



# Drift-implicit Euler scheme for sandwiched processes driven by Hölder noises

Giulia Di Nunno<sup>1,2</sup> · Yuliya Mishura<sup>3</sup> · Anton Yurchenko-Tytarenko<sup>1</sup>

Received: 20 April 2022 / Accepted: 24 September 2022  
© The Author(s) 2022

## Abstract

In this paper, we analyze the drift-implicit (or backward) Euler numerical scheme for a class of stochastic differential equations with unbounded drift driven by an arbitrary  $\lambda$ -Hölder continuous process,  $\lambda \in (0, 1)$ . We prove that, under some mild moment assumptions on the Hölder constant of the noise, the  $L^r(\Omega; L^\infty([0, T]))$ -approximation error converges to 0 as  $O(\Delta^\lambda)$ ,  $\Delta \rightarrow 0$ . To exemplify, we consider numerical schemes for the generalized Cox–Ingersoll–Ross and Tsallis–Stariolo–Borland models. The results are illustrated by simulations.

**Keywords** Sandwiched process · Unbounded drift · Hölder continuous noise · Numerical scheme

**Mathematics Subject Classification (2010)** 60H10; 60H35; 60G22; 91G30

## 1 Introduction

We analyze the *drift-implicit* (also known as *backward*) Euler numerical scheme for stochastic differential equations (SDEs) of the form

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + Z(t), \quad t \in [0, T], \quad (1.1)$$

where  $Z$  is a general  $\lambda$ -Hölder continuous noise,  $\lambda \in (0, 1)$ , and the drift  $b$  is unbounded and has one of the following two properties:

- (A)  $b(t, y)$  has an explosive growth of the type  $(y - \varphi(t))^{-\gamma}$  as  $y \downarrow \varphi(t)$ , where  $\varphi$  is a given Hölder continuous function of the same order  $\lambda$  as  $Z$  and  $\gamma > \frac{1}{\lambda} - 1$ ;
- (B)  $b(t, y)$  has an explosive growth of the type  $(y - \varphi(t))^{-\gamma}$  as  $y \downarrow \varphi(t)$  and an explosive decrease of the type  $-(\psi(t) - y)^{-\gamma}$  as  $y \uparrow \psi(t)$ , where  $\varphi$  and

✉ Anton Yurchenko-Tytarenko  
antony@math.uio.no

Extended author information available on the last page of the article.

$\psi$  are given Hölder continuous functions of the same order  $\lambda$  as  $Z$  such that  $\varphi(t) < \psi(t)$ ,  $t \in [0, T]$ , and  $\gamma > \frac{1}{\lambda} - 1$ .

The SDEs of this type were extensively studied in [14]. It was shown that the properties **(A)** or **(B)**, along with some relatively weak additional assumptions, ensure that the solution to (1.1) is bounded from below (*one-sided sandwich case*) by the function  $\varphi$  in the setting **(A)**, i.e.

$$Y(t) > \varphi(t), \quad t \in [0, T], \tag{1.2}$$

or stays between  $\varphi$  and  $\psi$  (*two-sided sandwich case*) in the setting **(B)**, i.e.

$$\varphi(t) < Y(t) < \psi(t), \quad t \in [0, T]. \tag{1.3}$$

We emphasize that the SDE type (1.1) includes and generalizes several widespread stochastic models. For example, the process given by

$$Y(t) = Y(0) - \int_0^t \frac{\kappa Y(s)}{1 - Y^2(s)} ds + Z(t), \quad t \in [0, T],$$

where  $Z$  is  $\lambda$ -Hölder continuous with  $\lambda > \frac{1}{2}$ , fits into the setting **(B)**, and can be regarded as a natural extension of the Tsallis–Stariolo–Borland (TSB) model employed in biophysics (for more details on the standard Brownian TSB model, see, e.g. [15, Subsection 2.3] or [16, Chapter 3 and Chapter 8]). Another important example is

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{Y^\gamma(s)} - \kappa_2 Y(s) \right) ds + Z(t), \quad t \in [0, T], \tag{1.4}$$

where  $Z$  is  $\lambda$ -Hölder continuous,  $\lambda \in (0, 1)$ , and  $\gamma > \frac{1}{\lambda} - 1$ . It can be shown (see [14, Subsection 4.2]) that, if  $\lambda > \frac{1}{2}$ , stochastic process  $X(t) := Y^{1+\gamma}(t)$  satisfies the SDE

$$X(t) = X(0) + (1+\gamma) \int_0^t (\kappa_1 - \kappa_2 X(s)) ds + \int_0^t X^\alpha(s) dZ(s), \quad t \in [0, T], \tag{1.5}$$

where  $\alpha := \frac{\gamma}{1+\gamma} \in (0, 1)$  and the integral w.r.t.  $Z$  exists as a pathwise limit of Riemann–Stieltjes integral sums. Equations of the type (1.5) are used in finance in the standard Brownian setting and are called *Chan–Karolyi–Longstaff–Sanders* (CKLS) or *constant elasticity of variance* (CEV) model (see, e.g. [4, 8, 9]). If  $\alpha = \frac{1}{2}$ , Eq. (1.5) is also known as the *Cox–Ingersoll–Ross* (CIR) equation, see, e.g. [10–12].

In this work, we develop a numerical approximation (both pathwise and in  $L^r(\Omega; L^\infty([0, T]))$ ) for sandwiched processes (1.1) which is similar to the *drift-implicit* (also known as *backward*) Euler scheme constructed for the classical Cox–Ingersoll–Ross process in [2, 3, 13] and extended to the case of the fractional Brownian motion with  $H > \frac{1}{2}$  in [18, 21, 22]. In this drift-implicit scheme, in order to generate  $\widehat{Y}(t_{k+1})$ , one has to solve the equation of the type

$$\widehat{Y}(t_{k+1}) = \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)) \tag{1.6}$$

with respect to  $\widehat{Y}(t_{k+1})$  which is in general a more computationally heavy problem in comparison to the standard Euler-type techniques (see, e.g. [14, Section 5]). However, this drift-implicit numerical method also has a substantial advantage: the

approximation  $\widehat{Y}$  maintains the property of being sandwiched, i.e. for all points  $t_k$  of the partition

$$\widehat{Y}(t_k) > \varphi(t_k)$$

in the setting **(A)** and

$$\varphi(t_k) < \widehat{Y}(t_k) < \psi(t_k)$$

in the case **(B)**. Having this in mind, we shall say that the drift-implicit scheme is *sandwich preserving*.

We note that a similar approximation scheme was studied in [21] and [18, 22] for processes of the type (1.4) driven by a fractional Brownian motion with  $H > 1/2$ . Our work can be seen as an extension of those. However, we emphasize that our results have several elements of novelty. In particular, the paper [21] discusses only pathwise convergence and not convergence in  $L^r(\Omega; L^\infty([0, T]))$ . The approach of [18] and [22] is very noise specific as both use Malliavin calculus techniques in the spirit of [19, Proposition 3.4] to estimate inverse moments of the considered process (which turns out to be crucial to control explosive growth of the drift). As a result, two limitations appear: a restrictive condition involving the time horizon  $T$  (see, e.g. [18, Eq. (8) and Remark 3.1]) and sensitivity to the choice of the noise, i.e. their method cannot be applied directly for drivers other than fBm with  $H > 1/2$ . This lack of flexibility in terms of the choice of the noise is a crucial disadvantage in, e.g. finance where modern empirical studies justify the use of fBm with extremely low Hurst index ( $H < 0.1$ ) [7] or even drivers with time-varying roughness [1]. Our approach makes use of [14, Theorem 3.2] based on the pathwise calculus and allows us to obtain strong convergence with no limitations on  $T$  for a substantially larger class of noises. In fact, we require only Hölder continuity of the noise and some moment condition on the corresponding Hölder coefficient which is often satisfied and shared by, e.g. *all* Hölder continuous Gaussian processes.

The paper is organized as follows. Section 2 describes the setting in detail and contains some necessary statements on the properties of the sandwiched processes. In Section 3, we give the convergence results in the setting **(B)** which turns out to be a bit simpler than **(A)** due to boundedness of the process. Section 4 extends the scheme to the setting **(A)**. In Section 5, we give some examples and simulations; in particular, we show that in some cases (e.g. for the generalized TSB and CIR models), Eq. (1.6) can be solved explicitly which drastically improves the computational efficiency of the algorithm.

## 2 Preliminaries and assumptions

Fix  $T > 0$  and define

$$\begin{aligned} \mathcal{D}_{a_1} &:= \{(t, y) \in [0, T] \times \mathbb{R}_+, y \in (\varphi(t) + a_1, \infty)\}, & a_1 &\geq 0, \\ \mathcal{D}_{a_1, a_2} &:= \{(t, y) \in [0, T] \times \mathbb{R}_+, y \in (\varphi(t) + a_1, \psi(t) - a_2)\}, \\ a_1, a_2 &\in \left[0, \frac{1}{2} \|\psi - \varphi\|_\infty\right). \end{aligned} \tag{2.1}$$

where  $\varphi, \psi \in C([0, T])$  are such that  $\varphi(t) < \psi(t), t \in [0, T]$ .

Throughout the paper, we will be dealing with a stochastic differential equation of the form

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + Z(t), \quad t \in [0, T]. \tag{2.2}$$

The noise  $Z = \{Z(t), t \in [0, T]\}$  is always assumed to satisfy the following conditions:

- (Z1)  $Z(0) = 0$  a.s.;
- (Z2)  $Z$  has a.s.  $\lambda$ -Hölder continuous paths for some  $\lambda \in (0, 1)$ , i.e. there exists a positive random variable  $\Lambda$  such that

$$|Z(t) - Z(s)| \leq \Lambda |t - s|^\lambda, \quad s, t \in [0, T], \quad a.s.$$

Given the noise  $Z$  satisfying (Z1)–(Z2), the initial value  $Y(0)$  and the drift  $b$  satisfy one of the two assumptions given below.

**Assumption A (One-sided sandwich case)** There exists a  $\lambda$ -Hölder continuous function  $\varphi: [0, T] \rightarrow \mathbb{R}$  with  $\lambda$  being the same as in (Z2) such that

- (A1)  $Y(0)$  is deterministic and  $Y(0) > \varphi(0)$ ,
- (A2)  $b: \mathcal{D}_0 \rightarrow \mathbb{R}$  is continuous and for any  $\varepsilon \in (0, 1)$ 

$$|b(t_1, y_1) - b(t_2, y_2)| \leq \frac{c_1}{\varepsilon^p} (|y_1 - y_2| + |t_1 - t_2|^\lambda), \quad (t_1, y_1), (t_2, y_2) \in \mathcal{D}_\varepsilon,$$
 where  $c_1 > 0$  and  $p > 1$  are some given constants and  $\lambda$  is from (Z2),

(A3)

$$b(t, y) \geq \frac{c_2}{(y - \varphi(t))^\gamma}, \quad (t, y) \in \mathcal{D}_0 \setminus \mathcal{D}_{y_*},$$

where  $y_*, c_2 > 0$  are some given constants and  $\gamma > \frac{1}{\lambda} - 1$  with  $\lambda$  being from (Z2),

- (A4) the partial derivative  $\frac{\partial b}{\partial y}$ , with respect to the spacial variable exists, is continuous and bounded from above, i.e.

$$\frac{\partial b}{\partial y}(t, y) < c_3, \quad (t, y) \in \mathcal{D}_0,$$

for some  $c_3 > 0$ .

**Assumption B (Two-sided sandwich case)** There exist  $\lambda$ -Hölder continuous functions  $\varphi, \psi: [0, T] \rightarrow \mathbb{R}$ ,  $\varphi(t) < \psi(t), t \in [0, T]$ , with  $\lambda$  being the same as in (Z2) such that

- (B1)  $Y(0)$  is deterministic and  $\varphi(0) < Y(0) < \psi(0)$ ,
- (B2)  $b: \mathcal{D}_{0,0} \rightarrow \mathbb{R}$  is continuous and for any  $\varepsilon \in \left(0, \min \left\{1, \frac{1}{2} \|\psi - \varphi\|_\infty\right\}\right)$

$$|b(t_1, y_1) - b(t_2, y_2)| \leq \frac{c_1}{\varepsilon^p} (|y_1 - y_2| + |t_1 - t_2|^\lambda), \quad (t_1, y_1), (t_2, y_2) \in \mathcal{D}_{\varepsilon,\varepsilon},$$

where  $c_1 > 0$  and  $p > 1$  are some given constants and  $\lambda$  is from (Z2),

(B3)

$$b(t, y) \geq \frac{c_2}{(y - \varphi(t))^\gamma}, \quad (t, y) \in \mathcal{D}_{0,0} \setminus \mathcal{D}_{y_*,0},$$

$$b(t, y) \leq -\frac{c_2}{(\psi(t) - y)^\lambda}, \quad (t, y) \in \mathcal{D}_{0,0} \setminus \mathcal{D}_{0,y_*},$$

where  $y_*, c_2 > 0$  are some given constants and  $\lambda > \frac{1}{\lambda} - 1$  with  $\lambda$  being from **(Z2)**,

- (B4)** the partial derivative  $\frac{\partial b}{\partial y}$ , with respect to the spacial variable exists, is continuous and bounded from above, i.e.

$$\frac{\partial b}{\partial y}(t, y) < c_3, \quad (t, y) \in \mathcal{D}_{0,0},$$

for some  $c_3 > 0$ .

Both Assumptions A and B along with **(Z1)–(Z2)** ensure that the SDE (2.2) has a unique solution. In the theorem below, we provide some relevant results related to sandwiched processes (see [14, Theorems 2.3, 2.5, 2.6, 3.1, and 3.2]).

**Theorem 2.1** *Let  $Z = \{Z(t), t \in [0, T]\}$  be a stochastic process satisfying **(Z1)–(Z2)**.*

- 1) *If the initial value  $Y(0)$  and the drift  $b$  satisfy assumptions **(A1)–(A3)**, then the SDE has a unique strong pathwise solution such that for all  $t \in [0, T]$*

$$Y(t) > \varphi(t) \quad a.s. \tag{2.3}$$

*Moreover, there exist deterministic constants  $L_1, L_2, L_3$ , and  $L_4 > 0$  depending only on  $Y(0)$ , the shape of  $b$  and  $\lambda$ , such that for all  $t \in [0, T]$ , the estimate (2.3) can be refined as follows:*

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \leq Y(t) \leq L_3 + L_4\Lambda \quad a.s., \tag{2.4}$$

*where  $\Lambda$  is from **(Z2)** and  $\gamma$  is from **(A3)**. In particular, if  $\Lambda$  is such that*

$$\mathbb{E} \left[ \Lambda^{\frac{r}{\gamma\lambda + \lambda - 1}} \right] < \infty \tag{2.5}$$

*for some  $r > 0$ , then*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \frac{1}{(Y(t) - \varphi(t))^r} \right] < \infty,$$

*and, if*

$$\mathbb{E}\Lambda^r < \infty \tag{2.6}$$

*for some  $r > 0$ , then*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |Y(t)|^r \right] < \infty.$$

- 2) *If the initial value  $Y(0)$  and the drift  $b$  satisfy assumptions **(B1)–(B3)**, then the SDE has a unique strong pathwise solution such that for all  $t \in [0, T]$*

$$\varphi(t) < Y(t) < \psi(t) \quad a.s. \tag{2.7}$$

Moreover, there exist deterministic constants  $L_1$  and  $L_2 > 0$  depending only on  $Y(0)$ , the shape of  $b$  and  $\lambda$ , such that for all  $t \in [0, T]$ , the estimate (2.7) can be refined as follows:

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \leq Y(t) \leq \psi(t) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \quad \text{a.s.}, \quad (2.8)$$

where  $\Lambda$  is from (Z2) and  $\gamma$  is from (B3). In particular, if  $\Lambda$  can be chosen in such a way that

$$\mathbb{E} \left[ \Lambda^{\frac{r}{\gamma\lambda + \lambda - 1}} \right] < \infty \quad (2.9)$$

for some  $r > 0$ , then

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \frac{1}{(Y(t) - \varphi(t))^r} \right] < \infty, \quad \mathbb{E} \left[ \sup_{t \in [0, T]} \frac{1}{(\psi(t) - Y(t))^r} \right] < \infty.$$

**Remark 2.2** Properties (2.3)–(2.4) and (2.7)–(2.8) hold on each  $\omega \in \Omega$  such that  $Z(\omega; t)$ ,  $t \in [0, T]$ , is Hölder continuous and we always consider only such  $\omega \in \Omega$  in all proofs with pathwise arguments. For notational simplicity, we will also omit  $\omega$  in brackets.

**Remark 2.3** Due to the property (2.7), the setting described in Assumption B will be referred to as the **two-sided sandwich case** since the solution is “sandwiched” between  $\varphi$  and  $\psi$  a.s. Similarly, the property (2.3) justifies the name **one-sided sandwich case** for the setting corresponding to Assumption A. In both cases A and B, the solution to (2.2) will be referred to as a **sandwiched process**.

**Remark 2.4** Note that assumptions (A4) and (B4) are not required for Theorem 2.1 to hold and will be used later on.

In what follows, conditions (2.5), (2.6), and (2.9) will play an important role since the  $L^r(\Omega; L^\infty([0, T]))$ -convergence of the approximation scheme will directly follow from the integrability of  $\Lambda$ . However, it should be noted that these conditions are not very restricting as indicated in the following example.

**Example 2.5 (Hölder Gaussian noises)** Let  $Z = \{Z(t), t \in [0, T]\}$  be an arbitrary Hölder continuous Gaussian process satisfying (Z1)–(Z2), e.g. standard or fractional Brownian motion. In this case, by [6], the random variable  $\Lambda$  from (Z2) can be chosen to have moments of all orders.

We now complete the Section with some examples of the sandwiched processes.

**Example 2.6 (Generalized CIR and CKLS/CEV models)** Let  $\varphi \equiv 0$ ,  $Z$  satisfy (Z1)–(Z2) with  $\lambda \in (0, 1)$  and  $Y(0)$ ,  $\kappa_1, \kappa_2 > 0$ ,  $\gamma > \frac{1}{\lambda} - 1$  be given. Then, by Theorem 2.1 (1), the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{Y^\gamma(s)} - \kappa_2 Y(s) \right) ds + Z(t), \quad t \in [0, T], \quad (2.10)$$

has a unique positive solution. Moreover, it can be shown (see [14, Subsection 4.2]) that, if  $\lambda > \frac{1}{2}$ , stochastic process  $X(t) := Y^{1+\gamma}(t)$ ,  $t \in [0, T]$ , a.s. satisfies the SDE of the form

$$X(t) = X(0) + (1 + \gamma) \int_0^t (\kappa_1 - \kappa_2 X(s)) ds + (1 + \gamma) \int_0^t X^\alpha(s) dZ(s), \quad t \in [0, T], \tag{2.11}$$

where  $\alpha := \frac{\gamma}{1+\gamma} \in (0, 1)$  and the integral w.r.t.  $Z$  exists a.s. as a pathwise limit of Riemann-Stieltjes integral sums. As mentioned already, the (2.11) appears in finance in the standard Brownian setting and is called *Chan–Karolyi–Longstaff–Sanders* (CKLS) or *constant elasticity of variance* (CEV) model (see, e.g. [4, 8, 9]). If  $\alpha = \frac{1}{2}$  (i.e. when  $\gamma = 1$ ), the (2.11) is also known as the *Cox–Ingersoll–Ross* (CIR) equation [10–12].

*Remark 2.7 (Connection with the classical Brownian CIR/CKLS models)*

- 1) If  $\gamma = 1$  in (2.10) (CIR case), Assumption (A3) demands  $Z$  to be Hölder continuous of order  $\lambda > \frac{1}{2}$ . That means that Example 2.6 does not cover the classical Brownian CIR model since the continuous modification of a standard Brownian motion has paths that are Hölder continuous only up to (but not including) the order  $1/2$ . However, it is still possible to establish a clear connection between our setting and the classical CIR model. Indeed, let  $\{W(t), t \in [0, T]\}$  be the continuous modification of a standard Brownian motion. Consider the CIR process  $X = \{X(t), t \in [0, T]\}$  defined by

$$dX(t) = a(b - X(t))dt + \sigma\sqrt{X(t)}dW(t), \quad X_0 > 0,$$

where  $a, b, \sigma > 0$  and  $2ab > \sigma^2$ . The latter condition ensures that  $X$  has positive paths a.s. and hence one can define  $Y := \sqrt{X}$ . By Itô’s formula,  $Y$  satisfies the SDE

$$dY(t) = \left( \frac{\kappa_1}{Y(t)} - \kappa_2 Y(t) \right) dt + \frac{\sigma}{2} dW(t), \quad Y_0 = \sqrt{X_0} > 0, \tag{2.12}$$

with  $\kappa_1 := \frac{4ab - \sigma^2}{8}$  and  $\kappa_2 := \frac{a}{2}$ , which has a type very similar to (2.10). The SDE (2.12) can then be used to define a drift-implicit Euler scheme of the form (1.6) which turns out to converge to the original process (2.12). For more details on the drift-implicit Euler scheme for the classical Brownian CIR process, see, e.g. [13].

- 2) If  $\gamma > 1$  in (2.10), Assumptions (Z1)–(Z2) and (A1)–(A4) allow  $Z$  to be a standard Brownian motion. However, in this case, one cannot use pathwise calculus to obtain (2.11) whereas the standard Itô’s formula shows that  $X := Y^{1+\gamma}$  does not coincide with the standard CKLS process. In order to cover the standard CKLS model, we have to modify the drift in (2.10) to compensate for the second order term in Itô’s formula as follows:

$$dY(t) = \left( \frac{\kappa_1}{Y^\gamma(t)} - \frac{\gamma\sigma^2}{2Y(t)} - \kappa_2 Y(t) \right) dt + \sigma dW(t). \tag{2.13}$$

The SDE (2.13) satisfies Assumption A and  $X := Y^{1+\gamma}$  is the solution to the SDE

$$X(t) = X(0) + (1 + \gamma) \int_0^t (\kappa_1 - \kappa_2 X(s)) ds + (1 + \gamma)\sigma \int_0^t X^\alpha(s) dW(s),$$

$$\alpha = \frac{\gamma}{1 + \gamma},$$

i.e.  $X := Y^{1+\gamma}$  is the classical CKLS process.

*Example 2.8 (Generalized TSB model)* Let  $\varphi \equiv -1$ ,  $\psi \equiv 1$ ,  $Y(0) \in (-1, 1)$ ,  $Z$  satisfy (Z1)–(Z2) with  $\lambda > \frac{1}{2}$  and  $\kappa > 0$ . Then, by Theorem 2.1 (2), the SDE of the form

$$Y(t) = Y(0) - \int_0^t \frac{\kappa Y(s)}{1 - Y^2(s)} ds + Z(t), \quad t \in [0, T], \tag{2.14}$$

has a unique solution such that  $-1 < Y(t) < 1$  for all  $t \in [0, T]$  a.s. In the standard Brownian setting, the SDE of the type (2.14) is known as the Tsallis–Stariolo–Borland (TSB) model and is used in biophysics (for more details, see, e.g. [15, Subsection 2.3] or [16, Chapter 3 and Chapter 8]).

*Example 2.9* For the given  $Z$  satisfying (Z1)–(Z2) with  $\lambda \in (0, 1)$ ,  $\lambda$ -Hölder continuous functions  $\varphi, \psi$ ,  $\varphi(t) < \psi(t)$ ,  $t \in [0, T]$ , and  $Y(0) \in (\varphi(0), \psi(0))$  consider the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{(Y(s) - \varphi(s))^\gamma} - \frac{\kappa_2}{(\psi(s) - Y(s))^\gamma} - \kappa_3 Y(s) \right) ds + Z(t),$$

$$t \in [0, T],$$

where  $\kappa_1, \kappa_2 > 0$ ,  $\kappa_3 \in \mathbb{R}$ , and  $\gamma > \frac{1}{\lambda} - 1$ . By Theorem 2.1 (2), this SDE has a unique solution such that  $\varphi(t) < Y(t) < \psi(t)$  a.s. Note that the TSB drift from (2.14) also has this shape with  $\varphi \equiv -1$ ,  $\psi \equiv 1$ ,  $\gamma = 1$ ,  $\kappa_1 = \kappa_2 = \frac{\kappa}{2}$ , and  $\kappa_3 = 0$  since

$$-\frac{\kappa y}{1 - y^2} = \frac{\kappa}{2} \left( \frac{1}{y + 1} - \frac{1}{1 - y} \right).$$

**Notation 2.10** In what follows,  $C$  denotes any positive deterministic constant that does not depend on the partition and the exact value of which is not relevant. Note that  $C$  may change from line to line (or even within one line).

### 3 The approximation scheme for the two-sided sandwich

We will start by considering the numerical scheme for the *two-sided sandwich case* which turns out to be slightly simpler due to boundedness of  $Y$ . Let the noise  $Z$  satisfy (Z1)–(Z2),  $Y(0)$  and  $b$  satisfy Assumption B and  $Y = \{Y(t), t \in [0, T]\}$  be the unique solution of the SDE (2.2). Consider a uniform partition  $\{0 = t_0 < t_1 < \dots < t_N = T\}$  of  $[0, T]$ ,  $t_k := \frac{Tk}{N}$ ,  $k = 0, 1, \dots, N$ , with the mesh  $\Delta_N := \frac{T}{N}$  such that

$$c_3 \Delta_N < 1, \tag{3.1}$$



where  $c_3$  is an upper bound for  $\frac{\partial b}{\partial y}$  from **(B4)**. Let us define  $\widehat{Y}(t)$  as follows:

$$\begin{aligned} \widehat{Y}(0) &= Y(0), \\ \widehat{Y}(t_{k+1}) &= \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)), \\ \widehat{Y}(t) &= \widehat{Y}(t_k), \quad t \in [t_k, t_{k+1}), \end{aligned} \tag{3.2}$$

where the second expression is considered as an equation with respect to  $\widehat{Y}(t_{k+1})$ .

*Remark 3.1* Equation with respect to  $\widehat{Y}(t_{k+1})$  from (3.2) has a unique solution such that  $\widehat{Y}(t_{k+1}) \in (\varphi(t_{k+1}), \psi(t_{k+1}))$ . Indeed, for any fixed  $t \in [0, T]$  and any  $z \in \mathbb{R}$ , consider the equation

$$y - b(t, y)\Delta_N = z \tag{3.3}$$

w.r.t.  $y$ . Assumption **(B4)** together with condition (3.1) imply that  $(y - b(t, y)\Delta_N)'_y > 0$  and, by **(B3)**,

$$\begin{aligned} y - b(t, y)\Delta_N &\rightarrow -\infty, & y &\rightarrow \varphi(t)+, \\ y - b(t, y)\Delta_N &\rightarrow \infty, & y &\rightarrow \psi(t) - . \end{aligned}$$

Thus, there exists a unique  $y \in (\varphi(t), \psi(t))$  satisfying (3.3).

*Remark 3.2* The value of  $\widehat{Y}(t)$  for  $t \in [0, T] \setminus \{t_0, \dots, t_N\}$  can also be defined via linear interpolation as

$$\widehat{Y}(t) = \frac{1}{\Delta_N} ((t_{k+1} - t)\widehat{Y}(t_k) + (t - t_k)\widehat{Y}(t_{k+1})), \quad t \in [t_k, t_{k+1}), \quad k = 0, \dots, N-1.$$

In such case, all results of this section hold with almost no changes in the proofs.

*Remark 3.3* The algorithms of the type (3.2) are sometimes called the **drift-implicit** [2, 3, 13] or **backward** [18] Euler approximation schemes.

Before presenting the main results of this section, we require some auxiliary lemmas. First of all, we note that the values  $\widehat{Y}(t_n)$ ,  $n = 0, 1, \dots, N$ , of the discretized process are bounded away from both  $\varphi$  and  $\psi$  by random variables that do not depend on the partition. Namely, we have the following result that can be regarded as a discrete modification of arguments in [14, Theorem 3.2].

**Lemma 3.4** *Let  $Z$  satisfy **(Z1)**–**(Z2)**, Assumption B hold and the mesh of the partition  $\Delta_N$  satisfy (3.1). Then there exist deterministic constants  $L_1$  and  $L_2 > 0$  depending only on  $Y(0)$ , the shape of the drift  $b$  and  $\lambda$ , such that*

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \leq \widehat{Y}(t_n) \leq \psi(t_n) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}, \quad n = 0, 1, \dots, N, \quad a.s.,$$

where  $\Lambda$  is from **(Z2)** and  $\gamma$  is from **(B3)**.

*Proof* We will prove that

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \leq \widehat{Y}(t_n), \quad n = 0, 1, \dots, N, \quad a.s. \quad (3.4)$$

by using the pathwise argument (see Remark 2.2). The other inequality can be derived in a similar manner. Recall that, by Assumption B,  $\varphi$  and  $\psi$  are  $\lambda$ -Hölder continuous, i.e. there exists  $K > 0$  such that

$$|\varphi(t) - \varphi(s)| + |\psi(t) - \psi(s)| \leq K|t - s|^\lambda, \quad t, s \in [0, T].$$

Denote also

$$\beta := \frac{\lambda^{\frac{\lambda}{1-\lambda}} - \lambda^{\frac{1}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}} > 0,$$

where  $c_2$  is from **(B3)**,

$$L_2 := K + (2\beta)^{\lambda-1} \left( \frac{(Y(0) - \varphi(0)) \wedge y_* \wedge (\psi(0) - Y(0))}{2} \right)^{1-\lambda-\gamma\lambda} > 0,$$

with the constants  $y_*$  and  $\gamma$  also from **(B3)**, and

$$\varepsilon := \frac{1}{(2\beta)^{\frac{1-\lambda}{\gamma\lambda + \lambda - 1}} (L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}.$$

Note that, with probability 1,

$$|\varphi(t) - \varphi(s)| + |\psi(t) - \psi(s)| + |Z(t) - Z(s)| \leq (L_2 + \Lambda)|t - s|^\lambda, \quad t, s \in [0, T],$$

and, furthermore, it is easy to check that  $\varepsilon < Y(0) - \varphi(0)$ ,  $\varepsilon < \psi(0) - Y(0)$ , and  $\varepsilon < y_*$ .

If  $\widehat{Y}(t_n) \geq \varphi(t_n) + \varepsilon$  for a particular  $n = 0, 1, \dots, N$ , then, by definition of  $\varepsilon$ , the bound of the type (3.4) holds automatically. Suppose that there exists  $n = 1, \dots, N$  such that  $\widehat{Y}(t_n) < \varphi(t_n) + \varepsilon$ . Denote by  $\kappa(n)$  the last point of the partition before  $t_n$  on which  $\widehat{Y}$  stays above  $\varepsilon$ , i.e.

$$\kappa(n) := \max\{k = 0, \dots, n - 1 \mid \widehat{Y}_{t_k} \geq \varphi(t_k) + \varepsilon\}$$

(note that such point exists since  $\widehat{Y}(t_0) - \varphi(0) = Y(0) - \varphi(0) > \varepsilon$ ). Then, for all  $k = \kappa(n) + 1, \dots, n$  we have that  $\widehat{Y}(t_k) < \varepsilon < y_*$  and therefore, using **(B3)**, we obtain that, with probability 1,

$$\begin{aligned} \widehat{Y}(t_n) - \varphi(t_n) &= \widehat{Y}(t_{\kappa(n)}) - \varphi(t_n) + \Delta_N \sum_{k=\kappa(n)+1}^n b(t_k, \widehat{Y}(t_k)) + Z(t_n) - Z(t_{\kappa(n)}) \\ &\geq \varepsilon + \varphi(t_{\kappa(n)}) - \varphi(t_n) + \frac{c_2}{\varepsilon^\gamma} (t_n - t_{\kappa(n)}) + Z(t_n) - Z(t_{\kappa(n)}) \\ &\geq \varepsilon + \frac{c_2}{\varepsilon^\gamma} (t_n - t_{\kappa(n)}) - (L_2 + \Lambda)(t_n - t_{\kappa(n)})^\lambda. \end{aligned}$$

Consider a function  $F_\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}$  such that

$$F_\varepsilon(t) = \varepsilon + \frac{c_2}{\varepsilon^\gamma} t - (L_2 + \Lambda)t^\lambda.$$

It is straightforward to verify that  $F_\varepsilon$  attains its minimum at

$$t_* := \left(\frac{\lambda}{c_2}\right)^{\frac{1}{1-\lambda}} \varepsilon^{\frac{\gamma}{1-\lambda}} (L_2 + \Lambda)^{\frac{1}{1-\lambda}}$$

and, taking into account the explicit form of  $\varepsilon$ ,

$$\begin{aligned} F_\varepsilon(t_*) &= \varepsilon + \frac{\lambda^{\frac{1}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}} \varepsilon^{\frac{\gamma\lambda}{1-\lambda}} (L_2 + \Lambda)^{\frac{1}{1-\lambda}} - \frac{\lambda^{\frac{\lambda}{1-\lambda}}}{c_2^{\frac{\lambda}{1-\lambda}}} \varepsilon^{\frac{\gamma\lambda}{1-\lambda}} (L_2 + \Lambda)^{\frac{1}{1-\lambda}} \\ &= \varepsilon - \beta \varepsilon^{\frac{\gamma\lambda}{1-\lambda}} (L_2 + \Lambda)^{\frac{1}{1-\lambda}} = \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}} \beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}} (L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \\ &= \frac{\varepsilon}{2}. \end{aligned}$$

Namely, even if  $\widehat{Y}(t_n) < \varphi(t_n) + \varepsilon$ , we still have that, with probability 1,

$$\widehat{Y}(t_n) - \varphi(t_n) \geq F_\varepsilon(t_n - t_{\kappa(n)}) \geq F_\varepsilon(t_*) = \frac{\varepsilon}{2},$$

and thus, with probability 1, for any  $n = 0, 1, \dots, N$

$$\begin{aligned} \widehat{Y}(t_n) &\geq \varphi(t_n) + \frac{\varepsilon}{2} = \varphi(t_n) + \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}} \beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}} (L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \\ &=: \varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \end{aligned}$$

where  $L_1 := \frac{1}{2^{\frac{\gamma\lambda}{\gamma\lambda+\lambda-1}} \beta^{\frac{1-\lambda}{\gamma\lambda+\lambda-1}}}$ . □

*Remark 3.5* It is clear that constants  $L_1$  and  $L_2$  in Lemma 3.4 can be chosen jointly for  $Y$  and  $\widehat{Y}$ , so that the inequalities

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \leq Y(t) \leq \psi(t) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \quad t \in [0, T],$$

and

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \leq \widehat{Y}(t_n) \leq \psi(t_n) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}, \quad n = 0, 1, \dots, N,$$

hold simultaneously with probability 1.

Next, we proceed with a simple property of the sandwiched process  $Y$  in (2.2).

**Lemma 3.6** *Let  $Z$  satisfy (Z1)–(Z2) and assumptions (B1)–(B3) hold.*

1) *There exists a positive random variable  $\Upsilon$  such that, with probability 1,*

$$|Y(t) - Y(s)| \leq \Upsilon |t - s|^\lambda, \quad t, s \in [0, T].$$

2) If, for some  $r \geq 1$ ,

$$\mathbb{E} \left[ \Lambda \frac{r \max\{p, \gamma\lambda + \lambda - 1\}}{\gamma\lambda + \lambda - 1} \right] < \infty, \tag{3.5}$$

where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (B2), and  $\gamma$  is from (B3), then one can choose  $\Upsilon$  such that

$$\mathbb{E}[\Upsilon^r] < \infty.$$

*Proof* Denote  $\phi(t) := \frac{1}{2}(\psi(t) + \varphi(t))$ ,  $t \in [0, T]$ . By (2.8),

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \leq Y(t) \leq \psi(t) - \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}, \quad t \in [0, T], \quad a.s.,$$

i.e. with probability 1  $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$ ,  $t \in [0, T]$ , where

$$\xi := \frac{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}{L_1} \tag{3.6}$$

and  $\mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$  is defined by (2.1). It is evident that  $(t, \phi(t)) \in \mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$ ,  $t \in [0, T]$ ; therefore, using (Z2), (B2), and (2.7), we can write that, with probability 1, for all  $0 \leq s < t \leq T$ :

$$\begin{aligned} |Y(t) - Y(s)| &\leq \int_s^t |b(u, Y(u))| du + |Z(t) - Z(s)| \\ &\leq \int_s^t |b(u, Y(u)) - b(u, \phi(u))| du + \int_s^t |b(u, \phi(u))| du + \Lambda(t - s)^\lambda \\ &\leq c_1 \xi^p \int_s^t |Y(u) - \phi(u)| du + \max_{u \in [0, T]} |b(u, \phi(u))| (t - s) + \Lambda(t - s)^\lambda \\ &\leq \left( c_1 \xi^p \|\psi - \varphi\|_\infty + \max_{u \in [0, T]} |b(u, \phi(u))| \right) (t - s) + \Lambda(t - s)^\lambda \\ &\leq C(\xi^p + \Lambda + 1)(t - s)^\lambda, \end{aligned} \tag{3.7}$$

where  $C$  is a positive constant. Now one can put

$$\Upsilon := C(\xi^p + \Lambda + 1) \tag{3.8}$$

and observe that the definition of  $\Upsilon$ , (3.5), and (3.6) implies that

$$\mathbb{E}[\Upsilon^r] < \infty. \quad \square$$

Next, using Lemma 3.4 and following the proof of Lemma 3.6, it is easy to obtain the following result.

**Corollary 3.7** *Let (Z1)–(Z2) and Assumption B hold. Then there exists a random variable  $\Upsilon$  independent of the partition such that with probability 1*

$$|\widehat{Y}(t_k) - \widehat{Y}(t_n)| \leq \Upsilon |t_k - t_n|^\lambda, \quad k, n = 0, \dots, N. \tag{3.9}$$

Furthermore, if (3.5) holds for some  $r \geq 1$ , then

$$\mathbb{E}[\Upsilon^r] < \infty.$$

Finally,  $\Upsilon$  can be chosen jointly for  $Y$  and  $\widehat{Y}$ , so that

$$|Y(t) - Y(s)| \leq \Upsilon |t - s|^\lambda, \quad t, s \in [0, T],$$

holds simultaneously with (3.9) with probability 1.

**Lemma 3.8** *Let  $Z$  satisfy (Z1)–(Z2), Assumption B hold and the mesh of the partition  $\Delta_N$  satisfy (3.1). Then*

- 1) *for any  $r \geq 1$ , there exists a positive random variable  $\mathcal{C}_1$  that does not depend on the partition such that*

$$\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \leq \mathcal{C}_1 \Delta_N^{\lambda r} \quad a.s.;$$

- 2) *if, additionally,*

$$\mathbb{E} \left[ \Lambda^{\frac{r(p+\max\{p,\gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}} \right] < \infty, \tag{3.10}$$

where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (B2), and  $\gamma$  is from (B3), then one can choose  $\mathcal{C}_1$  such that  $\mathbb{E}[\mathcal{C}_1] < \infty$ , i.e. there exists a deterministic constant  $C$  that does not depend on the partition such that

$$\mathbb{E} \left[ \sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* Fix  $\omega \in \Omega$  such that  $Z(\omega, t)$ ,  $t \in [0, T]$ , is Hölder continuous (for simplicity of notation, we will omit  $\omega$  in the brackets). Denote  $e_n := Y(t_n) - \widehat{Y}(t_n)$ ,  $\Delta Z_n := Z(t_n) - Z(t_{n-1})$ . Then

$$\begin{aligned} e_n &= Y(t_{n-1}) + \int_{t_{n-1}}^{t_n} b(s, Y(s)) ds + \Delta Z_n \\ &\quad - \widehat{Y}(t_{n-1}) - b(t_n, \widehat{Y}(t_n)) \Delta_N - \Delta Z_n \\ &= e_{n-1} + (b(t_n, Y(t_n)) - b(t_n, \widehat{Y}(t_n))) \Delta_N \\ &\quad + \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n))) ds. \end{aligned} \tag{3.11}$$

By the mean value theorem,

$$\begin{aligned} (b(t_n, Y(t_n)) - b(t_n, \widehat{Y}(t_n))) \Delta_N &= \frac{\partial b}{\partial y}(t_n, \Theta_n) \Delta_N (Y(t_n) - \widehat{Y}(t_n)) \\ &= \frac{\partial b}{\partial y}(t_n, \Theta_n) \Delta_N e_n \end{aligned}$$

with  $\Theta_n \in (Y(t_n) \wedge \widehat{Y}(t_n), Y(t_n) \vee \widehat{Y}(t_n))$ . Using this, we can rewrite (3.11) as follows:

$$\left( 1 - \frac{\partial b}{\partial y}(t_n, \Theta_n) \Delta_N \right) e_n = e_{n-1} + \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n))) ds, \tag{3.12}$$

where

$$1 - \frac{\partial b}{\partial y}(t_n, \Theta_n) \Delta_N > 1 - c_3 \Delta_N > 0$$

by **(B4)** and (3.1).

Next, denote

$$\zeta_0 := 1, \quad \zeta_n := \prod_{i=1}^n \left( 1 - \frac{\partial b}{\partial y}(t_i, \Theta_i) \Delta_N \right)$$

and define  $\tilde{e}_n := \zeta_n e_n$ . By multiplying both sides of (3.12) by  $\zeta_{n-1}$ , we obtain that

$$\tilde{e}_n = \tilde{e}_{n-1} + \zeta_{n-1} \int_{t_{n-1}}^{t_n} (b(s, Y(s)) - b(t_n, Y(t_n))) ds \tag{3.13}$$

and, expanding the terms  $\tilde{e}_{i-1}$  in (3.13) one by one,  $i = n, n - 1, \dots, 1$ , and taking into account that  $\tilde{e}_0 = 0$ , we obtain that

$$\tilde{e}_n = \sum_{i=1}^n \zeta_{i-1} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i))) ds.$$

Therefore,

$$e_n = \sum_{i=1}^n \frac{\zeta_{i-1}}{\zeta_n} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i))) ds.$$

Observe that, by assumption **(B4)** and (3.1), for any  $i, n \in \mathbb{N}, i < n$ ,

$$\begin{aligned} \frac{\zeta_k}{\zeta_n} &= \prod_{i=k+1}^n \left( 1 - \frac{\partial b}{\partial y}(t_i, \Theta_i) \Delta_N \right)^{-1} \leq \prod_{i=k+1}^N (1 - c_3 \Delta_N)^{-1} \\ &\leq (1 - c_3 \Delta_N)^{-N} = \left( 1 - \frac{c_3 T}{N} \right)^{-N} \rightarrow e^{c_3 T}, \quad N \rightarrow \infty, \end{aligned}$$

whence there exists a constant  $C$  that does not depend on  $i, n$  or  $N$  such that

$$\frac{\zeta_k}{\zeta_n} \leq C.$$

Using this, one can deduce that

$$\begin{aligned} |e_n|^r &\leq C \left| \sum_{i=1}^n \frac{\zeta_{i-1}}{\zeta_n} \int_{t_{i-1}}^{t_i} (b(s, Y(s)) - b(t_i, Y(t_i))) ds \right|^r \\ &\leq C \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))| ds \right)^r. \end{aligned}$$

Note that  $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$ , where  $\xi$  is defined by (3.6) and  $\mathcal{D}_{\frac{1}{\xi}, \frac{1}{\xi}}$  is defined via (2.1); hence, by (B2) as well as Lemma 3.6, we can deduce that

$$\begin{aligned} & \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))| ds \right)^r \\ & \leq C \xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r + C \xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |Y(s) - Y(t_i)| ds \right)^r \\ & \leq C \xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r + C \xi^{pr} \Upsilon^r \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r \\ & = C \xi^{pr} (1 + \Upsilon^r) \left( \sum_{i=1}^n \frac{1}{(1 + \lambda)} \Delta_N^{1+\lambda} \right)^r \\ & \leq C \xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r}. \end{aligned}$$

In other words, there exists a constant  $C$  that does not depend on the partition such that

$$|e_n|^r = |Y(t_n) - \widehat{Y}(t_n)|^r \leq C \xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r}$$

and, since the right-hand side of the relation above does not depend on  $n$  or  $N$ , we have

$$\sup_{n=0, \dots, N} |Y(t_n) - \widehat{Y}(t_n)|^r \leq C \xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r} =: C_1 \Delta_N^{\lambda r}. \tag{3.14}$$

It remains to notice that, by (3.6) and (3.8),

$$\mathbb{E} [\xi^{pr} (1 + \Upsilon^r)] < \infty$$

whenever (3.10) holds, which finally implies

$$\mathbb{E} \left[ \sup_{n=0, \dots, N} |Y(t_n) - \widehat{Y}(t_n)|^r \right] \leq \mathbb{E}[C_1] \Delta_N^{\lambda r} =: C \Delta_N^{\lambda r}. \quad \square$$

Now we are ready to proceed to the main results of this subsection.

**Theorem 3.9** *Let  $Z$  satisfy (Z1)–(Z2), Assumption B hold and the mesh of the partition  $\Delta_N$  satisfy (3.1). Then*

- 1) *for any  $r \geq 1$ , there exists a random variable  $C_2$  that does not depend on the partition such that*

$$\sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \leq C_2 \Delta_N^{\lambda r} \quad a.s.;$$

- 2) *if, additionally,*

$$\mathbb{E} \left[ \Lambda^{\frac{r(p + \max\{p, \gamma\lambda + \lambda - 1\})}{\gamma\lambda + \lambda - 1}} \right] < \infty,$$

where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (B2), and  $\gamma$  is from (B3), then one can choose  $C_2$  such that  $\mathbb{E}[C_2] < \infty$ , i.e. there exists a deterministic constant  $C$  that

does not depend on the partition such that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* Fix  $\omega \in \Omega$  such that  $Z(\omega, t), t \in [0, T]$ , is Hölder continuous (for simplicity of notation, we again omit  $\omega$  in the brackets) and consider an arbitrary  $t \in [0, T]$ . Denote

$$n(t) := \max\{n = 0, 1, \dots, N \mid t \geq t_n\},$$

i.e.  $t \in [t_{n(t)}, t_{n(t)+1})$ . Then

$$\begin{aligned} |Y(t) - \widehat{Y}(t)|^r &\leq C (|Y(t) - Y(t_{n(t)})|^r + |Y(t_{n(t)}) - \widehat{Y}(t_{n(t)})|^r) \\ &\leq C \Upsilon^r (t - t_{n(t)})^{\lambda r} + C(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda + \lambda - 1}} (1 + \Upsilon^r) \Delta_N^{\lambda r} \\ &\leq C \left( \Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda + \lambda - 1}} \right) \Delta_N^{\lambda r}, \end{aligned}$$

where we used Lemma 3.6 to estimate  $|Y(t) - Y(t_{n(t)})|^r$  and bound (3.14) to estimate  $|Y(t_{n(t)}) - \widehat{Y}(t_{n(t)})|^r$ . Therefore,

$$\sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \leq C \left( \Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda + \lambda - 1}} \right) \Delta_N^{\lambda r} =: C_2 \Delta_N^{\lambda r}.$$

Finally, using the same arguments as in Lemma 3.6 and Lemma 3.8, one can easily show that the condition

$$\mathbb{E} \left[ \Lambda^{\frac{r(p + \max\{p, \gamma\lambda + \lambda - 1\})}{\gamma\lambda + \lambda - 1}} \right] < \infty$$

implies that

$$\mathbb{E} \left[ \Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda + \lambda - 1}} \right] < \infty,$$

therefore

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \right] \leq C \left( \Upsilon^r + (1 + \Upsilon^r)(L_2 + \Lambda)^{\frac{pr}{\gamma\lambda + \lambda - 1}} \right) \Delta_N^{\lambda r}$$

for some constant  $C > 0$  that does not depend on the partition. □

**Theorem 3.10**

1) Let  $Z$  satisfy (Z1)–(Z2), Assumption B hold and the mesh of the partition  $\Delta_N$  satisfy (3.1). Then, for any  $r \geq 1$ , there exists a random variable  $C_3$  that does not depend on the partition such that

$$\sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \leq C_3 \Delta_N^{\lambda r} \quad a.s.$$

and

$$\sup_{n=0,1,\dots,N} \left| \frac{1}{\psi(t_n) - Y(t_n)} - \frac{1}{\psi(t_n) - \widehat{Y}(t_n)} \right|^r \leq C_3 \Delta_N^{\lambda r} \quad a.s.$$



2) If, additionally,

$$\mathbb{E} \left[ \Lambda^{\frac{r(2+p+\max\{p,\gamma\lambda+\lambda-1\})}{\gamma\lambda+\lambda-1}} \right] < \infty, \tag{3.15}$$

where  $\lambda$  and  $\Lambda$  are from **(Z2)**,  $p$  is from **(B2)**, and  $\gamma$  is from **(B3)**, then one can choose  $\mathcal{C}_3$  such that  $\mathbb{E}[\mathcal{C}_3] < \infty$ , i.e. there exists a deterministic constant  $C$  that does not depend on the partition such that

$$\mathbb{E} \left[ \sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \right] \leq C \Delta_N^{\lambda r}$$

and

$$\mathbb{E} \left[ \sup_{n=0,1,\dots,N} \left| \frac{1}{\psi(t_n) - Y(t_n)} - \frac{1}{\psi(t_n) - \widehat{Y}(t_n)} \right|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* By Remark 3.5 and estimate (3.14), with probability 1 for any  $n = 0, \dots, N$ :

$$\begin{aligned} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r &= \frac{|Y(t_n) - \widehat{Y}(t_n)|^r}{(Y(t_n) - \varphi(t_n))^r (\widehat{Y}(t_n) - \varphi(t_n))^r} \\ &\leq \frac{(L_2 + \Lambda)^{\frac{2r}{\gamma\lambda+\lambda-1}}}{L_1^{2r}} \sup_{n=0,1,\dots,N} |Y(t_n) - \widehat{Y}(t_n)|^r \\ &\leq C(L_2 + \Lambda)^{\frac{2r}{\gamma\lambda+\lambda-1}} \xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r} \\ &=: \mathcal{C}_3 \Delta_N^{\lambda r}. \end{aligned}$$

It remains to notice that, by (3.6) and (3.8), the condition (3.15) implies that  $\mathbb{E}[\mathcal{C}_3] < \infty$ . The second estimate can be obtained in a similar manner.  $\square$

### 4 One-sided sandwich case

The drift-implicit Euler approximation scheme described in Section 3 for the two-sided sandwich can also be adapted for the one-sided setting that corresponds to Assumption A on the SDE (1.1). However, in the two-sided sandwich case, the process  $Y$  was bounded (which was utilized, e.g. in Lemma 3.6) and, moreover, the behaviour of  $Y$  was similar near both  $\varphi$  and  $\psi$  so that it was sufficient to analyze only one of the bounds. In the one-sided case, each  $Y(t)$ , for  $t \in [0, T]$ , is not a bounded random variable; therefore, the approach from Section 3 has to be adjusted. For this, we will be using the inequalities (2.4).

Let the noise  $Z$  satisfy **(Z1)**–**(Z2)**,  $Y(0)$  and  $b$  satisfy Assumption A and  $Y = \{Y(t), t \in [0, T]\}$  be the unique solution of the SDE (2.2). In line with Section 3, we consider a uniform partition  $\{0 = t_0 < t_1 < \dots < t_N = T\}$  of  $[0, T]$ ,  $t_k := \frac{Tk}{N}$ ,  $k = 0, 1, \dots, N$ , with the mesh  $\Delta_N := \frac{T}{N}$  such that

$$c_3 \Delta_N < 1, \tag{4.1}$$

where  $c_3$  is an upper bound for  $\frac{\partial b}{\partial y}$  from assumption **(A4)**. The backward Euler approximation  $\widehat{Y}(t)$  is defined in a manner similar to (3.2), i.e.

$$\begin{aligned} \widehat{Y}(0) &= Y(0), \\ \widehat{Y}(t_{k+1}) &= \widehat{Y}(t_k) + b(t_{k+1}, \widehat{Y}(t_{k+1}))\Delta_N + (Z(t_{k+1}) - Z(t_k)), \\ \widehat{Y}(t) &= \widehat{Y}(t_k), \quad t \in [t_k, t_{k+1}), \end{aligned} \tag{4.2}$$

where the second expression is considered as an equation with respect to  $\widehat{Y}(t_{k+1})$ .

*Remark 4.1* Just as in the two-sided sandwich case, each  $\widehat{Y}(t_k)$ ,  $k = 1, \dots, N$ , is well defined since the equation

$$y - b(t, y)\Delta_N = z$$

has a unique solution w.r.t.  $y$  such that  $y > \varphi(t)$  for any fixed  $t \in [0, T]$  and any  $z \in \mathbb{R}$ . To understand this, note that assumption **(A4)** together with (4.1) imply that

$$(y - b(t, y)\Delta_N)'_y > 0. \tag{4.3}$$

Second, by **(A3)**,

$$y - b(t, y)\Delta_N \rightarrow -\infty, \quad y \rightarrow \varphi(t) +. \tag{4.4}$$

Next, by **(A2)**, for any  $(s, y_1), (s, y_2) \in \overline{\mathcal{D}}_1 := \{(u, y) \in [0, T] \times \mathbb{R}_+, y \in [\varphi(u) + 1, \infty)\}$ , we have that

$$|b(s, y_1) - b(s, y_2)| \leq c_1|y_1 - y_2|,$$

i.e.

$$\sup_{(s,y) \in \overline{\mathcal{D}}_1} \left| \frac{\partial b}{\partial y}(s, y) \right| < \infty.$$

Using this, **(A4)**, and the mean value theorem, for any positive  $y \geq \varphi(t) + 1$

$$\begin{aligned} b(t, y) &= b(t, \varphi(t) + 1) + \frac{\partial b}{\partial y}(t, \theta_y)(y - 1 - \varphi(t)) \\ &\leq \max_{s \in [0, T]} b(t, \varphi(t) + 1) + \max_{s \in [0, T]} |1 + \varphi(s)| \sup_{(s,y) \in \overline{\mathcal{D}}_1} \left| \frac{\partial b}{\partial y}(s, y) \right| + c_3 y \\ &=: C + c_3 y, \end{aligned}$$

whence

$$y - b(t, y)\Delta_N \geq -C\Delta_N + (1 - c_3\Delta_N)y \rightarrow \infty, \quad y \rightarrow \infty. \tag{4.5}$$

Existence and uniqueness of the solution then follows from (4.3)–(4.5).

*Remark 4.2* Similarly to the two-sided sandwich case, the value of  $\widehat{Y}(t)$  for  $t \in [0, T] \setminus \{t_0, \dots, t_N\}$  can also be defined via linear interpolation with no changes in formulations of the results and almost no variations in the proofs.

Our strategy for proving the convergence of  $\widehat{Y}$  to  $Y$  will be similar to what we have done in Section 3. Therefore, we will be omitting the details highlighting only the points which are different from the two-sided sandwich case. We start with some useful properties of  $\widehat{Y}$  and  $Y$ .

**Lemma 4.3** *Let  $Z$  satisfy (Z1)–(Z2), Assumption A hold and the mesh of the partition  $\Delta_N$  satisfy (4.1). Then there exist deterministic constants  $L_1, L_2 > 0$  depending only on  $Y(0)$ , the shape of the drift  $b$  and  $\lambda$ , such that*

$$\widehat{Y}(t_n) \geq \varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}} \quad a.s.,$$

where  $\Lambda$  is from assumption (Z2) and  $\gamma$  is from assumption (A3). Moreover, there exist constants  $L_3, L_4 > 0$  that also depend only on  $Y(0)$ , the shape of the drift  $b$  and  $\lambda$  such that

$$\widehat{Y}(t_n) \leq L_3 + L_4\Lambda, \quad n = 0, 1, \dots, N, \quad a.s.$$

for all partitions with the mesh satisfying  $\frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N < 1$  with  $c_1$  and  $p$  being from (A2).

*Proof* The proof of

$$\widehat{Y}(t_n) \geq \varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda + \lambda - 1}}}$$

is identical to the corresponding one in Lemma 3.4 and will be omitted. Let us prove that

$$\widehat{Y}(t_n) \leq L_3 + L_4\Lambda \quad a.s.$$

Fix  $\omega \in \Omega$  for which  $Z(\omega, t)$  is Hölder continuous, consider a partition with the mesh satisfying  $\frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N < 1$  and fix an arbitrary  $n = 0, 1, \dots, N - 1$ . Assume that  $\widehat{Y}(t_{n+1}) > \varphi(t_{n+1}) + (Y(0) - \varphi(0))$  (otherwise, the claim of the lemma holds automatically). Put

$$\kappa(n) := \max\{k = 0, 1, \dots, n \mid \widehat{Y}(t_k) \leq \varphi(t_k) + (Y(0) - \varphi(0))\}$$

and observe that  $(t_k, \widehat{Y}(t_k)) \in \mathcal{D}_{Y(0) - \varphi(0)}$  for any  $k = \kappa(n) + 1, \dots, n + 1$ , where  $\mathcal{D}_{Y(0) - \varphi(0)}$  is defined via (2.1). Next, by (A2), for any  $y \in \mathcal{D}_{Y(0) - \varphi(0)}$

$$\begin{aligned} |b(t, y)| - \left| b\left(t, \varphi(t) + (Y(0) - \varphi(0))\right) \right| &\leq \left| b(t, y) - b\left(t, \varphi(t) + (Y(0) - \varphi(0))\right) \right| \\ &\leq \frac{c_1}{(Y(0) - \varphi(0))^p} |y - \varphi(t) - (\varphi(0) - Y(0))| \\ &\leq \frac{c_1}{(Y(0) - \varphi(0))^p} |y| + \frac{c_1}{(Y(0) - \varphi(0))^p} |\varphi(t) + (\varphi(0) - Y(0))|, \end{aligned}$$

i.e. there exists a constant  $C > 0$  that does not depend on the partition such that

$$|b(t, y)| \leq C + \frac{c_1}{(Y(0) - \varphi(0))^p} |y|. \tag{4.6}$$

Next, observe that, for any  $k = \kappa(n) + 1, \dots, n + 1$ , we have

$$\begin{aligned} \widehat{Y}(t_k) &= \widehat{Y}(t_{\kappa(n)}) + \sum_{i=\kappa(n)+1}^k b(t_i, \widehat{Y}(t_i)) \Delta_N + Z(t_k) - Z(t_{\kappa(n)}) \\ &\leq \varphi(t_{\kappa(n)}) + (Y(0) - \varphi(0)) + \sum_{i=\kappa(n)+1}^k b(t_i, \widehat{Y}(t_i)) \Delta_N + \Lambda(t_k - t_{\kappa(n)})^\lambda \\ &\leq \left| \max_{s \in [0, T]} \varphi(s) + (Y(0) - \varphi(0)) \right| + T^\lambda \Lambda + \sum_{i=\kappa(n)+1}^k b(t_i, \widehat{Y}(t_i)) \Delta_N. \end{aligned}$$

Therefore, using (4.6) and

$$\widehat{Y}(t_k) > \varphi(t_k) \geq \min_{s \in [0, T]} \varphi(s),$$

one can write

$$\begin{aligned} |\widehat{Y}(t_k)| &\leq \left| \min_{s \in [0, T]} \varphi(s) \right| + \left| \max_{s \in [0, T]} \varphi(s) + (Y(0) - \varphi(0)) \right| + T^\lambda \Lambda + \sum_{i=\kappa(n)+1}^k |b(t_i, \widehat{Y}(t_i))| \Delta_N \\ &\leq \left| \min_{s \in [0, T]} \varphi(s) \right| + \left| \max_{s \in [0, T]} \varphi(s) + (Y(0) - \varphi(0)) \right| + T^\lambda \Lambda + C \sum_{i=\kappa(n)+1}^k \Delta_N \\ &\quad + \frac{c_1}{(Y(0) - \varphi(0))^p} \sum_{i=\kappa(n)+1}^k |\widehat{Y}(t_i)| \Delta_N, \end{aligned}$$

where  $C > 0$  is some positive constant that does not depend on the partition.

Now we want to apply the discrete version of the Gronwall inequality from [20, Lemma A.3]. In order to do that, we observe that

$$|\widehat{Y}(t_{\kappa(n)+1})| \leq C + T^\lambda \Lambda + \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N |\widehat{Y}(t_k)|,$$

and, for any  $k = \kappa(n) + 2, \dots, n + 1$ ,

$$|\widehat{Y}(t_k)| \leq C + T^\lambda \Lambda + \frac{c_1}{(Y(0) - \varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \Delta_N + \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N |\widehat{Y}(t_k)|.$$

Now, since  $\frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N < 1$ , we can write that

$$\left( 1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N \right) |\widehat{Y}(t_{\kappa(n)+1})| \leq C + T^\lambda \Lambda$$

and, for all  $k = \kappa(n) + 2, \dots, n + 1$ ,

$$\left( 1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N \right) |\widehat{Y}(t_k)| \leq C + T^\lambda \Lambda + \frac{c_1}{(Y(0) - \varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \Delta_N.$$

Put

$$N_0 := \min \left\{ N \geq 1 : \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N < 1 \right\} = \left\lceil \frac{T c_1}{(Y(0) - \varphi(0))^p} \right\rceil + 1$$

with  $[x]$  being the greatest integer less than or equal to  $x$  and observe that, for all  $N \geq N_0$ ,

$$1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N \geq 1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}.$$

Therefore,

$$\begin{aligned} |\widehat{Y}(t_{\kappa(n)+1})| &\leq \frac{C}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N} \Lambda \\ &\leq \frac{C}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}} \Lambda \\ &=: C_1 + C_2 \Lambda \end{aligned}$$

and, for all  $k = \kappa(n) + 2, \dots, n + 1$ ,

$$\begin{aligned} |\widehat{Y}(t_k)| &\leq \frac{C}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N} \Lambda \\ &\quad + \frac{c_1}{(Y(0) - \varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \frac{\Delta_N}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_N} \\ &\leq \frac{C}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}} + \frac{T^\lambda}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}} \Lambda \\ &\quad + \frac{c_1}{(Y(0) - \varphi(0))^p} \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \frac{\Delta_N}{1 - \frac{c_1}{(Y(0) - \varphi(0))^p} \Delta_{N_0}} \\ &=: C_1 + C_2 \Lambda + C_3 \sum_{i=\kappa(n)+1}^{k-1} |\widehat{Y}(t_i)| \Delta_N. \end{aligned}$$

Using a discrete version of the Gronwall inequality, we now obtain that for all  $k = \kappa(n) + 1, \dots, n + 1$

$$\begin{aligned} |\widehat{Y}(t_k)| &\leq (C_1 + C_2 \Lambda) \exp \left\{ C_3 \sum_{i=\kappa(n)+1}^{k-1} \Delta_N \right\} \leq (C_1 + C_2 \Lambda) \exp \{TC_3\} \\ &=: L_3 + L_4 \Lambda. \end{aligned}$$

which ends the proof. □

*Remark 4.4* It is clear that constants  $L_1, L_2, L_3$ , and  $L_4$  can be chosen jointly for  $Y$  and  $\widehat{Y}$ , so that the inequalities

$$\varphi(t) + \frac{L_1}{(L_2 + \Lambda)^\alpha} \leq Y(t) \leq L_3 + L_4 \Lambda, \quad t \in [0, T],$$

and

$$\varphi(t_n) + \frac{L_1}{(L_2 + \Lambda)^\alpha} \leq \widehat{Y}(t_n) \leq L_3 + L_4 \Lambda, \quad n = 0, 1, \dots, N,$$

hold simultaneously with probability 1.

Next, corresponding to Lemma 3.6 in the two-sided case,  $Y$  enjoys Hölder continuity with the Hölder constant being integrable provided that  $\Lambda$  has moments of sufficiently high order. This is summarized in the lemma below.

**Lemma 4.5** *Let  $Z$  satisfy (Z1)–(Z2) and assumptions (A1)–(A3) hold.*

1) *There exists a positive random variable  $\Upsilon$  such that with probability 1*

$$|Y(t) - Y(s)| \leq \Upsilon |t - s|^\lambda, \quad t, s \in [0, T].$$

2) *If, for some  $r \geq 1$ ,*

$$\mathbb{E} \left[ \Lambda^{\frac{r(p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}} \right] < \infty, \tag{4.7}$$

*where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (A2), and  $\gamma$  is from (A3), then one can choose  $\Upsilon$  such that*

$$\mathbb{E}[\Upsilon^r] < \infty.$$

*Proof* By (2.4),

$$Y(t) \geq \varphi(t) + \frac{L_1}{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}} \quad a.s.,$$

i.e. with probability 1  $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}}$ ,  $t \in [0, T]$ , where

$$\xi := \frac{(L_2 + \Lambda)^{\frac{1}{\gamma\lambda+\lambda-1}}}{L_1} \tag{4.8}$$

and  $\mathcal{D}_{\frac{1}{\xi}}$  is defined in (2.1). Denote  $\phi(t) := \varphi(t) + 1$  and notice that  $(t, \phi(t)) \in \mathcal{D}_{\frac{1}{\xi}}$ ,  $t \in [0, T]$ , since  $\frac{1}{\xi} \leq Y(0) - \varphi(0)$ . Thus, using the same arguments as applied in (3.7), we can write that, with probability 1, for any  $0 \leq s < t \leq T$ :

$$|Y(t) - Y(s)| \leq c_1 \xi^p \int_s^t |Y(u) - \phi(u)| du + \max_{u \in [0, T]} |b(u, \phi(u))|(t - s) + \Lambda(t - s)^\lambda,$$

where  $c_1$  is from (A2). Now, again by (2.4),

$$Y(t) \leq L_3 + L_4 \Lambda \quad a.s.,$$

hence with probability 1

$$\begin{aligned} |Y(t) - Y(s)| &\leq c_1 \xi^p \int_s^t |Y(u) - \phi(u)| du + \max_{u \in [0, T]} |b(u, \phi(u))|(t - s) + \Lambda(t - s)^\lambda \\ &\leq c_1 \xi^p (L_3 + L_4 \Lambda)(t - s) + c_1 \xi^p \max_{u \in [0, T]} |\phi(u)|(t - s) \\ &\quad + \max_{u \in [0, T]} |b(u, \phi(u))|(t - s) + \Lambda(t - s)^\lambda \\ &\leq C(1 + \xi^p \Lambda + \xi^p + \Lambda)(t - s)^\lambda, \end{aligned}$$

where  $C$  is a positive constant. Now one can put

$$\Upsilon := C(1 + \xi^p \Lambda + \xi^p + \Lambda) \tag{4.9}$$

and observe that

$$\mathbb{E}[\Upsilon^r] < \infty$$

whenever (4.7) holds. □

**Corollary 4.6** *Using Lemma 4.3 and following the proof of Lemma 4.5, it is easy to obtain that, for any partition with the mesh satisfying*

$$\max \left\{ c_3, \frac{c_1}{(Y(0) - \varphi(0))^p} \right\} \Delta_N < 1 \tag{4.10}$$

there is a random variable  $\Upsilon$  independent of the partition such that with probability 1

$$|\widehat{Y}(t_k) - \widehat{Y}(t_n)| \leq \Upsilon |t_k - t_n|^\lambda, \quad k, n = 0, \dots, N. \tag{4.11}$$

Furthermore, just like in Lemma 3.6, for  $r > 0$

$$\mathbb{E}[\Upsilon^r] < \infty$$

provided that

$$\mathbb{E} \left[ \Lambda^{\frac{r(p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}} \right] < \infty.$$

Finally, such  $\Upsilon$  can be chosen jointly for  $Y$  and  $\widehat{Y}$ , so that

$$|Y(t) - Y(s)| \leq \Upsilon |t - s|^\lambda, \quad t, s \in [0, T],$$

holds simultaneously with (4.11) with probability 1.

**Lemma 4.7** *Let  $Z$  satisfy (Z1)–(Z2), Assumption A hold and the mesh of the partition  $\Delta_N$  satisfy (4.1).*

1) *For any  $r \geq 1$ , there exists a positive random variable  $\mathcal{C}_4$  that does not depend on the partition such that*

$$\sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \leq \mathcal{C}_4 \Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E} \left[ \Lambda^{\frac{r(2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}} \right] < \infty, \tag{4.12}$$

where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (A2), and  $\gamma$  is from (A3), then one can choose  $\mathcal{C}_4$  such that  $\mathbb{E}[\mathcal{C}_4] < \infty$ , i.e. there exists a deterministic constant  $C$  that does not depend on the partition such that

$$\mathbb{E} \left[ \sup_{k=0,1,\dots,N} |Y(t_k) - \widehat{Y}(t_k)|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* Following the proof of Lemma 3.8, one can easily obtain that for any  $n = 0, 1, \dots, N$

$$|Y(t_n) - \widehat{Y}(t_n)| \leq C \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))| ds \right)^r.$$

Next, note that  $(t, Y(t)) \in \mathcal{D}_{\frac{1}{\xi}}$ , where  $\xi$  is defined by (4.8), so, by (A2) and Lemma 4.5,

$$\begin{aligned} & \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |b(s, Y(s)) - b(t_i, Y(t_i))| ds \right)^r \\ & \leq C\xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r + C\xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |Y(s) - Y(t_i)| ds \right)^r \\ & \leq C\xi^{pr} \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r + C\xi^{pr} \Upsilon^r \left( \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |s - t_i|^\lambda ds \right)^r \\ & = C\xi^{pr} (1 + \Upsilon^r) \left( \sum_{i=1}^n \frac{1}{(1 + \lambda)} \Delta_N^{1+\lambda} \right)^r \\ & \leq C\xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r}, \end{aligned}$$

i.e.

$$\sup_{n=0, \dots, N} |Y(t_n) - \widehat{Y}(t_n)|^r \leq C\xi^{pr} (1 + \Upsilon^r) \Delta_N^{\lambda r}. \tag{4.13}$$

In order to conclude the proof, it remains to notice that (4.8), (4.9), and (4.12) imply that

$$\mathbb{E} [\xi^{pr} (1 + \Upsilon^r)] < \infty.$$

□

Now we are ready to formulate the two main results of this section.

**Theorem 4.8** *Let  $Z$  satisfy (Z1)–(Z2), Assumption A hold and the mesh of the partition  $\Delta_N$  satisfy (4.10).*

1) *For any  $r \geq 1$ , there exists a random variable  $C_5$  that does not depend on the partition such that*

$$\sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \leq C_5 \Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E} \left[ \Delta^{\frac{r(2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}} \right] < \infty,$$

*where  $\lambda$  and  $\Delta$  are from (Z2),  $p$  is from (A2), and  $\gamma$  is from (A3), then one can choose  $C_5$  such that  $\mathbb{E}[C_5] < \infty$ , i.e. there exists a deterministic constant  $C$  that does not depend on the partition such that*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |Y(t) - \widehat{Y}(t)|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* The proof is similar to the one of Theorem 3.9 but instead of Lemmas 3.6, 3.8 and bound (3.14) one should apply Lemmas 4.5, 4.7 and bound (4.13). □



**Theorem 4.9** *Let  $Z$  satisfy (Z1)–(Z2), Assumption A hold and the mesh of the partition  $\Delta_N$  satisfy (4.10).*

1) *For any  $r \geq 1$ , there exists a random variable  $C_6$  that does not depend on the partition such that*

$$\sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \leq C_6 \Delta_N^{\lambda r} \quad a.s.$$

2) *If, additionally,*

$$\mathbb{E} \left[ \Lambda^{\frac{r(2+2p+\gamma\lambda+\lambda-1)}{\gamma\lambda+\lambda-1}} \right] < \infty, \tag{4.14}$$

where  $\lambda$  and  $\Lambda$  are from (Z2),  $p$  is from (A2), and  $\gamma$  is from (A3), then one can choose  $C_6$  such that  $\mathbb{E}[C_6] < \infty$ , i.e. there exists a deterministic constant  $C$  that does not depend on the partition such that

$$\mathbb{E} \left[ \sup_{n=0,1,\dots,N} \left| \frac{1}{Y(t_n) - \varphi(t_n)} - \frac{1}{\widehat{Y}(t_n) - \varphi(t_n)} \right|^r \right] \leq C \Delta_N^{\lambda r}.$$

*Proof* The proof is similar to Theorem 3.10 and is omitted. □

### 5 Examples and simulations

The algorithms presented in (3.2) and (4.2) imply that, in order to generate  $\widehat{Y}(t_{n+1})$ , one has to solve an equation that potentially can be challenging from the computational point of view. However, in some cases that are relevant for applications, this equation has a simple explicit solution.

Regarding the numerical examples that follow, we remark that:

- 1) all the simulations are performed in the R programming language on the system with Intel Core i9-9900K CPU and 64 Gb RAM;
- 2) in order to simulate paths of fractional Brownian motion, R package somebm is used;
- 3) in Example 5.3, discrete samples of the multifractional Brownian motion (mBm) values are simulated using the Cholesky decomposition of the corresponding covariance matrix (for covariance structure of the mBm, see, e.g. [5, Proposition 4]) and the R package nleqslv is used for solving (3.2) numerically.

*Example 5.1 (Generalized CIR processes)* Let  $\varphi \equiv 0$ ,  $Z$  satisfy (Z1)–(Z2) with  $\lambda \in (\frac{1}{2}, 1)$ ,  $Y(0), \kappa_1, \kappa_2 > 0, \gamma > \frac{1}{\lambda} - 1$  be given and  $\{Y(t), t \in [0, T]\}$  satisfy the SDE of the form

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{Y(s)} - \kappa_2 Y(s) \right) ds + Z(t), \quad t \in [0, T]. \tag{5.1}$$

This process fits into the framework of Section 4 and the equation for  $\hat{Y}(t_{k+1})$  from (4.2) reads as follows:

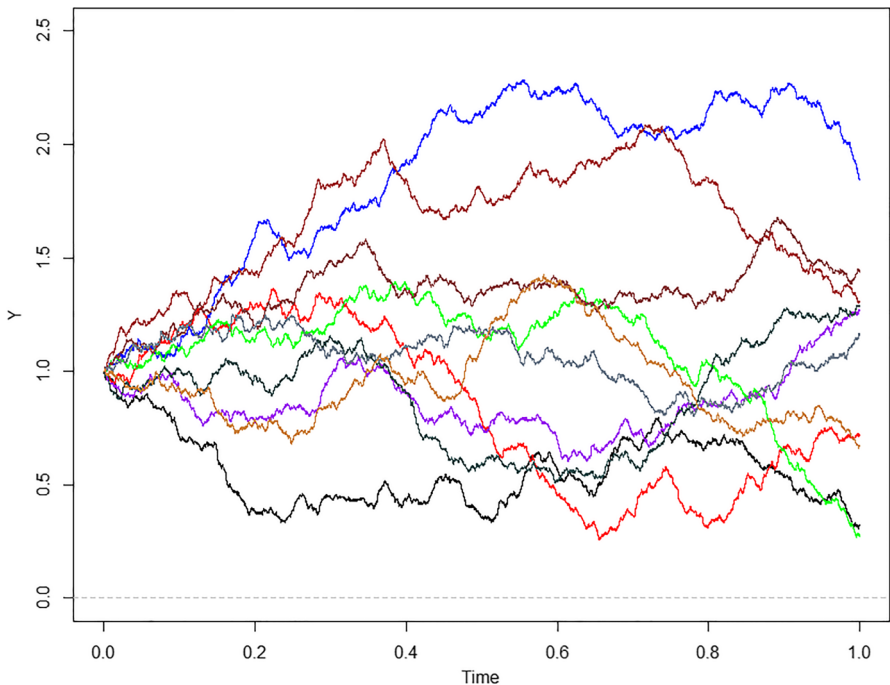
$$\hat{Y}(t_{k+1}) = \hat{Y}(t_k) + \left( \frac{\kappa_1}{\hat{Y}(t_{k+1})} - \kappa_2 \hat{Y}(t_{k+1}) \right) \Delta_N + Z(t_{k+1}) - Z(t_k).$$

It is easy to see that it has a unique positive solution

$$\hat{Y}(t_{k+1}) = \frac{\hat{Y}(t_k) + (Z(t_{k+1}) - Z(t_k)) + \sqrt{\left(\hat{Y}(t_k) + (Z(t_{k+1}) - Z(t_k))\right)^2 + 4\kappa_1 \Delta_N (1 + \kappa_2 \Delta_N)}}{2(1 + \kappa_2 \Delta_N)}.$$

Figure 1 contains 10 sample paths of the process (5.1) driven by a fractional Brownian motion with  $H = 0.7$ . In all simulation, we take  $N = 10000$ ,  $T = 1$ , and  $Y(0) = 1 = \kappa_1 = \kappa_2 = 1$ .

In order to illustrate the convergence, we also simulate the drift-implicit approximation  $\tilde{Y}$  with a small step size  $10^{-6}$  (it will serve as the “exact” solution). Then, using the same path of  $Z$ , we generate the drift-implicit Euler approximations with step sizes of the form  $1/N$ , where  $N$  runs over all divisors of  $10^6$ . Afterwards, we compute the  $L^\infty([0, T])$ -distances between the “exact” solution and its approximations with larger step sizes. This procedure is performed 10000 times and the mean square error of each  $L^\infty([0, T])$ -distance is computed. The resulting values



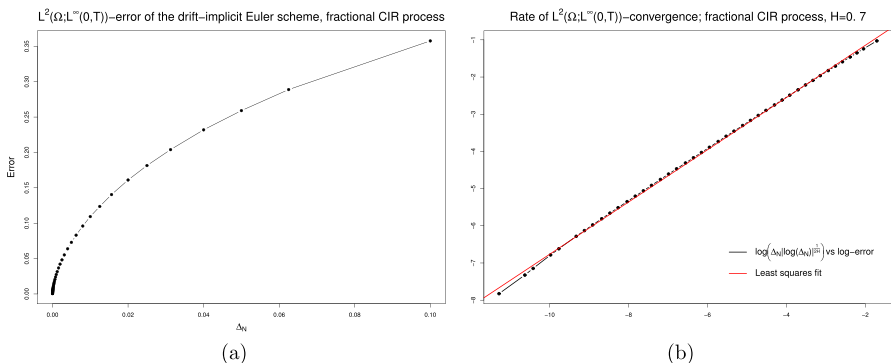
**Fig. 1** Ten sample paths of (5.1) generated using the drift-implicit Euler approximation scheme;  $N = 10000$ ,  $T = 1$ ,  $Y(0) = \kappa_1 = \kappa_2 = 1$ ,  $Z$  is a fractional Brownian motion with  $H = 0.7$

serve as consistent estimators of the corresponding  $L^2(\Omega; L^\infty([0, T]))$ -errors and are depicted on Fig. 2a.

Note that the drift-implicit Euler scheme for (5.1) driven by the fractional Brownian motion was the main subject of [18] and [22], but in both cases, the convergence of  $\widehat{Y}$  to  $Y$  is established only on  $[0, T]$  with  $T$  being small (see, e.g. [18, Eq. (8) and Remark 3.1]). Our results fill this gap and prove that convergence holds on arbitrary  $[0, T]$  for any model parameters. It should be noted though that the convergence rate in Theorem 4.8 is not optimal and can be improved for the fractional Brownian driver. It is well known that paths of a fractional Brownian motion are Hölder continuous up to (but not including) its Hurst index  $H$  and whence Theorem 4.8 indicates that the exact convergence speed of the drift-implicit Euler scheme is better than  $O(\Delta_N^\lambda)$  for any  $\lambda \in (0, H)$ . In turn, [18] uses the results on the modulus of the continuity of the fractional Brownian motion and establishes that the exact speed of convergence is  $O(\Delta_N^H \sqrt{|\log(\Delta_N)|})$  (provided that  $T$  is small enough). On Fig. 2b, we plot the values of  $\log(\Delta_N |\log(\Delta_N)|^{\frac{1}{2H}})$  against the logarithms of the corresponding  $L^2(\Omega; L^\infty([0, T]))$ -errors from Fig. 2a. The resulting points (depicted in black) turn out to be located along the line with the slope  $0.7022687 \approx 0.7 = H$  (depicted in red; least squares method was used to estimate the slope). This gives an empirical evidence to the conjecture that additional conditions on  $T$  in [18] can be lifted and the speed  $O(\Delta_N^H \sqrt{|\log(\Delta_N)|})$  is still preserved.

*Example 5.2 (Sandwiched process of the TSB type)* Consider a sandwiched SDE of the form

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{Y(s) - \varphi(s)} - \frac{\kappa_2}{\psi(s) - Y(s)} - \kappa_3 Y(s) \right) ds + Z(t), \quad t \in [0, T], \tag{5.2}$$



**Fig. 2** Convergence analysis of the drift-implicit Euler approximation scheme for (5.1);  $T = 1, Y(0) = \kappa_1 = \kappa_2 = 1, Z$  is a fractional Brownian motion with  $H = 0.7$ . On panel **a**,  $L^2(\Omega; L^\infty([0, T]))$ -errors are depicted. Panel **b** contains the values of  $\log(\Delta_N |\log(\Delta_N)|^{\frac{1}{2H}})$  plotted against the logarithms of the corresponding  $L^2(\Omega; L^\infty([0, T]))$ -errors (black) as well as the line fitted with the least squares method (red). The slope of the red line is  $0.7022687 \approx 0.7 = H$

where  $Z$  satisfies **(Z1)–(Z2)** with  $\lambda \in (\frac{1}{2}, 1)$ . This equation fits into the framework of Section 4 and the scheme (3.2) leads to  $N$  cubic equations of the form

$$\widehat{Y}^3(t_{n+1}) + B_{2,n}\widehat{Y}^2(t_{n+1}) + B_{1,n}\widehat{Y}(t_{n+1}) + B_{0,n} = 0, \quad n = 0, \dots, N - 1, \quad (5.3)$$

where

$$B_{0,n} := \frac{-\varphi(t_{n+1})\psi(t_{n+1})(\widehat{Y}(t_n) + \Delta Z_n) + \Delta_N(\kappa_1\psi(t_{n+1}) + \kappa_2\varphi(t_{n+1}))}{1 + \Delta_N\kappa_3},$$

$$B_{1,n} := \varphi(t_{n+1})\psi(t_{n+1}) + \frac{(\varphi(t_{n+1}) + \psi(t_{n+1}))(\widehat{Y}(t_n) + \Delta Z_n) - \Delta_N(\kappa_1 + \kappa_2)}{1 + \Delta_N\kappa_3},$$

$$B_{2,n} := -\varphi(t_{n+1}) - \psi(t_{n+1}) - \frac{\widehat{Y}(t_n) + \Delta Z_n}{1 + \Delta_N\kappa_3},$$

Note this equation can be solved explicitly using, e.g. the celebrated Cardano method. Namely, define

$$Q_{1,n} := B_{1,n} - \frac{B_{2,n}^2}{3}, \quad Q_{2,n} := \frac{2B_{2,n}^3}{27} - \frac{B_{2,n}B_{1,n}}{3} + B_{0,n}$$

and put

$$Q_n := \left(\frac{Q_{1,n}}{3}\right)^3 + \left(\frac{Q_{2,n}}{2}\right)^2,$$

$$\alpha_n := \sqrt[3]{-\frac{Q_{2,n}}{2} + \sqrt{Q_n}}, \quad \beta_n := \sqrt[3]{-\frac{Q_{2,n}}{2} - \sqrt{Q_n}},$$

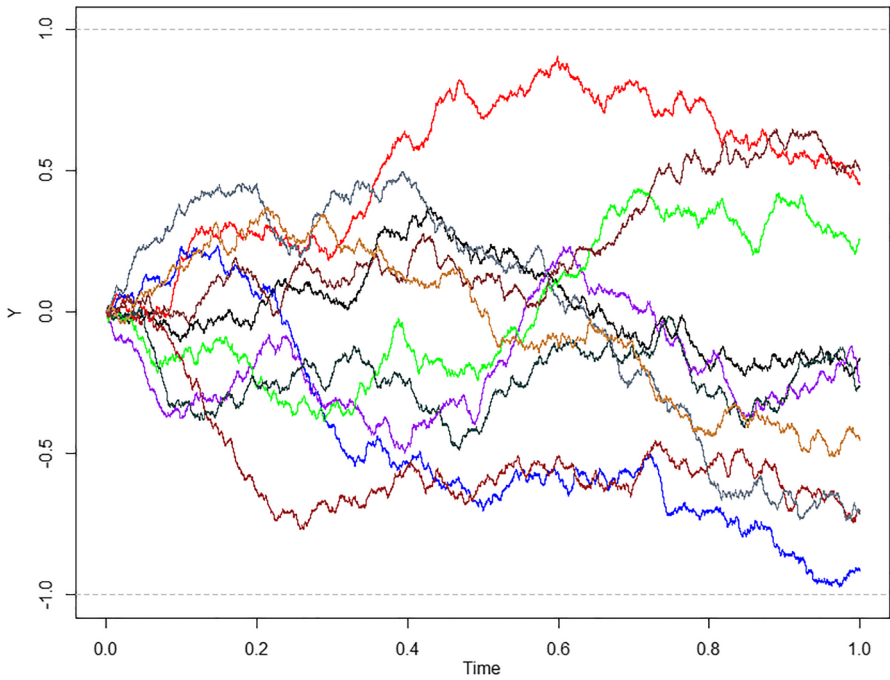
where among possible complex values of  $\alpha_n$  and  $\beta_n$ , one should take those for which  $\alpha_n\beta_n = -\frac{Q_{1,n}}{3}$ . Then the three roots of the cubic (5.3) are

$$y_{1,n} = \alpha_n + \beta_n, \quad y_{2,n} = -\frac{\alpha_n + \beta_n}{2} + i\frac{\alpha_n - \beta_n}{2}\sqrt{3},$$

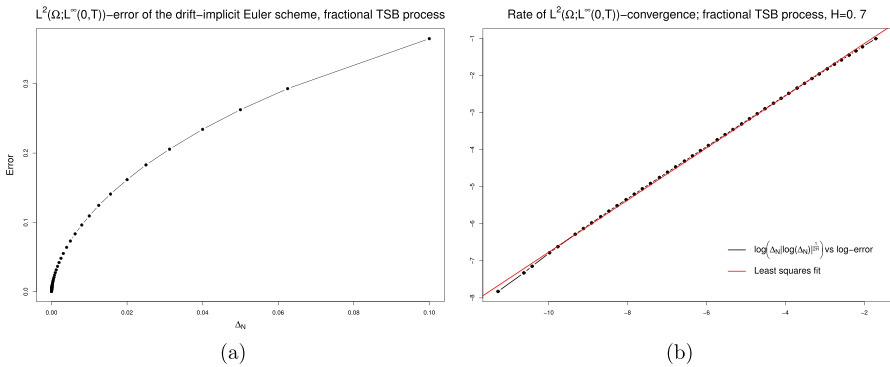
$$y_{3,n} = -\frac{\alpha_n + \beta_n}{2} - i\frac{\alpha_n - \beta_n}{2}\sqrt{3},$$

and  $\widehat{Y}(t_{n+1})$  is equal to the root which belongs to  $(\varphi(t_{n+1}), \psi(t_{n+1}))$  (note that there is exactly one root in that interval).

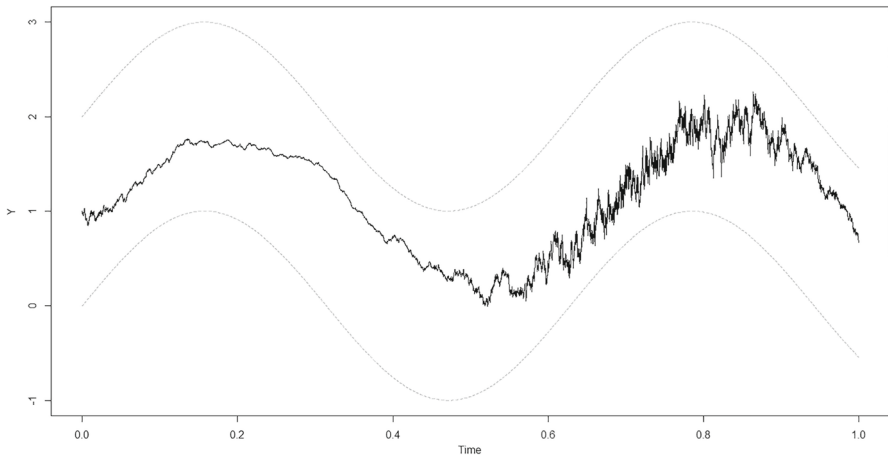
Figure 3 contains 10 sample paths of the process (5.2) driven by a fractional Brownian motion with  $H = 0.7$ . In all simulations, we take  $\varphi \equiv -1$ ,  $\psi \equiv 1$ ,  $N = 10000$ ,  $T = 1$ , and  $Y(0) = 0$ ,  $\kappa_1 = \kappa_2 = \frac{1}{2}$ ,  $\kappa_3 = 0$  (this case corresponds to the TSB equation described in Example 2.8). Simulation is performed by direct implementation of the Cardano’s method in R. On Fig. 4a, the  $L^2(\Omega; L^\infty([0, T]))$ -errors are depicted. Just as in Example 5.1, behaviour of the modulus of continuity of the fractional Brownian motion allows to suggest that the exact convergence speed of the numerical scheme is  $O\left(\Delta_N^H \sqrt{|\log(\Delta_N)|}\right)$ . Figure 4b gives an empirical evidence to this conjecture: the values of  $\log\left(\Delta_N |\log(\Delta_N)|^{\frac{1}{2H}}\right)$  plotted against the logarithms of the corresponding  $L^2(\Omega; L^\infty([0, T]))$ -errors (black) are located along the line (red) with the slope  $0.7033434 \approx 0.7 = H$  (least squares fit was used).



**Fig. 3** Ten sample paths of (5.2) generated using the drift-implicit Euler approximation scheme;  $\varphi \equiv -1$ ,  $\psi \equiv 1$ ,  $N = 10000$ ,  $T = 1$ ,  $Y(0) = 0$ ,  $\kappa_1 = \kappa_2 = \frac{1}{2}$ ,  $\kappa_3 = 0$ ,  $Z$  is a fractional Brownian motion with  $H = 0.7$



**Fig. 4** Convergence analysis of the drift-implicit Euler approximation scheme for (5.2);  $\varphi \equiv -1$ ,  $\psi \equiv 1$ ,  $T = 1$ ,  $Y(0) = 0$ ,  $\kappa_1 = \kappa_2 = \frac{1}{2}$ ,  $\kappa_3 = 0$ ,  $Z$  is a fractional Brownian motion with  $H = 0.7$ . On panel **a**,  $L^2(\Omega; L^\infty([0, T]))$ -errors are depicted. Panel **b** contains the values of  $\log(\Delta_N |\log(\Delta_N)|^{\frac{1}{2H}})$  plotted against the logarithms of the corresponding  $L^2(\Omega; L^\infty([0, T]))$ -errors (black) as well as the line fitted with the least squares method (red). The slope of the red line is  $0.7033434 \approx 0.7 = H$



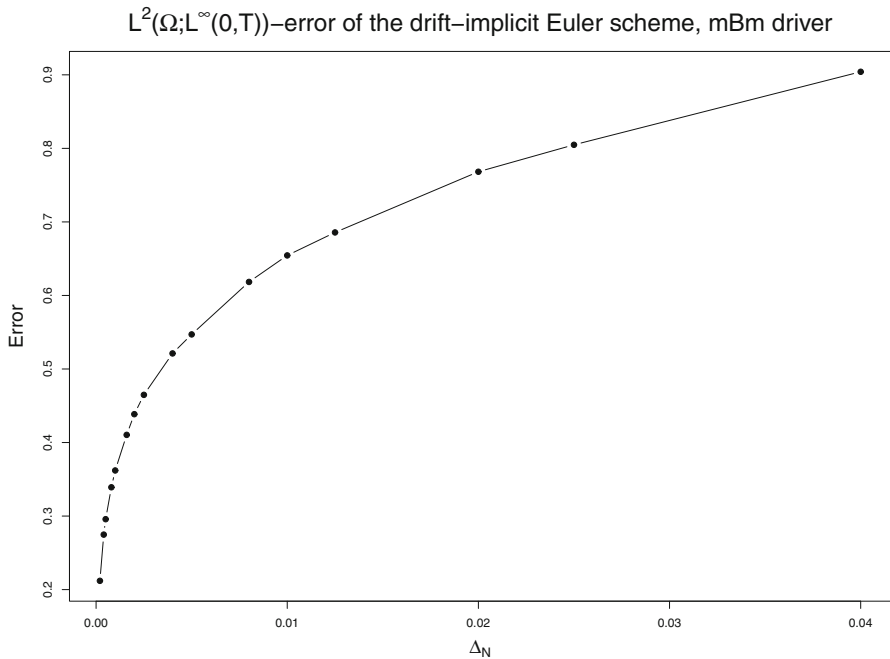
**Fig. 5** A sample path of (5.4) generated using the backward Euler approximation scheme;  $N = 10000$ ,  $T = 1$ ,  $Y(0) = 1$ ,  $\kappa_1 = \kappa_2 = 1$ ,  $\varphi(t) = \sin(10t)$ ,  $\psi(t) = \sin(10t) + 2$ ,  $Z$  is a multifractional Brownian motion with functional Hurst parameter  $H(t) = \frac{1}{5} \sin(2\pi t) + \frac{1}{2}$

In both Examples 5.1 and 5.2, equations for computing  $\widehat{Y}$  could be explicitly solved but the Hölder continuity of the noise could not be less than  $1/2$ . The next example shows that the drift-implicit Euler scheme can be applied in the rough case as well.

*Example 5.3 (Sandwiched process driven by multifractional Brownian motion)* Consider the sandwiched SDE of the form

$$Y(t) = Y(0) + \int_0^t \left( \frac{\kappa_1}{(Y(s) - \varphi(s))^4} - \frac{\kappa_2}{(\psi(s) - Y(s))^4} \right) ds + Z(t), \quad t \in [0, T]. \tag{5.4}$$

In this case, Theorem 2.1 guarantees existence and uniqueness of the solution for  $\lambda$ -Hölder  $Z$  with  $\lambda > \frac{1}{5}$  (note that this equation fits the framework of Example 2.9 from Section 2). On Fig. 5, one can see paths of the process (5.4) with  $\kappa_1 = \kappa_2 = 1$ ,  $\varphi(t) = \sin(10t)$ ,  $\psi(t) = \sin(10t) + 2$  driven by multifractional Brownian motion (mBm) with functional Hurst parameter  $H(t) = \frac{1}{5} \sin(2\pi t) + \frac{1}{2}$  (note that the lowest value of the functional Hurst parameter is  $H\left(\frac{3}{4}\right) = 0.3$ ; for more details on mBm, see [5] as well as [17, Lemma 3.1] for results on Hölder continuity of its paths). Figure 6 contains the  $L^2(\Omega; L^\infty([0, T]))$ -errors of approximation. Note a much slower rate of convergence: the multifractional Brownian motion  $Z$  under consideration is Hölder continuous up to the order 0.3; therefore, Theorem 3.9 guarantees convergence speed of only  $O(\Delta_N^\lambda)$  with  $\lambda \in (0, 0.3)$ .



**Fig. 6**  $L^2(\Omega; L^\infty([0, T]))$ -errors of the drift-implicit Euler approximation scheme for (5.4);  $T = 1$ ,  $Y(0) = 1$ ,  $\kappa_1 = \kappa_2 = 1$ ,  $\varphi(t) = \sin(10t)$ ,  $\psi(t) = \sin(10t) + 2$ ,  $Z$  is a multifractional Brownian motion with functional Hurst parameter  $H(t) = \frac{1}{5} \sin(2\pi t) + \frac{1}{2}$

**Acknowledgements** The present research is carried out within the frame and support of the ToppForsk project nr. 274410 of the Research Council of Norway with title STORM: Stochastics for Time-Space Risk Models. The second author was supported by Japan Science and Technology Agency CREST JPMJCR14D7, CREST Project reference number JPMJCR2115.

**Funding** Open access funding provided by University of Oslo (incl Oslo University Hospital)

**Data availability** The R code used to generate the sample paths from Section 5 is available from the corresponding author on reasonable request.

**Declarations**

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.


## References

1. Alfi, V., Coccetti, F., Petri, A., Pietronero, L.: Roughness and finite size effect in the NYSE stock-price fluctuations. *The European Phys. J. B* **55**(2), 135–142 (2007)
2. Alfonsi, A.: On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11**(4), 355–384 (2005)
3. Alfonsi, A.: Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. *Statist. Probab. Lett.* **83**(2), 602–607 (2013)
4. Andersen, L.B.G., Piterbarg, V.V.: Moment explosions in stochastic volatility models. *Finance Stochast.* **11**(1), 29–50 (2006)
5. Ayache, A., Cohen, S., Vehe, J.L.: The covariance structure of multifractional Brownian motion, with application to long range dependence. In: 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100), IEEE (2002)
6. Azmooedeh, E., Sottinen, T., Viitasaari, L., Yazigi, A.: Necessary and sufficient conditions for Hölder continuity of Gaussian processes. *Stat. Probability Lett.* **94**, 230–235 (2014)
7. Bayer, C., Friz, P., Gatheral, J.: Pricing under rough volatility. *Quant Finance* **16**(6), 887–904 (2016)
8. Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B.: An empirical comparison of alternative models of the short-term interest rate. *J. Finance* **47**(3), 1209–1227 (1992)
9. Cox, J.C.: The constant elasticity of variance option pricing model. *The J. Portfolio Manag.* **23**(5), 15–17 (1996)
10. Cox, J.C., Ingersoll, J.E., Ross, S.A.: A re-examination of traditional hypotheses about the term structure of interest rates. *J. Financ.* **36**(4), 769–799 (1981)
11. Cox, J.C., Ingersoll, J.E., Ross, S.A.: An intertemporal general equilibrium model of asset prices. *Econometrica* **53**(2), 363 (1985)
12. Cox, J.C., Ingersoll, J.E., Ross, S.A.: A theory of the term structure of interest rates. *Econometrica* **53**(2), 385 (1985)
13. Dereich, S., Neuenkirch, A., Szpruch, L.: An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proc. Royal Soc. A Math Phys. Eng. Sci.* **468**(2140), 1105–1115 (2012)
14. Di Nunno, G., Mishura, Y., Yurchenko-Tytarenko, A.: Sandwiched SDEs with unbounded drift driven by Hölder noises. *Adv. Appl. Probab.* **55**(3) (2022) (to appear)
15. Domingo, D., d’Onofrio, A., Flandoli, F.: Properties of bounded stochastic processes employed in biophysics. *Stoch. Anal. Appl.* **38**(2), 277–306 (2019)
16. d’Onofrio, A. (ed.): *Bounded Noises in Physics, Biology, and Engineering*. Springer New York, New York (2013)
17. Dozzi, M., Kozachenko, Y., Mishura, Y., Ralchenko, K.: Asymptotic growth of trajectories of multifractional Brownian motion, with statistical applications to drift parameter estimation. *Stat. Infer. Stoch. Process.* **21**(1), 21–52 (2018)
18. Hong, J., Huang, C., Kamrani, M., Wang, X.: Optimal strong convergence rate of a backward Euler type scheme for the Cox–Ingersoll–Ross model driven by fractional Brownian motion. *Stoch. Process. Appl.* **130**(5), 2675–2692 (2020)
19. Hu, Y., Nualart, D., Song, X.: A singular stochastic differential equation driven by fractional Brownian motion. *Statistics & Probability Letters* **78**(14), 2075–2085 (2008)
20. Kruse, R.: *Strong and weak approximation of semilinear stochastic evolution equations*. Springer International Publishing (2014)
21. Kubilius, K., Medžiūnas, A.: Positive solutions of the fractional SDEs with non-Lipschitz diffusion coefficient. *Mathematics* **9**, 1 (2020)
22. Zhang, S.-Q., Yuan, C.: Stochastic differential equations driven by fractional Brownian motion with locally Lipschitz drift and their implicit Euler approximation. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, pp 1–27 (2020)

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Affiliations

Giulia Di Nunno<sup>1,2</sup> · Yuliya Mishura<sup>3</sup> · Anton Yurchenko-Tytarenko<sup>1</sup> 

Giulia Di Nunno  
giulian@math.uio.no

Yuliya Mishura  
mys@univ.kiev.ua

- <sup>1</sup> Department of Mathematics, University of Oslo, Oslo, Norway
- <sup>2</sup> Department of Business and Management Science, NHH Norwegian School of Economics, Bergen, Norway
- <sup>3</sup> Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine