



## **Speeding up the Euler scheme for killed diffusions**

**LSE Research Online URL for this paper:** <http://eprints.lse.ac.uk/120789/>

Version: Published Version

---

### **Article:**

Cetin, Umut ORCID: 0000-0001-8905-853X and Hok, Julien (2024) Speeding up the Euler scheme for killed diffusions. *Finance and Stochastics*, 28 (3). 663 - 707. ISSN 0949-2984

<https://doi.org/10.1007/s00780-024-00534-4>

---

### **Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>



# Speeding up the Euler scheme for killed diffusions

Umut Çetin<sup>1</sup> · Julien Hok<sup>2</sup>

Received: 22 December 2021 / Accepted: 8 August 2023  
© The Author(s) 2024

## Abstract

Let  $X$  be a linear diffusion taking values in  $(\ell, r)$  and consider the standard Euler scheme to compute an approximation to  $\mathbb{E}[g(X_T)\mathbf{1}_{\{T < \zeta\}}]$  for a given function  $g$  and a deterministic  $T$ , where  $\zeta = \inf\{t \geq 0 : X_t \notin (\ell, r)\}$ . It is well known since Gobet (Stoch. Process. Appl. 87:167–197, 2000) that the presence of killing introduces a loss of accuracy and reduces the weak convergence rate to  $1/\sqrt{N}$  with  $N$  being the number of discretisations. We introduce a drift-implicit Euler method to bring the convergence rate back to  $1/N$ , i.e., the optimal rate in the absence of killing, using the theory of recurrent transformations developed in Çetin (Ann. Appl. Probab. 28:3102–3151, 2018). Although the current setup assumes a one-dimensional setting, multidimensional extension is within reach as soon as a systematic treatment of recurrent transformations is available in higher dimensions.

**Keywords** Diffusions with killing · Euler–Maruyama scheme · Drift-implicit scheme · Weak convergence · Recurrent transformations · Strict local martingales · Kato classes · Barrier options

**Mathematics Subject Classification** 60J60 · 60H35 · 91G80

**JEL Classification** C63 · G13

## 1 Introduction

Let  $X$  be a diffusion on some filtered probability space taking values in  $(\ell, r)$  and solving

$$X_t = x + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds, \quad t < \zeta, \quad (1.1)$$

---

✉ U. Çetin  
[u.cetin@lse.ac.uk](mailto:u.cetin@lse.ac.uk)

J. Hok  
[julienhok@yahoo.fr](mailto:julienhok@yahoo.fr)

<sup>1</sup> Department of Statistics, London School of Economics and Political Science, 10 Houghton St, London, WC2A 2AE, UK

<sup>2</sup> Investec Bank, 30 Gresham St, London EC2V 7QN, UK

where  $B$  is a Brownian motion and  $\zeta := \inf\{t \geq 0 : X_t \notin (\ell, r)\}$  is the first exit time from the interval  $(\ell, r)$ . The process is *killed* at  $\zeta$  and sent to a cemetery state.

Let us assume that at least one of the boundaries is accessible and  $\zeta$  is finite a.s., and consider  $\mathbb{E}[g(X_T)\mathbf{1}_{\{T < \zeta\}}]$  for a given function  $g$  and a deterministic  $T$ . Such computations appear very naturally in many applied problems of science, engineering and finance. For instance, in mathematical finance theory, such an expectation corresponds to the price of a barrier option with payoff  $g$  and maturity  $T$  written on a stock whose price process is given by  $X$ . The barrier feature renders the option worthless if the stock price hits one of the accessible boundaries before the maturity of the option.

A closed-form expression for  $\mathbb{E}[g(X_T)\mathbf{1}_{\{T < \zeta\}}]$  is rarely available even in this one-dimensional setting. Thus one needs to resort to an approximation scheme for an answer. Arguably the easiest approach is to run a standard Euler–Maruyama scheme on the SDE (1.1) by setting

$$\bar{X}_{t_{n+1}} = \bar{X}_{t_n} + \sigma(\bar{X}_{t_n})(B_{t_{n+1}} - B_{t_n}) + b(\bar{X}_{t_n})\frac{T}{N},$$

where  $\bar{X}_0 = x$ ,  $t_0 = 0$ ,  $N > 0$  is an integer,  $t_n = \frac{nT}{N}$  for  $n = 1, \dots, N$ , and compute  $\mathbb{E}[g(\bar{X}_T)\mathbf{1}_{\{T < \tau\}}]$ , where  $\tau$  is the first time that the discrete-time process  $(\bar{X}_{t_n})_{n=0,1,\dots,N}$  hits one of the barriers. Under standard regularity conditions on the diffusion process and  $g$ , such a scheme indeed converges as  $N \rightarrow \infty$ . However, it converges at a rate much slower than a standard Euler–Maruyama scheme applied to a diffusion process that is not killed at accessible boundaries.

Indeed it was shown by Gobet [21] that under standard hypotheses, the above scheme for the killed diffusion converges weakly at rate  $N^{-1/2}$  as opposed to  $N^{-1}$ , which is the rate of weak convergence for the Euler–Maruyama scheme in the absence of killing (see e.g. Talay and Tubaro [44] or Mikulevičius and Platen [34]). This rate is optimal since it is reached when  $X$  is a Brownian motion and  $g$  is an indicator function of a set strictly contained in  $(\ell, r)$  (see Siegmund and Yuh [43]).

Çetin [9] conjectured that using a *recurrent transformation* would bring the convergence rate back to  $N^{-1}$ . A recurrent transformation at heart is a change of measure that keeps the Markovian structure intact while transforming the process into a recurrent one. In particular,  $X$  never touches the boundaries of  $(\ell, r)$  under the new measure  $\mathbb{Q}$ . The article [9] shows that  $\mathbb{Q}$  is locally absolutely continuous with respect to the original measure  $\mathbb{P}$ , and  $X$  follows

$$dX_t = \sigma(X_t)dW_t + \left( b(X_t) + \sigma^2(X_t)\frac{h'}{h}(X_t) \right)dt \tag{1.2}$$

for some function  $h$  and a  $\mathbb{Q}$ -Brownian motion  $W$ . That the above claim was a conjecture and not following immediately from the standard results on Euler–Maruyama schemes is because  $\frac{h'}{h}$  is explosive near boundaries and is not Lipschitz, which is in fact needed for  $X$  not to touch the previously accessible boundaries after the measure change. This can create significant difficulties with approximation and may even lead to divergence (see e.g. the potential issues that may arise with non-Lipschitz drivers and methods on how to resolve them in Hutzenthaler et al. [27, 28]).

In this paper, we prove the above conjecture with a slight “twist”. Note that if one applies the Euler–Maruyama scheme naively to (1.2), one obtains as usual a Brownian motion with drift whose parameters change at the times of discretisation. This process will hit finite boundaries with positive probability and therefore will exit the state space of  $X$  with positive probability. One way to overcome this is to impose an ad hoc reflection on the boundaries. However, this will introduce a local-time term in computations requiring additional estimates on its convergence rate to 0. Moreover, it is far from obvious that reflection is the optimal resolution of problems arising from the discretised process exiting the domain.

We instead study a drift-implicit method that keeps the state space intact after discretisation. To see how, suppose that  $b \equiv 0$ , which can be obtained by changing the scale if necessary, and consider the backward Euler–Maruyama scheme

$$\widehat{X}_{t_{n+1}} = \widehat{X}_{t_n} + \sigma(\widehat{X}_{t_n})(B_{t_{n+1}} - B_{t_n}) + \frac{T}{N} \sigma^2(\widehat{X}_{t_n}) \frac{h'}{h}(\widehat{X}_{t_{n+1}}), \tag{1.3}$$

where  $h$  becomes a concave function. Note that differently than what one would expect from a backward scheme (see e.g. Mao and Szpruch [33], Alfonsi [3, 4] and Neuenkirch and Szpruch [36], to name a few), the  $\sigma^2$ -term in the drift of (1.2) is still evaluated at  $\widehat{X}_{t_n}$ . This stems from the fact that (1.2) with  $b \equiv 0$  should be viewed as a time-changed version of

$$dY_t = dW_t + \frac{h'}{h}(Y_t)dt,$$

where the time change is given by  $\int_0^t \sigma^2(Y_s)ds$ . We make extensive use of this correspondence in our proofs.

Our main result is Theorem 4.4 which proves that the rate of weak convergence of the above backward Euler–Maruyama scheme is  $N^{-1}$  under standard assumptions on the diffusion process. Moreover, there is no single function  $h$  that achieves this rate. We show that any nonnegative concave  $h$  vanishing at accessible boundaries can be used to obtain this convergence rate as long as it satisfies some mild growth conditions. Such functions are easy to construct, and we study in Sect. 5 constructions of some particular  $h$ -functions to compute approximate prices for barrier options in a Black–Scholes framework. In fact, we observe fast convergence in our numerical studies even in the absence of the growth conditions imposed by our theoretical analysis. Our numerical results are very promising, and error terms converge to 0 very rapidly even with a small number of iterations. Moreover, in the case of a particular local volatility model with double barriers, our method yields smaller error terms than the so-called *Brownian bridge method* when the number of discretisations is reasonably large.

We are not the first to consider implicit schemes for studying diffusions with infinite lifetime and taking values in a strict subset of  $\mathbb{R}$ . Alfonsi [3, 4] and Neuenkirch and Szpruch [36] consider such scalar processes whose SDE representation is given by

$$dY_t = dW_t + f(Y_t)dt, \tag{1.4}$$

with  $f$  satisfying the conditions of a Feller test ensuring that  $Y$  takes values in  $(\ell, r)$  (see also Dereich et al. [15] in the special case of the Cox–Ingersoll–Ross (CIR) process). The articles [4] and [36] show that the drift-implicit Euler scheme for  $Y$  converges strongly with rate  $N^{-1}$  if  $f$  satisfies certain integrability conditions including

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T (f'(Y_t))^2 dt \right] < \infty. \tag{1.5}$$

However, this condition cannot be satisfied by an  $h$  that paves the way for a recurrent transformation rendering  $X$  recurrent and following (1.2). (See Proposition C.1 in Appendix C for a proof in case  $b \equiv 0$  and  $\sigma \equiv 1$ .)

The estimates obtained in [4] and [36] rely on the Burkholder–Davis–Gundy (BDG) inequality and integrability of the relevant quadratic variation process, which require the corresponding local martingale to be a true martingale. As  $\frac{1}{h(X)}$  is a strict local submartingale under  $\mathbb{Q}$ , one needs to develop new techniques to arrive at the needed estimate for convergence theorems.

This brings to the fore another novelty of our paper. Given the impossibility of the use of the BDG inequality, we use potential-theoretic methods that yield the boundedness of inverse moments of  $h(X)$  under  $\mathbb{Q}$ , which is crucial for obtaining the weak convergence result in our paper (or strong convergence type results considered by [4, 36]). We use the theory of *Kato class* potentials to show the boundedness of the required moments. Kato potentials are one of the fundamental objects in the study of Schrödinger operators (see e.g. Aizenman and Simon [1], Cranston et al. [13], Chen [11] and Chen and Song [12]). We show in Theorem 2.6 that the additive functional  $dA_t = -\frac{1}{2} \frac{h''(X_t)}{h(X_t)} dt$  belongs to a particular Kato class defined in [11], which in turn yields the boundedness of the inverse moment of  $\frac{1}{h}(\widehat{X}_{t_n})$  (uniformly in  $N$ ) in conjunction with a comparison argument via Lemma 3.2. Potential theory also helps us to prove uniform bounds on the moments of integral functionals of  $h^{-2-p}(\widehat{X}_t)$  (see Theorem 3.3 for an exact description).

Our methodology offers hope to study the convergence rates for CIR processes or diffusions that live in a bounded interval or half space that do not satisfy (1.5). We show in this paper that if one considers the 3-dimensional Bessel process

$$dX_t = dW_t + \frac{1}{X_t} dt,$$

the implicit scheme in (1.3) converges weakly at rate  $N^{-1}$ . Clearly, (1.5) is violated since the reciprocal of a 3-dimensional Bessel process is a prime example of a strict local martingale. This process satisfies the conditions of Theorem 4.4 and one obtains the optimal convergence rate for sufficiently smooth  $g$ .

To the best of our knowledge, the use of the BDG inequality seems to be almost the only method to control the bounds of the moments in the literature concerning the numerical analysis of SDEs. The novel potential-theoretic approach taken in the present paper avoids the use of the BDG inequality in the computation of inverse moments and instead makes use of the concept of Kato classes. As a result, the appearance of local martingale terms does not introduce an extra difficulty in our framework. This

is an important contribution in its own right and has the potential to be useful in other contexts as well. We leave the investigation of the convergence rate for conservative diffusions on  $(0, \infty)$  satisfying (1.4) in the absence of condition (C.1) to future study.

Although our analysis assumes a one-dimensional framework, a close look into our technical analysis reveals that our convergence result does not depend heavily on this assumption apart from the comparison argument used in Lemma 3.2. In particular, it is relatively clear how to obtain a version of Theorem 2.9 in the multi-dimensional case using well-known potential-theoretic arguments on Kato classes. However, our main obstacle in extending our results to a multidimensional setting is the absence of a systematic study of recurrent transformations in higher dimensions. Also note that Lemma 3.2 is only used to obtain estimates on  $h(\widehat{X})$ , which is always a one-dimensional object with  $\widehat{X}$  referring to the continuous Euler scheme. Such a study and its applications to Euler methods for killed diffusions will be the subject of future research.

The outline of the paper is as follows. Section 2 fixes the setting, gives a brief summary of results for recurrent transformations needed for this paper together with novel inverse moment estimates, and introduces the backward Euler–Maruyama scheme that is tailored for our purposes. Section 3 obtains the moment estimates that are needed for the weak convergence analysis performed in Sect. 4. Theoretical results are confirmed via numerical studies in Sect. 5, and Sect. 6 concludes the paper.

## 2 Preliminaries

Let  $X$  be a regular diffusion on  $(\ell, r)$ , where  $-\infty \leq \ell < r \leq \infty$ . We assume that infinite boundaries are inaccessible, and if any of the boundaries are reached in finite time, the process is killed and sent to the cemetery state  $\Delta$ . This is the only instance when the process can be ‘killed’; we do not allow killing inside  $(\ell, r)$ . The set  $I$  consists of entrance boundaries and all points that can be reached in finite time starting from the interior of  $(\ell, r)$ . That is,  $I$  is the union of  $(\ell, r)$  with the regular, exit and entrance boundaries. The law induced on  $C(\mathbb{R}_+; I)$ , the space of  $I$ -valued continuous functions on  $\mathbb{R}_+ = [0, \infty)$ , by  $X$  with  $X_0 = x$  is denoted by  $P^x$  as usual, while  $\zeta$  corresponds to its lifetime, i.e.,  $\zeta := \inf\{t > 0 : X_t \notin (\ell, r)\}$ . We also introduce the set  $I_\Delta := I \cup \{\Delta\}$  and extend any  $I$ -valued Borel-measurable function  $f$  to  $I_\Delta$  by setting  $f(\Delta) = 0$  unless stated otherwise. The filtration  $(\mathcal{F}_t^0)_{t \geq 0}$  is the natural filtration of  $X$ ,  $\tilde{\mathcal{F}}_t$  is the universal completion of  $\mathcal{F}_t^0$ , and  $\mathcal{F}_t = \tilde{\mathcal{F}}_{t+}$  so that  $(\mathcal{F}_t)_{t \geq 0}$  is a right-continuous filtration. We also set  $\mathcal{F} := \bigvee_{t \geq 0} \mathcal{F}_t$ . We refer the reader to Borodin and Salminen [8, Chaps. I and II] for a summary of results and references on one-dimensional diffusions. The definitive treatment of such diffusions is of course in Itô and McKean [30, Chap. 3].

Since we are only concerned with the diffusion process until it is killed, we can assume without any loss of generality that  $X$  is on natural scale. The extra regularity conditions imposed in the following assumption are standard in the theory of Euler discretisations for SDEs.

**Assumption 2.1**  $X$  is a regular one-dimensional diffusion on  $(\ell, r)$  such that

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s, \quad t < \zeta,$$

where  $\sigma : (\ell, r) \rightarrow (0, \infty)$  is continuously differentiable with a bounded derivative,  $B$  is a standard Brownian motion and  $\zeta = \inf\{t > 0 : X_t \in \{\ell, r\}\}$ . Moreover,  $\sigma(\ell+)$  (resp.  $\sigma(r-)$ ) exists and is finite if  $\ell$  (resp.  $r$ ) is finite.

Note that the speed measure  $m$  associated with  $X$  is given by  $m(dx) = 2\sigma^{-2}(x)dx$  on the Borel subsets of  $(\ell, r)$ .

Since we are interested in diffusions with killing, the following assumption is needed to ensure that we are not dealing with a vacuum problem.

**Assumption 2.2**  $P^x[\zeta < \infty] > 0$  for each  $x \in (\ell, r)$ .

Let  $I_0$  be the set of points in  $I$  that can be reached from its interior in finite time. Note that under Assumptions 2.1 and 2.2, there are only two cases to consider:

Case 1. Both  $\ell$  and  $r$  are accessible, which in turn implies  $\ell$  and  $r$  are finite and  $I_0 = [\ell, r]$ .

Case 2. Only one of  $\ell$  and  $r$  is accessible, which can be assumed to be  $\ell$  without any loss of generality. In particular,  $I^0 = [\ell, r)$ .

As  $\ell$  is always finite as a result of the above convention, the following is also assumed for convenience.

**Assumption 2.3**  $\ell = 0$ .

As a transient diffusion on  $(0, r)$ ,  $X$  has a finite *potential density*  $u : (0, r)^2 \rightarrow \mathbb{R}_+$  with respect to its speed measure (see [8, Paragraph 11 in Sect. II.1]). That is, for any nonnegative and measurable  $f$  vanishing at accessible boundaries, we have

$$Uf(x) := \int_0^\infty E^x[f(X_t)]dt = \int_0^r f(y)u(x, y)m(dy).$$

The potential density is symmetric and explicitly known in terms of the scale function and speed measure of  $X$ . This leads for the potential density to the specification

$$u(x, y) = \begin{cases} (x \wedge y)(1 - \frac{x \vee y}{r}), & \text{if } r < \infty, \\ x \wedge y, & \text{otherwise.} \end{cases}$$

**Definition 2.4** Let  $\mathcal{S}$  be the space of continuous functions  $f : (0, r) \rightarrow (0, \infty)$  satisfying the integrability conditions  $\int_{(0,r)} f(y)m(dy) < \infty$  and  $\int_{(0,r)} yf(y)m(dy) < \infty$ . We define

$$\mathcal{H}_0 := \{h : h = Uf, f \in \mathcal{S}\}.$$

Moreover, we denote by  $\mathcal{H}$  the union of  $\mathcal{H}_0$  and the identity function if  $r = \infty$ . If  $r < \infty$ , then  $\mathcal{H} = \mathcal{H}_0$ .

Note that any  $h \in \mathcal{H}_0$  is a concave function that is twice continuously differentiable and satisfies on  $(\ell, r)$  that

$$\frac{1}{2}\sigma^2 h'' = -f.$$

The following lemma, whose proof is relegated to Appendix A, lists some important properties shared by the functions that belong to  $\mathcal{H}$ .

**Lemma 2.5** *Let  $h \in \mathcal{H}$ .*

1) *For any given  $z > 0$ , consider the function  $H$  defined by*

$$H(x) = x - z \frac{h'(x)}{h(x)}, \quad x \in (0, r).$$

*Then  $H$  is strictly increasing and  $H((0, r)) = \mathbb{R}$ .*

2) *The function  $h$  is increasing if  $r = \infty$ . However,  $h'$  is bounded. In particular, for  $h \in \mathcal{H}_0$ , we have*

$$h'(0) = \begin{cases} \int_0^\infty f(y)m(dy), & \text{if } r = \infty, \\ \int_0^\infty \frac{r-y}{r} f(y)m(dy) > 0, & \text{otherwise,} \end{cases}$$

$$h'(r) = \begin{cases} 0, & \text{if } r = \infty, \\ -\frac{1}{r} \int_0^r yf(y)m(dy) < 0, & \text{otherwise.} \end{cases}$$

3) *For any  $\alpha \geq 0$  and  $h \in \mathcal{H}_0$ , we have*

$$\int_{(0,r)} (u(y, y) \wedge 1) \frac{\alpha|h'(y)| - h''(y)}{h(y)} dy < \infty. \tag{2.1}$$

We are now ready to state the transformations that we use in the sequel.

**Theorem 2.6** *Suppose that Assumptions 2.1–2.3 are in force, and consider  $h \in \mathcal{H}$ . Then the following hold:*

1) *There exists a probability measure  $Q^{h,x}$  on  $\mathcal{F}$  that is locally absolutely continuous with respect to  $P^x$  such that*

$$dX_t = \sigma(X_t)dW_t + \sigma^2(X_t) \frac{h'(X_t)}{h(X_t)} dt \tag{2.2}$$

*and  $W$  is a  $Q^{h,x}$ -Brownian motion.*

2) *For any  $x \in (\ell, r)$ ,  $Q^{h,x}[\zeta < \infty] = 0$ .*

3) *Let  $g : I_0 \rightarrow \mathbb{R}$  be a continuous function vanishing at accessible boundaries. Then for any deterministic  $T > 0$ , we have*

$$E^x[g(X_T)\mathbf{1}_{\{T < \zeta\}}] = h(x)E^{h,x} \left[ \frac{g(X_T)}{h(X_T)} \exp \left( \frac{1}{2} \int_0^T \frac{\sigma^2(X_s)h''(X_s)}{h(X_s)} ds \right) \right], \tag{2.3}$$

*where  $E^{h,x}$  is the expectation operator associated with  $Q^{h,x}$ .*



**Proof** If  $h \in \mathcal{H}_0$ , the claims follow from Çetin [9, Theorem 3.2]. When  $r = \infty$  and  $h$  is the identity function, the stated transformation is the well-known Doob  $h$ -transform, and the reader is referred to Evans and Hening [18, Theorem 6.2] for a proof in a much more general setting.  $\square$

**Remark 2.7** When  $h \in \mathcal{H}_0$ , [9] shows that  $h(X) \exp(-\frac{1}{2} \int_0^\cdot \frac{\sigma^2(X_s)h''(X_s)}{h(X_s)} ds)$  is a  $P^x$ -martingale and  $X$  is a recurrent process under  $Q^{h,x}$ , where  $Q^{h,x}$  is as in Theorem 2.6. This measure transformation via the excessive function  $h$  and the multiplicative functional  $\exp(-\frac{1}{2} \int_0^\cdot \frac{\sigma^2(X_s)h''(X_s)}{h(X_s)} ds)$  is called a recurrent transformation.

Although similar to a Doob  $h$ -transform at heart, the concept of a recurrent transformation is fundamentally different from an  $h$ -transform. By an  $h$ -transform, we refer to a change of measure via an excessive function  $h : (0, r) \rightarrow \mathbb{R}_+$ . In particular,  $h(X)$  is at most a local martingale and can be a supermartingale that is not a local martingale. The latter case leads to a loss of mass after the change of measure that appears in the killing measure of the resulting diffusion (see Evans and Hening [18, Sect. 6] for details).

In contrast, first, a recurrent transformation always yields a recurrent diffusion while an  $h$ -transform yields a transient one. Indeed, when  $r = \infty$  and one uses the identity function in Theorem 2.6,  $Q^{\text{Id},x}[\lim_{t \rightarrow \infty} X_t = \infty] = 1$  since the corresponding scale function under  $Q^{\text{Id},x}$  is finite at  $\infty$ .

Second, the excessive function that is associated with an  $h$ -transform is in general not harmonic. That is,  $h(X)$  is a strict  $P^x$ -supermartingale. This leads to a killing, i.e., a loss of probability mass, under  $Q^{h,x}$  at some particular last passage time. We refer the reader to Çetin [9, Sect. 3.1] for more details on this point and the potential-theoretic connection between recurrent transformations and Doob  $h$ -transforms.

Finally, akin to what we are doing in the present paper, if one is interested in an  $h$ -transform that does not involve any killing as in the preceding paragraph and prevents the diffusion from hitting the boundaries, one needs to find an  $h$ -function such that  $h$  is harmonic, i.e.,  $h(X)$  is a  $P^x$ -local martingale, and  $h$  vanishes at accessible boundaries. However, the only harmonic functions of a diffusion on natural scale are affine functions. Combined with the requirement that  $h$  should vanish at 0 and  $r$ , this implies  $h \equiv 0$ . That is, there is no  $h$ -transform that yields a conservative diffusion that does not hit boundaries when  $r < \infty$ . On the other hand, any function  $h \in \mathcal{H}_0$  yields a recurrent diffusion process that avoids hitting the boundaries via a recurrent transformation.

In order to approximate the expectation on the right-hand side of (2.3), we use a backward Euler–Maruyama (BEM) scheme. Let  $N > 1$  be an integer and define  $t_n := \frac{n}{N}T$  for  $n = 0, \dots, N$ . Set  $\widehat{X}_0 = X_0$  and proceed inductively by setting

$$\widehat{X}_t = \widehat{X}_{t_n} + \sigma(\widehat{X}_{t_n})(W_t - W_{t_n}) + (t - t_n)\sigma^2(\widehat{X}_{t_n}) \frac{h'(\widehat{X}_t)}{h(\widehat{X}_t)} \tag{2.4}$$

for  $t \in (t_n, t_{n+1}]$  and  $n = 0, \dots, N - 1$ . Note that in view of Lemma 2.5, the mapping  $x \mapsto x - z \frac{h'}{h}(x)$  is one-to-one and onto for any given  $z > 0$ . Thus the above scheme is well defined since  $\sigma(x) > 0$  for all  $x \in (0, r)$ .

**Remark 2.8** One may also consider an exact simulation method as in Beskos and Roberts [6] to compute the right- or left-hand side of (2.3). However, the singularity of the drift term in (2.2) violates the regularity conditions assumed in [6] and its subsequent extensions. Thus one needs to consider the right-hand side and use the original diffusion after assuming enough regularity on  $\sigma$ . Provided we can perform an exact simulation, one can get an exact simulation of the minimum or maximum of a path together with its terminal value. In [6], for  $T = 1$  year, the time taking to run 50'000 simulations to compute the maximum of  $X$  on  $[0, 1]$  is about 3.1 seconds in their C-language program. To get a good pricing accuracy (a few bps in errors), one may need to run 0.5 or 1 million simulation paths, which may amount to 30 seconds to 1 minute. In our numerical calculations for the BEM method, we have used the Octave software, which is much slower than the C-language. On the other hand, the numerical experiments have shown that we needed very few discretisation time steps (20 or 30 steps) and very few paths to get comparable accuracy (see Sect. 5). As a result, our simulations took about 20 seconds to complete.

Nevertheless, the development of an exact simulation method for the recurrent transform to compute the right-hand side of (2.3) has the potential to significantly reduce the computation time by avoiding calculations using an inverse function. This interesting direction is left for future research.

As we shall see in Sect. 3, a crucial role will be played by diffusion processes on  $(0, r)$  of the type

$$dY_t = dW_t + \left( \frac{h'(Y_t)}{h(Y_t)} + c \right) dt, \quad t < \zeta(Y), \tag{2.5}$$

where  $\zeta(Y)$  denotes the first hitting time of 0 or  $r$ . Note that  $c = 0$  corresponds to the transformations defined in Theorem 2.6.

**Theorem 2.9** *Suppose that Assumptions 2.1–2.3 are in force,  $h \in \mathcal{H}$  and  $Y$  is a process defined by (2.5) with  $Y_0 = X_0$ . Assume further that  $c \leq 0$  if  $r = \infty$ , and  $c = 0$  if  $h(x) = x$  for all  $x$ . Then the following statements are valid:*

1)  $Q^{h, X_0}[\zeta(Y) = \infty] = 1$ .

2) *For any stopping time  $S$  that is  $Q^{h, X_0}$ -a.s. bounded, there exists a constant  $K$  that does not depend on  $X_0$  such that*

$$E^{h, X_0} \left[ \frac{1}{h(Y_S)} \right] < \frac{K}{h(X_0)}.$$

3) *For any  $t > 0$  and  $p \in [0, 1)$ ,*

$$E^{h, X_0} \left[ \int_0^t \frac{1}{h^{2+p}(Y_s)} ds \right] < \infty.$$

**Proof** 1) First observe that the scale function and speed measure for  $Y$  can be chosen as

$$s_y(x) = \int_d^x \frac{e^{-2cy}}{h^2(y)} dy, \quad m_y(dx) = 2h^2(x) \exp(2cx) dx,$$

where  $d \in (0, r)$ . Since  $s_y(0) = -\infty$ , 0 is an inaccessible boundary for  $Y$ . By the same token,  $r$  is also an inaccessible boundary when  $s_y(r) = \infty$ , which is valid when  $r < \infty$  or  $c \leq 0$ .

2) Define  $Z$  by

$$Z_t := \frac{1}{h(Y_t)} \exp\left(\frac{1}{2} \int_0^t \frac{2ch'(Y_s) + h''(Y_s)}{h(Y_s)} ds\right)$$

and note that  $Z$  is a nonnegative  $Q^{h, X_0}$ -local martingale by a straightforward application of Itô's formula. By Sharpe [42, Theorem 62.19], there exists a probability measure  $\tilde{P}$  such that

$$dY_t = d\beta_t + cdt, \quad t < \zeta(Y),$$

where  $\beta$  is a  $\tilde{P}$ -Brownian motion, and whenever  $S$  is a stopping time that is finite  $Q^{h, X_0}$ -a.s., one has

$$\begin{aligned} E^{h, X_0} \left[ \frac{1}{h(Y_S)} \right] &= \frac{1}{h(X_0)} \tilde{E} \left[ \mathbf{1}_{\{S < \zeta(Y)\}} \exp\left(-\frac{1}{2} \int_0^S \frac{2ch'(Y_s) + h''(Y_s)}{h(Y_s)} ds\right) \right] \\ &\leq \frac{1}{h(X_0)} \tilde{E} \left[ \mathbf{1}_{\{S < \zeta\}} \exp\left(\frac{1}{2} \int_0^S \frac{2(ch'(Y_s))^- - h''(Y_s)}{h(Y_s)} ds\right) \right], \end{aligned}$$

where  $x^-$  denotes the negative part of  $x$  and we drop the dependence on  $Y$  for  $\zeta$  to ease the exposition.

Suppose that  $S < R$   $Q^{h, X_0}$ -a.s., where  $R$  is a deterministic constant, and observe that  $\tilde{P}[S \geq R, S < \zeta] = 0$ . Thus

$$\begin{aligned} &\tilde{E} \left[ \mathbf{1}_{\{S < \zeta\}} \exp\left(\int_0^S \frac{2(ch'(Y_s))^- - h''(Y_s)}{2h(Y_s)} ds\right) \right] \\ &\leq \tilde{E} \left[ \exp\left(\int_0^{R \wedge \zeta} \frac{2(ch'(Y_s))^- - h''(Y_s)}{2h(Y_s)} ds\right) \right]. \end{aligned}$$

Let  $\mathcal{W}^{c, y}$  denote the law of the process  $\tilde{Y}$  starting at  $y$ , where  $d\tilde{Y}_t = d\beta_t + cdt$  and  $\tilde{Y}$  gets killed at hitting 0 or  $r$ . Thus

$$\tilde{E} \left[ \exp\left(\int_0^{R \wedge \zeta} \frac{2(ch'(Y_s))^- - h''(Y_s)}{2h(Y_s)} ds\right) \right] = \mathcal{W}^{c, X_0}[\exp(C_R)],$$

where  $C$  is the positive continuous additive functional of  $\tilde{Y}$  with

$$dC_t = \frac{1}{2} \frac{2(ch'(\tilde{Y}_t))^- - h''(\tilde{Y}_t)}{h(\tilde{Y}_t)} \mathbf{1}_{\{t < \tilde{\zeta}\}} dt.$$

Note that the potential function  $u_C$  of  $C$  is given by

$$u_C(x) = \mathcal{W}^x[C_\infty] = \int_0^r v(x, y) \mu_C(y) \frac{d\tilde{m}}{dy},$$

where  $v$  is the potential density of  $\tilde{Y}$ ,  $\mu_C(y) = \frac{1}{2} \frac{(2ch'(y))^- - h''(y)}{h(y)}$  and  $\tilde{m}$  is the associated speed measure of  $Y$ . As the scale function and speed measure of  $\tilde{Y}$  can be chosen as

$$\tilde{s}(x) = \frac{1 - e^{-2cx}}{2c}, \quad \tilde{m}(dx) = 2e^{2cx} dx,$$

where  $\tilde{s}(x) = x$  if  $c = 0$ , we obtain for  $x \leq y$  that

$$v(x, y) = \frac{\tilde{s}(x)(\tilde{s}(r) - \tilde{s}(y))}{\tilde{s}(r)},$$

with  $\frac{\tilde{s}(r) - \tilde{s}(y)}{\tilde{s}(r)}$  being interpreted as 1 if  $\tilde{s}(r) = \infty$ .

First observe that  $v(x, y) = u(x, y)$  if  $c = 0$ . On the other hand, if  $r = \infty$  and  $c < 0$ , then

$$v(y, y)e^{2cy} = \frac{e^{2cy} - 1}{2c} \leq y \wedge \frac{1}{2|c|}. \tag{2.7}$$

Similarly, for  $r < \infty$ ,

$$v(y, y)e^{2cy} = \frac{e^{-2cr}}{2c(1 - e^{-2cr})} (e^{2cy} - 1)(e^{2c(r-y)} - 1) \leq K(c, r)y \left(1 - \frac{y}{r}\right). \tag{2.8}$$

Thus for some  $K < \infty$ ,

$$\int_0^r v(y, y)\mu_C(y)2e^{2cy}dy \leq K \int_0^r (u(y, y) \wedge 1) \frac{(2ch'(y))^- - h''(y)}{h(y)} dy < \infty \tag{2.9}$$

by an application of (2.1) due to the bounds obtained via (2.7) and (2.8), and the assumption on the choice of  $c$  when  $r = \infty$ . As

$$u_C(x) \leq \int_0^r v(y, y)\mu_C(y)2e^{2cy} dy,$$

we deduce that  $u_C$  is bounded. Now consider a decreasing sequence  $(D_n)$  of subsets of  $(0, r)$  such that  $D_n \rightarrow \emptyset$ . Since

$$\int_0^r v(x, y)\mathbf{1}_{\{D_n\}}(y)\mu_C(y)2e^{2cy} \leq \int_0^r v(y, y)\mathbf{1}_{\{D_n\}}(y)\mu_C(y)2e^{2cy} dy$$

and the right-hand side converges to 0 by the dominated convergence theorem due to (2.9), we establish that  $\mu_C \in \mathbf{K}_1(\tilde{Y})$  (see Chen [11, Definition 2.2]) by [11, Proposition 2.4]. Therefore, by Chen and Song [12, Proposition 2.3] we arrive at the estimate  $\sup_{y \in (0, r)} \mathcal{W}^{c, y}[\exp(C_t)] \leq d_1 e^{d_2 t}$  for some constants  $d_1$  and  $d_2$ . This proves the claim.

3) Since the semigroup is self-dual with respect to the speed measure, denoting the measure  $dy2h^2(y)$  by  $\nu(dy)$  gives for any nonnegative measurable  $f$  that

$$\begin{aligned} & \int_0^r \nu(dy)e^{2cy} f(y) E^{h,y} \left[ \int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds \right] \\ &= \int_D \nu(dy)e^{2cy} \frac{1}{h^{2+p}(y)} E^{h,y} \left[ \int_0^t e^{-s} f(Y_s) ds \right] \\ &\leq \int_D dy \frac{2e^{2cy}}{h^p(y)} E^{h,y} \left[ \int_0^t f(Y_s) ds \right], \end{aligned}$$

where  $D := \{y : h(y) < 1 \wedge \frac{1}{2} \|h\|_\infty\}$ . In particular, when  $f(y) = q(\varepsilon, y, y^*)$  for some  $\varepsilon > 0$ , where  $q$  is the transition density of  $Y$  with respect to its speed measure, we obtain

$$\begin{aligned} & \int_0^r \nu(dy)e^{2cy} q(\varepsilon, y, y^*) E^{h,y} \left[ \int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds \right] \\ &\leq \int_D dy 2e^{2cy} h^{-p}(y) E^{h,y} [L_{t+\varepsilon}^{y^*}] \\ &\leq E^{h,y^*} [L_{t+\varepsilon}^{y^*}] \int_D dy 2e^{2cy} h^{-p}(y), \end{aligned}$$

where  $L^{y^*}$  is the diffusion local time with respect to the speed measure. Letting  $\varepsilon \rightarrow 0$ , we arrive at

$$E^{h,y^*} \left[ \int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds \right] \leq E^{h,y^*} [L_t^{y^*}] \int_D dy 2e^{2cy} h^{-p}(y) < \infty,$$

provided that  $y \mapsto E^{h,y} [\int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds]$  is lower semicontinuous. Note that the finiteness of the integral on the right-hand side follows from the fact that  $|h'(y)| \geq \alpha$  for some  $\alpha > 0$  on  $D$ .

Observe that

$$E^{h,y} \left[ \int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds \right] = \phi(y) - e^{-t} E^{h,y} [\phi(Y_t)],$$

where

$$\phi(y) := E^{h,y} \left[ \int_0^\infty ds \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} \right] = \int_D \frac{2e^{2cz} v_1(y, z)}{h^p(z)} dz,$$

where  $v_1$  is the 1-potential density of  $Y$ . Since  $v_1$  is jointly continuous (see [8, Paragraphs 10 and 11 in Chap. II]), the claimed semicontinuity follows. Finally, since

$$E^{h,y^*} \left[ \int_0^t \frac{1}{h^{2+p}(Y_s)} ds \right] \leq e^t E^{h,y^*} \left[ \int_0^t \frac{e^{-s} \mathbf{1}_{\{Y_s \in D\}}}{h^{2+p}(Y_s)} ds \right] + K$$

for some  $K$ , the claim follows from the arbitrariness of  $y^*$ . □

### 3 Moment estimates for the continuous BEM scheme

In this section, we obtain some moment estimates, including inverse ones, that are necessary to establish the weak rate of convergence (see Sect. 4). We start with the following consequence of Itô’s formula.

**Lemma 3.1** *Suppose that  $h \in \mathcal{H}$  is in  $C_b^2((0, r); (0, \infty))$  and such that  $h^{(3)}$  exists and satisfies the growth condition  $|h^{(3)}| \leq K(1 + h^{-p})$  for some constants  $K$  and  $p \in [0, 1)$ . Consider the BEM scheme defined by (2.4). Then for all  $t \in (t_n, t_{n+1}]$ ,*

$$d\widehat{X}_t = \frac{\sigma(\widehat{X}_{t_n})}{H_x(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dW_t + \frac{\sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \left( \frac{h'}{h}(\widehat{X}_t) + \mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t) \right) dt, \tag{3.1}$$

where

$$H(t_n, z; t, x) := x - \sigma^2(z)(t - t_n) \frac{h'}{h}(x),$$

$$\mu(t_n, z; t, x) := \left( H_x(t_n, z; t, x) - 1 \right) \frac{h'}{h}(x) + \frac{1}{2} \frac{\sigma^2(z)(t - t_n)}{H_x(t_n, z; t, x)} \left( \frac{h'}{h} \right)''(x).$$

Consider the sets  $O_1 := \{x : h'(x) > 0\}$  and  $O_2 := \{x : h'(x) < 0\}$ . Then

$$\inf_{x \in O_1} \mu(t_n, z; t, x) \geq c_1 \quad \text{and} \quad \sup_{x \in O_2} \mu(t_n, z; t, x) \leq c_2$$

for some constants  $c_1 \leq 0 \leq c_2$  that do not depend on  $t_n, t$  or  $z$ . In particular,  $c_1 = 0$  when  $h(x) = x$ .

**Proof** The decomposition (3.1) follows from Itô’s formula and straightforward calculations for the derivatives of the inverse function.

To prove the second assertion, first note that

$$H_x(t, x) - 1 := H_x(t_n, z; t, x) - 1 = -\sigma^2(z)(t - t_n) \left( \frac{h'}{h} \right)'(x) \geq 0,$$

where we drop the dependence on  $t_n$  and  $z$  to ease the exposition. Observe that

$$\mu = -\frac{\sigma^2(z)(t - t_n) \left( \frac{h'}{h} \right)'(x)}{H_x} \left( H_x \frac{h'}{h} - \frac{1}{2} \frac{\left( \frac{h'}{h} \right)''(x)}{\left( \frac{h'}{h} \right)'(x)} \right), \tag{3.2}$$

and that the claim follows immediately if  $h(x) = x$  since the term in parentheses in (3.2) becomes nonnegative. Thus it remains to show the assertion when  $h \in \mathcal{H}_0$ .

First consider the case  $r = \infty$ . Let  $u := \frac{h'}{h}$  and note that  $\lim_{x \rightarrow \infty} u'(x) = 0$  by Lemma 2.5. Moreover,  $|u'(x)| \leq Kx^{-2}$  for some  $K < \infty$ , which in turn implies

$$\lim_{x \rightarrow \infty} \frac{\log(-u'(x))}{x} = 0 = \lim_{x \rightarrow \infty} \frac{u''(x)}{u'(x)}, \tag{3.3}$$

where the second equality is an application of L'Hospital's rule. Thus

$$-\frac{1}{2} \frac{(\frac{h'}{h})''}{(\frac{h'}{h})'} > c \quad \text{on} \left( \frac{x^*}{2}, \infty \right)$$

for some  $c < 0$ , where  $x^* := \inf\{x : h'(x) = 0\} > 0$  by Lemma 2.5.

An alternative representation for  $\mu$  is given by

$$\mu = \sigma^2(z)(t - t_n) \left( -\frac{h'}{h} \left( \frac{h'}{h} \right)' \frac{1 + H_x}{H_x} + \frac{1}{2H_x} \frac{h'''h - h''h'}{h^2} \right). \tag{3.4}$$

Thus we are done if

$$w(t, x) := \frac{\sigma^2(z)(t - t_n)}{2H_x} \frac{h'''h - h''h'}{h^2}$$

is bounded from below on  $(0, \frac{x^*}{2})$ . Indeed, as  $h'$  is bounded away from 0 on this interval, the hypothesis on  $h'''$  implies that

$$w(t, x) \geq -K \frac{\sigma^2(z)(t - t_n) (\frac{h'}{h})^2}{1 + \sigma^2(z)(t - t_n) (\frac{h'}{h})^2},$$

leading to the desired lower bound.

When  $r < \infty$ , we have in particular that  $\sigma$  is bounded. Moreover, we have the estimate  $|w(t, x)| \leq K \frac{\sigma^2(z)(t-t_n)^{\frac{1}{h^2}}}{2H_x}$  for some constant  $K$  which renders  $w$  bounded. Observing that the remaining term in (3.4) has the correct sign completes the proof. □

The next result is a key comparison result that relates the inverse moments of the BEM scheme to those of the process (2.5) and thereby provide estimates that are valid uniformly in  $N$ .

**Lemma 3.2** *Suppose that  $h$  satisfies the conditions of Lemma 3.1,  $\sigma$  is bounded,  $r = \infty$  and consider the BEM scheme defined by (2.4). Then for any nondecreasing and measurable function  $\phi$  that does not change sign, we have*

$$E^{h, X_0}[\phi(\widehat{X}_{A_t^{-1}})] \geq E^{h, X_0}[\phi(Y_t)],$$

where  $Y$  is the process defined by (2.5) with  $c = c_1$  as in Lemma 3.1 and  $A$  is a continuous time change defined by  $A_0 = 0$  and

$$dA_t = \frac{\sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt, \quad t \in (t_n, t_{n+1}].$$

Moreover,  $Q^{h, X_0}[A_t \leq t \|\sigma\|_\infty^2] = 1$ .

**Proof** Consider the process  $\widehat{Y}$  defined by  $\widehat{Y}_t = \widehat{X}_{A_t^{-1}}$ . The Dambis–Dubins–Schwarz theorem (cf. Revuz and Yor [39, Theorem V.1.6]) yields

$$d\widehat{Y}_t = d\beta_t + \left( \frac{h'}{h}(\widehat{Y}_t) + \mu_t \right) dt, \quad t \in (t_n, t_{n+1}],$$

where  $\mu_t \geq c_1$  and  $\beta$  is a standard Brownian motion adapted to  $(\mathcal{F}_{A_t^{-1}})_{t \geq 0}$ . Then the comparison theorem for stochastic differential equations (cf. Çetin and Danilova [10, Theorem 2.10]) shows that

$$P^{h, X_0}[\widehat{Y}_t \geq Y_t, t \leq T] = 1,$$

where

$$Y_t = X_0 + \beta_t + \int_0^t \left( \frac{h'}{h}(Y_s) + c_1 \right) ds.$$

Since  $H_x \geq 1$ , it follows that  $A_t \leq t \|\sigma\|_\infty^2$ . This completes the proof. □

The main moment estimates are collected in the following theorem.

**Theorem 3.3** *Suppose that  $h$  satisfies the conditions of Lemma 3.1,  $\sigma$  is bounded and consider the BEM scheme defined by (2.4). Then for any  $T > 0$  and  $p \in [0, 1)$ , the following statements are valid:*

1) For each  $m \in \mathbb{N}$ ,

$$\sup_{t \leq T, N} E^{h, X_0} \left[ \frac{1}{h}(\widehat{X}_t) + \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n}) h^{-2-p}(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt + |\widehat{X}_t|^m \right] < \infty. \quad (3.5)$$

2) For each  $n$ ,

$$\operatorname{ess\,sup}_{\tau \in \mathcal{T}_n} E^{h, X_0} \left[ \frac{1}{h}(\widehat{X}_\tau) + X_\tau^m \middle| \mathcal{F}_{t_n} \right] < \infty, \quad (3.6)$$

where  $m \geq 0$  is an integer and

$$\mathcal{T}_n := \{ \tau : \tau \text{ is a stopping time valued in } [t_n, t_{n+1}] \text{ } \mathcal{Q}^{h, X_0}\text{-a.s.} \}.$$

Suppose further that  $p \leq \frac{1}{2}$  and  $\frac{h''}{h^{1-p}}$  is bounded. Then for each  $n \in \mathbb{N}$  and  $m \geq 0$ ,

$$\begin{aligned} & E^{h, X_0} \left[ \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \left( 1 - \exp \left( (s - t_n) \frac{\sigma^2 h''}{2h}(\widehat{X}_{t_n}) \right) \right) \frac{\sigma(\widehat{X}_{t_n})^2 (h^{-p}(\widehat{X}_s) + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \right] \\ & < \frac{KT}{N}, \end{aligned} \quad (3.7)$$

and  $K$  is independent of  $N$ .



The proof of Theorem 3.3 is lengthy and is relegated to Appendix B. We end this section with the following lemma that will be useful in our PDE approach to weak convergence rates in Sect. 4.

**Lemma 3.4** *Suppose that  $h$  satisfies the conditions of Lemma 3.1,  $\sigma$  is bounded and consider the BEM scheme defined by (2.4). Then for any  $T > 0$ , the following statements are valid:*

1) *Let  $p \in [0, 1)$  and  $m \geq 0$  be an integer. For each  $n$ ,*

$$\begin{aligned} & E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \left| \frac{h^{1-p}(\widehat{X}_t)(1 + \widehat{X}_t^m)\mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \right| dt \middle| \mathcal{F}_{t_n} \right] \\ & \leq \frac{KT}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n})(h^{-2-p}(\widehat{X}_t) + \widehat{X}_t^m)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \middle| \mathcal{F}_{t_n} \right], \end{aligned}$$

with  $K$  being a constant independent of  $n$ .

2) *Assume further that  $h \in C^4((0, r); (0, \infty))$ . Consider  $p \in [0, 1)$  and suppose*

$$\frac{|h^{(k)}|}{h} < \frac{K}{h^{k-2+p}}, \quad k \in \{2, 3, 4\},$$

for some  $K$ . Let  $f \in C^2((0, r); \mathbb{R})$  be a bounded function such that

$$|f^{(k)}(x)| \leq K(1 + x^m)h^{2-p-k}(x), \quad k \in \{1, 2\},$$

for some  $m \geq 0$ . Then for each  $n$  and  $t \in [t_n, t_{n+1}]$ ,

$$\begin{aligned} & \left| E^{h, X_0} \left[ f(\widehat{X}_t) \left( \frac{h''}{h}(\widehat{X}_{t_n}) - \frac{h''(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)h(\widehat{X}_t)} \right) \middle| \mathcal{F}_{t_n} \right] \right| \\ & \leq K E^{h, X_0} \left[ \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2(h^{-(2+p)}(\widehat{X}_s) + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \right] \\ & \quad - \frac{Kh''(\widehat{X}_{t_n})}{h(\widehat{X}_{t_n})} \\ & \quad \times E^{h, X_0} \left[ \int_{t_n}^t \frac{\sigma^2(\widehat{X}_{t_n})((h^{-p}(\widehat{X}_s) + \widehat{X}_s^m) + (s - t_n)(h^{-2}(\widehat{X}_s) + \widehat{X}_s^m))}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \right] \end{aligned}$$

for some constant  $K$  independent of  $n$ .

3) *Suppose  $f$  and  $h$  satisfy the conditions of part 2) and  $b \in C_b^2((0, r); \mathbb{R})$ . Then for each  $n$  and  $t \in [t_n, t_{n+1}]$ ,*

$$\begin{aligned} & \left| E^{h, X_0} \left[ f(\widehat{X}_t) \left( b(\widehat{X}_{t_n}) - \frac{b(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \right) \middle| \mathcal{F}_{t_n} \right] \right| \\ & \leq K E^{h, X_0} \left[ \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2(h^{-2-p}(\widehat{X}_s) + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \right] \end{aligned}$$

for some constant  $K$  independent of  $n$ .

**Proof** 1) It follows directly from the definition of  $\mu$  and the hypothesis on  $h'''$  that

$$h^{1-p}(\widehat{X}_t)|\mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)| \leq K\sigma^2(\widehat{X}_{t_n})(t - t_n)h^{-2-p}(\widehat{X}_t), \quad t \in [t_n, t_{n+1}],$$

for some  $K$ . Also note that if  $m \geq 1$ , there exist  $K$  and  $c \in (0, r)$  such that we have  $x^m h^{-(2+p)} \leq Kh^{m-(2+p)}$  for  $x \in [0, c]$ . Thus

$$x^m h^{-(2+p)} \leq K(x^m + h^{-(2+p)}). \tag{3.8}$$

2) Let  $\mu_s := \mu(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)$ ,  $u := \frac{h'}{h}$  and  $\eta_s := H_x(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)$ . Then Itô's formula yields

$$f(\widehat{X}_t) \left( \frac{h''}{h}(\widehat{X}_{t_n}) - \frac{h''(\widehat{X}_t)}{h(\widehat{X}_t)\eta_t^2} \right) = M_t + A_t,$$

where  $M$  is a local martingale with  $M_{t_n} = 0$  since  $\eta_{t_n} = 1$ , and

$$\begin{aligned} A_t = & \int_{t_n}^t \frac{\sigma^2(\widehat{X}_{t_n})f(\widehat{X}_s)}{2\eta_s^4} \left( \frac{2h''h'}{h^2}(\widehat{X}_s)\mu_s + \frac{(h'')^2 - hh^{(4)}}{h^2}(\widehat{X}_s) \right) ds \\ & - \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2 f(\widehat{X}_s) h^{(3)}}{\eta_s^4} \frac{h^{(3)}}{h}(\widehat{X}_s) \left( \mu_s + 2\sigma^2(\widehat{X}_{t_n})(s - t_n) \frac{u''(\widehat{X}_s)}{\eta_s} \right) ds \\ & - \int_{t_n}^t \frac{\sigma^4(\widehat{X}_{t_n})(s - t_n) f(\widehat{X}_s) h''(\widehat{X}_s)}{h(\widehat{X}_s)\eta_s^5} \left( 2\mu_s u''(\widehat{X}_s) + u^{(3)}(\widehat{X}_s) \right) ds \\ & - \int_{t_n}^t \frac{\sigma^4(\widehat{X}_{t_n})(s - t_n) f(\widehat{X}_s) h''(\widehat{X}_s)}{h(\widehat{X}_s)\eta_s^5} \frac{3\sigma^2(\widehat{X}_{t_n})(s - t_n)(u'')^2(\widehat{X}_s)}{\eta_s} ds \\ & + \int_{t_n}^t \left( \frac{h''}{h}(\widehat{X}_{t_n}) - \frac{h''(\widehat{X}_s)}{h(\widehat{X}_s)\eta_s^2} \right) \frac{\sigma^2(\widehat{X}_{t_n})}{\eta_s^2} \left( f'(\widehat{X}_s)(u(\widehat{X}_s) + \mu_s) + \frac{1}{2} f''(\widehat{X}_s) \right) ds \\ & + \int_{t_n}^t \frac{\sigma^2(\widehat{X}_{t_n}) f'(\widehat{X}_s)}{\eta_s^4} \left( \frac{h''h' - hh^{(3)}}{h^2}(\widehat{X}_s) \right. \\ & \quad \left. - 2 \frac{\sigma^2(\widehat{X}_{t_n})(s - t_n) u''(\widehat{X}_s) h''}{\eta_s h}(\widehat{X}_s) \right) ds. \end{aligned}$$

Observe that the hypothesis on  $h$  implies that

$$|u^{(k)}| \leq Kh^{-1-k}, \quad k \in \{0, 1, 2, 3\},$$

for some constant  $K$ . Moreover, for some other constant  $K$  that does not depend on  $s$ , we have  $|\mu_s| \leq K\sigma^2(\widehat{X}_{t_n})(s - t_n)h^{-3}$  and  $\sigma^2(\widehat{X}_{t_n})(s - t_n)h^{-2}\eta_s^{-1} \leq K$ . Thus combined with the assumption on  $f$ , we arrive at

$$\begin{aligned} |A_t| \leq & -K \frac{h''}{h}(\widehat{X}_{t_n}) \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2(1 + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} \left( h^{-p}(\widehat{X}_s)(1 + (s - t_n)h^{-2}(\widehat{X}_s)) \right) ds \\ & + K \int_{t_n}^t \frac{\sigma^2(\widehat{X}_{t_n})(1 + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)h^{2+p}(\widehat{X}_s)} ds \end{aligned}$$

for some constant  $K$ . This in particular implies that  $M$  is a true martingale since we can deduce from the estimates (3.5) and (3.6) that the set

$$\left\{ f(\widehat{X}_\tau) \frac{h''(\widehat{X}_\tau)}{h(\widehat{X}_\tau)\eta_\tau^2} : \tau \text{ is a stopping time valued in } (t_n, t_{n+1}] \right\}$$

is uniformly integrable as soon as we once again recall that  $|h''/h| < Kh^{-p}$  for some  $p < 1$ . Hence the claim holds in view of (3.8).

3) Applying Itô’s formula and repeating similar estimates yields the claim. □

### 4 Weak convergence of the BEM scheme

Consider on a filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \in [0, T]}, \mathbb{P})$  satisfying the usual conditions the stochastic differential equation

$$X_t = X_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t \mu(X_s) ds,$$

where  $X_0 \in (0, r)$ ,  $\sigma$  and  $\mu$  are bounded and Lipschitz on  $(0, r)$ , and  $\sigma(x) > \varepsilon$  for all  $x \in (0, r)$  and some  $\varepsilon > 0$ . Let  $\tau := \inf\{t \geq 0 : X_t \notin (0, r)\}$ . We are interested in a numerical approximation for  $\mathbb{E}[\tilde{g}(X_T)\mathbf{1}_{\{T < \tau\}}]$  for a sufficiently regular  $\tilde{g}$ .

Observe that by a Girsanov transformation, we can rewrite the above expression in terms of a diffusion process satisfying the conditions in earlier sections. Indeed, defining  $\mathbb{Q}$  on  $\mathcal{G}$  via

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T \frac{\mu(X_s)}{\sigma(X_s)} dW_s - \frac{1}{2} \int_0^T \frac{\mu^2(X_s)}{\sigma^2(X_s)} ds\right)$$

makes  $X$  solve  $dX_t = \sigma(X_t)dB_t$  for a  $\mathbb{Q}$ -Brownian motion  $B$ . Therefore,

$$\mathbb{E}[\tilde{g}(X_T)\mathbf{1}_{\{T < \tau\}}] = \exp(-F(X_0))\mathbb{E}^{\mathbb{Q}}\left[g(X_T) \exp\left(\int_0^T \sigma^2(X_t)b(X_t)dt\right)\mathbf{1}_{\{T < \tau\}}\right],$$

$$g(x) = \tilde{g}(x) \exp(F(x)),$$

$$F(x) = \int_c^x \frac{\mu(y)}{\sigma^2(y)} dy, \quad b = -\frac{1}{2}\left(\left(\frac{\mu}{\sigma^2}\right)' + \frac{\mu^2}{\sigma^4}\right),$$

and  $c \in (0, r)$ . Thus we may assume  $\mu \equiv 0$  and consider

$$\begin{aligned} & E^{X_0}\left[g(X_T) \exp\left(\int_0^T \sigma^2(X_t)b(X_t)dt\right)\mathbf{1}_{\{T < \tau\}}\right] \\ &= E^{h, X_0}\left[\frac{h(X_0)g(X_T)}{h(X_T)} \exp\left(\int_0^T \sigma^2(X_t)\left(b(X_t) + \frac{h''(X_t)}{2h(X_t)}\right)dt\right)\right], \end{aligned}$$

where  $X$  is a process satisfying Assumption 2.1,  $b$  is bounded,  $\varepsilon < \sigma < K\sigma$  and  $g$  is sufficiently regular.

We next introduce a set of conditions on the function  $h$  as well as the payoff  $g$  that will be needed for our analysis. The following proposition motivates some of the conditions stated in Assumption 4.2. Its proof is relegated to the end of this section.

**Proposition 4.1** *Suppose  $b \in C_b^4((0, r); \mathbb{R}), \sigma \in C_b^4((0, r)), h \in \mathcal{H}$  with*

$$\frac{|h^{(k)}|}{h} < \frac{K_h}{h^{k-2+p}}, \quad k \in \{2, 3, 4\},$$

for some  $K_h$  and  $p \in (0, 1), g \in C_b^6((0, r); \mathbb{R})$  is a bounded function with  $g^{(k)}(0) = 0$  (and  $g^{(k)}(r) = 0$  if  $r < \infty$ ) for  $k \in \{0, 1, 2, 3, 4\}$ , and define for  $t \leq T$

$$v(T - t, x) := E^{h,x} \left[ \frac{g(X_t)}{h(X_t)} \exp \left( \int_0^t \sigma^2(X_s) \left( b(X_s) + \frac{h''(X_s)}{2h(X_s)} \right) ds \right) \right]. \quad (4.1)$$

Then

$$v_t + \frac{\sigma^2}{2} v_{xx} + \sigma^2 \frac{h'}{h} v_x = -\sigma^2 v \left( b + \frac{h''}{2h} \right). \quad (4.2)$$

Moreover,  $v$  and  $v_t$  are uniformly bounded and there exists a constant  $K$  such that

$$\sup_{t \leq T} \left| \frac{\partial^k}{\partial x^k} v_t(t, x) \right| + \sup_{t \leq T} \left| \frac{\partial^k}{\partial x^k} v(t, x) \right| \leq K h^{2-p-k}(x), \quad k \in \{1, 2\}.$$

In view of Proposition 4.1 and for the convenience of the reader, we collect in Assumption 4.2 below all the assumptions needed to prove our convergence result.

**Assumption 4.2** The functions  $\sigma, b, h$  and  $g$  satisfy the following regularity conditions:

- 1)  $h \in \mathcal{H} \cap C^4((0, r); (0, \infty))$  is such that

$$\frac{|h^{(k)}|}{h} < \frac{K_h}{h^{p+k-2}}, \quad k \in \{2, 3, 4\},$$

for some  $K_h$  and  $p \in [0, \frac{1}{2}]$ .

- 2)  $\sigma \in C_b^2((0, r); (0, \infty))$  is bounded away from 0, i.e., there is some  $\varepsilon > 0$  such that  $\sigma(x) > \varepsilon$  for all  $x \in (0, r)$ .

- 3)  $b \in C_b^2((0, r); \mathbb{R})$ .

- 4)  $g \in C((0, r); \mathbb{R})$  is of polynomial growth with  $g(0) = 0$  (and  $g(r) = 0$  if  $r < \infty$ ).

5) The function  $v$  defined by (4.1) belongs to  $C^{1,4}((0, r); \mathbb{R})$  and satisfies (4.2) as well as the growth conditions

$$\sup_{t \leq T} \left| \frac{\partial^k}{\partial x^k} v_t(t, x) \right| + \sup_{t \leq T} \left| \frac{\partial^k}{\partial x^k} v(t, x) \right| \leq K(1 + x^m)h^{2-p-k}(x), \quad k \in \{1, 2\},$$

for some constant  $K$  and integer  $m \geq 0$ .

**Remark 4.3** The first condition on the derivatives of  $h$  is not restrictive for practical purposes. Indeed, if a given  $h \in \mathcal{H} \cap C^4((0, r); (0, \infty))$  does not satisfy this condition, one can always linearise this concave function near the boundaries at which  $h$  vanishes to obtain a new concave function satisfying the stated condition.

**Theorem 4.4** Consider the BEM scheme defined by (2.4) and the associated error

$$\begin{aligned} e(N) := & \frac{g(\widehat{X}_T)}{h(\widehat{X}_T)} \exp \left( \sum_{n=0}^{N-1} \frac{T}{N} \sigma^2(\widehat{X}_{t_n}) \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right) \\ & - \frac{g(X_T)}{h(X_T)} \exp \left( \int_0^T \sigma^2(X_t) \left( b(X_t) + \frac{h''(X_t)}{2h(X_t)} \right) dt \right). \end{aligned}$$

Then under Assumption 4.2,  $|E^{h, X_0}[e(N)]| \leq \frac{KT}{N}$  for some constant  $K$  independent of  $N$ .

**Proof** Let  $\pi_0(s) = 1$ ,

$$\pi_k(s) := \exp \left( \sum_{n=0}^{k-1} s \sigma^2(\widehat{X}_{t_n}) \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right), \quad k = 1, \dots, N,$$

with the convention that we set  $\pi_k := \pi_k(TN^{-1})$ , and observe that

$$\begin{aligned} & E^{h, X_0}[e(N)] \\ &= E^{h, X_0}[v(T, \widehat{X}_T)\pi_N] - v(0, X_0) \\ &= \sum_{n=0}^{N-1} E^{h, X_0}[v(t_{n+1}, \widehat{X}_{t_{n+1}})\pi_{n+1} - v(t_n, \widehat{X}_{t_n})\pi_n] \\ &= \sum_{n=0}^{N-1} E^{h, X_0} \left[ \pi_n \left( v(t_{n+1}, \widehat{X}_{t_{n+1}}) \right. \right. \\ & \quad \left. \left. \times \exp \left( \frac{T\sigma^2(\widehat{X}_{t_n})}{N} \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right) - v(t_n, \widehat{X}_{t_n}) \right) \right]. \end{aligned}$$

Next observe that

$$\begin{aligned}
 & E^{h, X_0} \left[ \pi_n \left( v(t_{n+1}, \widehat{X}_{t_{n+1}}) \right. \right. \\
 & \quad \left. \left. \times \exp \left( \frac{T \sigma^2(\widehat{X}_{t_n})}{N} \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right) - v(t_n, \widehat{X}_{t_n}) \right) \middle| \mathcal{F}_{t_n} \right] \\
 &= \pi_n E^{h, X_0} \left[ \left( v(t_{n+1}, \widehat{X}_{t_{n+1}}) \right. \right. \\
 & \quad \left. \left. \times \exp \left( \frac{T \sigma^2(\widehat{X}_{t_n})}{N} \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right) - v(t_n, \widehat{X}_{t_n}) \right) \middle| \mathcal{F}_{t_n} \right].
 \end{aligned}$$

Moreover, using Itô's formula and (4.2) after division by  $\sigma^2$  gives

$$\begin{aligned}
 & v(t_{n+1}, \widehat{X}_{t_{n+1}}) \exp \left( \frac{T \sigma^2(\widehat{X}_{t_n})}{N} \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} \right) \right) - v(t_n, \widehat{X}_{t_n}) \\
 &= M_{t_{n+1}} - M_{t_n} + \sum_{j=1}^3 I_j,
 \end{aligned}$$

where  $M$  is a local martingale and

$$\begin{aligned}
 I_1 &= \int_{t_n}^{t_{n+1}} \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \frac{\sigma^2(\widehat{X}_{t_n}) v_x(t, \widehat{X}_t) \mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt, \\
 I_2 &= \int_{t_n}^{t_{n+1}} \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \sigma^2(\widehat{X}_{t_n}) v_t(t, \widehat{X}_t) \\
 & \quad \times \left( \frac{1}{\sigma^2(\widehat{X}_{t_n})} - \frac{1}{\sigma^2(\widehat{X}_t) H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \right) dt, \\
 I_3 &= \int_{t_n}^{t_{n+1}} \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \sigma^2(\widehat{X}_{t_n}) v(t, \widehat{X}_t) \\
 & \quad \times \left( b(\widehat{X}_{t_n}) + \frac{h''(\widehat{X}_{t_n})}{2h(\widehat{X}_{t_n})} - \frac{b(\widehat{X}_t) + \frac{h''(\widehat{X}_t)}{2h(\widehat{X}_t)}}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \right) dt.
 \end{aligned}$$

First note that  $M$  is a true martingale due to (3.5) by the hypothesis on  $v$  and the boundedness of  $h$ . Moreover, Lemma 3.4 shows (for a generic constant  $K$  that may

change from line to line, but remains bounded uniformly in  $N$ ) that

$$\begin{aligned}
 & |E^{h, X_0}[I_1 + I_2 + I_3 | \mathcal{F}_{t_n}]| \\
 & \leq K \frac{T}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n})(h^{-2-p}(\widehat{X}_t) + \widehat{X}_t^m)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \middle| \mathcal{F}_{t_n} \right] \\
 & \quad + K E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} dt \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \sigma^2(\widehat{X}_{t_n}) \right. \\
 & \quad \quad \times \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2(h^{-2-p}(\widehat{X}_s) + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \left. \right] \\
 & \quad - K E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} dt \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \frac{\sigma^2 h''}{h}(\widehat{X}_{t_n}) \right. \\
 & \quad \quad \times \int_{t_n}^t \frac{\sigma(\widehat{X}_{t_n})^2(h^{-p}(\widehat{X}_s) + \widehat{X}_s^m)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \left. \right] \\
 & \quad - K E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} dt \frac{\pi_{n+1}(t - t_n)}{\pi_n(t - t_n)} \frac{\sigma^2 h''}{h}(\widehat{X}_{t_n}) \right. \\
 & \quad \quad \times \int_{t_n}^t \frac{\sigma^2(\widehat{X}_{t_n})(h^{-2}(\widehat{X}_s) + \widehat{X}_s^m)(s - t_n)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \middle| \mathcal{F}_{t_n} \left. \right] \\
 & \leq K \frac{T}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n})(h^{-2-p}(\widehat{X}_t) + \widehat{X}_t^m)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \middle| \mathcal{F}_{t_n} \right] \\
 & \quad + E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \left( 1 - \exp\left( (t - t_n) \frac{\sigma^2 h''}{2h}(\widehat{X}_{t_n}) \right) \right) \right. \\
 & \quad \quad \times \frac{\sigma(\widehat{X}_{t_n})^2(h^{-p}(\widehat{X}_t) + \widehat{X}_t^m)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \middle| \mathcal{F}_{t_n} \left. \right] \\
 & \quad + K \frac{T}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \frac{\sigma(\widehat{X}_{t_n})^2(h^{-2}(\widehat{X}_t) + \widehat{X}_t^m)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \middle| \mathcal{F}_{t_n} \right],
 \end{aligned}$$

where we have used the boundedness of  $\pi_{n+1}/\pi_n$  several times, and the last two lines follow from interchanging the order of integration on the third and fourth lines. This proves the assertion in view of Theorem 3.3 and in particular of (3.5) and (3.7), since the  $\pi_n$  are nonnegative and uniformly bounded and  $H_x \geq 1$ . □

We end this section with the

**Proof of Proposition 4.1** Note that  $v(T - t, x) = \frac{u(T-t, x)}{h(x)}$ , where

$$u(T - t, x) := E^x \left[ g(X_t) \exp \left( \int_0^t \sigma^2(X_s) b(X_s) ds \right) \mathbf{1}_{\{t < \zeta\}} \right].$$

Note that  $u(t, 0) = 0$  for  $t \leq T$ . Moreover, Ladyženskaja et al. [32, Theorem IV.5.2] yields that  $u$  is the unique solution of

$$u_t + \frac{1}{2}\sigma^2 u_{xx} + \sigma^2 u b = 0 \tag{4.3}$$

and that

$$\sup_{t \leq T, x \in (0, r)} \left| \frac{\partial^j}{\partial x^j} \frac{\partial^k}{\partial t^k} u \right| < \infty, \quad 0 \leq 2k + j \leq 5. \tag{4.4}$$

Also note that since

$$\begin{aligned} & E^x \left[ g(X_t) \exp \left( \int_0^t \sigma^2(X_s) b(X_s) ds \right) \mathbf{1}_{\{t < \zeta\}} \right] \\ &= g(x) + \frac{1}{2} E^x \left[ \int_0^t \sigma^2(X_u) \exp \left( \int_0^u \sigma^2(X_s) b(X_s) ds \right) \right. \\ &\quad \left. \times (g''(X_u) + 2g(X_u) b(X_u)) \mathbf{1}_{\{u < \zeta\}} du \right], \end{aligned}$$

we have

$$\begin{aligned} u_t(t, x) &= -\frac{1}{2} E^x \left[ \sigma^2(X_{T-t}) \exp(C_{T-t}) \right. \\ &\quad \left. \times (g''(X_{T-t}) + 2g(X_{T-t}) b(X_{T-t})) \mathbf{1}_{\{T-t < \zeta\}} \right], \end{aligned} \tag{4.5}$$

where  $C_t = \int_0^t \sigma^2(X_s) b(X_s) ds$ . In particular,  $u_t(\cdot, 0) = 0$ , which in turn implies  $u_{xx}(\cdot, 0) = 0$ . Analogous boundary conditions also hold at  $r$  if  $r$  is finite.

Let  $w := u_t$  and note that  $w$  solves (4.3) with the boundary condition  $w(t, 0) = 0$  and  $w(T, \cdot) = -\frac{1}{2}\sigma^2 g'' - \sigma^2 g b$ . Using the stochastic representation in (4.5) and analogous arguments, we again arrive at  $w_t$  vanishing at finite boundaries.

Using the PDE for  $u$ , it is straightforward to establish that  $v$  solves (4.2) and is bounded. Moreover, as  $v_x = \frac{hu_x - uh'}{h^2}$ , integration by parts gives

$$v_x(t, x) = \frac{\int_0^x (h(y)u_{xx}(t, y) - u(t, y)h''(y))dy}{h^2(x)}.$$

Since  $h'(0) < \infty$  and  $u$  and  $u_{xx}$  vanish at 0 and are jointly continuous near  $t = T$ , there exists a neighbourhood of 0 in which we have  $|h''|(y) \leq Kh^{1-p}(y) \leq K^2y$ ,  $|u(\cdot, y)| + |u_{xx}(\cdot, y)| < Ky$  (due to Lipschitz-continuity) and  $h(y) > cy$ . Thus whenever  $x$  belongs to this neighbourhood, we have

$$\frac{v_x(t, x)}{h^{1-p}(x)} \leq \frac{K \int_0^x (y(Ky + K^2y^{1-p}))dy}{c^{3-p}x^{3-p}} = \frac{K^2/3x^3 + K^3/(3-p)x^{3-p}}{c^{3-p}x^{3-p}}.$$

Thus  $v_x/h^{1-p}$  is bounded near 0. Analogous considerations when  $r < \infty$  show that the ratio is bounded over  $(0, r)$ .



Next observe that  $v_t$  is bounded since  $u_t$  vanishes at finite boundaries and  $u_{tx}$  is bounded. In particular,  $v_t h^p$  remains bounded near finite boundaries (uniformly in  $t$ ). Multiplying (4.2) by  $h^p$  and using the fact that  $v_x/h^{1-p}$  is bounded demonstrates that

$$\sup_{t \leq T, x \in (0, r)} |v_{xx}(t, x)h^p(x)| < \infty.$$

Finally, since  $v_t = \frac{w}{h}$ , repeating the above arguments and using the fact that  $w_{xx}$  vanishes at finite boundaries and is Lipschitz-continuous in view of (4.4), we deduce that  $v_{tx}/h^{1-p}$  is bounded. Similar arguments (due to the boundedness of  $w_{tx} = u_{tx}$  in view of (4.4)) also lead to

$$\sup_{t \leq T, x \in (0, r)} |v_{txx}(t, x)h^p(x)| < \infty. \quad \square$$

### 5 Numerical analysis

This section is dedicated to the numerical experiments illustrating the above technical analysis. As we shall see, one does not really need to satisfy all the conditions assumed in Theorem 4.4 in order to achieve the advertised convergence rate in practice. The experiments below compare our methodology developed in this paper to standard numerical approaches for pricing barrier options (see also the symmetrisation technique from Akahori and Imamura [2] and Imamura et al. [29] for another approach to barrier option pricing).

We consider the classical Black–Scholes model in the first part. As barrier option values are quite sensitive to the market skew/smile of volatility, the time-homogeneous *hyperbolic local volatility* model is also studied in the second part.

**Remark 5.1** In this one-dimensional setting, the value function of a “plain” knock-out option, i.e.,  $E^x[g(X_T)\mathbf{1}_{\{T < \zeta\}}]$ , can be found rather easily by applying a finite difference scheme to the associated PDE with vanishing boundary conditions at accessible boundaries.

The Monte Carlo BEM scheme introduced in this paper will have a clear advantage over the PDE method in higher dimensions. However, already in the one-dimensional case, it is quite flexible with respect to additional complications compared to the PDE method. In practice, a barrier-type payoff can be combined with various features like Asianing and forward starting. As an example, one can consider

$$\mathbf{1}_{\{\zeta > T\}} \left( \frac{1}{m} \sum_{i=1}^m X_{T_i} - X_{T_0} \right)^+$$

with  $0 < T_0 < T_1 < \dots < T_m = T$ , where the strike is not fixed today but at a future date  $T_0$ . A computation of the price of the above derivative can be made without much extra effort by using the BEM method, provided the discretisation includes the time points  $T_i$  for  $i = 0, \dots, m$ . While the pricing of such a derivative can

still be done via a PDE approach using a finite difference scheme, its implementation is relatively complex and involves essentially a 3-d PDE solver: one needs one dimension for the spot price  $X$ , one to capture the possible values of the Asianing  $\frac{1}{m} \sum_{i=1}^m X_{T_i}$ , and another to incorporate the possible values of the strike value at  $T_0$  (see e.g. Wilmott [45, Chap. 25], Musiela and Rutkowski [35, Sects. 6.2 and 7.1.10] or De Weert [14, Sect. 11]). Besides, this can be intensive in computation time.

### 5.1 Black–Scholes model for barrier options

For expository purposes, we assume the log-price  $X_t = \ln S_t$  under the risk-neutral probability  $\mathbb{P}$  is given by

$$X_t = x - \frac{1}{2}\sigma^2 t + \sigma W_t, \quad t < \zeta, \tag{5.1}$$

where  $\zeta = \inf\{t > 0 : X_t \in \{\ell, r\}\}$  and  $\sigma > 0$  is constant. Deterministic interest rates, a dividend yield or borrowing costs can be incorporated without difficulty. Above,  $\zeta$  represents the time of hitting a pre-specified barrier at which the option becomes worthless. Consequently, the value of the barrier option with payoff  $\tilde{g}$  is

$$\text{price} = \mathbb{E}^{\mathbb{P}}[\tilde{g}(X_T)\mathbf{1}_{\{\zeta > T\}}]. \tag{5.2}$$

To remove the drift in (5.1), we follow the Girsanov transformation, described at the beginning of Sect. 4, and obtain

$$dX_t = \sigma dW_t, \quad X_0 = x \tag{5.3}$$

under  $\mathbb{Q}$ , where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{\frac{1}{2}\sigma W_T - \frac{1}{8}\sigma^2 T} = e^{\frac{1}{2}(X_T - x + \frac{1}{2}\sigma^2 T) - \frac{1}{8}\sigma^2 T} = e^{-\frac{1}{2} + \frac{1}{8}\sigma^2 T} e^{\frac{1}{2}X_T}.$$

Consequently, with  $g(x) = \tilde{g}(x)e^{-\frac{1}{2}x}$ , we get

$$\text{price} = e^{\frac{1}{2}x - \frac{1}{8}\sigma^2 T} \mathbb{E}^{\mathbb{Q}}[g(X_T)\mathbf{1}_{\{\zeta > T\}}].$$

We shall perform a path transformation method as described earlier that either produces a recurrent process or generates a transient process with infinite lifetime (see Theorem 2.6). A lower barrier will be called  $\underline{s}$ , an upper barrier  $\bar{s}$ .

#### 5.1.1 Specification of the recurrent transformation

For a single barrier with  $\ell$  finite and  $r = \infty$ , we choose  $h(x) = e^{-\ell} - e^{-x}$  with  $h'(x) = e^{-x}$  and  $h''(x) = -e^{-x}$ . Note that with this choice of  $h$ , condition 1) of Assumption 4.2 is not satisfied for  $k = 2$  and any  $p \in [0, \frac{1}{2}]$ . Nevertheless, we apply the implicit scheme (2.4) so that the price (5.2) is approximated by

$$\text{price} \approx e^{\frac{1}{2}x - \frac{1}{8}\sigma^2 T} h(x) \mathbb{E}^{h,x} \left[ \frac{g}{h}(\widehat{X}_{t_N}) e^{\frac{\sigma^2}{2} \frac{T}{N} \sum_{n=0}^{N-1} \frac{h''}{h}(\widehat{X}_{t_n})} \right]$$

and still obtain the optimal convergence rate in the numerical experiments.

**Remark 5.2** In the Black–Scholes model with  $X_t = \ln S_t$ , the function  $H$  is identical at each time step and needs to be computed once. In the implementation, we introduce a dense grid covering the interval  $(\ell, r)$ , calculate the values of  $H$  on these points and  $H^{-1}$  is computed by piecewise constant approximation.

### 5.1.2 The transient transformation

In the single barrier case of a down-and-out option, we can also consider a transformation via  $h(x) = x - \ell$  when  $\ell$  is finite and  $r = \infty$ , as in Theorem 2.6. Under  $Q^{h,x}$ , the process  $X$  defined in (5.3) follows

$$dX_t = \sigma dW_t + \frac{\sigma^2}{X_t - \ell} dt, \quad X_0 = x.$$

One advantage of this transformation is that the inverse of the function  $H$  appearing in the implicit scheme (2.4) can be computed analytically and is given by

$$H^{-1}(x) = \frac{1}{2} \left( \sqrt{4\sigma^2 \frac{T}{N} + (x - \ell)^2} + x + \ell \right).$$

## 5.2 Down-and-out put option

For a down-and-out put barrier option, the payoff is given by  $\max(K - S_T)^+ \mathbf{1}_{\{\zeta > T\}}$ , where  $r = \infty$  and  $\ell = \log \underline{s}$ ,  $K$  is the option strike and  $T$  the maturity. As mentioned at the beginning of this section, to put our methodology in perspective, we have also implemented two other approaches to the numerical pricing of the barrier option:

- Standard Euler without hitting probability. This consists of discretising the SDE (5.1) according to the Euler scheme

$$\begin{cases} \widehat{X}_0 = \ln S_0, \\ \widehat{X}_{t_{i+1}} = \widehat{X}_{t_i} - \frac{1}{2}\sigma^2 \frac{T}{N} + \sigma(W_{t_{i+1}} - W_{t_i}). \end{cases} \tag{5.4}$$

and evaluating  $\tilde{g}(X_T) \mathbf{1}_{\{\zeta > T\}}$  by  $\tilde{g}(\widehat{X}_{t_N}) \mathbf{1}_{\{\zeta^N > T\}}$ , where

$$\zeta^N = \inf\{t_i > 0 : \widehat{X}_{t_i} \notin (\ell = \log \underline{s}, \infty)\}.$$

This numerical scheme for barrier option pricing has been studied in Gobet [21], where it was shown to have a convergence rate of  $\mathcal{O}(\frac{1}{\sqrt{N}})$ . This loss of accuracy is mainly due to the fact that it is possible for  $X$  to cross the barriers  $\ell$  or  $r$  at some time  $t$  between grid points  $t_i$  and  $t_{i+1}$  and never be below the barrier at any of the dates  $t_i$  for  $i = 1, \dots, N$ .

- Standard Euler with hitting probability. Although this is still based on the Euler scheme simulation (5.4), it applies a further correction to remove the barrier crossing biases via the conditional no-hitting probability  $\hat{p}_i$  using the Brownian bridge

technique (see e.g. [21, Sect. 1.1]). More precisely, the  $\hat{p}_i$  are defined and can be computed analytically as

$$\hat{p}_i := \mathbb{P}[\widehat{X}_t > \ell, \forall t \in [t_i, t_{i+1}] | \widehat{X}_{t_i} = x_i, \widehat{X}_{t_{i+1}} = x_{i+1}] = 1 - e^{-2 \frac{(x_i - \ell)(x_{i+1} - \ell)}{\sigma^2(t_{i+1} - t_i)}},$$

where the process  $(\widehat{X}_t)_{0 \leq t \leq T}$  is the continuous Euler scheme which interpolates  $(\widehat{X}_{t_i})_{0 \leq i \leq N}$  via

$$\widehat{X}_t = \widehat{X}_{t_i} - \frac{1}{2}\sigma^2(t - t_i) + \sigma(W_t - W_{t_i}), \quad t \in [t_i, t_{i+1}).$$

It then *corrects* the payoff  $\tilde{g}(X_T)\mathbf{1}_{\{\zeta > T\}}$  by considering instead  $\tilde{g}(\widehat{X}_{t_N}) \prod_{i=0}^{N-1} \hat{p}_i$ . As shown in Gobet [22], this bias correction brings the convergence rate back to the order  $N^{-1}$ , which is the rate of weak convergence for the Euler–Maruyama scheme in the absence of killing. Moreover, in this specific Black–Scholes implementation, the simulation is exact, i.e., no discretisation error occurs due to constant  $\sigma$ .

We next summarise the experiment details and comparison results.

### 5.2.1 Set of parameters

The numerical experiments are conducted using for the parameters the values  $S_0 = 1$ ,  $T = 1$  year,  $\ell = \log(\underline{x} = 0.8)$ ,  $r = \infty$  and  $\sigma = 20\%$ . For thoroughness, we have considered in-the-money ( $K = 1.2$ ), at-the-money ( $K = 1$ ) and out-of-the-money ( $K = 0.9$ ) options. To reduce statistical noise, the simulations are run with 1’000’000 Monte Carlo paths. The benchmark price is calculated analytically (see e.g. Haug [23, Sect. 4.17.1]).

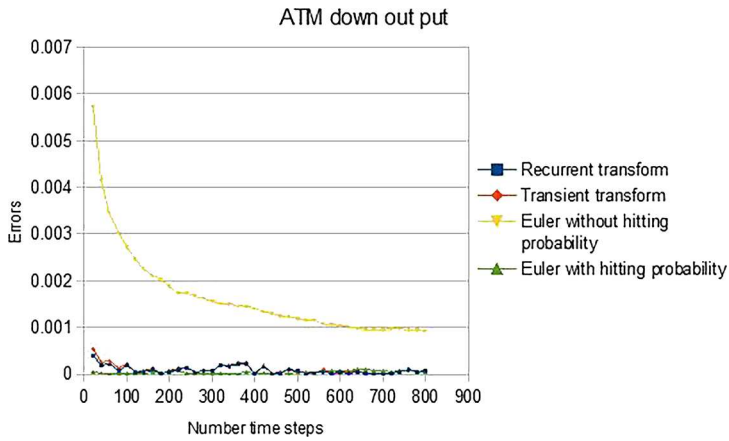
As our final results do not show any significant dependence on the moneyness of the option, we only report the results for at-the-money (ATM) options. In particular, the discrepancy between benchmark prices and the numerical value for ATM down-and-out put options is shown in Fig. 1. We have not observed any stability issues with any of our  $h$ -transformation schemes. As discussed earlier, the standard Euler scheme with the hitting probability method has no discretisation error. The discrepancy is therefore essentially the statistical noise.

Our numerical results show a rapid convergence of the numerical approximation of prices given by the recurrent and transient transforms via the implicit scheme and demonstrate clearly its effectiveness over the standard Euler scheme without any hitting probability correction. This confirms the findings of our theoretical analysis even without satisfying all the conditions of Theorem 4.4.

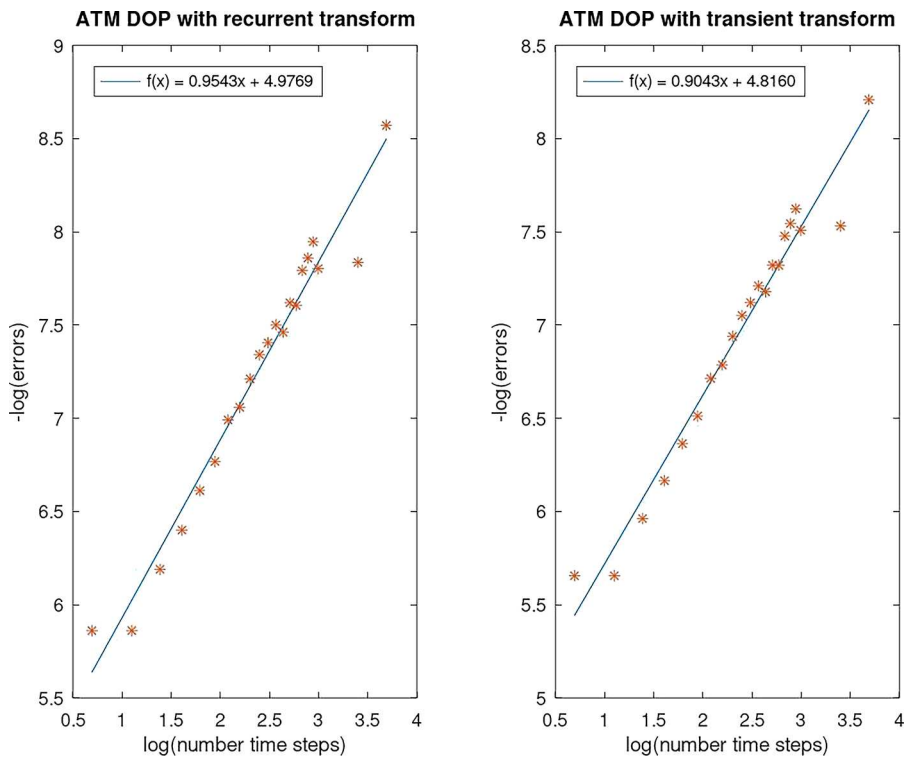
Moreover, the prices given by the recurrent and transient transforms are quite comparable as predicted by the theoretical analysis. Figure 2 show the log–log plot of the discrepancy associated to the recurrent and transient transforms, respectively, for an ATM down-and-out put option. The respective numerical rates of convergence observed are 0.95 and 0.9.

### 5.3 Time-homogeneous hyperbolic local volatility model

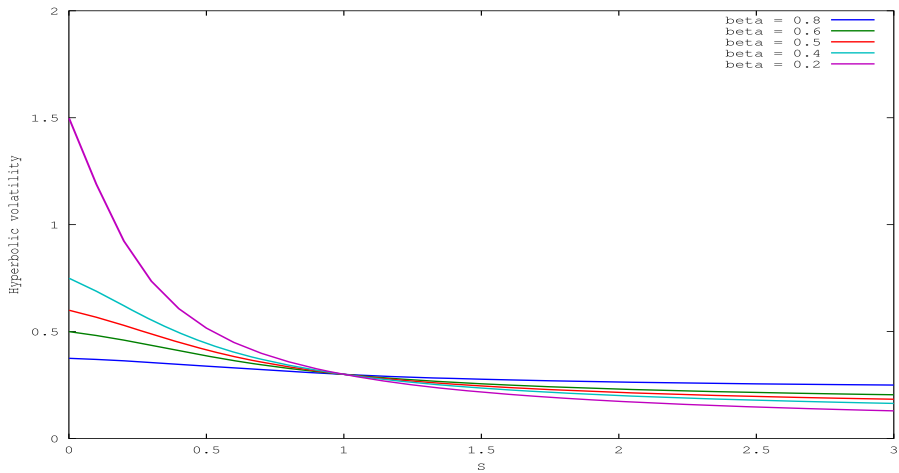
Empirical asset return distributions tend to exhibit fat tails (kurtosis) and skewness (asymmetric distribution). The *skew* or *smile* in implied volatility surfaces observed



**Fig. 1** Absolute discrepancy between the benchmark price for an ATM down-and-out put and those calculated with different numerical schemes when  $S_0 = 1$ ,  $K = 1$ ,  $T = 1$  year,  $\ell = \log(\underline{s} = 0.8)$ ,  $r = \infty$  and  $\sigma = 20\%$



**Fig. 2** Log-log plot of the absolute discrepancy for an ATM down-and-out put price with recurrent and transient transform numerical schemes when  $S_0 = 1$ ,  $K = 1$ ,  $T = 1$  year,  $\ell = \log(\underline{s} = 0.8)$ ,  $r = \infty$  and  $\sigma = 20\%$



**Fig. 3** Impact of the value  $\beta$  on the hyperbolic local volatility for fixed volatility level  $\nu = 0.3$

across various asset classes are a market reality (see e.g. Gatheral [19, Chap. 1], Overhaus et al. [37, Sect. 1.2] or Wilmott [46, Sect. 22.4]) and a manifestation of these stylised facts. We need more convenient models than *Gaussian* models for the asset  $S$  to reproduce more closely the implied volatility surfaces. Local volatility models, either parametric or non-parametric (see e.g. Dupire [17], Derman and Kani [16] or Rubinstein [41]) arguably capture the surface of implied volatilities more precisely than other approaches such as stochastic volatility models (see e.g. Ren et al. [38] or Romo [40]). Needless to say, the volatility surface has a significant impact on barrier option valuation.

For our analysis, we consider the time-homogeneous hyperbolic local volatility model (HLV) where the dynamics of the spot price under the risk-neutral measure is given by

$$dX_t = \sigma(X_t)dW_t, \quad X_0 = 1$$

with

$$\sigma(x) = \nu \left( \frac{(1 - \beta + \beta^2)}{\beta} x + \frac{(\beta - 1)}{\beta} \left( \sqrt{x^2 + \beta^2(1 - x)^2} - \beta \right) \right). \quad (5.5)$$

Here  $\nu > 0$  is the level of volatility, and  $\beta \in (0, 1]$  is the skew parameter. First introduced in Jäckel [31], this behaves similarly to the constant elasticity of variance (CEV), this model and has been widely used in quantitative finance for numerical experiments in Hok et al. [25, 26, 24]. A practical advantage of this model is that zero is not an attainable boundary, which in turn avoids some numerical instabilities present in the CEV model when the underlying asset price is close to zero (see e.g. Andersen and Andreasen [5]). It corresponds to the Black–Scholes model for  $\beta = 1$  and exhibits a skew for the implied volatility surface when  $\beta \neq 1$ . Figure 3 illustrates the impact of the parameter  $\beta$  on the skew of the volatility surface. We observe

that the skew increases significantly with decreasing value of  $\beta$ . For example with  $\nu = 0.3$ ,  $\beta = 0.2$ , the difference in volatility between strikes at 50% and at 100% is about 15%.

### 5.3.1 Down-and-up-out double barrier call option

In this implementation, we set  $h(x) = \frac{(x-\ell)(r-x)}{2(r-\ell)^2}$  with the functions  $h^{(1)}(x) = \frac{\ell+r-2x}{2(r-\ell)^2}$ ,  $h^{(2)}(x) = -\frac{1}{(r-\ell)^2}$  and  $h^{(3)}(x) = h^{(4)}(x) = 0$ . Note that with this choice of  $h$ , Condition 1) in Assumption 4.2 is not satisfied for  $k = 2$  and any  $p \in [0, \frac{1}{2}]$ . The associated BEM scheme is then solved by using a bisection method with Octave vectorisation for faster code execution. Consequently, the price is approximated by

$$\text{price} \approx h(x) \mathbb{E}^{h,x} \left[ \frac{(\widehat{X}_{t_N} - K)^+}{h(\widehat{X}_{t_N})} e^{\frac{1}{2} \frac{T}{N} \sum_{n=0}^{N-1} \sigma^2(\widehat{X}_{t_n}) \frac{h''}{h}(\widehat{X}_{t_n})} \right].$$

For comparison, we also compute the numerical price given by the standard Euler scheme (5.4) with series hitting probability correction, where the expression is given by an infinite series in Gobet [21]<sup>1</sup> as

$$\begin{aligned} \hat{p}_i &:= \mathbb{P}[\widehat{X}_t \in (\ell, r), \forall t \in [t_i, t_{i+1}] | \widehat{X}_{t_i} = x_i, \widehat{X}_{t_{i+1}} = x_{i+1}] \\ &= \mathbf{1}_{\{\ell < x_i, x_{i+1} < r\}} \sum_{n=-\infty}^{\infty} \left( e^{\frac{-2n(r-\ell)(n(r-\ell)+x_{i+1}-x_i)}{\sigma^2(t_{i+1}-t_i)}} - e^{\frac{-2(n(r-\ell)+x_i-r)(n(r-\ell)+x_{i+1}-r)}{\sigma^2(t_{i+1}-t_i)}} \right). \end{aligned} \tag{5.6}$$

Here  $\sigma$  is computed using the parametric local volatility function (5.5). A numerical study of (5.6) suggests that it suffices to calculate the leading two or three terms for most cases. To be conservative, in our tests, the  $\hat{p}_i$  are estimated using  $n$  from  $-5$  to 5. Experiment details and comparison results are described below.

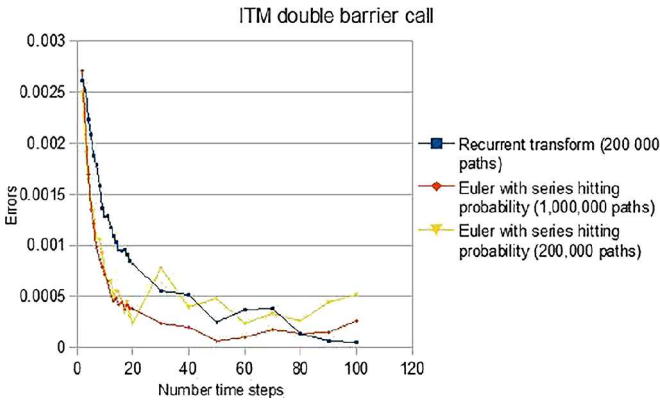
### 5.3.2 Set of parameters

The numerical experiments are conducted using for the parameters the values  $S_0 = 1$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.85$ ,  $\bar{s} = 1.25$ . For thoroughness, we consider in-the-money ( $K = 0.9$ ), at-the-money ( $K = 1$ ) and out-of-the-money ( $K = 1.05$ ) options. The benchmark prices for each numerical method are computed by the method itself with a very dense time grid and a high number of Monte Carlo paths.

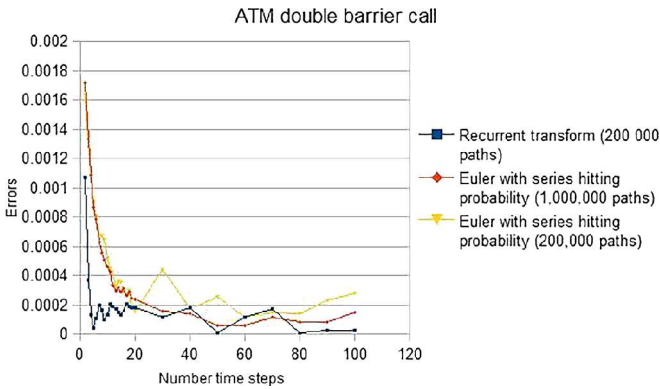
In this case, we observed some differences regarding the moneyness of the option in our numerical results. More precisely, the method performed relatively poorly for the ATM option. For this reason, we report below the results in all three cases and provide an explanation for the seemingly poor performance for the ATM option.

The discrepancies between benchmark prices and numerical methods for ITM, ATM and OTM double barrier call options are shown respectively in Figs. 4–6. Tables 1–3 provide the 95% confidence intervals associated to each considered numerical method. For the same number of MC paths, i.e., 200'000, the recurrent transform shows tighter confidence intervals, which match the confidence intervals of the

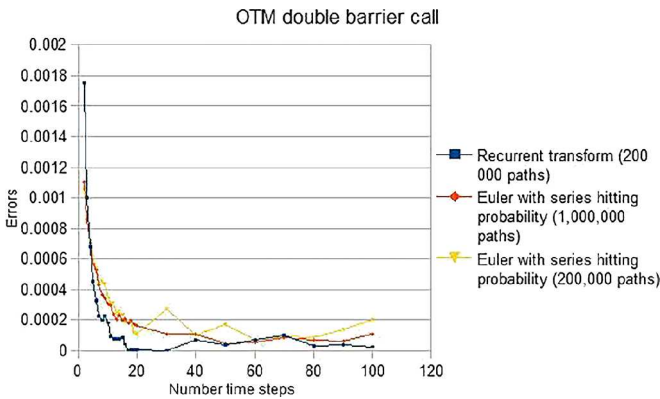
<sup>1</sup>with a typography correction.



**Fig. 4** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ITM double barrier call when  $S_0 = 1$ ,  $K = 0.9$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.85$ ,  $\bar{s} = 1.25$



**Fig. 5** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ATM double barrier call when  $S_0 = 1$ ,  $K = 1$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.85$ ,  $\bar{s} = 1.25$



**Fig. 6** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an OTM double barrier call when  $S_0 = 1$ ,  $K = 1.05$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.85$ ,  $\bar{s} = 1.25$



**Table 1** ITM double barrier call: numerical prices and confidence intervals as functions of the number of time steps. Recurrent transform with 200'000 paths, Euler with series hitting probability with 1'000'000 paths (Euler+ (1M)) and Euler with series hitting probability with 200'000 paths (Euler+ (200K)). Benchmark price = 0.0462

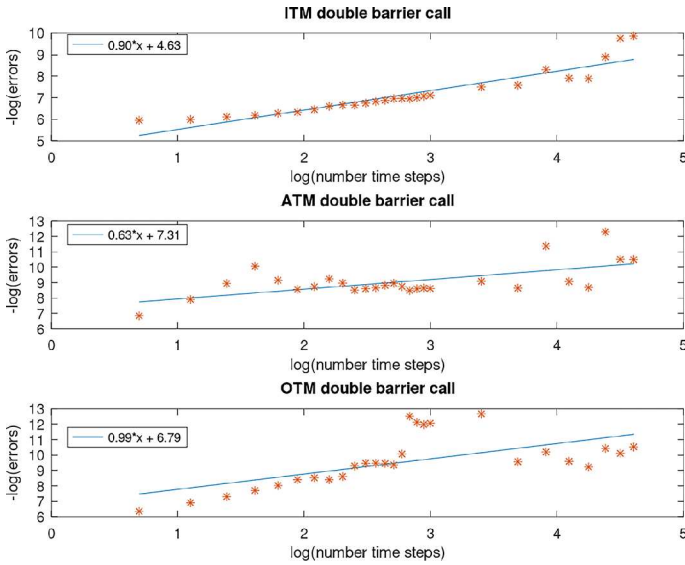
Time steps	Recurrent transform	Euler+ (1M)	Euler+ (200K)
2	0.0489 ± 0.00012	0.0435 ± 0.00012	0.0438 ± 0.00026
6	0.0481 ± 0.00014	0.0451 ± 0.00014	0.0449 ± 0.00031
10	0.0475 ± 0.00014	0.0456 ± 0.00014	0.0455 ± 0.00032
16	0.0472 ± 0.00015	0.0458 ± 0.00015	0.0458 ± 0.00033
20	0.0471 ± 0.00015	0.0459 ± 0.00015	0.0460 ± 0.00033
30	0.0468 ± 0.00015	0.0460 ± 0.00015	0.0455 ± 0.00034
40	0.0468 ± 0.00015	0.0461 ± 0.00015	0.0459 ± 0.00034
60	0.0466 ± 0.00016	0.0462 ± 0.00015	0.0460 ± 0.00035
100	0.0463 ± 0.00016	0.0460 ± 0.00016	0.0457 ± 0.00035

**Table 2** ATM double barrier call: numerical prices and confidence intervals as functions of the number of time steps. Recurrent transform with 200'000 paths, Euler with series hitting probability with 1'000'000 paths (Euler+ (1M)) and Euler with series hitting probability with 200'000 paths (Euler+ (200K)). Benchmark price = 0.0193

Time steps	Recurrent transform	Euler+ (1M)	Euler+ (200K)
2	0.0183 ± 0.00009	0.0176 ± 0.00006	0.0177 ± 0.00014
6	0.0194 ± 0.00010	0.0185 ± 0.00008	0.0185 ± 0.00017
10	0.0195 ± 0.00011	0.0189 ± 0.00008	0.0188 ± 0.00018
16	0.0195 ± 0.00011	0.0190 ± 0.00008	0.0190 ± 0.00018
20	0.0195 ± 0.00011	0.0191 ± 0.00008	0.0192 ± 0.00019
30	0.0194 ± 0.00011	0.0192 ± 0.00008	0.0189 ± 0.00019
40	0.0195 ± 0.00012	0.0192 ± 0.00009	0.0192 ± 0.00019
60	0.0194 ± 0.00012	0.0193 ± 0.00009	0.0192 ± 0.00019
100	0.0193 ± 0.00012	0.0192 ± 0.00009	0.0191 ± 0.00019

**Table 3** OTM double barrier call: numerical prices and confidence intervals as functions of the number of time steps. Recurrent transform with 200'000 paths, Euler with series hitting probability with 1'000'000 paths (Euler+ (1M)) and Euler with series hitting probability with 200'000 paths (Euler+ (200K)). Benchmark price = 0.0103

Time steps	Recurrent transform	Euler+ (1M)	Euler+ (200K)
2	0.0085 ± 0.00006	0.0092 ± 0.00004	0.0093 ± 0.00009
6	0.0100 ± 0.00008	0.0098 ± 0.00005	0.0098 ± 0.00011
10	0.0101 ± 0.00008	0.0100 ± 0.00005	0.0100 ± 0.00012
16	0.0103 ± 0.00008	0.0101 ± 0.00005	0.0101 ± 0.00012
20	0.0103 ± 0.00008	0.0102 ± 0.00005	0.0102 ± 0.00012
30	0.0103 ± 0.00008	0.0102 ± 0.00006	0.0101 ± 0.00012
40	0.0104 ± 0.00009	0.0102 ± 0.00006	0.0102 ± 0.00013
60	0.0104 ± 0.00009	0.0103 ± 0.00006	0.0103 ± 0.00013
100	0.0103 ± 0.00009	0.0102 ± 0.00006	0.0101 ± 0.00013

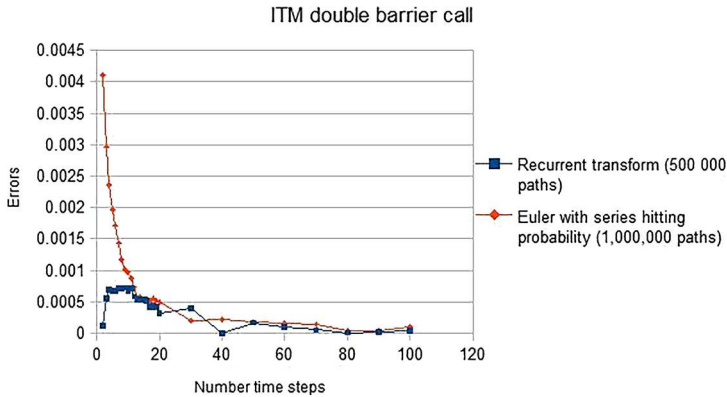


**Fig. 7** Log–log plot of the absolute discrepancy for double barrier call prices for ITM ( $K = 0.9$ ), ATM ( $K = 1$ ) and OTM ( $K = 1.05$ ) with the recurrent transform numerical scheme when  $S_0 = 1$ ,  $T = 1$  year,  $\underline{s} = 0.85$ ,  $\bar{s} = 1.25$ ,  $\nu = 20\%$  and  $\beta = 0.5$

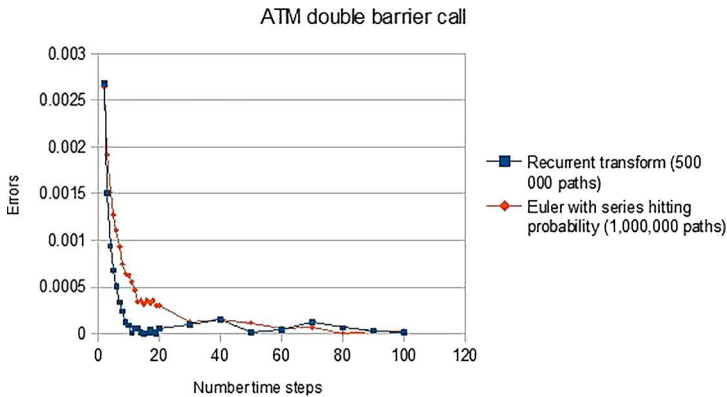
benchmark scheme with 1’000’000 paths. Overall, note that the size of the interval is about a few bps. We have not observed any stability issues with the recurrent transform scheme. Interestingly, our recurrent transformation has a much smaller error than the explicit Euler method with a hitting probability correction when the number of discretisations is reasonably large. More importantly, this outperformance is still valid even if the number of Monte Carlo simulations for the explicit Euler method is increased five times. Having said that, one should still treat such a conclusion with caution as our benchmark price and hitting probabilities are calculated by applying a truncation and thus subject to error. Nevertheless, the *outperformance* is still promising as our truncation is no coarser than the common industry practice.

Figure 7 shows the log–log plot of the discrepancy associated to the recurrent transform method for ITM, ATM and OTM double barrier call options. The numerical rates of convergence are respectively 0.91, 0.63 and 1, using 200’000 Monte Carlo simulations. Although the rate of convergence for the ATM option is far from the theoretical rate of 1, a closer look at Fig. 5 reveals a clue. Note that the error of the approximation converges very rapidly to zero after a few iterations, and further discretisations do not significantly alter the already very small error term. This indicates that the observed error in this case can be mostly attributed to the statistical noise, and the simple regression to obtain the convergence rate does not work well.

When we run the same experiment for the Euler scheme with a hitting probability correction with 200’000 Monte Carlo simulations, we observe a similar drop in the performance, and the convergence rates are found to be 0.50, 0.59 and 0.61, respectively. However, the convergence rates for the latter scheme increase to 0.83, 0.83 and 0.77, respectively, when the number of simulations are increased five times.



**Fig. 8** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ITM double barrier call when  $S_0 = 1$ ,  $K = 0.9$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.3$

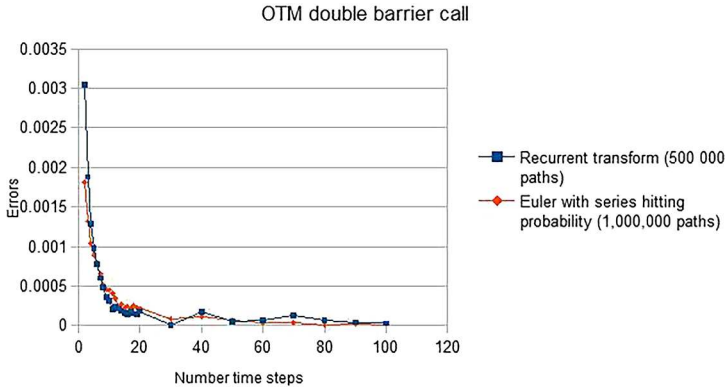


**Fig. 9** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ATM double barrier call when  $S_0 = 1$ ,  $K = 1$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.3$

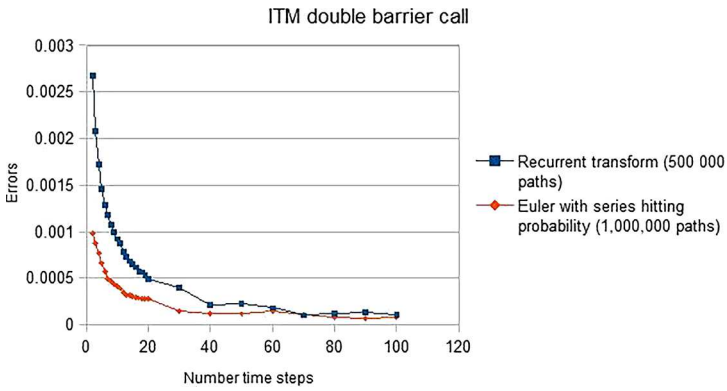
As complements, some variations of barrier levels are considered and their impact on numerical results is quantified:

- Low barrier  $\underline{s} = 0.8$  and high barrier  $\bar{s} = 1.3$ : widening of barrier levels.
- Low barrier  $\underline{s} = 0.8$  and high barrier  $\bar{s} = 1.15$ : tightening of barrier levels.

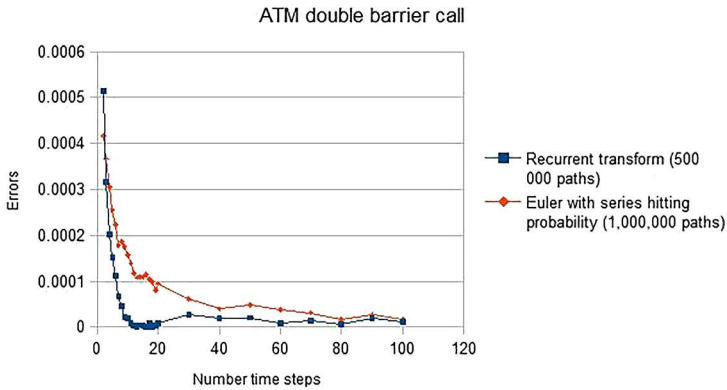
Here, we increase the number of MC samples from 200'000 to 500'000 for the BEM method and keep the same number of MC paths at 1'000'000 for the standard Euler scheme with series hitting probability. The 95% confidence intervals for both methods are comparable and of the order 2–3 bps. For each barrier level configuration, we have considered in-the-money ( $K = 0.9$ ), at-the-money ( $K = 1$ ) and out-of-the-money ( $K = 1.05$ ) options. The discrepancies between benchmark prices and numerical methods with respect to the number of time steps are presented in Figs. 8–13.



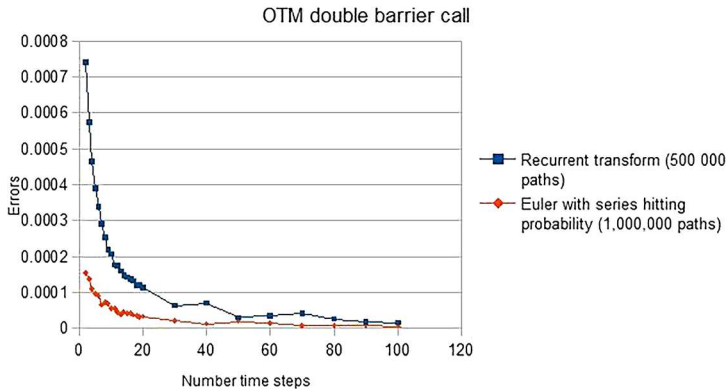
**Fig. 10** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an OTM double barrier call when  $S_0 = 1$ ,  $K = 1.05$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.3$



**Fig. 11** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ITM double barrier call when  $S_0 = 1$ ,  $K = 0.9$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.15$



**Fig. 12** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an ATM double barrier call when  $S_0 = 1$ ,  $K = 1$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.15$



**Fig. 13** Absolute discrepancy between the benchmark price and those calculated by different numerical schemes for an OTM double barrier call when  $S_0 = 1$ ,  $K = 1.05$ ,  $\nu = 20\%$ ,  $\beta = 0.5$ ,  $T = 1$  year,  $\underline{s} = 0.8$ ,  $\bar{s} = 1.15$

Overall, both methods show comparable convergence results which are in accordance with the theoretical analysis.

## 6 Conclusion

We have introduced a novel backward Euler–Maruyama method to increase the weak convergence rate of approximations in the presence of killing. The numerical experiments confirm our theoretical result that the convergence rate is of the order  $1/N$ , where  $N$  is the number of discretisation steps. This corresponds to an order-1 weak convergence rate, which is the best rate that one can achieve (see Gobet [21]).

Moreover, the numerical studies suggest that one does not need a large  $N$  to obtain a sufficiently close approximations as all numerical studies indicate error terms diminishing very rapidly with a small number of iterations. The numerical experiments also suggested that our method outperforms the *Brownian bridge method* in certain cases, although such a statement currently does not have any theoretical backing.

We suggest a couple of interesting avenues for future research in addition to the extension of the BEM scheme to higher dimensions as discussed in the introduction:

- The transform method with BEM has the potential to achieve a higher order of weak convergence. Indeed, one possibility of improvement is to combine the BEM scheme with the Romberg extrapolation method (see e.g. Glasserman [20, Sect. 6.2.4]). Some preliminary tests are encouraging, showing that the Romberg method generates significantly lower errors compared to the Euler hitting probability method and BEM. We believe that we can expand further the weak error to justify the use of the Romberg method, and a higher order weak convergence analysis is currently under investigation.

- Exact simulation of diffusions. Especially in a multidimensional setting, it is a challenge to keep the simulated values of  $X$  in its domain even though it should not touch its boundaries in theory. The implicit scheme considered in this paper is one way out. Exact simulation, introduced in Beskos and Roberts [6] for

the one-dimensional case, may be another way to resolve this issue. It involves a rejection-sampling algorithm and, when applicable, returns *exact* draws from any finite-dimensional distribution of the solution to the SDE. The method has been further extended to multivariate diffusions in Blanchet and Zhang [7], although some open questions remain regarding the speed of convergence of the algorithm. It will be interesting to study exact simulations for the recurrent transformations that may lead to a decrease in computation time by avoiding implicit schemes, especially in higher dimensions.

### Appendix A: Proof of Lemma 2.5

**Proof** 1) Since  $h$  is concave,  $H'(x) > 1$ , which shows the desired strict monotonicity. If  $r = \infty$  and  $h(x) = x$ , it is immediate that  $H((0, \infty)) = \mathbb{R}$ .

Next suppose  $h \in \mathcal{H}_0$ . Then dominated convergence implies  $h(0) = 0$  as well as  $h(r) = 0$  if  $r < \infty$  as the potential density vanishes at finite endpoints. Moreover, as  $h$  is strictly concave and never vanishes in the interior of the state space,  $h'(0) > 0$ . Thus

$$\lim_{x \rightarrow 0} \frac{h'(x)}{h(x)} = \infty.$$

This proves the desired range for  $H$  when  $r = \infty$ . Indeed, in this case,  $h$  is increasing which in turn yields

$$\frac{h'(x)}{h(x)} \leq \frac{h'(1)}{h(1)}, \quad x \geq 1.$$

If  $r < \infty$ , similar considerations imply  $h'(r) < 0$  and therefore

$$\lim_{x \rightarrow r} \frac{h'(x)}{h(x)} = -\infty.$$

This completes the proof of the first assertion.

2) If  $r = \infty$  and  $h(x) = x$ , then  $h'(x) = 1$  for all  $x \geq 0$ . If  $h \in \mathcal{H}_0$ , then

$$h'(x) = \int_x^\infty f(y)m(dy),$$

which is nonnegative and finite by the assumption on  $f$ . In particular, we obtain that  $h'(0) = \int_0^\infty f(y)m(dy)$  and  $h'(\infty) = 0$ .

If  $r < \infty$ , then

$$h(x) = \frac{r-x}{r} \int_0^x yf(y)m(dy) + x \int_x^r \frac{r-y}{r} f(y)m(dy).$$

Thus

$$h'(x) = \int_x^r f(y)m(dy) - \frac{1}{r} \int_0^r yf(y)m(dy).$$

This yields the desired boundedness and the boundary levels for the derivatives.

3) First suppose  $r < \infty$ . Since  $h'$  does not vanish at the boundaries, the ratio  $u(y, y)/h(y)$  is bounded. Moreover,

$$-\int_{(0,r)} h''(y)dy = \int_{(0,r)} f(y)m(dy).$$

This proves the claim when  $r < \infty$  since  $h'$  is clearly integrable by the finiteness of  $h$  and it changes sign only once.

Now suppose that  $r = \infty$ . Thus  $u(y, y) = y$  and

$$-\int_{(0,\infty)} y \frac{h''(y)}{h(y)} dy = \int_{(0,1)} y \frac{f(y)}{h(y)} m(dy) + \int_{(1,\infty)} y \frac{f(y)}{h(y)} m(dy).$$

The first integral on the right-hand side is finite since  $f$  is  $m$ -integrable and  $y/h(y)$  is bounded on  $[0, 1]$  as  $h'(0) > 0$ . The second integral is also finite since  $h(\infty) > 0$  and  $\int yf(y)m(dy) < \infty$  by assumption. Moreover, noting that  $h' \geq 0$ , we get

$$\begin{aligned} \int_0^\infty (y \wedge 1) \frac{h'(y)}{h(y)} dy &= \int_0^1 \frac{yh'(y)}{h(y)} dy + \int_1^\infty \frac{h'(y)}{h(y)} dy \\ &\leq K \int_0^1 h'(y)dy + \log \frac{h(\infty)}{h(1)} < \infty. \end{aligned} \quad \square$$

### Appendix B: Proof of Theorem 3.3

The proof is divided into several steps, considering first the case  $r = \infty$  and making use of the comparison result used in the proof of Lemma 3.2. In what follows,  $K$  denotes a generic constant independent of  $N$ .

1) First suppose  $r = \infty$ . Since  $1/h$  is decreasing, Lemma 3.2 and Theorem 2.9 imply  $\sup_{t \leq T, N} E^{h, X_0}[\frac{1}{h}(\widehat{X}_t)] < \infty$ . Moreover, Lemma 3.2 also yields that

$$\begin{aligned} E^{h, X_0} \left[ \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n})h^{-2-p}(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \right] &\leq E^{h, X_0} \left[ \int_0^{A_T} \frac{1}{h^{2+p}(Y_t)} dt \right] \\ &\leq E^{h, X_0} \left[ \int_0^{\|\sigma\|_\infty^2 T} \frac{1}{h^{2+p}(Y_t)} dt \right] < \infty, \end{aligned}$$

where  $Y$  is a process with the same law as the process in Theorem 2.9 with  $c = c_1$  and the last inequality follows from Theorem 2.9. Similarly, by considering instead the time change

$$dA_t = \frac{\sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt, \quad t \in (t_n, t_{n+1}), \quad A_{t_n} = t_n,$$

we obtain  $h^{-p}(\widehat{X}_\tau) \leq h^{-p}(Y_{A_\tau})$ , where  $Y$  is a process such that  $Y_{t_n} = \widehat{X}_{t_n}$  and

$$dY_t = d\beta_t + \left( \frac{h'}{h}(Y_t) + c_1 \right) dt, \quad t \geq t_n,$$

with  $\beta$  a standard Brownian motion. Consequently, Theorem 2.9 yields

$$\operatorname{ess\,sup}_{\tau \in \mathcal{T}_n} E^{h, X_0} \left[ \frac{1}{h}(\widehat{X}_\tau) \middle| \mathcal{F}_{t_n} \right] < \infty$$

since  $A_\tau \leq t_n + \|\sigma\|_\infty^2 \frac{T}{N}$  a.s. for  $\tau \in \mathcal{T}_n$ .

2) For the case  $r < \infty$ , we first define the constants  $x_1 := \inf\{x \geq 0 : h'(x) = 0\}$  and  $x_2 := \inf\{x \geq x_1 : h'(x) < 0\}$ . Then there exist functions  $h_1$  and  $h_2$  such that  $h = h_1 h_2$  and  $h_1$  (resp.  $h_2$ ) is nondecreasing (resp. nonincreasing) and constant on  $(x_1, r)$  (resp.  $(0, x_2)$ ). Now define the processes  $\widehat{Y}^i$ , where  $\widehat{Y}_0^i = X_0$  and

$$\begin{aligned} d\widehat{Y}_t^i &= \frac{\sigma(\widehat{X}_{t_n})}{H_x(t_n, \widehat{X}_{t_n}; t, \widehat{Y}_t^i)} dW_t \\ &+ \frac{\sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{Y}_t^i)} \left( \frac{h'_i}{h_i}(\widehat{Y}_t^i) + c_i \right) dt, \quad t \in (t_n, t_{n+1}]. \end{aligned}$$

Applying Itô's formula to  $((x_2 - \widehat{Y}_t^1)^+)^2$  and  $((x_2 - \widehat{X}_t)^+)^2$ , the comparison theorem used in the proof of Lemma 3.2 shows that

$$P^{h, X_0}[\widehat{Y}_t^1 \wedge x_2 \leq \widehat{X}_t \wedge x_2, t \leq T] = 1.$$

An analogous argument shows that we also have

$$P^{h, X_0}[\widehat{Y}_t^2 \vee x_1 \geq \widehat{X}_t \vee x_1, t \leq T] = 1.$$

As  $h_1$  is nondecreasing,  $h_2$  is nonincreasing and  $\frac{h}{h_1}$  (resp.  $\frac{h}{h_2}$ ) is constant on  $(0, x_2)$  (resp.  $(x_1, r)$ ), the above comparisons imply that we have  $\frac{1}{h(\widehat{X}_t \wedge x_2)} \leq \frac{1}{h(\widehat{Y}_t^1 \wedge x_2)}$  and  $\frac{1}{h(\widehat{X}_t \vee x_1)} \leq \frac{1}{h(\widehat{Y}_t^2 \vee x_1)}$ . So the same time-change argument as in the proof of Lemma 3.2 yields that

$$\sup_{t \leq T, N} E^{h, X_0} \left[ \frac{1}{h}(\widehat{X}_t) \right] < \infty$$

by another application of Theorem 2.9. More precisely, note that Lemma 3.2 is not immediately applicable to  $h_2$  and  $\widehat{Y}^2$  since the former is decreasing and the latter's state space is not  $(0, \infty)$ . However, if one considers instead the function  $h_2^r(x) = h_2(r - x)$  and the process  $\widehat{Y}^{2,r} = r - \widehat{Y}^2$ , the new function  $h_2^r$  becomes increasing on the state space of  $\widehat{Y}^{2,r}$ , which now equals  $(0, \infty)$ . By applying Lemma 3.2 to  $h_2^r$  and  $\widehat{Y}^{2,r}$ , we obtain the estimate above. Note that this reasoning actually implies that

$$\sup_{t \leq T, N} E^{h, X_0} \left[ \frac{1}{h(\widehat{X}_t)} \right] \leq \frac{K'}{h(X_0)}$$



for some  $K'$  that depends only on  $T$ . Similarly, in view of Theorem 2.9 again,

$$\sup_N E^{h, X_0} \left[ \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{\sigma^2(\widehat{X}_{t_n}) h^{-2-p}(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt \right] < \infty.$$

Finally, analogous considerations also yield  $\text{ess sup}_{\tau \in \mathcal{T}_n} E^{h, X_0} [\frac{1}{h}(\widehat{X}_\tau) | \mathcal{F}_{t_n}] < \infty$ .

3) We now show the boundedness of the moments. Note that there is nothing to show when  $r < \infty$ . So assume that  $r = \infty$ . Recall that

$$\widehat{X}_t = \widehat{X}_{t_n} + \sigma^2(\widehat{X}_{t_n})(t - t_n) \frac{h'}{h}(\widehat{X}_t) + \sigma(\widehat{X}_{t_n})(W_t - W_{t_n}).$$

Thus

$$E^{h, X_0}[\widehat{X}_t] \leq E^{h, X_0}[\widehat{X}_{t_n}] + K(t - t_n)$$

for some  $K$  due to the boundedness of  $\sigma$  and  $h'$  as well as the uniform bound on the inverse moment of  $h(\widehat{X}_t)$ . This shows that  $\sup_{t \leq T, M} E^{h, X_0}[\widehat{X}_t] \leq X_0 + KT$ . Now suppose that  $E(m) := \sup_{t \leq T, N} E^{h, X_0}[\widehat{X}_t^m] < \infty$  and deduce from (3.1) that

$$\begin{aligned} d\widehat{X}_t^{m+1} &= dZ_t + (m + 1) \frac{\widehat{X}_t^m \sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \left( \frac{h'}{h}(\widehat{X}_t) + \mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t) \right) dt \\ &\quad + \frac{1}{2} m(m + 1) \frac{\widehat{X}_t^{m-1} \sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt, \end{aligned}$$

where  $Z$  is a local martingale. Next observe that for  $m \geq 1$ ,

$$x^m/h \leq K(1 + x^{m-1}) \tag{B.1}$$

as  $h(0) = 0, h'(0) > 0$  and  $h'/h \leq \frac{1}{x}$ . This last identity follows from the fact that

$$h(x) = - \int_0^x yh''(y)dy + xh'(x).$$

Moreover, the representation (3.2) of  $\mu$  and (3.3) show that

$$|\mu| \leq K(H_x + 1) \frac{1}{h} \tag{B.2}$$

since the term in front of the parentheses in (3.2) is bounded.

Now observe that  $\tau_k := \inf\{t \geq t_n : \widehat{X}_t \geq k\}, k \in \mathbb{N}$ , is a localising sequence for  $Z$ . Therefore a standard localisation argument, (B.1) and (B.2) together imply for  $t \in (t_n, t_{n+1}]$  that in view of Fatou's lemma,

$$E^{h, X_0}[\widehat{X}_t^{m+1}] \leq E^{h, X_0}[\widehat{X}_{t_n}^{m+1}] + (t - t_n)KE(m - 1)$$

for some constant  $K$ , which in turn yields

$$E(m + 1) \leq X_0^{m+1} + KTE(m - 1).$$

Finally, note that this in particular implies that  $Z$  is a true martingale. Thus for  $\tau \in \mathcal{T}_n$  and  $m \geq 2$ , we obtain

$$\widehat{X}_\tau^m \leq \widehat{X}_{t_n}^m + M_\tau + K \int_{t_n}^{\tau} \widehat{X}_t^{m-1} dt.$$

Taking conditional expectations yields (3.6) because

$$E^{h, X_0}[\widehat{X}_\tau^m | \mathcal{F}_{t_n}] \leq \widehat{X}_{t_n}^m + K E^{h, X_0} \left[ \int_{t_n}^{\tau} \widehat{X}_t^{m-1} dt \middle| \mathcal{F}_{t_n} \right]. \tag{B.3}$$

To establish (3.7), we need the following result.

**Lemma B.1** *Suppose that  $h$  satisfies the conditions of Lemma 3.1,  $\sigma$  is bounded and consider the BEM scheme defined by (2.4). For any  $p \in [0, 1)$ , any  $n$  and  $t_n \leq s \leq t < t_{n+1}$ , we have*

$$E^{h, X_0}[h^{-p}(\widehat{X}_t) | \mathcal{F}_s] \leq h^{-p}(\widehat{X}_s) \exp(K(t - s))$$

for some constant  $K > 0$  that is independent of  $n$ .

**Proof** Let  $\mu_t := \mu(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)$ . A straightforward application of Itô’s formula yields

$$\begin{aligned} & h^{-p}(\widehat{X}_u) - h^{-p}(\widehat{X}_s) \\ &= M_u - M_s \\ & \quad - \int_s^u \frac{\sigma^2(\widehat{X}_{t_n}) p h^{-p}(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} \left( \frac{2\mu_t h'(\widehat{X}_t) + h''(\widehat{X}_t)}{2h(\widehat{X}_t)} + \frac{1-p}{2} \left( \frac{h'}{h}(\widehat{X}_t) \right)^2 \right) dt \\ & \leq M_u - M_s - \int_s^u \frac{\sigma^2(\widehat{X}_{t_n})}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} p h^{-p}(\widehat{X}_t) \left( -\frac{\alpha_1}{h(\widehat{X}_t)} + \frac{(1-p)\alpha_2}{h^2(\widehat{X}_t)} \right) dt \end{aligned}$$

where  $M$  is a local martingale and  $\alpha_1$  and  $\alpha_2$  are positive constants depending on the bounds on  $h'$  and  $h''$  since  $\mu_t > c_1$  (resp.  $\mu_t < c_2$ ) whenever  $h'(\widehat{X}_t) > 0$  (resp.  $h'(\widehat{X}_t) < 0$ ) by Lemma 3.1 and  $h'$  never vanishes at the same time as  $h$ . Thus there exists a constant  $K$  that depends only on  $h$ ,  $p$  and  $c_1$  and  $c_2$  such that

$$h^{-p}(\widehat{X}_u) - h^{-p}(\widehat{X}_s) \leq M_u - M_s + \int_s^u K \frac{\sigma^2(\widehat{X}_{t_n}) h^{-p}(\widehat{X}_t)}{H_x^2(t_n, \widehat{X}_{t_n}; t, \widehat{X}_t)} dt,$$

since  $-\alpha_1 x + (1-p)\alpha_2 x^2$  is bounded from below.

Next note that  $\tau_k := \inf\{t \geq t_n : \widehat{X}_t < 1/k\}, k \in \mathbb{N}$ , is a localising sequence for  $M$ . Using the optional stopping theorem, Fatou’s lemma and monotone convergence, we thus arrive at

$$\begin{aligned} & E^{h, X_0}[h^{-p}(\widehat{X}_t) | \mathcal{F}_s] \\ & \leq h^{-p}(\widehat{X}_s) + K E^{h, X_0} \left[ \int_s^t h^{-p}(\widehat{X}_u) du \middle| \mathcal{F}_s \right], \quad t_n \leq s \leq t \leq t_{n+1}, \end{aligned}$$

for some constant  $K$  in view of the boundedness of  $\sigma$ . We deduce the claim by Gronwall’s lemma.  $\square$

Now we return to the proof of the estimate (3.7). Observe that the hypothesis on  $h''$  implies

$$1 - \exp\left((s - t_n)\sigma^2(\widehat{X}_{t_n})\frac{h''}{2h}(\widehat{X}_{t_n})\right) \leq K \frac{T}{N} \frac{1}{h^p}(\widehat{X}_{t_n})$$

for some  $K > 0$ . Without loss of generality, we also suppose that  $h \leq 1$ . Thus

$$\begin{aligned} & E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \left( 1 - \exp\left((s - t_n)\sigma^2(\widehat{X}_{t_n})\frac{h''}{2h}(\widehat{X}_{t_n})\right) \right) \frac{\sigma^2(\widehat{X}_{t_n})h^{-p}(\widehat{X}_s)}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \right] \\ & \leq K \frac{T}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} h^{-p}(\widehat{X}_{t_n})h^{-p}(\widehat{X}_s) ds \right] \\ & \leq K \frac{T}{N} E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} h^{-2p}(\widehat{X}_{t_n}) ds \right] \\ & \leq K \frac{T^2}{N^2} E^{h, X_0} [h^{-1}(\widehat{X}_{t_n})], \end{aligned}$$

where the second line follows from Lemma B.1 and  $H_x \geq 1$ .

Next suppose  $m \geq 1$ . Note that calculations similar to those leading to (B.3) yield

$$E^{h, X_0} [\widehat{X}_t^m | \mathcal{F}_{t_n}] \leq \widehat{X}_{t_n}^m + K E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \widehat{X}_s^{m-1} ds \mid \mathcal{F}_{t_n} \right].$$

Thus the elementary inequality  $x^{m-1} \leq 1 + x^m$  and Gronwall’s lemma imply that

$$E^{h, X_0} [\widehat{X}_t^m | \mathcal{F}_{t_n}] \leq K \left( \widehat{X}_{t_n} + \frac{T}{N} \right).$$

Therefore we get

$$\begin{aligned} & E^{h, X_0} \left[ \int_{t_n}^{t_{n+1}} \left( 1 - \exp\left((s - t_n)\sigma^2(\widehat{X}_{t_n})\frac{h''}{2h}(\widehat{X}_{t_n})\right) \right) \frac{\sigma^2(\widehat{X}_{t_n})\widehat{X}_s^m}{H_x^2(t_n, \widehat{X}_{t_n}; s, \widehat{X}_s)} ds \right] \\ & \leq K \frac{T}{N} E^{h, X_0} \left[ \left( \widehat{X}_{t_n} + \frac{T}{N} \right) \left( 1 - \exp\left(-\frac{T}{N} \frac{a}{h}(\widehat{X}_{t_n})\right) \right) \right] \\ & \leq K \frac{T^2}{N^2} E^{h, X_0} \left[ \left( \widehat{X}_{t_n} + \frac{T}{N} \right) h^{-1}(\widehat{X}_{t_n}) \right] \\ & \leq K \frac{T^2}{N^2} E^{h, X_0} [\widehat{X}_{t_n} + h^{-1}(\widehat{X}_{t_n})], \end{aligned}$$

where the last line follows from (3.8). Combining the above estimates, we arrive at the claimed result (3.7) via (3.5).  $\square$

### Appendix C: Relationship with Alfonsi [4] and Neuenkirch and Szpruch [36]

The purpose of this short section is to show that the results of [4] and [36] do not apply to the simulation of the recurrent diffusion process in (1.2). We impose below Assumptions 2.1–2.3.

**Proposition C.1** *Let  $X$  be given by (1.2), where  $h$  is a function satisfying the conditions of Çetin [9, Theorem 3.2]. Suppose further that  $b \equiv 0$  and  $\sigma \equiv 1$ . Then*

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \left( \left( \frac{h'}{h} \right)' (X_t) \right)^2 dt \right] = \infty.$$

**Proof** Suppose to the contrary that

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \left( \left( \frac{h'}{h} \right)' (X_t) \right)^2 dt \right] < \infty. \tag{C.1}$$

First observe that

$$d \left( \frac{1}{h} (X_t) \right) = -\frac{h'}{h^2} (X_t) dW_t + dC_t, \tag{C.2}$$

where  $C$  is an adapted, continuous and increasing process. Moreover,  $h$  is concave since  $X$  is on natural scale under  $\mathbb{P}$ , and we also have

$$\left( \frac{h'}{h} \right)' = \frac{h''}{h} - \left( \frac{h'}{h} \right)^2.$$

In particular,  $\left| \left( \frac{h'}{h} \right)' \right| > \left( \frac{h'}{h} \right)^2$ . Thus

$$\mathbb{E}^{\mathbb{Q}} \left[ \left( \frac{h'}{h} \right)^4 (X_t) \right] \leq \mathbb{E}^{\mathbb{Q}} \left[ \left( \left( \frac{h'}{h} \right)' (X_t) \right)^2 \right]$$

Also note that  $h'$  never vanishes at the boundary points where  $h$  does, and it does vanish at  $r$  if  $r = \infty$ , where  $h$  does not. Consequently,

$$\frac{(h')^2}{h^4} \leq K_1 \left( \frac{h'}{h} \right)^4 + K_2$$

for some  $K_1$  and  $K_2$ , which in conjunction with (C.1) implies that

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_0^T \left( \frac{h'}{h^2} \right)^2 (X_t) dt \right] < \infty.$$

Therefore, the local martingale term in (C.2) is a true  $\mathbb{Q}$ -martingale, which in turn yields that  $\frac{1}{h} \exp(-A)$  is a true  $\mathbb{Q}$ -martingale, where  $dA_t = -\frac{1}{2} \frac{h''(X_t)}{h(X_t)} dt$ . In particular,  $\mathbb{P} \ll \mathbb{Q}$  when restricted to  $\mathcal{F}_t$ . This further implies that  $\mathbb{P} \approx \mathbb{Q}$  when restricted

to  $\mathcal{F}_t$  since  $\mathbb{Q} \ll \mathbb{P}$  on the same restriction by [9, Theorem 3.2]. However, this is a contradiction since for any  $t > 0$ ,  $\mathbb{P}[\zeta < t] > 0$  while  $\mathbb{Q}[\zeta < t] = 0$  because  $X$  is a recurrent process under  $\mathbb{Q}$ .  $\square$

## Declarations

**Competing Interests** 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. Aizenman, M., Simon, B.: Brownian motion and Harnack inequality for Schrödinger operators. *Commun. Pure Appl. Math.* **35**, 209–273 (1982)
2. Akahori, J., Imamura, Y.: On a symmetrization of diffusion processes. *Quant. Finance* **14**, 1211–1216 (2014)
3. Alfonsi, A.: On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11**, 355–384 (2005)
4. Alfonsi, A.: Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. *Stat. Probab. Lett.* **83**, 602–607 (2013)
5. Andersen, L., Andreasen, J.: Volatility skews and extensions of the LIBOR market model. *Appl. Math. Finance* **7**, 1–32 (2000)
6. Beskos, A., Roberts, G.O.: Exact simulation of diffusions. *Ann. Appl. Probab.* **15**, 2422–2444 (2005)
7. Blanchet, J., Zhang, F.: Exact simulation for multivariate Itô diffusions. *Adv. Appl. Probab.* **52**, 1003–1034 (2020)
8. Borodin, A.N., Salminen, P.: *Handbook of Brownian Motion—Facts and Formulae*, 2nd edn. Birkhäuser, Basel (2002)
9. Çetin, U.: Diffusion transformations, Black–Scholes equation and optimal stopping. *Ann. Appl. Probab.* **28**, 3102–3151 (2018)
10. Çetin, U., Danilova, A.: *Dynamic Markov Bridges and Market Microstructure: Theory and Applications*. Springer, Berlin (2018)
11. Chen, Z.-Q.: Gaugeability and conditional gaugeability. *Trans. Am. Math. Soc.* **354**, 4639–4679 (2002)
12. Chen, Z.-Q., Song, R.: General gauge and conditional gauge theorems. *Ann. Probab.* **30**, 1313–1339 (2002)
13. Cranston, M., Fabes, E., Zhao, Z.: Conditional gauge and potential theory for the Schrödinger operator. *Trans. Am. Math. Soc.* **307**, 171–194 (1988)
14. De Weert, F.: *Exotic Options Trading*. Wiley, New York (2011)
15. Dereich, S., Neuenkirch, A., Szpruch, L.: An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proc. R. Soc. A, Math. Phys. Eng. Sci.* **468**, 1105–1115 (2012)
16. Derman, E., Kani, I.: Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility. *Int. J. Theor. Appl. Finance* **1**, 61–110 (1998)
17. Dupire, B.: Pricing with a smile. *Risk* **7**, 18–20 (1994)
18. Evans, S.N., Hening, A.: Markov processes conditioned on their location at large exponential times. *Stoch. Process. Appl.* **129**, 1622–1658 (2019)
19. Gatheral, J.: *The Volatility Surface: A Practitioner's Guide*. Wiley, New York (2011)
20. Glasserman, P.: *Monte Carlo Methods in Financial Engineering*. Springer, Berlin (2013)

21. Gobet, E.: Weak approximation of killed diffusion using Euler schemes. *Stoch. Process. Appl.* **87**, 167–197 (2000)
22. Gobet, E.: Euler schemes and half-space approximation for the simulation of diffusion in a domain. *ESAIM Probab. Stat.* **5**, 261–297 (2001)
23. Haug, E.G.: *The Complete Guide to Option Pricing Formulas*. McGraw-Hill, New York (2007)
24. Hok, J., Kucherenko, S.: Pricing and risk analysis in hyperbolic local volatility model with quasi-Monte Carlo. *Wilmott Mag.* **113**, 62–69 (2021)
25. Hok, J., Ngare, P., Papapantoleon, A.: Expansion formulas for European quanto options in a local volatility FX-LIBOR model. *Int. J. Theor. Appl. Finance* **21**, 1850017 (2018)
26. Hok, J., Tan, S.-H.: Calibration of local volatility model with stochastic interest rates by efficient numerical PDE methods. *Decis. Econ. Finance* **42**, 609–637 (2019)
27. Hutzenthaler, M., Jentzen, A., Kloeden, P.E.: Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. A, Math. Phys. Eng. Sci.* **467**, 1563–1576 (2011)
28. Hutzenthaler, M., Jentzen, A., Kloeden, P.E.: Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22**, 1611–1641 (2012)
29. Imamura, Y., Ishigaki, Y., Okumura, T.: A numerical scheme based on semi-static hedging strategy. *Monte Carlo Methods Appl.* **20**, 223–235 (2014)
30. Itô, K., McKean, H.P. Jr.: *Diffusion Processes and Their Sample Paths*, 2nd edn. Springer, Berlin (1974)
31. Jäckel, P.: Quanto skew (2009). Available online at <http://www.jaeckel.org/QuantoSkew.pdf>
32. Ladyženskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N.: *Linear and Quasilinear Equations of Parabolic Type*. Translations of Mathematical Monographs, vol. 23. Am. Math. Soc., Providence (1968)
33. Mao, X., Szpruch, L.: Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Comput. Appl. Math.* **238**, 14–28 (2013)
34. Mikulevičius, R., Platen, E.: Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.* **151**, 233–239 (1991)
35. Musiela, M., Rutkowski, M.: *Martingale Methods in Financial Modelling*, 2nd edn. Springer, Berlin (2006)
36. Neuenkirch, A., Szpruch, L.: First order strong approximations of scalar SDEs defined in a domain. *Numer. Math.* **128**, 103–136 (2014)
37. Overhaus, M., Bermúdez, A., Buehler, H., Ferraris, A., Jordinson, C., Lamouar, A.: *Equity Hybrid Derivatives*. Wiley, New York (2007)
38. Ren, Y., Madan, D., Qian, M.: Calibrating and pricing with embedded local volatility models. *Risk* **20**, 138–143 (2007)
39. Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin (1999)
40. Romo, J.M.: The quanto adjustment and the smile. *J. Futures Mark.* **9**, 877–908 (2012)
41. Rubinstein, M.: Implied binomial trees. *J. Finance* **49**, 771–818 (1994)
42. Sharpe, M.: *General Theory of Markov Processes*. Academic Press, Boston (1988)
43. Siegmund, D., Yuh, Y.-S.: Brownian approximations to first passage probabilities. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **59**, 239–248 (1982)
44. Talay, D., Tubaro, L.: Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8**, 483–509 (1990)
45. Wilmott, P.: *Paul Wilmott on Quantitative Finance*, 2nd edn. Wiley, New York (2006)
46. Wilmott, P.: *Paul Wilmott Introduces Quantitative Finance*. Wiley, New York (2007)