - Y_s| \, ds \right] \right) \\ & + \frac{3t}{\epsilon} \|b^N - b\|_{C_T C^{-\beta}} \left(\kappa^2 \|b^N - b\|_{C_T C^{-\beta}} + (2\kappa)^{2\alpha'} \|u\|_{C_T C^{1+\alpha'}}^2 \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'-1} \right) \\ & + \frac{3}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x (s, \psi (s, Y_s^N)) - u_x (s, \psi (s, Y_s))|^2 \, ds \right]. \end{aligned} \tag{51}$$

We take $\epsilon = \|b^N - b\|_{C_T C^{-\beta}} < 1$ and choose $\zeta \in (0, 1)$ such that $\alpha'\zeta > 1/2$. Recall that $u_x(s, \cdot)$ is α' -Hölder continuous with the constant $\|u\|_{C_T C^{1+\alpha'}}$, as $\alpha' < 1 - \hat{\beta}$, (see Sect. 2.4) and $\psi(s, \cdot)$ is 2-Lipschitz, so

$$|u_x (s, \psi (s, Y_s^N)) - u_x (s, \psi (s, Y_s))| \leq \|u\|_{C_T C^{1+\alpha'}} 2^{\alpha'} |Y_s^N - Y_s|^{\alpha'}.$$

This bound and the observation that $\epsilon^\zeta > \epsilon$ yield

$$\begin{aligned} & \frac{3}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{Y_s^N - Y_s > \epsilon\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x(s, \psi(s, Y_s^N)) - u_x(s, \psi(s, Y_s))|^2 ds \right] \\ & \leq \frac{3}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{\epsilon < Y_s^N - Y_s \leq \epsilon^\zeta\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x(s, \psi(s, Y_s^N)) - u_x(s, \psi(s, Y_s))|^2 ds \right] \\ & \quad + \frac{3}{\epsilon} \mathbb{E} \left[\int_0^t \mathbb{1}_{\{Y_s^N - Y_s > \epsilon^\zeta\}} e^{1 - \frac{Y_s^N - Y_s}{\epsilon}} |u_x(s, \psi(s, Y_s^N)) - u_x(s, \psi(s, Y_s))|^2 ds \right] \\ & \leq \frac{3}{\epsilon} t \|u\|_{C_T C^{1+\alpha'}}^2 2^{2\alpha'} \epsilon^{2\alpha'\zeta} + \frac{3}{\epsilon} t e^{1-\epsilon^\zeta-1}, \end{aligned}$$

where in the last inequality we used that $\|u_x\|_{L^\infty} \leq 1/2$. We insert this bound into (51) and recall that $\epsilon = \|b^N - b\|_{C_T C^{-\beta}}$ to obtain

$$\begin{aligned} \mathbb{E} [L_t^0(Y^N - Y)] & \leq (4 + 8\lambda\kappa t + 3\kappa^2 t) \|b^N - b\|_{C_T C^{-\beta}} + 4\lambda \left(\mathbb{E} \left[\int_0^t |Y_s^N - Y_s| ds \right] \right) \\ & \quad + 3t \left((2\kappa)^{2\alpha'} \|u\|_{C_T C^{1+\alpha'}}^2 \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'-1} \right) \\ & \quad + 3t \|u\|_{C_T C^{1+\alpha'}}^2 2^{2\alpha'} \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'\zeta-1} + \frac{3t}{\|b^N - b\|_{C_T C^{-\beta}}} e^{1 - \|b^N - b\|_{C_T C^{-\beta}}^{\zeta-1}} \\ & \leq 3T \|u\|_{C_T C^{1+\alpha'}}^2 2^{2\alpha'} \|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'\zeta-1} + 4\lambda \left(\mathbb{E} \left[\int_0^T |Y_s^N - Y_s| ds \right] \right) + o \left(\|b^N - b\|_{C_T C^{-\beta}}^{2\alpha'\zeta-1} \right). \end{aligned}$$

Taking $\zeta = \alpha/\alpha'$ completes the proof. \square

Proof of Proposition 3.3. Our arguments are inspired by the proof of Proposition 3.1 from [9]. For our arguments we fix ρ that satisfies the conditions of Lemma 4.1 and denote by κ the constant from Lemma 4.2. By the definition of ψ, ψ^N , Lemma 4.3, and the fact that ψ is 2-Lipschitz, we have

$$\begin{aligned} |X_t^N - X_t| & = |\psi^N(t, Y_t^N) - \psi(t, Y_t)| \leq |\psi^N(t, Y_t^N) - \psi(t, Y_t^N)| + |\psi(t, Y_t^N) - \psi(t, Y_t)| \\ & \leq 2\kappa \|b^N - b\|_{C_T C^{-\beta}} + 2 |Y_t^N - Y_t|. \end{aligned} \quad (52)$$

For the term $|Y^N - Y|$ we use Itô-Tanaka's formula and take expectations of both sides:

$$\begin{aligned} \mathbb{E} [|Y_t^N - Y_t|] & = \mathbb{E} |Y_0^N - Y_0| + \mathbb{E} \left[\frac{1}{2} L_t^0(Y^N - Y) \right] \\ & \quad + \lambda \mathbb{E} \left[\int_0^t \text{sgn}(Y^N - Y) (u^N(s, \psi^N(s, Y_s^N)) - u(s, \psi(s, Y_s))) ds \right], \end{aligned}$$

where the stochastic integral disappears due to the boundedness of u_x and u_x^N . For the first term we use that $Y_0 = x + u(0, x)$, $Y_0^N = x + u^N(0, x)$, and Lemma 4.1 with any $0 < \alpha < 1 - \beta$, to conclude that

$$|u^N(0, x) - u(0, x)| \leq \|u^N - u\|_{C_T C^{1+\alpha}}^{(\rho)} \leq o \left(\|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1} \right),$$

where we used that $\|b^N - b\|_{C_T C^{-\beta}} = o(\|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1})$, because $2\hat{\alpha} - 1 < 1$ and $\|b^N - b\|_{C_T C^{-\beta}} < 1$ as assumed in the statement of the proposition. The second term is bounded by Proposition 4.7 with $\alpha = \hat{\alpha}$. For the third term we employ a bound from Lemma 4.4 again with $\alpha = \hat{\alpha}$ and the fact that $|\text{sgn}(x)| \leq 1$. In summary, we obtain

$$\mathbb{E} [|Y_t^N - Y_t|] \leq o \left(\|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1} \right) + (A/2 + \lambda) \mathbb{E} \left[\int_0^t |Y_s^N - Y_s| ds \right] + B/2 \|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1}.$$

Finally, using Gronwall's lemma we get the following bound

$$\mathbb{E} [|Y_t^N - Y_t|] \leq B/2 \|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1} e^{(A/2+\lambda)t} + o\left(\|b^N - b\|_{C_T C^{-\beta}}^{2\hat{\alpha}-1}\right).$$

We take expectation of both sides of (52), take the supremum over $t \in [0, T]$ and insert the above bound to conclude. \square

5. NUMERICAL IMPLEMENTATION

In this section we describe an implementation of our numerical scheme and analyse the results obtained. Our implementation was done in the Python programming language and can be found in [6]. We recall that numerical implementation proceeds in two steps. First we approximate the drift b with b^N as in (23). Then we apply the classical Euler–Maruyama scheme (25) to approximate the SDE with drift b^N . For the numerical example we will consider a time homogeneous drift, *i.e.*, $b \in \mathcal{C}^{(-\hat{\beta})+}$.

Since the drift b is a Schwartz distribution in $\mathcal{C}^{(-\hat{\beta})+}$, producing a numerical approximation of it poses some challenges. We cannot simply discretise b and then convolve it with the heat kernel to get b^N as in (22), see also (12), because the discretization of a distribution is not meaningful. Instead, we use the fact that an element of $\mathcal{C}^{(-\hat{\beta})+}$ can be obtained as the distributional derivative of a function h in $\mathcal{C}^{(-\hat{\beta}+1)+}$, and that the derivative commutes with the heat kernel as explained in (13). Indeed, since $b \in \mathcal{C}^{-\hat{\beta}+\epsilon}$ for some $\epsilon > 0$ and $-\hat{\beta} \in (-1/2, 0)$, we have $\gamma := -\hat{\beta} + \epsilon + 1 > 0$ and $h \in \mathcal{C}^\gamma$ is a function. Using (13) allows us to discretise the function h first, then convolve it with the heat kernel, which has a smoothing effect, and finally take the derivative.

Without loss of generality we will pick a function h that is constant outside of a compact set, which implies that the distribution h' is supported on the same compact. For example one can choose $L > 0$ large enough so that the solution (X_t) of the SDE with the drift $b\mathbb{1}_{[-L, L]}$ replacing b will, with a large probability, stay within the interval $[-L, L]$ for all times $t \in [0, T]$, rendering numerically irrelevant the fact that the drift b has been cut outside the compact support $[-L, L]$.

The simplest example of a drift $b \in \mathcal{C}^{\gamma-1}$, $\gamma \in (1/2, 1)$, would be given by the derivative of the locally γ -Hölder continuous function $|x|^\gamma$, smoothly cut outside the compact set $[-L, L]$. However, this function is not differentiable in $x = 0$ only. Instead, for our numerical implementation we consider $h \in \mathcal{C}^\gamma$, which is capable of fully encompassing the rough nature of the drift $b = h'$. This can be obtained if the function h has a ‘rough’ behaviour in (almost) each point of the interval $[-L, L]$. In this section, we take as h a transformation of a trajectory of a fractional Brownian motion (fBm)² $(B_x^H)_{x \geq 0}$. Indeed, it is known that \mathbb{P} -almost all paths of a fBm B^H are locally γ -Hölder continuous for any $\gamma < H$, but paths are almost nowhere differentiable, see Section 1.7 from [3]. We construct a compactly supported function h as follows. We take a path $(B_x^H(\omega))_{x \in [0, 2L]}$ with Hurst index $H = -\hat{\beta} + 1 + 2\epsilon \in (1/2, 1)$ for some small $\epsilon > 0$, which is thus locally γ -Hölder continuous with $\gamma = -\hat{\beta} + 1 + \epsilon$. The constraint $H \in (1/2, 1)$ ensures that $\hat{\beta} \in (0, 1/2)$, as needed in Assumption 2.2. We define function $g : \mathbb{R} \rightarrow \mathbb{R}$ as

$$g(x) = \left(B_x^H - \frac{B_{2L}^H}{2L} x \right) \mathbb{1}_{\{x \in [0, 2L]\}}(x),$$

which ensures that $g(0) = g(2L) = 0$. This transformation, inspired by the so called Brownian bridge, ensures that g is continuous and keeps the same regularity as the fBm B^H . Finally, for convenience, we translate g so that it is supported in $[-L, L]$ rather than $[0, 2L]$. This is done for the sake of symmetry, since we choose to start our SDE from an initial condition close to 0. We choose L large enough so that the paths of (X_t) stay in the strip $[-L, L]$ with high probability up to our terminal time T . In summary, the function h is constructed as

$$h(x) = \left(B_{x+L}^H(\omega) - \frac{B_{2L}^H(\omega)}{2L} (x - L) \right) \mathbb{1}_{\{x \in [-L, L]\}}(x), \quad (53)$$

²A fractional Brownian motion $(B_x^H)_{x \geq 0}$ with Hurst parameter $H \in (0, 1)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a centered Gaussian stochastic process with covariance given by $\mathbb{E}(B_x^H B_y^H) = \frac{1}{2}(|x|^{2H} + |y|^{2H} - |x - y|^{2H})$. When $H = \frac{1}{2}$ we recover a Brownian motion.

and we have $h \in \mathcal{C}^\gamma$ with $\gamma = -\hat{\beta} + \varepsilon + 1$, so that $b := h' \in \mathcal{C}^{(-\hat{\beta})+}$, and both are supported on $[-L, L]$. By a slight abuse of notation we denote $B^H(x) = h(x)$ so that $b = (B^H)'$.

To compute b^N we apply the semigroup $P_{1/N}$ to b , as in (22), which is equivalent to a convolution with the heat kernel $p_{1/N}$. Since the derivative commutes with the convolution as recalled in (13) we get

$$b^N(x) := (P_{1/N}b)(x) = (p_{1/N} * (B^H)')(x) = (p'_{1/N} * B^H)(x) = \int_{-\infty}^{\infty} p'_{1/N}(y)B^H(x - y) dy. \tag{54}$$

As the derivative of the heat kernel is $p'_{1/N}(y) = -\frac{y}{1/N}p_{1/N}(y)$, we have

$$b^N(x) = - \int_{-\infty}^{\infty} B^H(x - y) \frac{y}{1/N} p_{1/N}(y) dy. \tag{55}$$

In order to approximate the above integral numerically, we create a uniform discretisation of the interval $[-L, L]$ with $2M + 1$ points, denoted by $\mathfrak{M} = \{x_{-M}, \dots, x_M\}$, with the mesh size $\delta := \frac{L}{M}$. We sample the fBm $B^H(x)$ in $x \in \mathfrak{M}$ and extend it to the real line as a piecewise constant function \hat{B}^H as follows:

$$\hat{B}^H(z) = \sum_{j=-M}^{M-1} B^H(x_j) \mathbb{1}_{\{z \in [x_j - \frac{\delta}{2}, x_j + \frac{\delta}{2}]\}}(z), \quad z \in \mathbb{R}. \tag{56}$$

Inserting (56) into (55) and performing the change of variable $z = x_i - y$ we obtain a numerical approximation of b^N at any point $x_i \in \mathfrak{M}$:

$$\begin{aligned} b^N(x_i) &\approx - \int_{-\infty}^{\infty} \hat{B}^H(x_i - y) \frac{y}{1/N} p_{1/N}(y) dy \\ &= - \int_{-\infty}^{\infty} \sum_{j=-M}^M B^H(x_j) \mathbb{1}_{\{z \in [x_j - \frac{\delta}{2}, x_j + \frac{\delta}{2}]\}} \frac{x_i - z}{1/N} p_{1/N}(x_i - z) dz \\ &= - \sum_{j=-M}^M B^H(x_j) \int_{x_j - \delta/2}^{x_j + \delta/2} \frac{x_i - z}{1/N} p_{1/N}(x_i - z) dz \\ &= - \sum_{j=-M}^M B^H(x_j) \int_{-\delta/2}^{\delta/2} \frac{x_i - x_j - z'}{1/N} p_{1/N}(x_i - x_j - z') dz', \end{aligned} \tag{57}$$

where we performed a second change of variable $z' = z - x_j$ in the last equality. Let us denote the integral in the last line of (57) as

$$\mathcal{I}^N(y) := \int_{-\delta/2}^{\delta/2} \frac{y - z'}{1/N} p_{1/N}(y - z') dz', \tag{58}$$

so that equation (57) reads

$$b^N(x_i) \approx - \sum_{j=-M}^M B^H(x_j) \mathcal{I}^N(x_i - x_j) =: - (B^H * \mathcal{I}^N)(x_i), \tag{59}$$

where $*$ denotes the discrete convolution between vectors $(B^H(x_i))_{i=-M}^M$ and $(\mathcal{I}^N(x_i))_{i=-2M}^{2M}$. The latter vector needs to be formally defined for a δ -discretisation of the interval $[-2L, 2L]$, however, in practice, the values of $\mathcal{I}^N(y)$ decrease quickly to 0 as $|y|$ increases (this is illustrated in Figure 1 for two values of $1/N$), so it is enough to use the values of $(\mathcal{I}^N(x_i))$ for $x_i \in \mathfrak{M}$. We will use the above approximation (59) in our numerical computations.

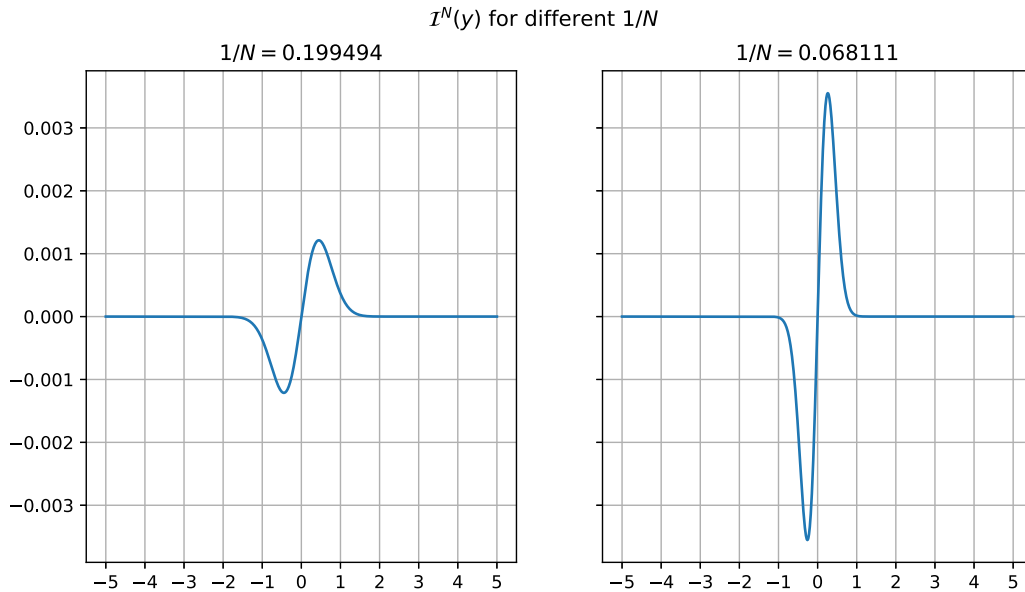


FIGURE 1. From left to right: $\mathcal{I}^N(y)$ for $y \in [-5, 5]$ and $1/N = 0.199494, 0.068111$ respectively.

TABLE 1. Minimum amount of points $2M + 1$ needed in the discretisation \mathfrak{M} when $m = 2^{12}$ and $L = 5$. This varies according to the variance $1/N$, which in turn is a function of $\hat{\beta}$.

| $\hat{\beta}$ | 10^{-6} | $1/16$ | $1/8$ | $1/4$ | $1/2$ |
|-----------------------------------|-----------|----------|----------|----------|----------|
| $1/N(m) = m^{-\eta(\hat{\beta})}$ | 0.001288 | 0.001398 | 0.001489 | 0.001768 | 0.003906 |
| $2M + 1$ | 279 | 269 | 261 | 239 | 161 |

At this stage a few remarks regarding the size of \mathfrak{M} are in order. The drift b^N is implemented with a discrete convolution given in equation (59), where one of the convoluting terms is \mathcal{I}^N , which is depicted in Figure 1. The peaks of \mathcal{I}^N depend on the magnitude of $1/N$; indeed they become higher and thinner as the variance $1/N$ of the heat kernel $p_{1/N}$ decreases. We choose $N = N(m) = m^\eta$ for $\eta > 0$ given in equation (33), so as the number of steps m in the Euler scheme increases, the variance $1/N(m) = m^{-\eta}$ decreases, leading to a vector $\mathcal{I}^{N(m)}$ of mostly zeros up to numerical precision. If the discretization \mathfrak{M} is too coarse compared to $1/N(m)$, the vector $\mathcal{I}^{N(m)}$ does not retain enough information. To avoid this issue we have to make sure that there are sufficiently many points in \mathfrak{M} , i.e., that $2M$ is large enough compared to the size of the support $[-L, L]$. More specifically, we require that the distance δ between points in the grid is less than the standard deviation of the heat kernel $\sqrt{1/N(m)}$, i.e., $\delta < \sqrt{1/N(m)} = m^{-\eta/2}$ for every m considered in numerical computations, which leads to

$$2M > 2Lm^{\eta/2}. \tag{60}$$

As an example, let us consider a discretisation of the interval $[-5, 5]$ (hence $L = 5$) and use $m = 2^{12}$ points in the Euler–Maruyama scheme. We will need at least $2M + 1 > 2Lm^{\eta/2} + 1 = 10 \cdot 2^{6\eta} + 1$ points in \mathfrak{M} , where the value η depends on $\hat{\beta}$ and it is given explicitly in equation (33). The smallest admissible number of points $2M + 1$ of the discretisation \mathfrak{M} for varying $\hat{\beta}$ is given in Table 1.

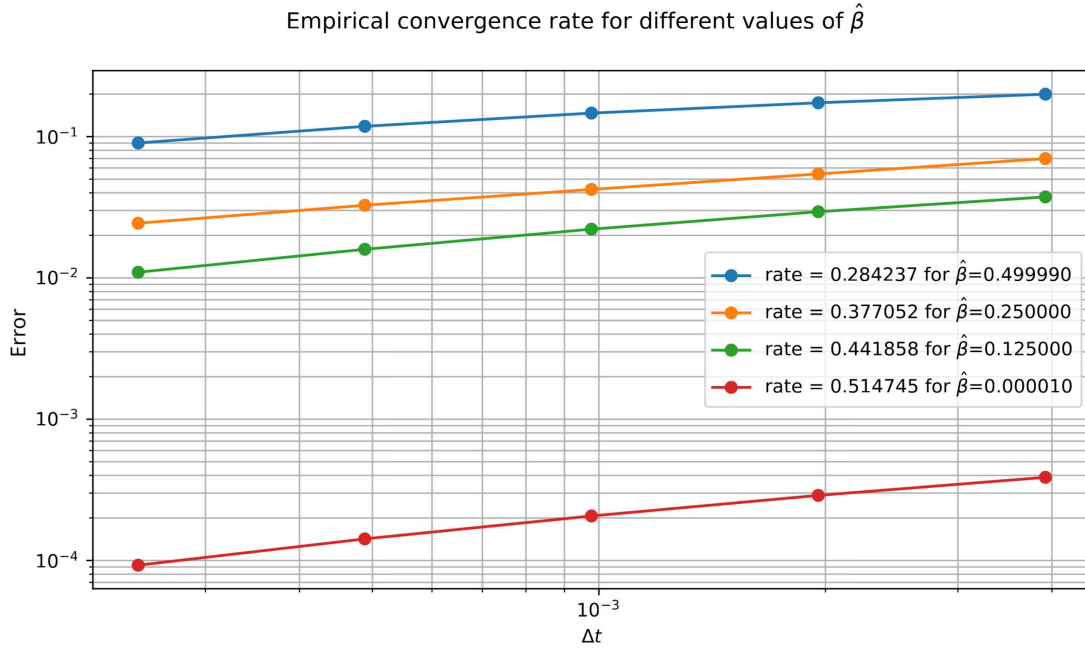


FIGURE 2. Examples of empirical convergence rates for $\hat{\beta} = \varepsilon, 1/8, 1/4, 1/2 - \varepsilon$, with $\varepsilon = 10^{-6}$, obtained running an Euler scheme with $Q = 10^4$ sample paths, $T = 1$, $m = 2^{15}$ points for the proxy of the real solution and $m_i = 2^{8+i}$ for $i = 0, \dots, 4$, for the approximated solutions. The empirical convergence rate r for each $\hat{\beta}$ is the slope of its corresponding line, which is plotted in a log-log graph.

TABLE 2. Average of empirical convergence rates obtained using 50 Euler–Maruyama approximations with 10^4 sample paths for each $\hat{\beta}$, the conjecture rate of $1/2 - \hat{\beta}/2$ and the theoretical rate from Theorem 3.4. All values rounded to 5 decimal places. Same values are plotted in Figure 3.

| $\hat{\beta}$ | $\varepsilon = 10^{-6}$ | 1/16 | 1/8 | 1/4 | 3/8 | 7/16 | 1/2 - ε |
|----------------------------|-------------------------|---------|---------|---------|---------|---------|---------------------|
| Average of empirical rates | 0.50814 | 0.48542 | 0.44063 | 0.36960 | 0.30478 | 0.29048 | 0.28275 |
| $1/2 - \hat{\beta}/2$ | 0.49999 | 0.46875 | 0.43750 | 0.37500 | 0.31250 | 0.28125 | 0.25000 |
| Theoretical rate | 0.16666 | 0.13243 | 0.10000 | 0.04545 | 0.01111 | 0.00270 | 0.00000 |

We now summarise the procedure to compute numerically an approximation of the drift $b^{N(m)}$.

- (1) Fix $\hat{\beta}$, L , and the number m of steps of the Euler scheme. Compute the smallest integer M that satisfies (60). Define the discretisation \mathfrak{M} and the corresponding mesh size δ .
- (2) Simulate a single path of fBm on the interval $[0, 2L]$ with mesh size δ and with a given Hurst parameter $H = -\hat{\beta} + 1 + \epsilon$ for some small $\epsilon > 0$.
- (3) Transform the path of fBm into a bridge on $[-L, L]$ by applying the transformation (53) to get a vector $B^H(x_i)$ for $x_i \in \mathfrak{M}$.
- (4) Compute numerically the integral $\mathcal{I}^{N(m)}(x_i)$ for $x_i \in \mathfrak{M}$.
- (5) Perform the discrete convolution $-(B^H * \mathcal{I}^{N(m)})$ as in (59) to approximate numerically $b^{N(m)}(x_i)$ for $x_i \in \mathfrak{M}$.
- (6) Extend $b^{N(m)}(x)$ for all $x \in [-L, L]$ by linear interpolation.

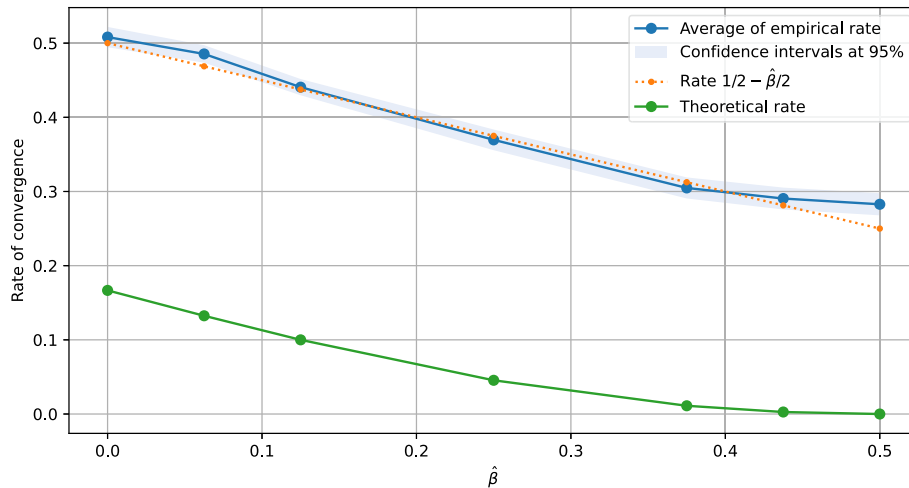


FIGURE 3. Plot of average of empirical convergence rates obtained as average of 50 Euler–Maruyama approximations with 10^4 sample paths for each $\hat{\beta}$. The plot contains also the 95% confidence interval for the average, the conjectured rate of $1/2 - \hat{\beta}/2$ and the theoretical rate obtained in Theorem 3.4.

Once we have a numerical approximation of the drift coefficient, the remaining step is to apply the standard Euler–Maruyama scheme. To calculate the empirical rate of convergence of the numerical scheme we must have approximations with an increasing number of steps m as well as a proxy of the real solution, since the real solution is unknown in a closed form. The strong error of the scheme is calculated by Monte Carlo approximation of the L^1 norm of the difference between those approximations and the proxy at time T . The procedure reads as follows:

- (1) Choose a “large” m for the proxy of the real solution, and $m_i \ll m$, $i = 0, \dots, I$, for the approximations and such that each m_i divides m . Choose a number of sample paths $Q \in \mathbb{N}$ sufficiently large.
- (2) Denote by $\Delta t_i = T/m_i$ the time-step for the approximated solutions, where T is the terminal time.
- (3) Define \mathfrak{M} with M points, where M and m satisfy (60). Generate one fBm path on the discrete grid \mathfrak{M} .
- (4) Run the Euler–Maruyama scheme for the proxy solution and for the approximated solutions up to time T , with the same Q Brownian motion paths.
- (5) Compute the strong error between the proxy solution corresponding to m and the approximations corresponding to m_i , by calculating a Monte Carlo average of the absolute differences between computed solutions at time T across the Q sample paths. Denote this approximations of the strong error by ϵ_i , $i = 0, \dots, I$.
- (6) As we expect $\epsilon_i \approx cm_i^{-r} = c(\Delta t_i)^r$, where r is the convergence rate, we compute the empirical rate r by performing a linear regression of $\log_{10}(\epsilon_i)$ on $\log_{10}(\Delta t_i)$, $i = 0, \dots, I$.

In Figure 2 we plot the empirical convergence rate we obtained for different choices of the smoothness parameter $\hat{\beta}$, that is for $\hat{\beta} = \varepsilon, 1/8, 1/4, 1/2 - \varepsilon$, with $\varepsilon = 10^{-6}$, and with the rest of the parameters as indicated in the caption to Figure 2. Note that as $\hat{\beta}$ grows and the drift becomes more rough, the empirical convergence rate becomes smaller and at the same time the strong error increases. Note that the empirical convergence rate is close to $1/2$ when $\hat{\beta} \approx 0$, which agrees with the theoretical results obtained by Dareiotis *et al.* [8] in the realm of measurable functions, see also [4]. Indeed, they show a strong convergence rate of $1/2 + \alpha/2$ when $b \in C^\alpha$ for $\alpha \in [0, 1)$, which reduces to $1/2$ for measurable functions ($\alpha = 0$). On the other hand, for $\beta \rightarrow 1/2$ we have an empirical convergence rate close to $1/4$. We now refer to the comment made in Remark 3.5. The nearly linear dependence of the log-error on $\log(\Delta t)$ suggests that the bound in Proposition 3.3 holds. We are not able to directly verify the condition $\|b^N - b\|_{C_T C^{-\beta}} < 1$ because the drift b is not available.

Finally, we performed a further experiment to better compare the empirical rate with the theoretical results. Since the drift of the SDE is obtained running a single path of a fBm, and clearly there is randomness there, we decided to run the algorithm for 50 different paths, for each value $\hat{\beta}$, and then we computed the average of the empirical convergence rates as well as its 95% confidence interval. We compared this with the theoretical rate obtained in Theorem 3.4 and with the conjecture that the rate should be $1/2 - \hat{\beta}/2$. The latter would be the natural extension of the results of [8] if they could be extended into the case of distributions, in particular with $-\hat{\beta} \in (-1/2, 0)$ which is the case we treat here. We collected the results in Table 2 below and also plotted them in Figure 3. This experiment strongly suggests that our theoretical result is not optimal, and that the convergence rate indeed could be $1/2 - \hat{\beta}/2$. Further studies are needed to prove or disprove this conjecture.

ACKNOWLEDGMENTS

The second author acknowledges partial financial support under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU – Project Title “Non-Markovian Dynamics and Non-local Equations” – 202277N5H9 – CUP: D53D23005670006 – Grant Assignment Decree No. 973 adopted on June 30, 2023, by MUR.

DATA AVAILABILITY STATEMENT

The research code described in Section 5 of this article is available on Zenodo. See the repository “Implementation of the Numerical Methods from ‘Convergence rate of numerical solutions to SDEs with distributional drifts in Besov spaces’”, with DOI [10.5281/zenodo.8239428](https://doi.org/10.5281/zenodo.8239428), and with identifier [6] in the references list.

REFERENCES

- [1] H. Bahouri, J.-Y. Chemin and R. Danchin, Fourier Analysis and Nonlinear Partial Differential Equations. Vol. 343 of *Grundlehren Der Mathematischen Wissenschaften*. Springer Berlin Heidelberg (2011).
- [2] R.F. Bass and Z.-Q. Chen, Brownian motion with singular drift. *Ann. Probab.* **31** (2003) 791–817.
- [3] F. Biagini, Y. Hu, B. Øksendal and T. Zhang, Stochastic Calculus for Fractional Brownian Motion and Applications. *Probability and its Applications*. Springer (2008).
- [4] O. Butkovsky, K. Dareiotis and M. Gerencsér, Approximation of SDEs: a stochastic sewing approach. *Probability Theory and Related Fields* **181** (2021) 975–1034.
- [5] G. Cannizzaro and K. Chouk, Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. *Ann. Probab.* **46(3)** (2018) 1710–1763.
- [6] L.M. Chaparro Jáquez, E. Issoglio and J. Palczewski, Implementation of the numerical scheme from “Convergence rate of numerical solutions to SDEs with distributional drifts in Besov spaces”. <https://zenodo.org/record/8239428>.
- [7] P.-E. Chaudru de Raynal and S. Menozzi, On multidimensional stable-driven stochastic differential equations with besov drift. *Electron. J. Probab.* **27** (2022) 1–52.
- [8] K. Dareiotis, M. Gerencsér and K. Lê, Quantifying a convergence theorem of Gyöngy and Krylov. *Ann. Appl. Probab.* **33** (2023) 2291–2323.
- [9] T. De Angelis, M. Germain and E. Issoglio, A numerical scheme for stochastic differential equations with distributional drift. *Stoch. Process. App.* **154** (2022) 55–90.
- [10] F. Delarue and R. Diel, Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields* **165** (2016) 1–63.
- [11] F. Flandoli, F. Russo and J. Wolf, Some SDEs with distributional drift Part I: General calculus. *Osaka J. Math.* **40** (2003) 493–542.
- [12] F. Flandoli, F. Russo and J. Wolf, Some SDEs with distributional drift. Part II: Lyons-Zheng structure, Itô’s formula and semimartingale characterization. *Random Oper. Stoch. Equ.* **12** (2004) 145–184.
- [13] F. Flandoli, E. Issoglio and F. Russo, Multidimensional stochastic differential equations with distributional drift. *Trans. Am. Math. Soc.* **369** (2017) 1665–1688.
- [14] L. Goudenège, E.M. Haress and A. Richard, Numerical approximation of SDEs with fractional noise and distributional drift. *Stoch. Process. App.* **181** (2025) 104533.

- [15] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes. Vol. 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam (1981).
- [16] E. Issoglio and F. Russo, A PDE with drift of negative Besov index and linear growth solutions. *Differ. Integral Equ.* **37** (2024) 585–622.
- [17] E. Issoglio and F. Russo, Stochastic differential equations with singular coefficients: the martingale problem view and the stochastic dynamics view. *J. Theor. Probab.* **37** (2024) 2352–2393.
- [18] E. Issoglio, S. Pagliarani, F. Russo and D. Trevisani, Degenerate McKean-Vlasov equations with drift in anisotropic negative Besov spaces. Preprint [arXiv:2401.09165](https://arxiv.org/abs/2401.09165) (2024).
- [19] B. Jourdain and S. Menozzi, Convergence rate of the Euler–Maruyama scheme applied to diffusion processes with $L^Q - L^p$ drift coefficient and additive noise. *Ann. Appl. Probab.* **34** (2024) 1663–1697.
- [20] A. Kohatsu-Higa, A. Lejay and K. Yasuda, Weak rate of convergence of the Euler–Maruyama scheme for stochastic differential equations with non-regular drift. *J. Comput. Appl. Math.* **326** (2017) 138–158.
- [21] K. Lé, A stochastic sewing lemma and applications. *Electron. J. Probab.* **25** (2020) 1–55.
- [22] G. Leobacher and M. Szölgényi, A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Anna. Appl. Probab.* **27** (2017) 2383–2418.
- [23] G. Leobacher and M. Szölgényi, Convergence of the Euler–Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numer. Math.* **138** (2018) 219–239.
- [24] A. Neuenkirch and M. Szölgényi, The Euler–Maruyama scheme for SDEs with irregular drift: convergence rates via reduction to a quadrature problem. *IMA J. Numer. Anal.* **41** (2020) 1164–1196.
- [25] A. Neuenkirch, M. Szölgényi and L. Szpruch, An adaptive Euler–Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis. *SIAM J. Numer. Anal.* **57** (2019) 378–403.
- [26] E. Platen and P.E. Kloeden, Numerical Solution of Stochastic Differential Equations. Springer (1992).
- [27] P. Przybyłowicz and M. Szölgényi, Existence, uniqueness, and approximation of solutions of jump-diffusion SDEs with discontinuous drift. *Appl. Math. Comput.* **403** (2021) 126191.
- [28] M. Renardy and R.C. Rogers, An Introduction to Partial Differential Equations. *Texts in Applied Mathematics*, 2nd edition. Springer (2004).
- [29] M. Szölgényi, Stochastic differential equations with irregular coefficients: mind the gap! *Int. Math. Nachr.* **246** (2021) 43–56.
- [30] H. Triebel, Theory of Function Spaces. *Modern Birkhäuser Classics*. Birkhäuser Verlag (2010).



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.