OPTIMAL STRONG APPROXIMATION OF THE ONE-DIMENSIONAL SQUARED BESSEL PROCESS

MARIO HEFTER AND ANDRÉ HERZWURM

ABSTRACT. We consider the one-dimensional squared Bessel process given by the stochastic differential equation (SDE)

$$dX_t = 1\,dt + 2\sqrt{X_t}\,dW_t, \quad X_0 = x_0, \quad t \in [0,1],$$

and study strong (pathwise) approximation of the solution X at the final time point $t = 1$. This SDE is a particular instance of a Cox-Ingersoll-Ross (CIR) process where the boundary point zero is accessible. We consider numerical methods that have access to values of the driving Brownian motion W at a finite number of time points. We show that the polynomial convergence rate of the n-th minimal errors for the class of adaptive algorithms as well as for the class of algorithms that rely on equidistant grids are equal to infinity and $1/2$, respectively. This shows that adaption results in a tremendously improved convergence rate. As a by-product, we obtain that the parameters appearing in the CIR process affect the convergence rate of strong approximation.

1. INTRODUCTION

In recent years, strong approximation of stochastic differential equations (SDEs) has intensively been studied for SDEs of the form

$$dX_t = (a - bX_t)\,dt + \sigma\sqrt{X_t}\,dW_t, \quad X_0 = x_0, \quad t \geq 0, \tag{1}$$

with a one-dimensional Brownian motion W, and $a, x_0 \geq 0$, $b \in \mathbb{R}$, and $\sigma > 0$. These SDEs are known to have a unique non-negative strong solution. Such SDEs were proposed in [11] as a model for short-term interest rates. The solution is called Cox-Ingersoll-Ross (CIR) process. Moreover, CIR processes are used as the volatility process in the Heston model [18].

Strong approximation is of particular interest due to the multi-level Monte Carlo technique, see [14, 15, 17]. For an optimality result of this technique applied to quadrature problems of the form $\mathrm{E}(f(X))$ with $f\colon C([0,1]) \to \mathbb{R}$, we refer to [12]. In mathematical finance, the functional f often represents a discounted payoff of some derivative and $\mathrm{E}(f(X))$ is the corresponding price.

In [1], various numerical schemes have been proposed and numerically tested for the SDE (1) with different choices of the corresponding parameters. These numerical results indicate a convergence at a polynomial rate, which depends on the parameters a, σ. More precisely, the empirical convergence rate is monotonically decreasing in the quotient $\sigma^2/(2a)$ for all numerical schemes that have been tested. Polynomial convergence rates for strong approximation of (1) have been proven by [4, 13, 2, 25, 19], where either a global or the final time error w.r.t. the L_p-norm is studied. All these results only hold for some parameter range within $\sigma^2/(2a) < 2$ and share the same monotonicity, see Figure 1. For an overview of numerical schemes

Key words and phrases. Cox-Ingersoll-Ross process; strong approximation; n-th minimal error; adaptive algorithm; reflected Brownian motion.

and results on strong convergence without a rate we refer to [13] and the references therein.

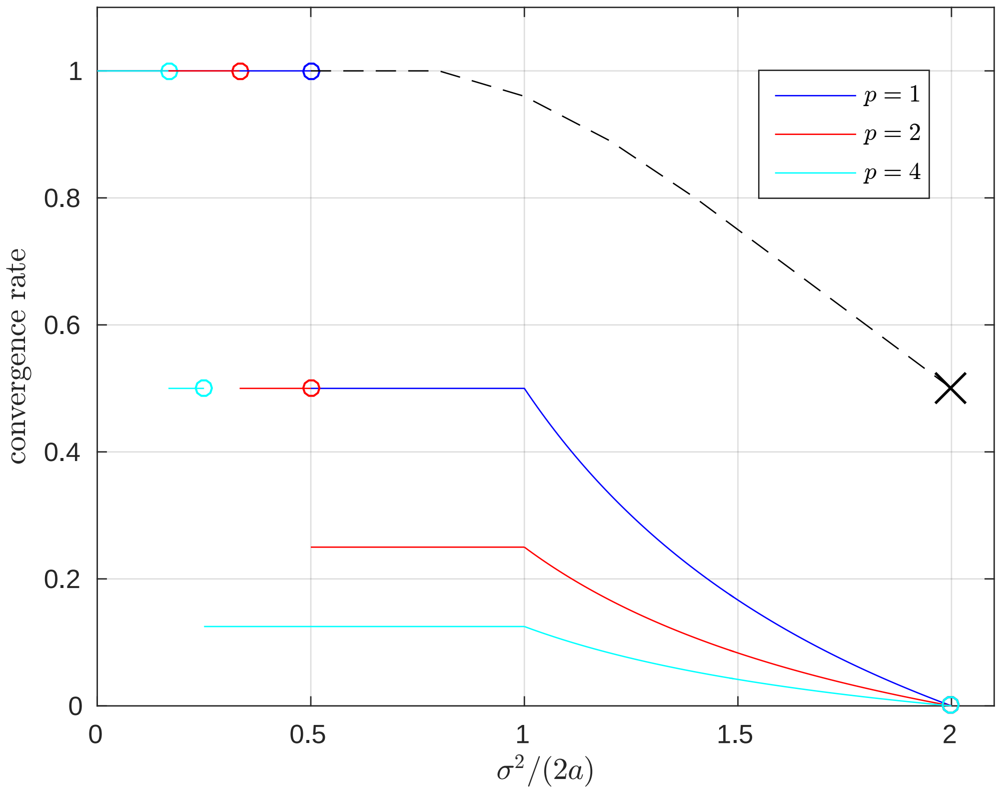

FIGURE 1. Convergence rates of the known (best) upper bounds from [13, 2, 25, 19] are shown for different values of $p \in [1, \infty[$ with error given by (4). All these upper bounds are obtained for the drift-implicit Euler scheme. The dashed line shows the corresponding empirical convergence rates from [1] for $p = 1$.

In this paper we consider the particular case of the SDE (1) with

$$\sigma^2/(2a) = 2.$$

By rescaling with $(2/\sigma)^2$, we thus may restrict ourselves to SDEs of the form

$$dX_t = (1 - bX_t)\, dt + 2\sqrt{X_t}\, dW_t, \quad X_0 = x_0, \quad t \geq 0, \tag{2}$$

with $x_0 \geq 0$ and $b \in \mathbb{R}$. Furthermore, we focus on the particular instance of (2) with $b = 0$, i.e.,

$$dX_t = 1\, dt + 2\sqrt{X_t}\, dW_t, \quad X_0 = x_0, \quad t \geq 0. \tag{3}$$

Its solution is called the square of a 1-dimensional Bessel process. For a detailed study of (squared) Bessel processes we refer to [26, Chap. XI].

In the context of (strong) approximation of SDEs, the majority of numerical methods in the literature are non-adaptive [22]. A non-adaptive algorithm uses a fixed discretisation of the driving Brownian motion whereas adaptive algorithms may sequentially choose the evaluation points. The most frequently studied methods in the class of non-adaptive algorithms are Euler or Milstein-type methods that are based on values of the driving Brownian motion on an equidistant grid. For various strong approximation problems of SDEs satisfying standard assumptions, adaption does not help up to a multiplicative constant, see, e.g., [24]. In particular, [23] showed for strong approximation of scalar SDEs at the final time point that no adaptive method can be better (up to a multiplicative constant) than the classical Milstein scheme. Let us stress that these standard assumptions are not fulfilled by (3) since the diffusion coefficient is not even locally Lipschitz continuous.

In contrast to that, the main result of this paper is that adaptive methods are far superior to methods that are based on an equidistant grid for strong approximation of the solution X_1 of (3). For this, we determine the polynomial convergence rate of the corresponding n-th minimal errors, which will be introduced below.

Let X_1 be the solution of (3) at time $t = 1$, and let $p \in [1, \infty[$. The error of an approximation $\widehat{X}_1$ of X_1 is defined by

$$e_p(\widehat{X}_1) = \left(\mathrm{E}\left(\left|X_1 - \widehat{X}_1\right|^p\right)\right)^{1/p}. \tag{4}$$

At first, we consider the class of methods that only use values of the driving Brownian motion W on an equidistant grid with n points given by

$$\mathfrak{C}^{\mathrm{eq}}(n) = \left\{\widehat{X}_1 = \Phi(W_{\frac{1}{n}}, W_{\frac{2}{n}}, \ldots, W_1) : \ \Phi \colon \mathbb{R}^n \to \mathbb{R} \text{ Borel-measurable}\right\}.$$

The corresponding n-th minimal error for the approximation of X_1 is given by

$$e_p^{\mathrm{eq}}(n) = \inf\left\{e_p(\widehat{X}_1) : \ \widehat{X}_1 \in \mathfrak{C}^{\mathrm{eq}}(n)\right\}. \tag{5}$$

Roughly speaking, $e_p^{\mathrm{eq}}(n)$ is the error of the best algorithm for the approximation of X_1 w.r.t. the L_p-norm that only uses $W_{\frac{1}{n}}, W_{\frac{2}{n}}, \ldots, W_1$. Clearly, Euler and Milstein-type schemes fit into this class of algorithms. In the case $p = 2$, the optimal approximation is given by the conditional expectation of X_1 given the σ-algebra generated by $W_{\frac{1}{n}}, W_{\frac{2}{n}}, \ldots, W_1$.

The class of adaptive methods that use values of the driving Brownian motion W at n sequentially chosen points is given by

$$\begin{aligned} \mathfrak{C}^{\mathrm{ad}}(n) = \Big\{\widehat{X}_1 = \Phi(W_{t_1}, \ldots, W_{t_n}) : \ & \Phi \colon \mathbb{R}^n \to \mathbb{R} \text{ Borel-measurable}, \\ & t_1 \in [0,1], \\ & t_2 = \varphi_2(W_{t_1}), \ \varphi_2 \colon \mathbb{R} \to [0,1] \text{ Borel-measurable}, \\ & \vdots \\ & t_n = \varphi_n(W_{t_1}, \ldots, W_{t_{n-1}}), \ \varphi_n \colon \mathbb{R}^{n-1} \to [0,1] \text{ Borel-measurable}\Big\}. \end{aligned}$$

Here, in contrast to the class $\mathfrak{C}^{\mathrm{eq}}$, the k-th evaluation site t_k may depend on the previous $k-1$ observations of W. Moreover, considering the particular choice of constant mappings $\varphi_k = k/n$ yields $\mathfrak{C}^{\mathrm{eq}}(n) \subseteq \mathfrak{C}^{\mathrm{ad}}(n)$ for all $n \in \mathbb{N}$. The n-th minimal error for the approximation of X_1 for the class of adaptive methods is given by

$$e_p^{\mathrm{ad}}(n) = \inf\left\{e_p(\widehat{X}_1) : \ \widehat{X}_1 \in \mathfrak{C}^{\mathrm{ad}}(n)\right\}. \tag{6}$$

We clearly have $e_p^{\mathrm{ad}}(n) \leq e_p^{\mathrm{eq}}(n)$ for all $n \in \mathbb{N}$.

In the following we present our main results. We write $a_n \preccurlyeq b_n$ for sequences of non-negative reals a_n and b_n if there exists a constant $c > 0$ such that $a_n \leq c \cdot b_n$ for all $n \in \mathbb{N}$. Moreover, we write $a_n \asymp b_n$ if $a_n \preccurlyeq b_n$ and $b_n \preccurlyeq a_n$.

We show that the polynomial convergence rate of the n-th minimal error e_p^{eq} is equal to $1/2$ for all $p \in [1, \infty[$. More precisely, Corollary 1 yields

$$e_p^{\mathrm{eq}}(n) \asymp n^{-1/2} \tag{7}$$

for all $p \in [1, \infty[$. Of course, the constants hidden in the "$\asymp$"-notation may depend on p. Furthermore, the corresponding upper bound is attained by the drift-implicit Euler scheme $\widehat{X}_1^{n,\mathrm{imp}}$, see Theorem 2 and (23). In the more general case of the SDE (2), the drift-implicit Euler scheme is given by

$$\widehat{X}_{\frac{k+1}{n}}^{n,\mathrm{imp}} = \left(\frac{\sqrt{\widehat{X}_{\frac{k}{n}}^{n,\mathrm{imp}}} + \left(W_{\frac{k+1}{n}} - W_{\frac{k}{n}}\right) + \sqrt{\left(\sqrt{\widehat{X}_{\frac{k}{n}}^{n,\mathrm{imp}}} + \left(W_{\frac{k+1}{n}} - W_{\frac{k}{n}}\right)\right)^2}}{2 + \frac{b}{n}} \right)^2 \tag{8}$$

for $k = 0, \ldots, n-1$ and $\widehat{X}_0^{n,\mathrm{imp}} = x_0$. Let us mention that the drift-implicit Euler scheme is actually proposed for the SDE (1) with parameters satisfying $\sigma^2/(2a) < 2$, see [1]. Nevertheless, it is still well defined in the limiting case $\sigma^2/(2a) = 2$. Note that the upper bound from (7) is the first strong convergence result with a positive rate in the case $\sigma^2/(2a) = 2$, cf. Figure 1.

For adaptive algorithms the situation is rather different. Corollary 2 shows that

$$e_p^{\mathrm{ad}}(n) \preccurlyeq n^{-q} \tag{9}$$

for all $p \in [1, \infty[$ and for all $q \in [1, \infty[$. Hence the polynomial convergence rate of the n-th minimal error e_p^{ad} is equal to infinity. More precisely, for every $q \in [1, \infty[$ we construct an adaptive algorithm that converges (at least) at a polynomial rate q, see Theorem 4. In fact, numerical experiments suggest an exponential decay, see Figure 3. Moreover, such algorithms can be easily implemented on a computer with number of operations of order n^2. Combining (7) and (9) establishes our claim that adaptive algorithms are far superior to non-adaptive algorithms that are based on equidistant grids for strong approximation of (3). Let us stress that this is the first result on SDEs where adaption results in an improved convergence rate compared to methods that are based on equidistant grids.

A key step for the proofs of (7) and (9) consists of identifying the pathwise solution of (3), see Proposition 1, and link this problem to global optimization under the Wiener measure. Let us mention that the analysis of the adaptive algorithm in Theorem 4 heavily relies on results of [9].

Although we have shown that adaptive algorithms are far superior to methods that are based on an equidistant grid for a particular choice of the parameters of SDE (1), it is open whether this superiority also holds for more general parameter constellations.

We now turn to the more general case of the SDE (2) with $b \in \mathbb{R}$. Moreover, we consider a stronger error criterion which is pathwise given by the supremum norm. In this case we obtain

$$\left(\mathrm{E} \left(\sup_{0 \le t \le 1} \left| X_t - \overline{X}_t^n \right|^p \right) \right)^{1/p} \preccurlyeq n^{-1/2} \cdot \sqrt{\ln(1+n)} \tag{10}$$

for all $p \in [1, \infty[$, where $\overline{X}^n$ denotes a projected equidistant Euler scheme, see Remark 7. This scheme coincides with the drift-implicit Euler scheme for $b = 0$. Let us stress that this error bound is the first strong convergence result with a positive rate in the case $\sigma^2/(2a) = 2$ and arbitrary $b \in \mathbb{R}$, cf. Figure 1. At present, we have only shown the upper bound (10). Nevertheless, we expect this upper bound to be sharp even for adaptive algorithms.

Let us briefly comment on some consequences of the results presented above for strong approximation of CIR processes.

It is well-known that the parameters a, b, and σ in SDE (1) have an influence on the behavior of its solution. For instance, the solution remains strictly positive (the boundary point 0 is inaccessible) if and only if the so-called Feller condition $\sigma^2/(2a) \leq 1$ is satisfied. As illustrated in Figure 1, the drift-implicit Euler scheme converges at least with rate 1 if $\sigma^2/(2a) < \min(2/(3p), 1/2)$, see [2, 25], and so does the corresponding n-th minimal error for methods using an equidistant grid. Hence the quotient $\sigma^2/(2a)$ affects the convergence rate of the n-th minimal error for equidistant methods since it drops down to $1/2$ for $\sigma^2/(2a) = 2$ and $b = 0$ according to (7).

In contrast to the known upper bounds, cf. Figure 1, the convergence rate of the drift-implicit Euler scheme for (3) does not depend on the L_p-norm appearing in the error criterion, see (7).

Let us comment on lower bounds for strong approximation of SDEs at the final time point based on the values of the driving Brownian motion. In [10], a two-dimensional SDE is presented where the corresponding convergence rate is shown to be 1/2. In contrast to rate 1 for smooth scalar SDEs [23], the difficulty in [10] arises from the presence of Lévy areas. More recently, the existence of SDEs with smooth coefficients has been shown where the corresponding n-th minimal error converges arbitrarily slow to zero, see [16, 20]. It is crucial that these SDEs are multi-dimensional. Apart from (7), we are not aware of any other lower bound with convergence rate less than 1 for a scalar SDE.

This paper is organized as follows. In Section 2 we derive an explicit representation of the solution of (3) and the more general case of (2). Using this representation we show sharp upper and lower bounds of e_p^{eq} in Section 3. In Section 4 we consider a particular adaptive method that achieves an arbitrarily high polynomial convergence rate. Finally, we illustrate our results by numerical experiments.

2. Squared Bessel Process of Dimension One

In this section, we will derive an explicit expression for the strong solution of (3) by using basic results about reflected SDEs. Subsequently, we will extend this technique to the more general case of SDE (2).

In the following let $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete probability space and let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration on this space satisfying the usual conditions.

Given $x_0 \geq 0$ and a Brownian motion B w.r.t. $(\mathcal{F}_t)_{t\geq 0}$, we define

$$W_t = \int_0^t \operatorname{sgn}(B_s + \sqrt{x_0})\, dB_s \tag{11}$$

for all $t \geq 0$ with $\operatorname{sgn} = 1_{\{x>0\}} - 1_{\{x\leq 0\}}$. Then, W is a Brownian motion w.r.t. $(\mathcal{F}_t)_{t\geq 0}$. Indeed, the quadratic variation of W satisfies

$$[W]_t = \int_0^t \operatorname{sgn}(B_s + \sqrt{x_0})^2\, ds = \int_0^t 1\, ds = t,$$

and thus Lévy's characterization can be applied. Now, consider the SDE (3) where the driving Brownian motion W has the particular form given in (11). Due to this construction of W, we see that the solution of (3) is given by

$$X_t = (B_t + \sqrt{x_0})^2, \tag{12}$$

since $\sqrt{X_t} = \left|B_t + \sqrt{x_0}\right|$ and hence

$$\begin{aligned} 2\int_0^t \sqrt{X_s}\, dW_s &= 2\int_0^t |B_t + \sqrt{x_0}| \cdot \operatorname{sgn}(B_s + \sqrt{x_0})\, dB_s \\ &= 2\int_0^t (B_s + \sqrt{x_0})\, dB_s \\ &= B_t^2 - t + 2\sqrt{x_0}\, B_t. \end{aligned}$$

Moreover, Tanaka's formula [21, Prop. III.6.8] given by

$$|B_t - a| = |a| + \int_0^t \operatorname{sgn}(B_s - a)\, dB_s + 2L_t^B(a), \quad a \in \mathbb{R},$$

where $L^B(a)$ denotes the local time of B in a, yields for $a = -\sqrt{x_0}$ that

$$\begin{aligned} |B_t + \sqrt{x_0}| &= \sqrt{x_0} + W_t + 2L_t^B(-\sqrt{x_0}) \\ &= \sqrt{x_0} + W_t + \max\left(0, \sup_{0\le s\le t} -(\sqrt{x_0} + W_s)\right), \end{aligned}$$

where the second equality follows from Skorokhod's lemma, see [21, Lem. III.6.14]. Finally, using $\sup(-A) = -\inf(A)$ for $A \subseteq \mathbb{R}$ leads to the solution of (3) given by

$$X_t = \left((W_t + \sqrt{x_0}) + \left(\inf_{0\le s\le t} W_s + \sqrt{x_0}\right)^-\right)^2, \tag{13}$$

where $x^- = -\min(0, x)$ denotes the negative part of x. We stress that the explicit solution (13) of the SDE (3) holds for any Brownian motion W and does not depend on the particular construction given in (11), see [21, Cor. V.3.23], since pathwise uniqueness and strong existence holds for the SDE (3). Hence the unique strong solution of the SDE (3) is given by (13).

Remark 1. It is well-known that the solution of the SDE (3) can be expressed in terms of B in (12), see, e.g., [26, Ex. IX.3.16]. However, we are not aware of a result regarding the explicit form of the strong solution given by (13).

In the context of SDEs, the somehow explicit solution of X by means of B in (12) is rather useless for strong approximation. The concept of strong solutions entails a functional dependence of the solution process and the input Brownian motion appearing in the SDE, which is W in our case. We thus seek to "construct" the solution X out of W.

Remark 2. Equation (12) clearly yields

$$(X_t)_{t\ge0} \overset{\mathrm{d}}{=} \left((W_t + \sqrt{x_0})^2\right)_{t\ge0},$$

cf. [21, Thm. III.6.17]. However, this equation is only valid in the distributional sense and does not hold pathwise. Note that the Brownian path attains its running minimum whenever the solution hits 0. More precisely, we have

$$X_t = 0 \quad \Leftrightarrow \quad W_t \le -\sqrt{x_0} \ \wedge\ W_t = \inf_{0\le s\le t} W_s, \tag{14}$$

cf. Figure 2.

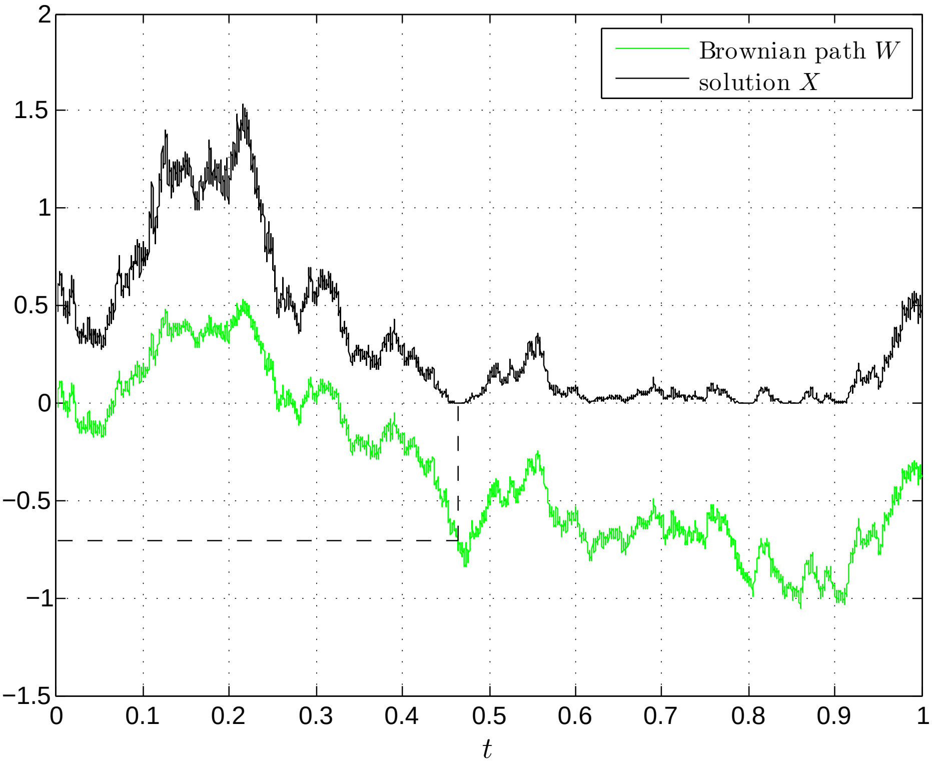

FIGURE 2. Brownian path and corresponding solution (13) with initial condition $x_0 = 0.5$. The solution hits the boundary 0 according to (14).

Remark 3. The expression

$$(W_t + \sqrt{x_0}) + \left(\inf_{0 \le s \le t} W_s + \sqrt{x_0}\right)^-$$

appearing in (13) is known as a reflected Brownian motion (with reflecting barrier at 0) starting in $\sqrt{x_0} \ge 0$, see [29].

We now turn to the more general case of SDE (2) with arbitrary $b \in \mathbb{R}$.

Proposition 1. *The unique strong solution of the SDE* (2) *is given by*

$$X_t = \left(u_t + e^{-\frac{b}{2}t}\left(\inf_{0 \le s \le t} e^{\frac{b}{2}s}\, u_s\right)^-\right)^2,$$

where u denotes the unique strong solution of the SDE

$$du_t = -\frac{b}{2}\, u_t\, dt + dW_t, \quad u_0 = \sqrt{x_0}, \quad t \ge 0. \tag{15}$$

In particular, the unique strong solution of the SDE (3) *is given by* (13).

Remark 4. Note that the solution to the linear SDE (15) is called Ornstein-Uhlenbeck process and can be solved explicitly by

$$u_t = e^{-\frac{b}{2}t}\left(\sqrt{x_0} + \int_0^t e^{\frac{b}{2}s}\, dW_s\right), \quad t \geq 0, \tag{16}$$

see, e.g., [21, Ex. V.6.8)].

Proof of Proposition 1. Analogous to the above derivation, we start with a Brownian motion B and consider the SDE

$$du_t^B = -\frac{b}{2}\, u_t^B\, dt + dB_t, \quad u_0^B = \sqrt{x_0}, \quad t \geq 0.$$

Moreover, we assume that the Brownian motion appearing in (2) has the particular form

$$W_t = \int_0^t \operatorname{sgn}(u_s^B)\, dB_s, \quad t \geq 0.$$

Then, Itô's formula shows

$$\begin{aligned} d\left(\left(u_t^B\right)^2\right) &= \left(1 - b\left(u_t^B\right)^2\right) dt + 2\, u_t^B\, dB_t \\ &= \left(1 - b\left(u_t^B\right)^2\right) dt + 2\left|u_t^B\right| dW_t. \end{aligned}$$

Hence the solution of SDE (2) is given by

$$X_t = \left(u_t^B\right)^2, \quad t \geq 0. \tag{17}$$

On the other hand, the Tanaka-Meyer formula [21, Thm. III.7.1(v)] applied to

$$\tilde{u}_t^B = u_t^B e^{\frac{b}{2}t}$$

combined with the explicit expression (16) yields

$$\begin{aligned} \left|\tilde{u}_t^B\right| &= \sqrt{x_0} + \int_0^t e^{\frac{b}{2}s} \operatorname{sgn}\left(\tilde{u}_s^B\right) dB_s + 2\Lambda_t^{\tilde{u}_B}(0) \\ &= \sqrt{x_0} + \int_0^t e^{\frac{b}{2}s}\, dW_s + 2\Lambda_t^{\tilde{u}_B}(0), \end{aligned}$$

where $\Lambda^{\tilde{u}_B}(0)$ denotes the semimartingale local time of $\tilde{u}^B$ at 0. Thus, Skorokhod's lemma [21, Lem. III.6.14] shows

$$\left|\tilde{u}_t^B\right| = \sqrt{x_0} + \int_0^t e^{\frac{b}{2}s}\, dW_s + \left(\inf_{0\leq s\leq t} \sqrt{x_0} + \int_0^s e^{\frac{b}{2}u}\, dW_u\right)^-,$$

and consequently

$$\left|u_t^B\right| = u_t + e^{-\frac{b}{2}t}\left(\inf_{0\leq s\leq t} e^{\frac{b}{2}s}\, u_s\right)^-.$$

It remains to apply [21, Cor. V.3.23]. □

Remark 5. Consider the situation of the proof of Proposition 1. The Tanaka-Meyer formula applied to u_t^B yields

$$\begin{aligned} \left|u_t^B\right| &= \sqrt{x_0} - \frac{b}{2}\int_0^t u_s^B \operatorname{sgn}\left(u_s^B\right) ds + \int_0^t \operatorname{sgn}\left(u_s^B\right) dB_s + 2\Lambda_t^{u_B}(0) \\ &= \sqrt{x_0} - \frac{b}{2}\int_0^t \left|u_s^B\right| ds + \int_0^t dW_s + 2\Lambda_t^{u_B}(0). \end{aligned}$$

Hence $Z_t = \left|u_t^B\right|$ is the solution of the reflected SDE on the domain $D =]0,\infty[$ given by

$$dZ_t = -\frac{b}{2}\, Z_t\, dt + dW_t + dK_t, \quad Z_0 = \sqrt{x_0}, \quad t \geq 0, \tag{18}$$

where K is a process of bounded variation with variation increasing only when $Z_t \in \partial D = \{0\}$. For details on reflected SDEs we refer to [29]. In view of (17), we can express the solution X to the SDE (2) by

$$X_t = (Z_t)^2, \quad t \geq 0.$$

3. Equidistant Methods for SDE (3)

As shown in the previous section, the SDE (3) admits the explicit solution (13). This immediately links the problem of approximating SDE (3) to global optimization under the Wiener measure. We refer to [27] for results on global optimization under the Wiener measure.

In this section we show sharp (up to constants) upper and lower bounds for e_p^{eq} for all $p \geq 1$. Recall that e_p^{eq} denotes the n-th minimal error corresponding to the approximation of SDE (3) and algorithms that are based on equidistant grids, see (5). We show that the convergence rate of e_p^{eq} is equal to $1/2$. Moreover, we show that the drift-implicit Euler scheme converges with the optimal rate $1/2$.

For the proofs we exploit results from [3] on the asymptotic error distribution of the infimum of a Brownian motion approximated by equidistant points on the unit interval. For $n \in \mathbb{N}$ we define

$$\delta_n = \sqrt{n} \cdot \left(\min_{0 \leq k \leq n} W_{\frac{k}{n}} - \inf_{0 \leq s \leq 1} W_s \right), \tag{19}$$

where W denotes a standard Brownian motion. The following result is due to [3, Thm. 1 and Lem. 6].

Theorem 1 ([3])**.**

(i) The sequence $(\delta_n)_{n\in\mathbb{N}}$ converges in distribution.

(ii) For all $p \in [1,\infty[$ the sequence $(\delta_n^p)_{n\in\mathbb{N}}$ is uniformly integrable. In particular, for all $p \in [1,\infty[$ we have

$$\left(\mathrm{E}\left(\left| \min_{0 \leq k \leq n} W_{\frac{k}{n}} - \inf_{0 \leq s \leq 1} W_s \right|^p \right) \right)^{1/p} \preccurlyeq n^{-1/2}. \tag{20}$$

Remark 6. In [3], the limiting distribution of the sequence $(\delta_n)_{n\in\mathbb{N}}$ is given explicitly by means of three-dimensional Bessel processes.

Recall that the solution of (3) is given by $X_1 = (Y_1)^2$ with

$$Y_1 = (W_1 + \sqrt{x_0}) + \left(\inf_{0 \leq s \leq 1} W_s + \sqrt{x_0} \right)^- . \tag{21}$$

Moreover, for $n \in \mathbb{N}$ we define the approximation $\widehat{X}_1^{(n)}$ of X_1 by

$$\widehat{X}_1^{(n)} = \left(\widehat{Y}_1^{(n)} \right)^2,$$

where

$$\widehat{Y}_1^{(n)} = (W_1 + \sqrt{x_0}) + \left(\min_{0 \leq k \leq n} W_{\frac{k}{n}} + \sqrt{x_0} \right)^-$$

serves as an approximation of Y_1. Here, the global infimum is simply replaced by the discrete minimum over n equidistant knots.

The following upper bound is a consequence of Theorem 1(ii).

Theorem 2. *For every $1 \leq p < \infty$ we have*

$$e_p\left(\widehat{X}_1^{(n)}\right) \preccurlyeq n^{-1/2}.$$

In particular, the n-th minimal error satisfies

$$e_p^{\mathrm{eq}}(n) \preccurlyeq n^{-1/2}.$$

Proof. At first, note that $0 \leq \widehat{Y}_1^{(n)} \leq Y_1$ and

$$\left|Y_1 - \widehat{Y}_1^{(n)}\right| = \left(\inf_{0\leq s\leq 1} W_s + \sqrt{x_0}\right)^- - \left(\min_{0\leq k\leq n} W_{\frac{k}{n}} + \sqrt{x_0}\right)^- \leq \min_{0\leq k\leq n} W_{\frac{k}{n}} - \inf_{0\leq s\leq 1} W_s.$$

Hence we get

$$\left|X_1 - \widehat{X}_1^{(n)}\right| = \left(Y_1 + \widehat{Y}_1^{(n)}\right) \cdot \left|Y_1 - \widehat{Y}_1^{(n)}\right| \leq 2Y_1 \cdot \left(\min_{0\leq k\leq n} W_{\frac{k}{n}} - \inf_{0\leq s\leq 1} W_s\right).$$

Finally, Cauchy-Schwarz inequality yields

$$\begin{aligned}\mathrm{E}\left(\left|X_1 - \widehat{X}_1^{(n)}\right|^p\right) &\leq 2^p \cdot \left(\mathrm{E}\left(Y_1^{2p}\right) \cdot \mathrm{E}\left(\left|\min_{0\leq k\leq n} W_{\frac{k}{n}} - \inf_{0\leq s\leq 1} W_s\right|^{2p}\right)\right)^{1/2} \\ &\preccurlyeq n^{-p/2}\end{aligned}$$

due to (20) and

$$\mathrm{E}\left(Y_1^r\right) < \infty, \quad 1 \leq r < \infty,$$

since $Y_1 \stackrel{\mathrm{d}}{=} \left|W_1 + \sqrt{x_0}\right|$, see Remark 2. ☐

The natural extension of $\widehat{X}_1^{(n)}$ to an approximation on the whole equidistant grid with mesh size $1/n$ is given by

$$\widehat{X}_{\frac{k}{n}}^{(n)} = \left(\widehat{Y}_{\frac{k}{n}}^{(n)}\right)^2, \quad k = 0, \dots, n,$$

where

$$\widehat{Y}_{\frac{k}{n}}^{(n)} = \left(W_{\frac{k}{n}} + \sqrt{x_0}\right) + \left(\min_{0\leq l\leq k} W_{\frac{l}{n}} + \sqrt{x_0}\right)^-, \quad k = 0, \dots, n.$$

Let us stress that this scheme can be expressed by the following Euler-type scheme

$$\widehat{Y}_{\frac{k+1}{n}}^{(n)} = \max\left(\widehat{Y}_{\frac{k}{n}}^{(n)} + \left(W_{\frac{k+1}{n}} - W_{\frac{k}{n}}\right), 0\right), \quad k = 0, \dots, n-1, \tag{22}$$

and $\widehat{Y}_0^{(n)} = \sqrt{x_0}$. In the context of reflected SDEs, scheme (22) is simply the projected Euler scheme for a reflected Brownian motion. Due to

$$\max(x, 0) = \frac{x + |x|}{2} = \frac{x + \sqrt{x^2}}{2}, \quad x \in \mathbb{R},$$

the drift-implicit Euler scheme (8) for $b = 0$ reads

$$\widehat{X}_{\frac{k+1}{n}}^{n,\mathrm{imp}} = \left(\max\left(\sqrt{\widehat{X}_{\frac{k}{n}}^{n,\mathrm{imp}}} + \left(W_{\frac{k+1}{n}} - W_{\frac{k}{n}}\right), 0\right)\right)^2, \quad k = 0, \dots, n-1. \tag{23}$$

Thus it coincides with the squared version of (22), i.e.,

$$\widehat{X}^{n,\mathrm{imp}}_{\frac{k}{n}} = \widehat{X}^{(n)}_{\frac{k}{n}}, \quad k = 0, \dots, n.$$

Remark 7. We briefly comment on results for the more general case of SDE (2) with arbitrary $b \in \mathbb{R}$. Moreover, we consider a stronger global error criterion where pathwise the global error is measured in the supremum norm. Up to a logarithmic factor we will obtain the same error bound as in Theorem 2.

For $n \in \mathbb{N}$ we denote by $\overline{Z}^n = (\overline{Z}^n)_{0 \le t \le 1}$ the projected Euler scheme with n steps associated to the reflected SDE (18) up to time $t = 1$, see [28]. More precisely, $\overline{Z}^n$ is defined by

$$\begin{aligned} \overline{Z}^n_0 &= \sqrt{x_0}, \\ \overline{Z}^n_{\frac{k+1}{n}} &= \max\left(\overline{Z}^n_{\frac{k}{n}} - \frac{b}{2} \cdot \overline{Z}^n_{\frac{k}{n}} \cdot \frac{1}{n} + \left(W_{\frac{k+1}{n}} - W_{\frac{k}{n}}\right), 0\right), \quad k = 0, \dots, n-1, \end{aligned}$$

and piecewise constant interpolation, i.e.,

$$\overline{Z}^n_t = \overline{Z}^n_{\frac{k}{n}}, \quad t \in [k/n, (k+1)/n[, \quad k = 0, \dots, n-1.$$

The solution X to SDE (2) is then approximated by

$$\overline{X}^n_t = \left(\overline{Z}^n_t\right)^2, \quad t \in [0,1].$$

At the grid points, this scheme coincides with the drift-implicit Euler scheme (23) if $b = 0$. Similar to the proof of Theorem 2, we obtain

$$\left(\mathrm{E}\left(\sup_{0 \le t \le 1} \left|X_t - \overline{X}^n_t\right|^p\right)\right)^{1/p} \preccurlyeq n^{-1/2} \cdot \sqrt{\ln(1+n)}$$

for all $p \in [1, \infty[$ due to [28, Cor. 2.5, Cor. 2.6 and Thm. 3.2(i)] and Remark 5. Let us stress that this error bound is the first strong convergence result with a positive rate in the case $\sigma^2/(2a) = 2$, cf. Figure 1.

We now turn to the question whether an algorithm can do better than $\widehat{X}^{(n)}_1$ in an asymptotic sense if this algorithm has the same information about the Brownian motion as $\widehat{X}^{(n)}_1$.

Recall the definition of the n-th minimal error e^{eq}_p given in (5). The proof of the following theorem is postponed to Section 5.

Theorem 3. *For all* $p \in [1, \infty[$ *we have*

$$e^{\mathrm{eq}}_p(n) \succcurlyeq n^{-1/2}.$$

Combining Theorem 2 and Theorem 3 yields the following asymptotic behavior of the n-th minimal error.

Corollary 1. *For all* $p \in [1, \infty[$ *we have*

$$e^{\mathrm{eq}}_p(n) \asymp n^{-1/2}.$$

In particular, the drift-implicit Euler scheme (23) *is asymptotically optimal.*

Remark 8. In Corollary 1 we obtain the same rate as in [27] for global optimization. In [27], the author studies optimal approximation of the time point where a Brownian motion attains its maximum, and provides a detailed analysis of general non-adaptive algorithms that do not necessarily rely on equidistant grids.

Remark 9. If we allow for more information about the Brownian path than just point evaluations $W_{\frac{1}{n}}, \ldots, W_1$, the situation may change completely. For instance, if we consider algorithms that have access to the final value W_1 and to the infimum $\inf_{0\leq s\leq 1} W_s$ of the Brownian path, the problem of strong approximation of X_1 becomes trivial. Let us stress that the joint distribution of $(\inf_{0\leq s\leq 1} W_s, W_1)$ has an explicit representation by means of a Lebesgue density, see, e.g., [6, p. 154].

4. Adaptive Methods for SDE (3)

In this section we present an adaptive algorithm for the approximation of the solution of SDE (3) based on sequential observations $W_{t_1}, W_{t_2}, \ldots$ of the Brownian motion W. In contrast to Section 3, here the points $t_1, t_2, \ldots$ are chosen adaptively, i.e., the k-th evaluation site t_k is a measurable function of $W_{t_1}, \ldots, W_{t_{k-1}}$. From Proposition 1 it is clear that the actual task consists of the approximation of the global infimum $\inf_{s\in[0,1]} W_s$. For this we use the adaptive algorithm from [9], see also [7, 8]. This algorithm approximates $\inf_{s\in[0,1]} W_s$ by the discrete minimum $\min_{0\leq k\leq n} W_{t_k}$. In the following we describe the adaptive choice of $t_1, t_2, \ldots, t_n$.

The first observation is non-adaptively chosen to be the endpoint, i.e., $t_1 = 1$. Moreover, for notational convenience we define $t_0 = 0$.

Let $n \in \mathbb{N}$, and consider the $(n+1)$-th step of the algorithm where t_{n+1} will be chosen based on the previous observations $W_{t_1}, \ldots, W_{t_n}$. For this, we denote the ordered first n evaluation sites by

$$0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = 1,$$

such that $\{t_0, \ldots, t_n\} = \{t_0^{(n)}, \ldots, t_n^{(n)}\}$. Moreover, we assume that we have made the following observations

$$W_{t_0^{(n)}} = y_0^{(n)} = 0, \quad W_{t_1^{(n)}} = y_1^{(n)}, \quad \ldots, \quad W_{t_n^{(n)}} = y_n^{(n)}, \tag{24}$$

and we denote the corresponding discrete minimum by

$$m^{(n)} = \min_{0\leq k\leq n} y_k^{(n)}.$$

Conditioned on (24), we have independent Brownian bridges from $y_{k-1}^{(n)}$ to $y_k^{(n)}$ on the subinterval $[t_{k-1}^{(n)}, t_k^{(n)}]$, for $k \in \{1, \ldots, n\}$. In the following we denote a Brownian bridge from x to y on $[0, T]$ with $x, y \in \mathbb{R}$ and $T > 0$ by $B^{x,T,y}$.

The basic idea of the adaptive algorithm is a simple greedy strategy: The next observation is taken at the midpoint of the subinterval where the probability that the corresponding Brownian bridge undershoots the current discrete minimum minus some threshold $\varepsilon^{(n)} > 0$ is maximal. More precisely, we split the interval according to

$$k^* = \operatorname{argmax}_{1\leq k\leq n} \mathrm{P}\left(\inf_{0\leq s\leq T_k^{(n)}} B_s^{y_{k-1}^{(n)}, T_k^{(n)}, y_k^{(n)}} \leq m^{(n)} - \varepsilon^{(n)}\right) \tag{25}$$

with $T_k^{(n)} = t_k^{(n)} - t_{k-1}^{(n)}$, and evaluate W at

$$t_{n+1} = \left(t_{k^*}^{(n)} + t_{k^*-1}^{(n)}\right)/2.$$

We note that the infimum of a Brownian bridge $(B_t^{x,T,y})_{t\in[0,T]}$ satisfies

$$\mathrm{P}\left(\inf_{0\le s\le T} B_s^{x,T,y} < z\right) = \exp\left(-\frac{2}{T}\,(x-z)(y-z)\right)$$

for $z \le \min(x,y)$, see [6, p. 67]. Hence (25) reduces to maximizing

$$k^* = \mathrm{argmax}_{1\le k\le n}\ \frac{T_k^{(n)}}{\left(y_{k-1}^{(n)} - m^{(n)} + \varepsilon^{(n)}\right)\cdot\left(y_k^{(n)} - m^{(n)} + \varepsilon^{(n)}\right)}.$$

Finally, we have to specify the threshold $\varepsilon^{(n)}$. This threshold is chosen to be

$$\varepsilon^{(n)} = \sqrt{\lambda\, h^{(n)}\, \ln(1/h^{(n)})},$$

where $h^{(n)} = \min_{1\le k\le n} T_k^{(n)}$ denotes the length of the smallest subinterval at step n and where $\lambda \in [1,\infty[$ is some prespecified parameter. Let us stress that all above (adaptive) quantities depend on the choice of the parameter λ, although it is not explicitly indicated.

This amounts to a family of adaptive algorithms defined by

$$\widehat{X}_{\mathrm{ad},\lambda}^{(n)} = \left(W_{t_1} + \sqrt{x_0} + \left(\min_{0\le k\le n} W_{t_k} + \sqrt{x_0}\right)^-\right)^2 \tag{26}$$

for the approximation of (13), where the adaptively chosen points $t_1,\dots,t_n$ depend on the prespecified choice of $\lambda \in [1,\infty[$.

Remark 10. A straightforward implementation of the algorithm (26) on a computer requires operations of order n^2.

The following result is an immediate consequence of [9, Thm. 1].

Theorem 4. *For all $p \in [1,\infty[$ and for all $q \in [1,\infty[$ there exists $\lambda \in [1,\infty[$ such that*

$$e_p\left(\widehat{X}_{\mathrm{ad},\lambda}^{(n)}\right) \preccurlyeq n^{-q}.$$

Remark 11. The analysis in [9] shows that

$$\lambda \ge 144\cdot(1+2pq)$$

is sufficient to obtain a convergence order $q \in [1,\infty[$ for the L_p-norm in Theorem 4. However, numerical experiments indicate an exponential decay even for small values of λ, see Figure 3.

Recall the definition of the n-th minimal error e_p^{ad} given in (6).

Corollary 2. *For all $p \in [1,\infty[$ and for all $q \in [1,\infty[$ we have*

$$e_p^{\mathrm{ad}}(n) \preccurlyeq n^{-q}.$$

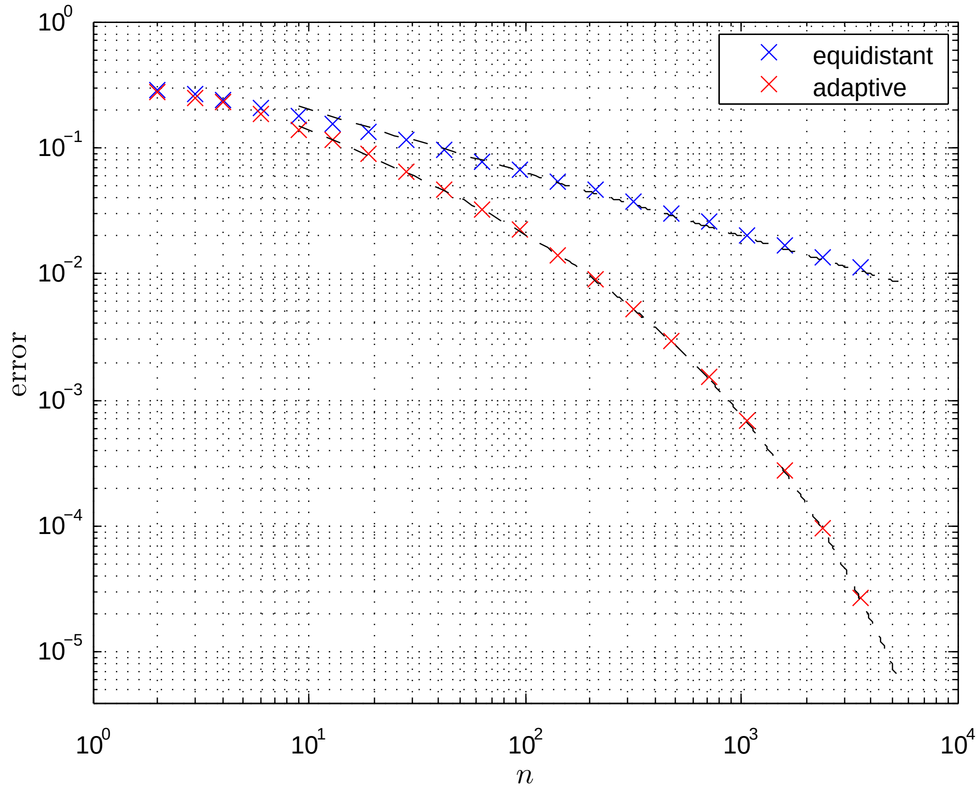

FIGURE 3. Numerical results for the drift-implicit Euler scheme (23) (blue) and the adaptive algorithm (26) with $\lambda = 4$ (red). The error given by $e_p(\cdot)$ with $p = 2$ is estimated based on 10^4 samples. The dashed lines show $0.63 \cdot n^{-1/2}$ and $0.64 \cdot n^{-0.54} \exp(-0.095\sqrt{n})$.

5. Proofs

In this section we provide the proof of Theorem 3, which relies on Lemma 2 and Lemma 3.

For $n \in \mathbb{N}$ we define $i_n^* \in \{0, \dots, n\}$ to be an index that satisfies

$$W_{\frac{i_n^*}{n}} = \min_{0 \le i \le n} W_{\frac{i}{n}}.$$

Note that i_n^* is (almost surely) uniquely defined.

Lemma 1. *For $z \ge 0$ we have*

$$\inf_{n \in \mathbb{N}} \mathrm{P}\left(\left| W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}} \right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -z \right) > 0.$$

Proof. Let $n \in \mathbb{N}$. Due to conditional independence we get

$$\begin{aligned}
&\mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -z\right) \\
&= \sum_{i=0}^{n} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -z \,\middle|\, i_n^* = i\right) \cdot \mathrm{P}\left(i_n^* = i\right) \\
&= \sum_{i=0}^{n} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}} \,\middle|\, i_n^* = i\right) \cdot \mathrm{P}\left(W_{\frac{i_n^*}{n}} \le -z \,\middle|\, i_n^* = i\right) \cdot \mathrm{P}\left(i_n^* = i\right).
\end{aligned}$$

Moreover, straightforward calculations show

$$\begin{aligned}
\mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}} \,\middle|\, i_n^* = i\right) &= \mathrm{P}\left(\left|W_{\frac{1}{n}}\right| \le \frac{1}{\sqrt{n}} \,\middle|\, W_{\frac{1}{n}} \ge 0,\ \ldots,\ W_{\frac{n-i}{n}} \ge 0\right) \\
&= \mathrm{P}\left(|W_1| \le 1 \,\middle|\, W_1 \ge 0,\ \ldots,\ W_{n-i} \ge 0\right) \\
&\ge \mathrm{P}(0 \le W_1 \le 1)
\end{aligned}$$

for $i \in \{0, \ldots, n\}$. Thus we have

$$\begin{aligned}
&\mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -z\right) \\
&\qquad \ge \mathrm{P}(0 \le W_1 \le 1) \cdot \mathrm{P}\left(W_{\frac{i_n^*}{n}} \le -z\right) \ge \mathrm{P}(0 \le W_1 \le 1) \cdot \mathrm{P}\left(W_1 \le -z\right) > 0,
\end{aligned}$$

which completes the proof. □

Lemma 2. *There exist $n_0 \in \mathbb{N}$ and $0 < \varepsilon_0 < 1$ such that*

$$\mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0},\ W_{\frac{i_n^*}{n}} \le W_1 - \varepsilon_0\right) \ge \varepsilon_0$$

for all $n \ge n_0$.

Let us mention that the events

$$W_{\frac{i_n^*}{n}} \le -\sqrt{x_0} \quad \wedge \quad W_{\frac{i_n^*}{n}} \le W_1 - \varepsilon_0$$

simply ensure that reflection occurs and that the discrete minimum (and thus the global infimum) is not attained at the final time point $t = 1$.

Proof of Lemma 2. Donsker's invariance principle and [5, Thm. 2.7] yield

$$\left(W_{\frac{i_n^*}{n}} - W_1\right)_n \xrightarrow{\mathrm{d}} \inf_{0 \le s \le 1} W_s - W_1.$$

Thus the portmanteau theorem implies

$$\liminf_{n\to\infty} \mathrm{P}\left(W_{\frac{i_n^*}{n}} - W_1 \le -\varepsilon\right) \ge \mathrm{P}\left(\inf_{0 \le s \le 1} W_s - W_1 < -\varepsilon\right)$$

for $\varepsilon > 0$. Moreover, since $\mathrm{P}\left(\inf_{0\le s\le 1} W_s - W_1 < 0\right) = 1$, there exists $\varepsilon_0 > 0$ such that

$$\begin{aligned}
&\liminf_{n\to\infty} \mathrm{P}\left(W_{\frac{i_n^*}{n}} - W_1 \le -\varepsilon_0\right) \\
&\qquad \ge 1 - \frac{1}{2} \cdot \inf_{n \in \mathbb{N}} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0}\right)
\end{aligned}$$

due to Lemma 1. Hence we get

$$\begin{aligned}
&\liminf_{n\to\infty} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0},\ W_{\frac{i_n^*}{n}} \le W_1 - \varepsilon_0\right) \\
&\ge \liminf_{n\to\infty} \left(\mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0}\right) + \mathrm{P}\left(W_{\frac{i_n^*}{n}} \le W_1 - \varepsilon_0\right) - 1\right) \\
&\ge \inf_{n\in\mathbb{N}} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0}\right) + \liminf_{n\to\infty} \mathrm{P}\left(W_{\frac{i_n^*}{n}} \le W_1 - \varepsilon_0\right) - 1 \\
&\ge \frac{1}{2} \cdot \inf_{n\in\mathbb{N}} \mathrm{P}\left(\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \le \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \le -\sqrt{x_0}\right) > 0
\end{aligned}$$

due to Lemma 1. □

Lemma 3. *Let $0 < \varepsilon_0 < 1$ be according to Lemma 2. Then there exists a constant $c_0 > 0$ such that*

$$\mathrm{P}\left(\left|X_1 - \widehat{X}_1^{(n)}\right| \le \frac{c_0}{2\sqrt{n}}\right) \ge 1 - \frac{\varepsilon_0}{4}$$

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

Proof. According to Theorem 1(i) and the portmanteau theorem, we have

$$\liminf_{n\to\infty} \mathrm{P}\left(\delta_n \le c\right) \ge \mathrm{P}(\delta < c)$$

for all $c \in \mathbb{R}$, where δ denotes the limit of $(\delta_n)_{n\in\mathbb{N}}$ given by (19). In particular, there exists $c > 0$ such that

$$\mathrm{P}\left(\min_{0\le k\le n} W_{\frac{k}{n}} - \inf_{0\le s\le 1} W_s \le \frac{c}{\sqrt{n}},\ Y_1 \le c\right) \ge 1 - \frac{\varepsilon_0}{4}$$

for all $n \in \mathbb{N}$ and Y_1 given by (21). Finally, noting that

$$\left|X_1 - \widehat{X}_1^{(n)}\right| \le 2Y_1 \cdot \left(\min_{0\le k\le n} W_{\frac{k}{n}} - \inf_{0\le s\le 1} W_s\right),$$

cf. Theorem 2, implies

$$\left\{\min_{0\le k\le n} W_{\frac{k}{n}} - \inf_{0\le s\le 1} W_s \le \frac{c}{\sqrt{n}},\ Y_1 \le c\right\} \subseteq \left\{\left|X_1 - \widehat{X}_1^{(n)}\right| \le \frac{c_0}{2\sqrt{n}}\right\}$$

with $c_0 = 4c^2$. □

Proof of Theorem 3. At first, note that Jensen's inequality implies

$$e_1^{\mathrm{eq}}(n) \le e_p^{\mathrm{eq}}(n)$$

for all $n \in \mathbb{N}$ and $p \in [1, \infty[$. Thus it suffices to consider $p = 1$, i.e., we will show the existence of a constant $c_1 > 0$ such that

$$\mathrm{E}\left(\left|X_1 - \widehat{X}_1\right|\right) \ge c_1 \cdot n^{-1/2}$$

for all $n \in \mathbb{N}$ and for all random variables $\widehat{X}_1$ that are measurable w.r.t. the σ-algebra $\mathcal{A}_n$ generated by $W_{\frac{1}{n}}, W_{\frac{2}{n}}, \ldots, W_1$.

Let $n_0 \in \mathbb{N}$, $\varepsilon_0 > 0$, and $c_0 > 0$ be according to Lemma 2 and Lemma 3, respectively. Without loss of generality, we may assume that $n \ge n_0$. In the following, we consider two cases separately.

At first, suppose that

$$\mathrm{P}\left(\left|\widehat{X}_1 - \widehat{X}_1^{(n)}\right| > \frac{c_0}{\sqrt{n}}\right) \geq \frac{\varepsilon_0}{2}.$$

By using reverse triangle inequality

$$\left|X_1 - \widehat{X}_1\right| \geq \left|\widehat{X}_1^{(n)} - \widehat{X}_1\right| - \left|X_1 - \widehat{X}_1^{(n)}\right|,$$

we get

$$\left\{\left|\widehat{X}_1^{(n)} - \widehat{X}_1\right| > \frac{c_0}{\sqrt{n}}\right\} \cap \left\{\left|X_1 - \widehat{X}_1^{(n)}\right| \leq \frac{c_0}{2\sqrt{n}}\right\} \subseteq \left\{\left|X_1 - \widehat{X}_1\right| > \frac{c_0}{2\sqrt{n}}\right\},$$

and thus

$$\begin{aligned}\mathrm{P}\left(\left|X_1 - \widehat{X}_1\right| > \frac{c_0}{2\sqrt{n}}\right) &\geq \mathrm{P}\left(\left|\widehat{X}_1^{(n)} - \widehat{X}_1\right| > \frac{c_0}{\sqrt{n}},\ \left|X_1 - \widehat{X}_1^{(n)}\right| \leq \frac{c_0}{2\sqrt{n}}\right)\\ &\geq \frac{\varepsilon_0}{2} + \left(1 - \frac{\varepsilon_0}{4}\right) - 1\end{aligned}$$

due to Lemma 3. This yields

$$\mathrm{E}\left(\left|X_1 - \widehat{X}_1\right|\right) \geq \frac{c_0}{2\sqrt{n}} \cdot \mathrm{P}\left(\left|X_1 - \widehat{X}_1\right| > \frac{c_0}{2\sqrt{n}}\right) \geq \frac{c_0\,\varepsilon_0}{8} \cdot \frac{1}{\sqrt{n}}. \tag{27}$$

Now suppose that

$$\mathrm{P}\left(\left|\widehat{X}_1 - \widehat{X}_1^{(n)}\right| \leq \frac{c_0}{\sqrt{n}}\right) > 1 - \frac{\varepsilon_0}{2},$$

and define

$$A_n = \left\{\left|W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*+1}{n}}\right| \leq \frac{1}{\sqrt{n}},\ W_{\frac{i_n^*}{n}} \leq -\sqrt{x_0},\ W_{\frac{i_n^*}{n}} \leq W_1 - \varepsilon_0,\ \left|\widehat{X}_1 - \widehat{X}_1^{(n)}\right| \leq \frac{c_0}{\sqrt{n}}\right\}.$$

Let us stress that $A_n \in \mathcal{A}_n$ and

$$\mathrm{P}(A_n) > \varepsilon_0 + \left(1 - \frac{\varepsilon_0}{2}\right) - 1 = \frac{\varepsilon_0}{2}$$

due to Lemma 2. Moreover, we observe that reverse triangle inequality

$$\left|X_1 - \widehat{X}_1\right| \geq \left|X_1 - \widehat{X}_1^{(n)}\right| - \left|\widehat{X}_1^{(n)} - \widehat{X}_1\right|$$

yields

$$\left\{\left|X_1 - \widehat{X}_1^{(n)}\right| \geq \frac{2c_0}{\sqrt{n}}\right\} \cap \left\{\left|\widehat{X}_1 - \widehat{X}_1^{(n)}\right| \leq \frac{c_0}{\sqrt{n}}\right\} \subseteq \left\{\left|X_1 - \widehat{X}_1\right| \geq \frac{c_0}{\sqrt{n}}\right\}.$$

Combining this with

$$\begin{aligned}\left|X_1 - \widehat{X}_1^{(n)}\right| &= \left(W_{\frac{i_n^*}{n}} - \inf_{0\leq s\leq 1} W_s\right) \cdot \left(W_1 - W_{\frac{i_n^*}{n}} + W_1 - \inf_{0\leq s\leq 1} W_s\right)\\ &\geq \left(W_{\frac{i_n^*}{n}} - \inf_{0\leq s\leq 1} W_s\right) \cdot 2\varepsilon_0\end{aligned}$$

on A_n, we obtain

$$A_n \cap \left\{W_{\frac{i_n^*}{n}} - \inf_{0\leq s\leq 1} W_s \geq \frac{c_0}{\varepsilon_0\sqrt{n}}\right\} \subseteq \left\{\left|X_1 - \widehat{X}_1\right| \geq \frac{c_0}{\sqrt{n}}\right\}.$$

Furthermore, we have

$$\inf_{0\le s\le 1} W_s \le W_{\frac{i_n^*}{n}+\frac{1}{2n}}$$

on A_n, since the discrete minimum is not attained at $t=1$, and thus

$$A_n \cap \left\{ W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*}{n}+\frac{1}{2n}} \ge \frac{c_0}{\varepsilon_0\sqrt{n}} \right\} \subseteq \left\{ \left| X_1 - \widehat{X}_1 \right| \ge \frac{c_0}{\sqrt{n}} \right\}.$$

This yields

$$\begin{aligned} \mathrm{P}\left(\left| X_1 - \widehat{X}_1 \right| \ge \frac{c_0}{\sqrt{n}} \right) &\ge \mathrm{P}\left(A_n \cap \left\{ W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*}{n}+\frac{1}{2n}} \ge \frac{c_0}{\varepsilon_0\sqrt{n}} \right\} \right) \\ &= \mathrm{E}\left(1_{A_n} \cdot \mathrm{P}\left(W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*}{n}+\frac{1}{2n}} \ge \frac{c_0}{\varepsilon_0\sqrt{n}} \,\middle|\, \mathcal{A}_n \right) \right) \end{aligned}$$

due to $A_n \in \mathcal{A}_n$. Conditioned on $W_{\frac{i}{n}} = y_i$ and $W_{\frac{i+1}{n}} = y_{i+1}$ for $y_i, y_{i+1} \in \mathbb{R}$, we have

$$W_{\frac{i}{n}+\frac{1}{2n}} \sim \mathcal{N}((y_i + y_{i+1})/2, 1/(4n))$$

according to the Brownian bridge construction of W. Hence we get

$$1_{A_n} \cdot \mathrm{P}\left(W_{\frac{i_n^*}{n}} - W_{\frac{i_n^*}{n}+\frac{1}{2n}} \ge \frac{c_0}{\varepsilon_0\sqrt{n}} \,\middle|\, \mathcal{A}_n \right) = 1_{A_n} \cdot f\left(W_{\frac{i_n^*+1}{n}} - W_{\frac{i_n^*}{n}} \right),$$

where $f : \mathbb{R} \to \mathbb{R}$ is given by

$$f(x) = \mathrm{P}\left(\frac{Z}{\sqrt{4n}} \ge \frac{c_0}{\varepsilon_0\sqrt{n}} + \frac{x}{2} \right)$$

with $Z \sim \mathcal{N}(0,1)$. Finally, using

$$1_{A_n} \cdot f\left(W_{\frac{i_n^*+1}{n}} - W_{\frac{i_n^*}{n}} \right) \ge 1_{A_n} \cdot f\left(\frac{1}{\sqrt{n}} \right) = 1_{A_n} \cdot \mathrm{P}\left(Z \ge \frac{2c_0}{\varepsilon_0} + 1 \right),$$

we obtain

$$\mathrm{P}\left(\left| X_1 - \widehat{X}_1 \right| \ge \frac{c_0}{\sqrt{n}} \right) \ge \mathrm{P}(A_n) \cdot \mathrm{P}\left(Z \ge \frac{2c_0}{\varepsilon_0} + 1 \right) \ge \frac{\varepsilon_0}{2} \cdot \mathrm{P}\left(Z \ge \frac{2c_0}{\varepsilon_0} + 1 \right)$$

and hence

$$\mathrm{E}\left(\left| X_1 - \widehat{X}_1 \right| \right) \ge \frac{c_0}{\sqrt{n}} \cdot \mathrm{P}\left(\left| X_1 - \widehat{X}_1 \right| \ge \frac{c_0}{\sqrt{n}} \right) \ge \frac{c_0\,\varepsilon_0}{2\sqrt{n}} \cdot \mathrm{P}\left(Z \ge \frac{2c_0}{\varepsilon_0} + 1 \right). \tag{28}$$

Combining (27) and (28) completes the proof. □

Acknowledgement

We thank James M. Calvin, Martin Hutzenthaler, and Klaus Ritter for valuable discussions and comments.

References

[1] Aurélien Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, 11, 2005.

[2] Aurélien Alfonsi. Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process. *Statistics and Probability Letters*, 83(2):602–607, 2013.

[3] Søren Asmussen, Peter Glynn, and Jim Pitman. Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann. Appl. Probab.*, 5(4): 875–896, 1995.

[4] Abdel Berkaoui, Mireille Bossy, and Awa Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 1 2008.

[5] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999.

[6] Andrei N. Borodin and Paavo Salminen. *Handbook of Brownian Motion – Facts and Formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002.

[7] James M. Calvin. Average performance of a class of adaptive algorithms for global optimization. *Ann. Appl. Probab.*, 7(3):711–730, 1997.

[8] James M. Calvin. A one-dimensional optimization algorithm and its convergence rate under the Wiener measure. *J. Complexity*, 17(2):306–344, 2001.

[9] James M. Calvin, Mario Hefter, and André Herzwurm. Adaptive approximation of the minimum of Brownian motion. *ArXiv e-prints*, 1601.01276, 2016.

[10] J. M. C. Clark and R. J. Cameron. The maximum rate of convergence of discrete approximations for stochastic differential equations. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 162–171. Springer, Berlin-New York, 1980.

[11] John C. Cox, Jonathan E. Ingersoll Jr., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[12] Jakob Creutzig, Steffen Dereich, Thomas Müller-Gronbach, and Klaus Ritter. Infinite-dimensional quadrature and approximation of distributions. *Foundations of Computational Mathematics*, 9(4):391–429, 2009.

[13] Steffen Dereich, Andreas Neuenkirch, and Lukasz Szpruch. An Euler-type method for the strong approximation of the Cox-Ingersoll-Ross process. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 468(2140):1105–1115, 2012.

[14] Michael B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[15] Michael B. Giles. Multilevel Monte Carlo methods. *Acta Numer.*, 24:259–328, 2015.

[16] Martin Hairer, Martin Hutzenthaler, and Arnulf Jentzen. Loss of regularity for Kolmogorov equations. *Ann. Probab.*, 43(2):468–527, 2015.

[17] Stefan Heinrich. Monte Carlo complexity of global solution of integral equations. *J. Complexity*, 14(2):151–175, 1998.

[18] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[19] Martin Hutzenthaler, Arnulf Jentzen, and Marco Noll. Strong convergence rates and temporal regularity for Cox-Ingersoll-Ross processes and Bessel processes with accessible boundaries. *ArXiv e-prints*, mar 2014.

[20] Arnulf Jentzen, Thomas Müller-Gronbach, and Larisa Yaroslavtseva. On stochastic differential equations with arbitrary slow convergence rates for strong approximation. *ArXiv e-prints*, jun 2015.

[21] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1988.

[22] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics 23. Berlin: Springer, 2010.

[23] Thomas Müller-Gronbach. Optimal pointwise approximation of SDEs based on Brownian motion at discrete points. *Ann. Appl. Probab.*, 14(4):1605–1642, 2004.

[24] Thomas Müller-Gronbach and Klaus Ritter. Minimal errors for strong and weak approximation of stochastic differential equations. In *Monte Carlo and quasi-Monte Carlo methods 2006*, pages 53–82. Springer, Berlin, 2008.

[25] Andreas Neuenkirch and Lukasz Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, pages 1–34, 2014.

[26] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999.

[27] Klaus Ritter. Approximation and optimization on the Wiener space. *J. Complexity*, 6(4):337–364, 1990.

[28] Leszek Słomiński. Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Process. Appl.*, 94(2):317–337, 2001.

[29] Hiroshi Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.

Fachbereich Mathematik, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany

E-mail address: hefter@mathematik.uni-kl.de

Fachbereich Mathematik, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany

E-mail address: herzwurm@mathematik.uni-kl.de