

# Exact simulation of the first-passage time of diffusions

S. Herrmann<sup>1</sup> and C. Zucca<sup>2</sup>

<sup>1</sup>Institut de Mathématiques de Bourgogne (IMB) - UMR 5584, CNRS,  
Université de Bourgogne Franche-Comté, F-21000 Dijon, France  
Samuel.Herrmann@u-bourgogne.fr

<sup>2</sup>Department of Mathematics 'G. Peano',  
University of Torino, Via Carlo Alberto 10, 10123 Turin, Italy,  
cristina.zucca@unito.it

May 22, 2017

## Abstract

Since diffusion processes arise in so many different fields, efficient techniques for the simulation of sample paths, like discretization schemes, represent crucial tools in applied probability. Such methods permit to obtain approximations of the first-passage times as a by-product. For efficiency reasons, it is particularly challenging to simulate directly this hitting time by avoiding to construct the whole paths. In the Brownian case, the distribution of the first-passage time is explicitly known and can be easily used for simulation purposes. The authors introduce a new rejection sampling algorithm which permits to perform an exact simulation of the first-passage time for general one-dimensional diffusion processes. The efficiency of the method, which is essentially based on Girsanov's transformation, is described through theoretical results and numerical examples.

**Key words and phrases:** first-passage time, Brownian motion, diffusion processes, Girsanov's transformation, exact simulation, randomized algorithm.

**2010 AMS subject classifications:** primary 65C05 ; secondary: 65N75, 60G40.

## Introduction

First-passage times of diffusion processes appear in several fields, i.e. economics [21], mathematical finance [23, 28], neuroscience [9, 34], physics [1], psychology [29] and reliability theory [30]. In neuroscience particular diffusion processes model the evolution of the membrane potential and the first-passage times of the considered process through a suitable boundary describes the neuronal firing times (cf. [9, 34]). Likewise, in metrology or in quality control, diffusion processes are used to describe the error of some instrument which is required to be bounded. The first-passage time through a suitable boundary reproduces the

first time the error becomes out of control [40]. In this setting, rapid detection of anomalies corresponds to recognize the optimal stopping time of a diffusion process [41]. The importance of practical applications dealing with the first passage time motivated many mathematical studies on the subject.

To introduce the corresponding mathematical problem, let us consider the stochastic process  $(X_t, t \geq 0)$  the solution of the SDE:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x \in \mathbb{R},$$

where  $(B_t, t \geq 0)$  stands for the standard one-dimensional Brownian motion. We denote by  $\tau_L$  the first time the diffusion reaches the level  $L$ :

$$\tau_L := \inf\{t \geq 0 : X_t = L\}. \quad (0.1)$$

The study of the hitting times  $\tau_L$  and of its approximations for general diffusion processes is an active area of research.

Explicit expressions or even asymptotic solutions to boundary crossing problems are only available in rare situations which moreover do not correspond to interesting instances for applications. Consequently studies on this subject have been directed to the development of alternative methods like numerical or simulative ones.

As far as numerical methods are concerned, they are efficient but need specific algorithms for each function of the hitting time  $\tau_L$ . Most of the studies focus on the probability density function  $p$  defined by  $p(t)dt = \mathbb{P}(\tau_L \in dt)$  which satisfies Volterra-type integral equations. It suffices therefore to approximate certain integral. Examples of studies in this direction are Durbin [11, 12], Ferebee [13], Ricciardi et al. [33]; Giorno et al. [14], Sacerdote and Tomassetti [36]. The numerical approach proposed in [8] seems to be particularly efficient. This type of approach is also used in [3, 35] to deal with first passage times of a two-dimensional correlated diffusion process. For particular families of diffusion which can be constructed using simple transformations of a Brownian motion, Pötzelberger and Wang [37, 32, 38] proposed a different approach based on the explicit Brownian crossing probabilities.

The second important course of action uses simulations to determine the hitting times  $\tau_L$ . These methods often require a strong computational effort and are affected by a non negligible approximation error. Since stopping times for time discrete Markov processes can be quite easily simulated, a natural way to deal with diffusion processes is to introduce a discretization scheme for the corresponding stochastic differential equation. However, only an upper bound of the stopping time can be determined through this approximation. It is therefore important to improve the algorithm. In [7] and [19] a shift of the boundary is proposed to improve the approximation, while in [15, 16, 17] the proposed method compute on each small interval of the time discretization, the probability for the process to cross the boundary in this time window. Such a method needs precise asymptotics of hitting probabilities for pinned diffusions. Gobet [18] extended the study to a multidimensional diffusion and described the error for both the discrete and the continuous Euler schemes. Precise asymptotics for general pinned diffusions are pointed out by Baldi and Caramellino [2] permitting to deal with a larger class of diffusions.

Let us just notice that besides the simulated approximation methods there also exist simulated exact methods dealing with the sample of diffusion paths on a

fixed time interval, first introduced by Beskos & Roberts [6]. Several modifications of this algorithm have been proposed [4, 5, 25]. However, the two main approaches are not exclusive: recently Herrmann and Tanré [20] proposed a new simulation method which has no link with the two approaches just described. It is based on an iterative algorithm and leads to an efficient approximation of Brownian hitting times for curved boundaries, and, by extension, can deal with diffusion processes expressed as functionals of a Brownian motion.

This paper is a contribution in the framework of the simulated exact methods. Here we present an efficient method for simulating directly the first-passage time  $\tau_L$  without building the whole paths of the diffusion process and without any approximation error. The method can be summarized as follows: we need an algorithm producing a random variable  $Y$  such that  $Y$  and  $\tau_L$  are identically distributed. In particular, this approach permits to reduce the approximation error of  $\mathbb{E}[\psi(\tau_L)]$ , where  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  is any measurable and bounded function, to the only Monte-Carlo one. Our approach rests on two important facts: it is easy to simulate the first-passage time of the standard Brownian motion on one hand and Girsanov's transformation allows to link the Brownian paths to any diffusion paths on the other hand. Combining these features, we highlight the following acceptance-rejection algorithm (A0):

- Step 1. Simulate  $T_L$  the first hitting time of the Brownian motion
- Step 2. If the event  $A$  is true, then  $Y = T_L$  otherwise go to Step 1.

The first challenge is to find an appropriate event  $A$ . Subsequently we present alternatives in order to reduce the averaged number of steps of this acceptance/rejection algorithm. Our main idea originates in two different mathematical studies; let us now focus our attention on them.

- Ichiba & Kardaras [22] proposed an efficient estimation of the first-passage density using on one hand Girsanov's transformation and on the second hand a Monte Carlo procedure in order to estimate the Radon-Nikodym derivative which involves a Bessel bridge of dimension  $\delta = 3$ . This interesting approach permits to deal with a large family of diffusion processes but introduces a Monte-Carlo approximation error.
- The second pillar is the procedure of exact simulation introduced by Beskos & Roberts [6]. Using a judicious rejection sampling algorithm, they simulate in a surprisingly simple way a diffusion process on any finite time interval. They propose also to use their algorithm in order to simulate extremes and hitting times by building a skeleton of the whole trajectory.

The algorithm proposed here, based on these two pillars, outperform [6] when the objective is the simulation of  $\tau_L$  instead of the whole diffusion paths.

The material is organized as follows: after reminding the famous Girsanov transformation, we present in details the rejection sampling algorithm in Section 1 and prove that the outcome  $Y$  has the expected distribution. In Section 2, we focus our attention on the efficiency of the procedure: the number of steps is described and upper-bounds are computed. Unfortunately our procedure can not be applied to any diffusion processes since it requests different technical

assumptions. We discuss in Section 3 and Section 4 two modifications of the main algorithm in order to weaken these assumptions. Finally, for completeness, numerical examples permit to appreciate and illustrate the efficiency of the method (Section 5).

## 1 Rejection sampling method for first-passage time of diffusions

Let us first consider  $(X_t, t \geq 0)$  the solution of the following stochastic differential equation:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x \in \mathbb{R}, \quad (1.1)$$

where  $(B_t, t \geq 0)$  stands for the standard one-dimensional Brownian motion. We denote by  $\tau_L$  the first time the diffusion reaches the level  $L$  as defined in (0.1). It suffices to assume the existence of a unique weak solution to the equation, see for instance [26] for the corresponding conditions. In order to simplify the presentation of all algorithms, we restrict our study to the constant diffusion case:  $\sigma(x) \equiv 1$ . Using the *Lamperti transform*, we observe that this restriction is not sharp at all. Indeed if  $X$  is solution to (1.1) then by fixing  $x \in \mathbb{R}$ , we define

$$Y_t = \eta(X_t) := \int_x^{X_t} \frac{1}{\sigma(u)} du$$

which satisfies therefore the equation:  $dY_t = b_0(Y_t) dt + dB_t$  with

$$b_0(x) = \frac{b \circ \eta^{-1}(x)}{\sigma \circ \eta^{-1}(x)} - \frac{1}{2} \sigma' \circ \eta^{-1}(x).$$

That is why, from now on,  $X$  stands for the unique solution of

$$dX_t = b(X_t)dt + dB_t, \quad t \geq 0, \quad X_0 = x. \quad (1.2)$$

and the level to reach satisfies  $L > x$ . Let us note that, for the particular Brownian case:  $b(x) \equiv 0$ , the first-passage time is extremely simple to simulate since  $\tau_L$  and  $(L - x)^2/G^2$  have the same distribution where  $G$  is a standard gaussian random variable. The aim is to use this simple form when considering the general diffusion case and the important tool for such a strategy is Girsanov's formula. We first recall it before introducing the main rejection algorithm.

### 1.1 Girsanov's transformation

Let us first introduce 3 important expressions related to the drift term  $b$  :

$$\beta(x) = \int_0^x b(y)dy, \quad p(x) = \int_0^x e^{-\beta(y)} dy \quad \text{and} \quad \gamma := \frac{b^2 + b'}{2}. \quad (1.3)$$

**Assumption 1.1.** *The drift term  $b \in \mathcal{C}^1(]-\infty, L])$ .*

Girsanov's transformation permits to link the diffusion process  $X$  solution of (1.2) to the Brownian motion (a slight modification of Proposition 2.1 in [22]).

**Proposition 1.1.** *Under Assumption 1.1, for any bounded measurable function  $\psi : \mathbb{R} \rightarrow \mathbb{R}$ , we obtain*

$$\mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L < \infty\}}] = \mathbb{E}_{\mathbb{Q}}[\psi(\tau_L)\eta(\tau_L)] \exp\left\{\beta(L) - \beta(x)\right\}, \quad (1.4)$$

where  $\mathbb{P}$  (resp.  $\mathbb{Q}$ ) corresponds to the distribution of the diffusion  $X$  (resp. the Brownian motion  $B$ ) and

$$\eta(t) := \mathbb{E}\left[\exp - \int_0^t \gamma(L - R_s)ds \mid R_t = L - x\right]. \quad (1.5)$$

Here  $(R_t, t \geq 0)$  stands for a 3-dimensional Bessel process with  $R_0 = 0$ .

*Proof.* Let us denote  $\zeta$  the explosion time of the diffusion  $X$  defined by (1.2). The Feller test for explosion (see for instance Proposition 5.22 in [26]) claims that Assumption 1.1 leads to  $\tau_L < \zeta$ ,  $\mathbb{P}$ -almost surely. By applying successively Girsanov's transformation (Proposition 1.7.5.4 in [24]) and Itô's formula, we obtain

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L < \infty\}}] &= \mathbb{E}_{\mathbb{Q}}\left[\psi(\tau_L) \exp\left(\int_0^{\tau_L} b(B_s)dB_s - \frac{1}{2} \int_0^{\tau_L} b^2(B_s)ds\right)\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\psi(\tau_L) \exp\left(\beta(B_{\tau_L}) - \beta(x) - \int_0^{\tau_L} \gamma(B_s)ds\right)\right]. \end{aligned}$$

Under  $\mathbb{Q}$ ,  $(B_t)$  corresponds to a one-dimensional Brownian motion with starting point  $x$  and  $\tau_L = \inf\{t \geq 0 : B_t = L\} < \infty$  a.s. (recurrence of the Brownian paths). Hence  $B_{\tau_L} = L$  and consequently

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L < \infty\}}] &= e^{\beta(L) - \beta(x)} \mathbb{E}_{\mathbb{Q}}\left[\psi(\tau_L) \exp - \int_0^{\tau_L} \gamma(B_s)ds\right] \\ &= e^{\beta(L) - \beta(x)} \mathbb{E}_{\mathbb{Q}}\left[\psi(\tau_L)\eta(\tau_L)\right], \end{aligned}$$

where

$$\eta(t) := \mathbb{E}_{\mathbb{Q}}\left[\exp - \int_0^t \gamma(B_s)ds \mid \tau_L = t\right].$$

In order to conclude it suffices to note that, given  $\{\tau_L = t\}$ ,  $R_s := L - B_{t-s}$  is a Bessel bridge [39] for  $s \in [0, t]$ .  $\square$

## 1.2 Description of the algorithm

The rejection sampling algorithm for the simulation of  $\tau_L$  is based on Proposition 1.1. We need to introduce the following

**Assumption 1.2.** *The function  $\gamma$  defined by (1.3) is non-negative and the first-passage time satisfies  $\tau_L < \infty$  a.s.*

These main conditions permit the function  $\eta$  defined by (1.5) to belong to the interval  $[0, 1]$  and therefore to represent a probability of rejection in the procedure.

**Remark 1.2.** Let us just note that  $\tau_L < \infty$  a.s. if the drift term satisfies one of the additional conditions: either  $\lim_{x \rightarrow -\infty} p(x) = -\infty$  (see (1.3) for the definition of  $p$ ) or  $\lim_{x \rightarrow -\infty} v(x) < \infty$  where

$$v(x) = \int_0^x \int_0^y \frac{2p'(y)}{p'(z)} dz dy, \quad (1.6)$$

according to the Feller test for explosion (Proposition 5.5.32 in [26]).

Let us now focus our attention to a first theoretical description of the acceptance-rejection algorithm.

**ALGORITHM (A0).**

Let us fix the level  $L > x$ .

Step 1: Simulate a non-negative random variable  $T$  with p.d.f.  $f_T$ .

Step 2: Simulate a 3-dimensional Bessel process  $(R_t)$  on the time interval  $[0, T]$  with endpoint  $R_T = L - x$ . We define by  $D_{R,T}$  the stochastic domain:

$$D_{R,T} := \left\{ (t, v) \in [0, T] \times \mathbb{R}_+ : v \leq \gamma(L - R_t) \right\}.$$

This domain depends on both random elements  $(R_t)$  and  $T$ .

Step 3: Simulate a Poisson point process  $N$  on the state space  $[0, T] \times \mathbb{R}_+$ , independent of the Bessel process, whose intensity measure is the Lebesgue one.

Step 4: If  $N(D_{R,T}) = 0$  then set  $Y = T$  otherwise go to Step 1.

Outcome: the random variable  $Y$ .

We shall discuss later on, how this algorithm can be applied in practice.

**Theorem 1.3.** If  $\gamma$  is a non-negative function then the p.d.f.  $f_Y$  of the outcome variable  $Y$  satisfies

$$f_Y(t) = \frac{1}{\Xi} \eta(t) f_T(t), \quad (1.7)$$

where  $\eta$  is given by (1.5) and  $\Xi$  stands for the normalization coefficient

$$\Xi := \int_0^\infty \eta(t) f_T(t) dt.$$

In particular, under Assumption 1.1 & 1.2, if  $T/(L-x)^2$  has the same distribution as  $1/G^2$  where  $G$  is a standard Gaussian r.v., then  $Y$  and  $\tau_L$ , defined by (0.1) & (1.2), are identically distributed.

Let us observe that in the particular case  $T \sim \frac{(L-x)^2}{G^2}$ , (1.4) leads to the following identify:

$$\Xi = \exp - \left\{ \beta(L) - \beta(x) \right\}.$$

*Proof.* The arguments presented here are quite classical for a rejection sampling method. We introduce  $\mathcal{I}$  the number of iterations in Algorithm (A0). For each

iteration  $\mathcal{I} = i$ , we denote by  $T^{(i)}$ ,  $R_t^{(i)}$  and  $N^{(i)}$  the generated random variables or processes. Of course each iteration corresponds to independent variables. Let  $\psi$  be a measurable bounded function. We get

$$\mathbb{E}[\psi(Y)] = \sum_{i=1}^{\infty} \mathbb{E}[\psi(Y)1_{\{\mathcal{I}=i\}}] = \sum_{i=1}^{\infty} \mathbb{E}\left[\psi(T^{(i)})1_{E_1 \cap \dots \cap E_{i-1} \cap \bar{E}_i}\right],$$

where  $E_i := \left\{N^{(i)}(D_{R^{(i)}, T^{(i)}}) \neq 0\right\}$  and  $\bar{E}_i = \Omega \setminus E_i$ . Hence, using the independence property, we obtain

$$\begin{aligned} \mathbb{E}[\psi(Y)] &= \sum_{i=1}^{\infty} \mathbb{E}\left[\psi(T^{(i)})1_{\bar{E}_i}\right] \mathbb{P}(E_1) \dots \mathbb{P}(E_{i-1}) \\ &= \sum_{i=1}^{\infty} \mathbb{E}\left[\psi(T^{(1)})1_{\bar{E}_1}\right] \mathbb{P}(E_1)^{i-1} = \mathbb{E}\left[\psi(T^{(1)})1_{\bar{E}_1}\right] (1 - \mathbb{P}(E_1))^{-1}. \end{aligned}$$

From now on, we delete the index for notational simplicity. Therefore

$$\mathbb{E}[\psi(Y)] = \frac{1}{\mathbb{P}(N(D_{R,T}) = 0)} \mathbb{E}\left[\psi(T)\mathbb{P}(N(D_{R,T}) = 0|T)\right].$$

The distribution of the Poisson point process leads to:

$$\begin{aligned} \mathbb{P}(N(D_{R,T}) = 0|T) &= \mathbb{E}\left[\mathbb{P}(N(D_{R,T}) = 0|R, T)|T\right] = \mathbb{E}\left[\exp -\lambda(D_{R,T})|T\right] \\ &= \mathbb{E}\left[\exp -\int_0^T \gamma(L - R_t)dt \middle| R_T = L - x, T\right] = \eta(T). \end{aligned}$$

We deduce

$$\mathbb{E}[\psi(Y)] = \frac{1}{\mathbb{P}(N(D_{R,T}) = 0)} \mathbb{E}[\psi(T)\eta(T)].$$

For  $\psi \equiv 1$ , we obtain  $\mathbb{P}(N(D_{R,T}) = 0) = \mathbb{E}[\eta(T)]$  which implies the statement of the proposition.  $\square$

Algorithm (A0) seems quite difficult to achieve in practice since we do not know how to simulate a Poisson point process on the state space  $[0, T] \times \mathbb{R}_+$  on one hand and since the domain  $D_{R,T}$  depends on the whole trajectory of  $(R_t)$  on the other hand. In order to overcome the first difficulty, we reduce the domain to a bounded one introducing the following assumption:

**Assumption 1.3.** *there exists a constant  $\kappa > 0$  such that*

$$0 \leq \gamma(x) \leq \kappa, \quad \text{for all } x \in ]-\infty, L]. \quad (1.8)$$

It suffices therefore to simulate the Poisson point process on the reduced space  $[0, T] \times [0, \kappa]$ . The second difficulty is not insurmountable. In fact we don't actually need to precisely describe the entire domain  $D_{R,T}$ : we only need to know if the points of the Poisson process belong to it. This can fortunately be obtained in a quite simple way, that's why we propose two modifications of the theoretical Algorithm (A0) – Algorithm (A1) and (A2) – which are easy to implement. These modifications essentially concern the way used to simulate the Poisson point process.

ALGORITHM (A1).

Let us fix the level  $L > x$ .

Step 1: Simulate a non-negative random variable  $T$  with p.d.f.  $f_T$ .

Initialization:  $\beta = (0, 0, 0)$ ,  $\mathcal{W} = 0$ ,  $\mathcal{E}_0 = 0$  and  $\mathcal{E}_1$  an exponentially distributed random variable of average  $1/\kappa$ .

Step 2: While  $\mathcal{E}_1 \leq T$  and  $\mathcal{W} = 0$ ,

Step 2.1: Simulate three independent random variables: a 3-dimensional gaussian random variable  $G$ , an exponentially distributed variable  $e$  with average  $1/\kappa$  and a standard uniform random variable  $V$ .

Step 2.2: Compute

$$\beta \leftarrow \frac{T - \mathcal{E}_1}{T - \mathcal{E}_0} \beta + \sqrt{\frac{(T - \mathcal{E}_1)(\mathcal{E}_1 - \mathcal{E}_0)}{T - \mathcal{E}_0}} G.$$

Change the value of  $\mathcal{W}$ ,  $\mathcal{E}_0$  and  $\mathcal{E}_1$  in the following way:

if  $\kappa V \leq \gamma \left( L - \|\mathcal{E}_1(L - x)(1, 0, 0)/T + \beta\| \right)$  then  $\mathcal{W} = 1$  else  $\mathcal{W} = 0$ ,

$$\mathcal{E}_0 \leftarrow \mathcal{E}_1 \text{ and } \mathcal{E}_1 \leftarrow \mathcal{E}_1 + e.$$

Step 3: If  $\mathcal{W} = 0$  then set  $Y = T$  otherwise go to Step 1.

Outcome: the r.v.  $Y$ .

In order to present the second modified algorithm, we introduce different notations. Let us consider a finite set  $\mathcal{S} \subset \mathbb{R} \times \mathbb{R}^3$  with the property: if both  $(t, x)$  and  $(t, y)$  belong to  $\mathcal{S}$  then  $x = y$ . Let us moreover define  $\underline{\mathcal{S}} = \inf\{t : (t, x) \in \mathcal{S}\}$  and  $\overline{\mathcal{S}} = \sup\{t : (t, x) \in \mathcal{S}\}$ . For any  $t \in [\underline{\mathcal{S}}, \overline{\mathcal{S}}]$  we introduce  $\overleftarrow{t} := (\overleftarrow{t}_1, \overleftarrow{t}_2)$  corresponding to the couple  $(s, x) \in \mathcal{S}$  such that  $\overleftarrow{t}_1 = \sup\{s : (s, x) \in \mathcal{S}, s \leq t\}$ . Similarly we define  $\overrightarrow{t} := (\overrightarrow{t}_1, \overrightarrow{t}_2)$  the couple  $(s, x) \in \mathcal{S}$  such that  $\overrightarrow{t}_1 = \inf\{s : (s, x) \in \mathcal{S}, s > t\}$ .

ALGORITHM (A2).

Let us fix the level  $L > x$ .

Step 1: Initialization:  $\mathcal{E} = 0$  and  $\mathcal{W} = 0$ . Simulate a non-negative random variable  $T$  with p.d.f.  $f_T$ . Set  $\mathcal{S} = \{(0, (0, 0, 0)), (T, (0, 0, 0))\}$ .

Step 2: While  $\mathcal{E} \leq \kappa$  and  $\mathcal{W} = 0$ ,

Step 2.1: Simulate two independent random variables: an exponentially distributed variable  $e$  with average  $1/T$  and a uniformly distributed variable  $U$  on the interval  $[0, T]$ . We change the value of  $\mathcal{E}$ :  $\mathcal{E} \leftarrow \mathcal{E} + e$ .

Step 2.2: If  $\mathcal{E} \leq \kappa$  then we simulate a 3-dimensional gaussian random variable  $G$  with covariance  $\text{Id}_3$  and we define:

$$\beta = \overleftarrow{U}_2 + \frac{\overrightarrow{U}_2 - \overleftarrow{U}_2}{\overrightarrow{U}_1 - \overleftarrow{U}_1} (U - \overleftarrow{U}_1) + \sqrt{\frac{(\overrightarrow{U}_1 - U)(U - \overleftarrow{U}_1)}{\overrightarrow{U}_1 - \overleftarrow{U}_1}} G.$$



We increase the set  $\mathcal{S}$  in the following way:  $\mathcal{S} \leftarrow \mathcal{S} \cup \{(U, \beta)\}$  and achieve the test:

if  $\mathcal{E} \leq \gamma\left(L - \|U(L-x)(1,0,0)/T + \beta\|\right)$  then  $\mathcal{W} = 1$  else  $\mathcal{W} = 0$ .

Step 3: If  $\mathcal{W} = 0$  then  $Y = T$  otherwise return to Step 1.

Outcome:  $Y$ .

**Proposition 1.4.** Under Assumption 1.1–1.3, if  $T/(L-x)^2$  has the same distribution as  $1/G^2$  where  $G$  is a standard Gaussian r.v., then  $Y$ , the outcome of Algorithm (A1) (respectively Algorithm (A2)) and  $\tau_L$ , defined by (0.1) & (1.2), are identically distributed.

*Proof.* Algorithms (A1) and (A2) are just two different modifications of Algorithm (A0).

*Step 1.* Let us first present a quite different expression for  $\eta(t)$  defined by (1.5) (these arguments were already developed in [22]). Let us define

$$\mathcal{Z}(y) := \mathbb{E}\left[F(R_u, 0 \leq u \leq t) \mid R_t = y\right], \quad y \geq 0, \quad (1.9)$$

for any non-negative functional  $F$ . Since the Bessel process  $(R_t, t \geq 0)$  has the same distribution as the norm of the 3-dimensional Brownian motion  $(B_t, t \geq 0)$ , we obtain:

$$\mathcal{Z}(y) := \mathbb{E}\left[F(\|B_u\|, 0 \leq u \leq t) \mid \|B_t\| = y\right] = \mathbb{E}\left[G(B_t) \mid \|B_t\| = y\right],$$

where

$$G(x) := \mathbb{E}\left[F(\|B_u\|, 0 \leq u \leq t) \mid B_t = x\right], \quad x \in \mathbb{R}^3.$$

The Brownian path can be decomposed as follows:

$$B_u = \frac{u}{t} B_t + \sqrt{t} \beta_{u/t}, \quad 0 \leq u \leq t.$$

Here  $(\beta_s, 0 \leq s \leq 1)$  stands for a standard 3-dimensional Brownian bridge with  $\beta_0 = \beta_1 = 0$ . We obtain therefore a new expression for the functional  $G$ :

$$G(x) = \mathbb{E}\left[F\left(\left\|\frac{u}{t}x + \sqrt{t}\beta_{u/t}\right\|, 0 \leq u \leq t\right)\right]$$

Since the Brownian bridge is rotationally invariant, we observe  $G(x) = G(\|x\|e_1)$  where  $e_1 = (1, 0, 0)$ . We deduce

$$\mathcal{Z}(y) = \mathbb{E}\left[F\left(\left\|\frac{u}{t}y e_1 + \sqrt{t}\beta_{u/t}\right\|, 0 \leq u \leq t\right)\right], \quad (1.10)$$

and in particular:

$$\eta(t) = \mathbb{E}\left[\exp - \int_0^t \gamma\left(L - \left\|\frac{u}{t}(L-x)e_1 + \sqrt{t}\beta_{u/t}\right\|\right) du\right].$$

*Step 2.* Using the new expression of the weight  $\eta(t)$ , we propose the same kind of rejection algorithm as (A1). It suffices therefore to consider the domain:

$$D_{\beta,T} = \left\{ (t, y) \in [0, T] \times \mathbb{R}_+ : y \leq \gamma \left( L - \frac{t}{T} (L - x) e_1 + \sqrt{T} \beta_{t/T} \right) \right\} \quad (1.11)$$

and the associated event  $N(D_{\beta,T}) = 0$  where  $N$  is a Poisson point process on the state space  $[0, T] \times [0, \kappa]$ . As usual, there exist three methods of simulation for the Poisson process.

1. One method consists in first sampling the number of points in the rectangle using the Poisson distribution and secondly in choosing their position uniformly on the rectangle. An algorithm based on such a procedure would be time consuming and requires a huge memory space. It is not reasonable in practice.
2. The second method consists in the simulation of the point process in the infinite domain  $[0, T] \times \mathbb{R}_+$ : we simulate a sequence of independent couples  $(e_k, U_k)_{k \geq 1}$ ,  $e_k$  and  $U_k$  being independent random variables.  $e_k$  is exponentially distributed with average  $1/T$  and  $U_k$  is uniformly distributed on the interval  $[0, T]$  for any  $k \geq 1$ . Let us define:

$$\eta := \inf \left\{ n \geq 1 : \sum_{k=1}^n e_k > \kappa \right\}.$$

The point process therefore corresponds to the couples  $(e_1, U_1), \dots, (e_1 + e_2 + \dots + e_{\eta-1}, U_{\eta-1})$  belonging to the domain  $[0, T] \times [0, \kappa]$ . Let us note that  $\eta = 1$  corresponds to the case where no point of the Poisson process belongs to the domain. Algorithm (A2) is based on this simulation.

3. A third method consists in using similar arguments as those just described but we interchange the space and the time variables. That means that a cumulated sum of exponentially distributed variables with average  $1/\kappa$  until it overcomes the level  $T$  are used to describe the time variable and uniform variables on the interval  $[0, \kappa]$  for the space variable: we obtain Algorithm (A1).

Once the Poisson process is realized on the domain  $[0, T] \times [0, \kappa]$ , it suffices to test each point in order to observe if it belongs or not to the domain  $D_{\beta,T}$  defined by (1.11). For this test we use the classical Brownian bridge simulation as described in the fourth step of Algorithm (A1) or Step 2.2 in Algorithm (A2).  $\square$

## 2 Efficiency of the algorithms

Let us denote by  $\mathcal{I}$  the number of iterations observed in order to simulate the hitting time  $\tau_L$  defined by (0.1) & (1.2) and  $\mathcal{N}_1, \dots, \mathcal{N}_{\mathcal{I}}$  the numbers of random points (Poisson process) used for each iteration. Of course the efficiency is directly linked to

$$\mathcal{N}_{\Sigma} = \mathcal{I} + \mathcal{N}_1 + \dots + \mathcal{N}_{\mathcal{I}}, \quad (2.1)$$

the total number of random variables. It is therefore quite challenging to modify the algorithms in order to reduce it. Focusing on the efficiency, we first describe

the number  $\mathcal{I}$ , secondly we propose some modification in order to reduce it and finally present upper-bound for the total number of random variables  $\mathcal{N}_\Sigma$ .

## 2.1 Number of iterations

Let us first note that the distribution of the number of iterations  $\mathcal{I}$  (number of random variates  $T$  produced in order to obtain the outcome  $Y$ ) does not depend on the choice of the algorithm (A0), (A1) or (A2). The variable  $\mathcal{I}$  is exponentially large as soon as  $L - x$  becomes large:

**Proposition 2.1.** *Under Assumptions 1.1–1.3, the following upper-bound holds:*

$$\mathbb{E}[\mathcal{I}] = \exp\{\beta(L) - \beta(x)\} \leq \exp((L - x)\sqrt{2\kappa}). \quad (2.2)$$

*Proof.* Let us just note that  $\mathcal{I}$  is geometrically distributed with

$$\mathbb{P}(\mathcal{I} = 1) = \mathbb{P}(N(D_{R,T}) = 0) = \mathbb{E}_x[\eta(\tau_L)].$$

Since  $\tau_L < \infty$  a.s., (1.4) leads to  $\mathbb{E}[\mathcal{I}] = \mathbb{E}[\eta(\tau_L)]^{-1} = \exp\{\beta(L) - \beta(x)\}$ . Moreover, due to the condition  $\gamma \leq \kappa$ , we obtain

$$\mathbb{E}_x[\eta(\tau_L)] \geq \mathbb{E}_x[\exp -\kappa\tau_L] = \mathbb{E}\left[\exp -\frac{\kappa(L-x)^2}{G^2}\right] = \exp\left\{-\frac{\kappa(L-x)^2}{G^2}\right\},$$

where  $G$  is a standard gaussian r.v. □

In the simulation framework, such an exponential large number of iterations seems crippling in practice for large values of  $L$  and stimulates one to find interesting modifications of Algorithms (A1)–(A2).

### Space splitting

In order to reduce the number of iterations, we could replace the simulation of  $\tau_L$  for a diffusion starting in  $x$  (denoted here by  $\tau(x \rightarrow L)$ ) by the simulation of  $k$  independent first-passage times  $\tau(x + (i-1)(L-x)/k \rightarrow x + i(L-x)/k)$  for  $i = 1, \dots, k$ . It suffices then to take the sum of all these quantities. The global number of iterations for this new algorithm, called Algorithm (A1)<sub>split</sub> or (A2)<sub>split</sub>, becomes:

$$\begin{aligned} \mathbb{E}[\mathcal{I}_{\text{split}}] &= \sum_{i=1}^k \exp\left\{\beta\left(x + i(L-x)/k\right) - \beta\left(x + (i-1)(L-x)/k\right)\right\} \\ &\leq k \exp\left(\frac{(L-x)}{k}\sqrt{2\kappa}\right) \leq (\lfloor(L-x)\sqrt{2\kappa}\rfloor + 1)e. \end{aligned}$$

The previous inequality is related to the particular choice  $k = (\lfloor(L-x)\sqrt{2\kappa}\rfloor + 1)$ . Such a procedure permits to replace an exponentially large number of iterations into a linear one, as  $L$  becomes large.

### Shifting the function $\gamma$

An other way to reduce  $\mathcal{I}$  is to decrease the parameter  $\kappa$ . Such a trick is possible in the following situation.

**Assumption 2.1.** *There exist two constants  $\kappa > \gamma_0 > 0$  such that*

$$\kappa \geq \gamma(x) \geq \gamma_0, \quad \forall x \in ]-\infty, L].$$

The modification is based on the following observation: (1.4) can be replaced by

$$\mathbb{E}_{\mathbb{P}}[\psi(\tau_L)1_{\{\tau_L < \infty\}}] = \mathbb{E}_{\mathbb{Q}}[\psi(\tau_L)\eta_0(\tau_L)e^{-\gamma_0\tau_L}] \exp\left\{\beta(L) - \beta(x)\right\}, \quad (2.3)$$

$\eta_0(t)$  being defined in a similar way as  $\eta(t)$  in (1.5) just by replacing  $\gamma(t)$  by the shifted function  $\gamma(t) - \gamma_0$ . In other words, the upper-bound parameter  $\kappa$  can be replaced by the shifted value  $\kappa - \gamma_0$  and consequently the number of iterations is reduced.

**Proposition 2.2.** *Let Algorithm (A1)<sub>shift</sub>, respectively (A2)<sub>shift</sub>, defined in a similar way as Algorithm (A1), resp. (A2), just by replacing the function  $\gamma(\cdot)$  by the shifted function  $\gamma(\cdot) - \gamma_0$ . Under Assumption 1.1, 1.2 and 2.1, if  $T$  is an inverse gaussian random variable  $\text{IG}(\frac{L-x}{\sqrt{2\gamma_0}}, (L-x)^2)$  then the outcome of the modified algorithm  $Y$  and  $\tau_L$  defined by (0.1) & (1.2) are identically distributed. Moreover this modification reduces the number of iterations:*

$$\mathbb{E}[\mathcal{I}_{\text{shift}}] = \mathbb{E}[\mathcal{I}]e^{-(L-x)\sqrt{2\gamma_0}} \quad \text{for } x < L. \quad (2.4)$$

*Proof.* Let us consider a non-negative bounded function  $f$  and introduce

$$E_f := \mathbb{E}_x \left[ f(\tau_L) \exp -\gamma_0\tau_L \right],$$

where  $\tau_L$  is the Brownian first-passage time, that is  $\tau_L$  and  $(L-x)^2/G^2$  are identically distributed (here  $G$  stands for a standard gaussian r.v.). Hence

$$\begin{aligned} E_f &= \int_0^\infty f(t) \frac{L-x}{\sqrt{2\pi t^3}} \exp -\left\{ \frac{(L-x)^2}{2t} + \gamma_0 t \right\} dt \\ &= e^{-\sqrt{2\gamma_0}(L-x)} \int_0^\infty f(t) \frac{L-x}{\sqrt{2\pi t^3}} \exp -\left\{ \frac{(L-x - \sqrt{2\gamma_0}t)^2}{2t} \right\} dt \end{aligned}$$

We deduce that  $E_f/E_1 = \mathbb{E}[f(T)]$  where  $T \sim \text{IG}(\frac{L-x}{\sqrt{2\gamma_0}}, (L-x)^2)$ . Let us denote by  $f_T$  the *p.d.f.* of the variable  $T$ . By Theorem 1.3, we get

$$\begin{aligned} \mathbb{E}[\psi(Y)] &= \left( \int_0^\infty \psi(t)\eta_0(t)f_T(t) dt \right) \left( \int_0^\infty \eta_0(t)f_T(t) dt \right)^{-1} \\ &= \frac{E_{\psi\eta_0}}{E_1} \left( \frac{E_{\eta_0}}{E_1} \right)^{-1} = \frac{E_{\psi\eta_0}}{\mathbb{E}_x[\eta(\tau_L)]}. \end{aligned}$$

Let us recall that  $\mathbb{E}_x[\eta(\tau_L)] = \exp -\{\beta(L) - \beta(x)\}$ . Then (2.3) permits to point out the required distribution. Let us just notice that using the inverse gaussian variable increases the conditional probability of acceptance:

$$\mathbb{P}(\mathcal{I}_{\text{mod}} = 1|T) = \mathbb{P}(N(D_{R,T}) = 0|T) = \eta_0(T) > \eta(T).$$

Consequently the probability of acceptance/rejection at each step changes:  $\mathcal{I}_{\text{mod}}$  is geometrically distributed and

$$\begin{aligned} \mathbb{P}(\mathcal{I}_{\text{shift}} = 1) &= \mathbb{P}(N(D_{R,T}) = 0) = \mathbb{E}[\eta_0(T)] = \frac{\mathbb{E}[\eta_0(\tau_L)e^{-\gamma_0\tau_L}]}{\mathbb{E}[e^{-\gamma_0\tau_L}]} \\ &= \mathbb{E}[\eta(\tau_L)] \exp\left\{(L-x)\sqrt{2\gamma_0}\right\} = \exp -\left\{\beta(L) - \beta(x) - (L-x)\sqrt{2\gamma_0}\right\}. \end{aligned}$$

□

**Remark 2.3.** Using a many-to-one transformation, Michael, Schucany and Haas introduced a simple generator of inverse gaussian distributions (see, for instance, [10] p.149). We can apply the following procedure to simulate  $T \sim \text{IG}(\mu, \lambda)$ : let  $N$  be a standard gaussian r.v. and  $U$  uniformly distributed on  $[0, 1]$ , independent of  $N$ .

1. Set  $X = \mu + \frac{\mu^2 N^2}{2\lambda} - \frac{\mu}{2\lambda} \sqrt{4\mu\lambda N^2 + \mu^2 N^4}$ .
2. If  $U \leq \frac{\mu}{\mu+X}$  then  $T = X$  else  $T = \frac{\mu^2}{X}$ .

## 2.2 Efficiency of Algorithm (A1)

In this section, we are looking for an upper-bound of the number of random variates  $\mathcal{N}_\Sigma$  used in order to simulate one variate  $\tau_L$ . First we shall focus on the first iteration in Algorithm (A1) and secondly adapt the procedure to the whole number of variates.

**Proposition 2.4.** Let  $\gamma$  satisfy Assumptions 1.1–1.3. We assume moreover the existence of  $C_\gamma > 0$  and  $r < 1$  such that

$$\inf_{y \leq z \leq L} \gamma(z) \geq C_\gamma |y|^{-r}, \quad \text{for all } y \leq -1. \quad (2.5)$$

Then there exist two constants  $M_{\gamma,1} > 0$  and  $M_{\gamma,2} > 0$  such that the number of random points in the first iteration of Algorithm (A1), satisfies

$$\mathbb{E}[\mathcal{N}_1] \leq M_{\gamma,1} + \kappa M_{\gamma,2}(x^2 + (L-x)^{(1+r)/2}), \quad \text{for } x < L.$$

*Proof.* Let us denote by  $\mathbb{P}_c$  (resp.  $\mathbb{E}_c$ ) the conditional probability measure (resp. expectation) given both  $\tau_L = T$  and the Bessel bridge trajectory  $(R_t, 0 \leq t \leq T)$ . We observe that

$$\mathbb{P}_c(\mathcal{N}_1 = 1) = \mathbb{P}_c(\epsilon_1 > T) + \mathbb{P}_c(U_1 < \gamma(L - R_{\epsilon_1}), \epsilon_1 \leq T),$$

where  $(U_k)_k$  is a sequence of uniformly distributed r.v. on  $[0, \kappa]$  and  $(\epsilon_k)_k$  is a sequence of exponentially distributed r.v. with average  $\frac{1}{\kappa}$ , both sequences being independent. We denote  $\mathcal{E}_k = \epsilon_1 + \dots + \epsilon_k$ . Hence

$$\mathbb{P}_c(\mathcal{N}_1 = 1) = e^{-\kappa T} + \mathbb{E}_c \left[ \frac{\gamma}{\kappa} (L - R_{\epsilon_1}) 1_{\{\epsilon_1 \leq T\}} \right] = e^{-\kappa T} + \int_0^T \gamma(L - R_t) e^{-\kappa t} dt.$$

By similar computations, for  $k \geq 1$ ,

$$\begin{aligned} \mathbb{P}_c(\mathcal{N}_1 > k) &= \mathbb{P}_c(U_1 \geq \gamma(L - R_{\epsilon_1}), \dots, U_k \geq \gamma(L - R_{\mathcal{E}_k}), \mathcal{E}_k \leq T) \\ &= \frac{1}{\kappa^k} \mathbb{E}_c \left[ \prod_{i=1}^k (\kappa - \gamma(L - R_{\mathcal{E}_i})) 1_{\{\mathcal{E}_k \leq T\}} \right] \\ &= \frac{1}{\kappa^k} \mathbb{E}_c \left[ \prod_{i=1}^{k-1} (\kappa - \gamma(L - R_{\mathcal{E}_k U^{(i)}})) (\kappa - \gamma(L - R_{\mathcal{E}_k})) 1_{\{\mathcal{E}_k \leq T\}} \right], \end{aligned}$$

where  $(U^{(1)}, \dots, U^{(k)})$  stands for a set of reordered uniform random variables and  $\mathcal{E}_k$  is gamma distributed  $\gamma(k, \kappa)$ . Hence

$$\begin{aligned} \mathbb{P}_c(\mathcal{N}_1 > k) &= \mathbb{E} \left[ \left( 1 - \int_0^1 \frac{\gamma}{\kappa}(L - R_{\xi_s}) ds \right)^{k-1} \left( 1 - \frac{\gamma}{\kappa}(L - R_{\xi}) \right) 1_{\{\xi \leq T\}} \right] \\ &= \int_0^T \left( 1 - \int_0^1 \frac{\gamma}{\kappa}(L - R_{us}) ds \right)^{k-1} \left( 1 - \frac{\gamma}{\kappa}(L - R_u) \right) \frac{u^{k-1} \kappa^k}{(k-1)!} e^{-\kappa u} du \end{aligned}$$

Consequently, using a change of variable, we obtain

$$\begin{aligned} \mathbb{E}_c[\mathcal{N}_1] &= \sum_{k=0}^{\infty} \mathbb{P}_c(\mathcal{N}_1 > k) \\ &= 1 + \int_0^T \left( \kappa - \gamma(L - R_u) \right) \exp \left\{ -u \int_0^1 \gamma(L - R_{us}) ds \right\} du \\ &= 1 + \int_0^T \left( \kappa - \gamma(L - R_u) \right) \exp \left\{ - \int_0^u \gamma(L - R_w) dw \right\} du. \quad (2.6) \end{aligned}$$

By integration by parts, (2.6) becomes  $\mathbb{E}_c[\mathcal{N}_1] = H_T + \kappa I_T$  where

$$H_T := e^{-\int_0^T \gamma(L - R_w) dw}, \quad I_T := \int_0^T \exp \left\{ - \int_0^u \gamma(L - R_w) dw \right\} du. \quad (2.7)$$

Let us first study the expression  $I_T$ . Using the scaling property of the Bessel bridge, we have

$$I_T = T \int_0^1 \exp \left\{ -T \int_0^u \gamma(L - \sqrt{T} \hat{R}_w^T) dw \right\} du,$$

where  $(\hat{R}_w^T, 0 \leq w \leq 1)$  is a 3-dimensional Bessel bridge with condition  $\hat{R}_1^T = (L - x)/\sqrt{T}$ . Let us observe that

$$\hat{R}_w \leq_{st} \frac{L - x}{\sqrt{T}} + \bar{R}_w, \quad 0 \leq w \leq 1. \quad (2.8)$$

Here  $(\bar{R})$  is a standard 3-dimensional Bessel bridge. This stochastic ordering result can be proven using the following arguments: first the Bessel bridge  $\hat{R}^T$  has the same distribution as the Euclidian norm of  $\frac{(L-x)w}{\sqrt{T}} e_1 + \beta_w$  where  $\beta_w$  is a standard 3-dimensional Brownian bridge, see (1.9) and (1.10). Secondly the triangle inequality leads to

$$\left\| \frac{(L-x)w}{\sqrt{T}} e_1 + \beta_w \right\| \leq \left\| \frac{(L-x)w}{\sqrt{T}} e_1 \right\| + \|\beta_w\| \leq \frac{L-x}{\sqrt{T}} + \|\beta_w\|.$$

In order to prove the stochastic ordering (2.8) it suffices then to note that  $\|\beta_w\|$  and  $\bar{R}_w$  are identically distributed.

Let us now introduce some parameter  $\alpha \in ]0, 1[$  and define

$$\gamma_T = \inf_{y \in [x - T^{1/2+\alpha}, L]} \gamma(y). \quad (2.9)$$

Then

$$I_T \leq T \int_0^1 e^{-T\gamma_T u} du \mathbb{P}\left(\sup_{u \in [0,1]} \sqrt{T} \hat{R}_u^T \leq T^{1/2+\alpha} + (L-x)\right) \\ + T \mathbb{P}\left(\sup_{u \in [0,1]} \sqrt{T} \hat{R}_u^T > T^{1/2+\alpha} + (L-x)\right).$$

By (2.8) we obtain:

$$I_T \leq \frac{1}{\gamma_T} 1_{\{T \geq (x+1)^{2/(1+2\alpha)}\}} + T 1_{\{T < (x+1)^{2/(1+2\alpha)}\}} + T \mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right) \\ \leq \frac{1}{\gamma_T} 1_{\{T \geq (x+1)^{2/(1+2\alpha)}\}} + (x+1)^2 + T \mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right). \quad (2.10)$$

Using the agreement formula (see, for instance, Theorem 2 in [31]), we obtain

$$\mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right) = C_3 \mathbb{E}\left[\sqrt{\bar{\tau}} 1_{\{\bar{\tau} < T^{-2\alpha}\}}\right].$$

Here  $C_3 = \sqrt{2}/\Gamma(3/2)$  and  $\bar{\tau} = \tau + \hat{\tau}$  where  $\tau$  is the first hitting time of the level 1 for a 3-dimensional Bessel process and  $\hat{\tau}$  an independent copy of  $\tau$ . Combining classical inequalities and the explicit expression of the Laplace transform corresponding to the distribution of  $\tau$  (see [27]) leads to

$$\mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right) \leq C_3 T^{-\alpha} \mathbb{P}\left(\exp -\lambda \bar{\tau} > \exp -\lambda T^{-2\alpha}\right) \\ \leq C_3 T^{-\alpha} e^{\lambda T^{-2\alpha}} \mathbb{E}[e^{-\lambda \bar{\tau}}] = C_3 T^{-\alpha} e^{\lambda T^{-2\alpha}} \frac{(2\lambda)^{1/2}}{C_3^2 I_{1/2}^2(\sqrt{2\lambda})},$$

for any  $\lambda > 0$ .  $I_\nu$  stands for the Bessel function of the first kind. In particular  $I_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sinh x$ . The particular choice  $\lambda = T^{2\alpha}/2$  leads to

$$\mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right) \leq \frac{\sqrt{e\pi}}{2\sqrt{2}} \frac{\pi T^\alpha}{2 \sinh^2(T^\alpha)}.$$

There exists a constant  $C_\alpha > 0$  such that

$$T \mathbb{P}\left(\sup_{u \in [0,1]} \bar{R}_u > T^\alpha\right) \leq \frac{\sqrt{e}}{2} \left(\frac{\pi}{2}\right)^{3/2} \sup_{y \geq 0} \frac{y^{1+\alpha}}{\sinh^2(y^\alpha)} \leq C_\alpha.$$

Let us go back to the upper-bound (2.10). By definition,  $\gamma_T \geq C_\gamma |x - T^{1/2+\alpha}|^{-r}$  as soon as  $T \geq (x+1)^{2/(1+2\alpha)}$  and therefore

$$I_T \leq \frac{1}{C_\gamma} |x - T^{1/2+\alpha}|^r 1_{\{T \geq (x+1)^{2/(1+2\alpha)}\}} + (x+1)^2 + C_\alpha.$$

Let us recall that  $T$  is the Brownian first-passage time through the level  $L$  starting in  $x$ :  $T$  and  $\frac{(L-x)^2}{G^2}$  are identically distributed,  $G$  being a standard

gaussian variable. Taking now the expectation with respect to  $T$ , we get

$$\begin{aligned}
\mathbb{E}[\mathcal{N}_1] &\leq 1 + \kappa \mathbb{E}[I_T] \\
&\leq 1 + \frac{\kappa}{C_\gamma} \mathbb{E}[|x - T^{1/2+\alpha}|^r \mathbf{1}_{\{T \geq (x+1)^{2/(1+2\alpha)}\}}] + \kappa(x+1)^2 + \kappa C_\alpha \\
&\leq 1 + \frac{\kappa 2^r}{C_\gamma} \left( |x|^r + \mathbb{E}[T^{(1/2+\alpha)r}] \right) + \kappa(x+1)^2 + \kappa C_\alpha \\
&= 1 + \frac{\kappa 2^r}{C_\gamma} \left( |x|^r + |L-x|^{(1+2\alpha)r} \mathbb{E}[G^{-(1+2\alpha)r}] \right) + \kappa(x+1)^2 + \kappa C_\alpha.
\end{aligned}$$

Since  $r < 1$ , it is possible to choose  $\alpha \in ]0, 1[$  small enough such that  $(1+2\alpha)r < 1$  and therefore  $\mathbb{E}[G^{-(1+2\alpha)r}] < \infty$ . Let us choose for instance  $\alpha = \frac{1-r}{4r}$  that means  $(1+2\alpha)r = \frac{1+r}{2}$   $\square$

**Proposition 2.5.** *Let us consider Algorithm (A1)<sub>shift</sub> with input variable  $T \sim \text{IG}(\frac{L-x}{\sqrt{2\gamma_0}}, (L-x)^2)$  as in Proposition 2.2. Then the number of random points associated to the first iteration satisfies:*

$$\mathbb{E}[\mathcal{N}_1] \leq 1 + \kappa \frac{L-x}{\sqrt{2\gamma_0}}.$$

*Proof.* The arguments are similar to those developed in the proof of Proposition 2.4. The formula (2.7) leads to  $\mathbb{E}_c[\mathcal{N}_1] \leq 1 + \kappa T$ . Since  $T$  is an inverse gaussian r.v. we obtain the announced upper-bound:

$$\mathbb{E}[\mathcal{N}_1] \leq 1 + \kappa \mathbb{E}[T] = 1 + \kappa \frac{L-x}{\sqrt{2\gamma_0}} < \infty. \quad \square$$

The main description of the algorithm efficiency encompasses both the information about the first iteration and the description of the number of iterations.

**Theorem 2.6.** *Let us denote by  $\mathcal{N}_\Sigma$  the total number of random points used in the simulation of the diffusion first-passage time.*

1. *Under condition (2.5), the number of points of Algorithm (A1), in the context of Proposition 1.4, satisfies:*

$$\mathbb{E}[\mathcal{N}_\Sigma] \leq 2 \left( M_{\gamma,1} + \kappa M_{\gamma,2} (x^2 + (L-x)^{(1+r)/2}) \right) e^{\beta(L) - \beta(x)},$$

$M_{\gamma,1}$  and  $M_{\gamma,2}$  being the constants introduced in Proposition 2.4.

2. *The number of points of the modified algorithm (A1)<sub>shift</sub>, in the context of Proposition 2.2, satisfies:*

$$\mathbb{E}[\mathcal{N}_\Sigma] \leq 2 \left( 1 + \kappa \frac{L-x}{\sqrt{2\gamma_0}} \right) e^{\beta(L) - \beta(x) - (L-x)\sqrt{2\gamma_0}}.$$

Let us notice in particular that the average number of random points is finite which actually plays an important role for numerical purposes.



*Proof.* We shall just prove the first item, the second one is based on similar arguments. Let us recall that  $\mathcal{I}$  is the number of iterations in Algorithm (A1) and  $A$  stands for the event *Acceptance* in the acceptance-rejection method. Therefore

$$\begin{aligned}\mathbb{E}[\mathcal{N}_\Sigma] &= \mathbb{E}[\mathcal{N}_1 + \dots + \mathcal{N}_\mathcal{I}] = \sum_{k=1}^{\infty} \mathbb{E}[\mathcal{N}_1 + \dots + \mathcal{N}_k | \mathcal{I} = k] \mathbb{P}(\mathcal{I} = k) \\ &= \sum_{k=1}^{\infty} \left( (k-1) \mathbb{E}[\mathcal{N}_1 | A^c] + \mathbb{E}[\mathcal{N}_1 | A] \right) \mathbb{P}(\mathcal{I} = k).\end{aligned}\quad (2.11)$$

Let us denote by  $p = e^{-(\beta(L) - \beta(x))}$  and observe that  $\mathbb{P}(A) = p$ . Moreover we get the following upper bound

$$\mathbb{E}[\mathcal{N}_1 | A^c] = \frac{\mathbb{E}[\mathcal{N}_1 1_{A^c}]}{1-p} \leq \frac{\mathbb{E}[\mathcal{N}_1]}{1-p}.$$

We deduce that (2.11) becomes

$$\mathbb{E}[\mathcal{N}_\Sigma] \leq \mathbb{E}[\mathcal{N}_1] \left( \sum_{k=1}^{\infty} \left\{ \frac{k-1}{1-p} + \frac{1}{p} \right\} (1-p)^{k-1} p \right) = \frac{2}{p} \mathbb{E}[\mathcal{N}_1].$$

Proposition 2.4 permits to conclude the proof.  $\square$

### 2.3 Efficiency of Algorithm (A2)

In this section,  $\mathcal{N}_1$  still represents the number of random points used for the first iteration of the algorithm.

**Theorem 2.7.** *Let  $\gamma$  satisfies Assumptions 1.1–1.3 and the particular condition (2.5). We consider Algorithm (A2) with  $T$  corresponding to the Brownian first-passage time to the level  $L$ , with  $x < L$ . Then there exist two constants  $M_{\gamma,1} > 0$  and  $M_{\gamma,2} > 0$  such that the total number of random points  $\mathcal{N}_\sigma$  in Algorithm (A2) satisfies*

$$\mathbb{E}[\mathcal{N}_\Sigma] \leq 2 \left( M_{\gamma,1} + \kappa M_{\gamma,2} (x^2 + (L-x)^{(1+r)/2}) \right) e^{\beta(L) - \beta(x)}.$$

The statement of Theorem 2.7 is similar to Theorem 2.6 but the arguments of proof are slightly different.

*Proof.* We recall that  $\mathbb{P}_c$  stands for the conditional probability given the first-passage time  $T$  and the Bessel bridge  $(R_t)$  ending with  $R_T = L - x$ . We observe that

$$\mathbb{P}_c(\mathcal{N}_1 > 1) = \mathbb{P}_c(\gamma(L - R_{U_1}) \leq e_1 \leq \kappa), \quad (2.12)$$

where  $U_1$  is uniformly distributed on the interval  $[0, T]$  and  $e_1$  is an exponentially distributed random variable independent of  $U_1$  and with average  $1/T$ . Hence (2.12) becomes

$$\begin{aligned}\mathbb{P}_c(\mathcal{N}_1 > 1) &= \int_0^T \int_0^\infty 1_{\{\gamma(L - R_u) \leq t \leq \kappa\}} e^{-Tt} du dt \\ &= \frac{1}{T} \int_0^T \left( e^{-T\gamma(L - R_u)} - e^{-T\kappa} \right) du = \mathbb{E}_c[e^{-T\gamma(L - R_{U_1})}] - e^{-T\kappa}.\end{aligned}\quad (2.13)$$

Let us generalize these computations. We introduce  $(e_n, n \geq 0)$  a sequence of exponentially distributed random variables with average  $1/T$ ,  $\mathcal{E}_n := e_1 + \dots + e_n$  and  $(U_n, n \geq 0)$  a sequence of uniform distributed independent r.v. on the interval  $[0, T]$ , all r.v. being independent. Moreover let us denote by  $F_c(t) = \gamma(L - R_t)$  for any  $t \in [0, T]$  (deterministic function under the conditional distribution  $\mathbb{P}_c$ ) and  $G_c(y) = \frac{1}{T} \int_0^T 1_{\{y \geq F_c(t)\}} dt$ . For  $k > 1$ , we get

$$\mathbb{P}_c(\mathcal{N}_1 > k) = \mathbb{P}_c\left(F_c(U_1) \leq \mathcal{E}_1, F_c(U_2) \leq \mathcal{E}_2, \dots, F_c(U_k) \leq \mathcal{E}_k < \kappa\right). \quad (2.14)$$

Let us note that, under  $\mathbb{P}_c$ , we have on one hand  $\mathcal{E}_k \sim \gamma(k, \frac{1}{T})$  and on the other hand:

$$\left(\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_{k-1}, \mathcal{E}_k\right) \sim \left(V^{(1)}\mathcal{E}_k, V^{(2)}\mathcal{E}_k, \dots, V^{(k-1)}\mathcal{E}_k, \mathcal{E}_k\right)$$

where  $(V^{(1)}, V^{(2)}, \dots, V^{(k-1)})$  are ordered variables corresponding to the sequence of uniform random on  $[0, 1]$ :  $(V_1, V_2, \dots, V_{k-1})$ .

$$\begin{aligned} \mathbb{P}_c(\mathcal{N}_1 > k) &= \mathbb{P}_c(F_c(U_1) \leq V^{(1)}\mathcal{E}_k, F_c(U_2) \leq V^{(2)}\mathcal{E}_k, \dots, F_c(U_k) \leq \mathcal{E}_k < \kappa) \\ &= \mathbb{E}_c \left[ G_c(V^{(1)}\mathcal{E}_k) G_c(V^{(2)}\mathcal{E}_k) \dots G_c(V^{(k-1)}\mathcal{E}_k) G_c(\mathcal{E}_k) 1_{\{\mathcal{E}_k \leq \kappa\}} \right] \\ &= \mathbb{E}_c \left[ G_c(V_1\mathcal{E}_k) G_c(V_2\mathcal{E}_k) \dots G_c(V_{k-1}\mathcal{E}_k) G_c(\mathcal{E}_k) 1_{\{\mathcal{E}_k \leq \kappa\}} \right] \\ &= \mathbb{E}_c \left[ \mathcal{G}_c^{k-1}(\mathcal{E}_k) G_c(\mathcal{E}_k) 1_{\{\mathcal{E}_k \leq \kappa\}} \right]. \end{aligned}$$

Here  $\mathcal{G}_c$  stands for  $\mathcal{G}_c(y) = \int_0^1 G_c(ty) dt$ .

$$\mathbb{P}_c(\mathcal{N}_1 > k) = \frac{T^k}{\Gamma(k)} \int_0^\kappa x^{k-1} \mathcal{G}_c^{k-1}(x) G_c(x) e^{-Tx} dx.$$

We deduce

$$\begin{aligned} \mathbb{E}_c[\mathcal{N}_1] &= \sum_{k \geq 0} \mathbb{P}_c(\mathcal{N}_1 > k) = 1 + T \sum_{k \geq 0} \int_0^\kappa \frac{(xT\mathcal{G}_c(x))^k}{k!} G_c(x) e^{-Tx} dx \\ &= 1 + T \int_0^\kappa G_c(x) \exp\left\{xT(\mathcal{G}_c(x) - 1)\right\} dx. \end{aligned}$$

Taking the expectation with respect to both  $T$  and the Bessel bridge  $(R_t)$ , we obtain

$$\mathbb{E}[\mathcal{N}_1] = 1 + \int_0^\kappa \mathbb{E} \left[ T G_c(x) \exp\left\{xT(\mathcal{G}_c(x) - 1)\right\} \right] dx. \quad (2.15)$$

Let us now consider some upper-bound. First we note that  $0 \leq G_c(x) \leq 1$  for any  $x \in \mathbb{R}$ . Let us now consider  $\mathcal{G}_c(x)$  using a change of variable:

$$\begin{aligned} \mathcal{G}_c(x) &= \frac{1}{T} \int_0^1 \left( \int_0^T 1_{\{ux \geq \gamma(L - R_t)\}} dt \right) du \\ &\triangleq \int_0^1 \int_0^1 1_{\{ux \geq \gamma(L - \sqrt{T}\hat{R}_w^T)\}} dw du, \end{aligned}$$

where  $(\hat{R}_w^T, 0 \leq w \leq 1)$  is a 3-dimensional Bessel bridge with condition  $\hat{R}_1^T = (L - x)/\sqrt{T}$ , as considered in the proof of Proposition 2.4. For  $x > 0$ ,

$$\mathcal{G}_c(x) \leq 1 - \frac{1}{x} \inf_{y \in [L - \sqrt{T}\Pi_T, L]} \gamma(y), \quad \text{with } \Pi_T = \sup_{w \in [0, 1]} \hat{R}_w^T.$$

By (2.15) and the previous inequalities, we obtain:

$$\mathbb{E}[\mathcal{N}_1] \leq 1 + \int_0^\kappa \mathbb{E} \left[ T \exp -T \inf_{y \in [L - \sqrt{T}\Pi_T, L]} \gamma(y) \right] dx. \quad (2.16)$$

Let us just note that the variable  $\Pi_T$  depends on  $x$ . Using the obvious upper-bound  $T e^{-T\gamma_T} \leq \int_0^1 T e^{-T\gamma_T u} du$  and the definition of  $\gamma_T$  in (2.9), we observe that all arguments developed in the proof of Proposition 2.4 and in Theorem 2.6, from (2.9) on, can be adapted to (2.16).  $\square$

### 3 When the diffusion hits the level $L$ before a given fixed time $t_0$

Both Algorithm (A1) and (A2) concern first-passage times for the diffusion  $(X_t, t \geq 0)$  defined in (1.2) under some particular conditions. The boundedness of  $\gamma := \frac{b^2 + b'}{2}$  on  $] -\infty, L]$  is indisputably the most restrictive condition (Assumption 1.3). Unfortunately the result presented so far cannot be generalized to the particular situation when  $\gamma$  takes some negative values. Nevertheless we can adapt the algorithms by considering  $\tau_L | \tau_L \leq t_0$  instead of  $\tau_L$ , defined by (0.1), for any fixed time  $t_0$ .

Let us consider a modification of the theoretical Algorithm (A<sub>0</sub>). The idea is to focus our attention to random times which are a.s. smaller than some given value  $t_0$ .

**Assumption 3.1.** *There exists two constants  $\kappa > 0 > -m$  such that*

$$-m \leq \gamma(x) \leq \kappa, \quad \forall x \leq L.$$

ALGORITHM (A3).

Let us fix the level  $L > x$ ,  $x$  being the initial value of the diffusion.

Step 1: Simulate a non-negative random variable  $T$  with p.d.f.  $f_T$  of support  $[0, t_0]$ .

Step 2: Simulate a 3-dimensional Bessel process  $(R_t)$  on the time interval  $[0, T]$  with endpoint  $R_T = L - x$ . We define by  $D_{R,T}^m$  the stochastic domain satisfying:

$$D_{R,T}^m := \left\{ (t, v) \in [0, T] \times \mathbb{R}_+ : v \leq \gamma(L - R_t) + \frac{mt_0}{T} \right\}.$$

Step 3: Simulate a Poisson point process  $N$  on the state space  $[0, T] \times \mathbb{R}_+$ , independent of the Bessel process, whose intensity measure is the Lebesgue one.

Step 4: If  $N(D_{R,T}^m) = 0$  then set  $Y = T$  otherwise go to Step 1.

Outcome: the random variable  $Y$ .

It is obvious that the support of the random variable  $Y$  is contained in  $[0, t_0]$  by construction.

**Proposition 3.1.** *Let  $\gamma$  a function satisfying Assumptions 1.1, 1.2 & 3.1 and  $G$  a standard gaussian r.v. If  $T$  has the same distribution as the conditional distribution of  $(L - x)^2/G^2$  given  $(L - x)^2/G^2 \leq t_0$ , then the random variable  $Y$  obtained by Algorithm (A3) has the same distribution as  $\tau_L$  given  $\tau_L \leq t_0$ .*

*Proof. Step 1.* Let us first describe the dependence of the distribution of the outcome  $Y$  with respect to the distribution of the input  $T$ . Let us also note that, due to the condition on the function  $\gamma$ ,  $\gamma(L - R_t) + \frac{mt_0}{T} \geq 0$  for any  $t$  and  $T \leq t_0$ . We deduce that

$$\begin{aligned} \mathbb{P}(N(D_{R,T}^m) = 0 | R, T) &= \exp - \int_0^T \left( \gamma(L - R_t) + \frac{mt_0}{T} \right) dt \\ &= e^{-mt_0} \exp - \int_0^T \gamma(L - R_t) dt. \end{aligned}$$

Following the arguments developed in Theorem 1.3, we obtain, for any non-negative or bounded function  $\psi$ ,

$$\begin{aligned} \mathbb{E}[\psi(Y)] &= \frac{1}{\mathbb{P}(N(D_{R,T}^m) = 0)} \mathbb{E} \left[ \psi(T) \mathbb{P}(N(D_{R,T}^m) = 0 | T) \right] \\ &= \frac{\mathbb{E}[\psi(T) e^{-mt_0} \eta(T)]}{\mathbb{E}[e^{-mt_0} \eta(T)]} = \frac{\mathbb{E}[\psi(T) \eta(T)]}{\mathbb{E}[\eta(T)]}. \end{aligned} \quad (3.1)$$

where  $\eta$  is defined in (1.5). In other words, the *p.d.f.* of the outcome  $Y$  is given by

$$f_Y(t) = \frac{1}{\Xi} \eta(t) f_T(t),$$

where  $\Xi$  is the normalization coefficient. The result is quite similar to the statement of Theorem 1.3. The differences are the following: on one hand, the function  $\gamma$  can take negative values, on the other hand the support of the distribution of  $T$  has to be compact.

*Step 2.* Let us now link the conditional distributions of  $Y$  and  $\tau_L$ . Girsanov's transformation, see Proposition 1.1, implies

$$\mathbb{E}_{\mathbb{P}}[\psi(\tau_L) 1_{\{\tau_L \leq t_0\}}] = \mathbb{E}_{\mathbb{Q}}[\psi(\tau_L) 1_{\{\tau_L \leq t_0\}} \eta(\tau_L)] \exp \left\{ \beta(L) - \beta(x) \right\}, \quad (3.2)$$

where  $\mathbb{P}$  (resp.  $\mathbb{Q}$ ) corresponds to the distribution of the diffusion  $X$  (resp. the Brownian motion  $B$ ). The particular case  $\psi \equiv 1$  leads to:

$$\mathbb{P}(\tau_L \leq t_0) = \mathbb{E}_{\mathbb{Q}}[1_{\{\tau_L \leq t_0\}} \eta(\tau_L)] \exp \left\{ \beta(L) - \beta(x) \right\}. \quad (3.3)$$

Combining the ratio between (3.2) and (3.3) on one side and the equation (3.1) on the other side leads to

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\psi(\tau_L) | \tau_L \leq t_0] &= \frac{\mathbb{E}_{\mathbb{Q}}[\psi(\tau_L) 1_{\{\tau_L \leq t_0\}} \eta(\tau_L)]}{\mathbb{E}_{\mathbb{Q}}[1_{\{\tau_L \leq t_0\}} \eta(\tau_L)]} = \frac{\mathbb{E}_{\mathbb{Q}}[\psi(\tau_L) \eta(\tau_L) | \tau_L \leq t_0]}{\mathbb{E}_{\mathbb{Q}}[\eta(\tau_L) | \tau_L \leq t_0]} \\ &= \frac{\mathbb{E}[\psi(T) \eta(T)]}{\mathbb{E}[\eta(T)]} = \mathbb{E}[\psi(Y)], \end{aligned}$$

where  $T$  has the same law as  $\tau_L | \tau_L \leq t_0$  under the probability measure  $\mathbb{Q}$  (Brownian hitting time). In order to conclude, it suffices to note that, under  $\mathbb{Q}$ , the random variable  $\tau_L$  and  $(L - x)^2/G^2$  are identically distributed.  $\square$

Of course, for simulation purposes, we don't use directly Algorithm (A3) since it is impossible to simulate the Poisson point process  $N$  on the whole unbounded state space  $[0, T] \times \mathbb{R}_+$ . So we apply exactly the same procedures developed in Section 1 which transform Algorithm (A0) into (A1) or (A2). Such procedures modify Algorithm (A3) and permit to easily execute numerical simulations (Section 5.2).

## 4 Unbounded drift terms

The aim of this section is to handle with drift terms  $b$  which do not satisfy the condition  $\gamma(x) \leq \kappa$  for any  $x \leq L$ . From now on, we shall just assume that

$$\gamma(x) \geq 0, \quad \forall x \leq L. \quad (4.1)$$

The algorithms (A1) and (A2) cannot be used in such a situation since the domain associated to the Poisson point process is unbounded. One way to overcome this difficulty is to replace the drift term of the diffusion by a modified version: for a given large parameter  $\rho$ ,

$$b_\rho(x) = \begin{cases} b(x) & \text{if } -\rho \leq x \leq L \\ b(-\rho) + b'(-\rho)(x + \rho)e^{x+\rho} & \text{if } x < -\rho. \end{cases} \quad (4.2)$$

In other words, it suffices to keep the function  $b$  on the interval  $[-\rho, L]$  and to extend this function in a regular way to  $] -\infty, L]$  in order to get a bounded function. Here  $\lim_{x \rightarrow -\infty} b_\rho(x) = b(-\rho)$ , the boundedness being verified.

Obviously  $b_\rho$  is a  $C^1$ -continuous function and both Algorithm (A1) and (A2) can be applied to the diffusion associated to this new drift term. The diffusion is denoted  $X^\rho$ .

**Proposition 4.1.** *Let the function  $\gamma$  satisfy Assumption 1.1. Moreover we assume that  $\lim_{x \rightarrow -\infty} p(x) = -\infty$  where  $p$  is defined in (1.3). Then the distribution of  $\tau_L^\rho$  converges towards the distribution of  $\tau_L$  as  $\rho \rightarrow \infty$ . Moreover the Kolmogorov distance satisfies:*

$$d(\tau_L, \tau_L^\rho) := \sup \left\{ |F_{\tau_L^\rho}(t) - F_{\tau_L}(t)| : t \in \mathbb{R}_+ \right\} = \mathcal{O} \left( -p(-\rho) \right) \text{ as } \rho \rightarrow \infty.$$

Here  $F_Y(x)$  stands for the cumulative distribution function of the r.v.  $Y$ .

In practice, in order to approximate  $\tau_L$  we simulate  $\tau_L^\rho$  using one of the algorithms presented before, with the parameter  $\rho$  large enough. Let us notice that the condition  $\lim_{x \rightarrow -\infty} p(x) = -\infty$  appearing in the statement of Proposition 4.1 implies Assumption 1.2 as mentioned in Remark 1.2.

*Proof.* The arguments are quite straightforward: let us consider  $x \in \mathbb{R}_+$ , then

$$F_{\tau_L}(t) = \mathbb{P}(\tau_L \leq t, \tau_L \leq \tau_{-\rho}) + \mathbb{P}(\tau_L \leq t, \tau_L > \tau_{-\rho}).$$

Since  $X^\rho$  and  $X$  have the same distribution before exiting from the interval  $[-\rho, L]$  (same drift term on this space subset), we obtain

$$\begin{aligned} |F_{\tau_L}(t) - F_{\tau_L^\rho}(t)| &= |\mathbb{P}(\tau_L \leq t, \tau_L > \tau_{-\rho}) - \mathbb{P}(\tau_L^\rho \leq t, \tau_L^\rho > \tau_{-\rho}^\rho)| \\ &\leq 2\mathbb{P}(\tau_L > \tau_{-\rho}) = 2 \frac{p(L) - p(x)}{p(L) - p(-\rho)} \sim -\frac{2}{p(-\rho)}. \end{aligned}$$

□

The efficiency of this algorithm is already described in Section 2 and numerical results are presented in Section 5.3.

## 5 Examples and numerics

In this section we study the performances of the different algorithms proposed through the paper in order to simulate  $\tau_L$  the first-passage time of the diffusion starting in  $x$  with  $x < L$ . Let us just give a quick review in the following table describing which algorithm should be used when the function  $\gamma$  satisfies particular bounds on the interval  $] -\infty, L]$ .

Condition on $\gamma$	r.v. simulated	Algorithm and statement
$0 \leq \gamma(x) \leq \kappa$	$\tau_L$	(A1) or (A2) in Proposition 1.4
$0 < \gamma_0 \leq \gamma(x) \leq \kappa$	$\tau_L$	(A1) <sub>shift</sub> or (A2) <sub>shift</sub> in Prop. 2.2
$-m \leq \gamma(x) \leq \kappa$	$\tau_L$ given $\tau_L \leq t_0$	(A3) in Prop. 3.1
$0 \leq \gamma(x)$	$\tau_L^\rho$ (approx.)	(A1) $^\rho$ or (A2) $^\rho$ in Prop. 4.1

Let us recall that the efficiency of all these algorithms can be increased using the *space splitting* argument proposed in Section 2.1.

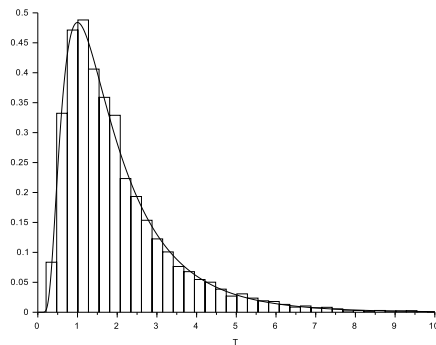
In order to illustrate results stem from simulations, we consider the first hitting time of a Brownian motion with drift  $\mu$ . In this case, a closed form of the *p.d.f.* for  $\tau_L$  is available (see, for instance [26] p. 197):

$$\mathbb{P}(\tau_L \in dt) = \frac{L-x}{\sqrt{2\pi t^3}} \exp -\frac{(L-x-\mu t)^2}{2t} dt, \quad t > 0.$$

Since the level  $L$  is larger than the starting position  $x$ , Assumption 1.2 will be satisfied if and only if  $\mu > 0$ .

In the opposite figure, there is a comparison between the theoretical probability density function of  $\tau_L$  and its normalized histogram (10 000 simulations) obtained by Algorithm (A1).

Here  $b(x) \equiv 1$ ,  $x = 0$  and level  $L = 2$ .



## 5.1 Algorithms (A1) and (A2)

We shall compare Algorithm (A1) and (A2), both of them simulating exactly the first-passage time  $\tau_L$  for a diffusion process satisfying Assumption 1.1–1.3.

**Example 1.** We consider the following stochastic differential equation:

$$dX_t = (2 + \sin(X_t)) dt + dB_t, \quad t \geq 0, \quad X_0 = 0. \quad (5.1)$$

We first observe that  $\gamma(x) = (b^2(x) + b'(x))/2 = ((2 + \sin(x))^2 + \cos(x))/2$  satisfies  $0 \leq \gamma \leq 5$ . We use Algorithm (A1) in order to produce an histogram of the hitting time distribution (Figure 1). The total number of random variables

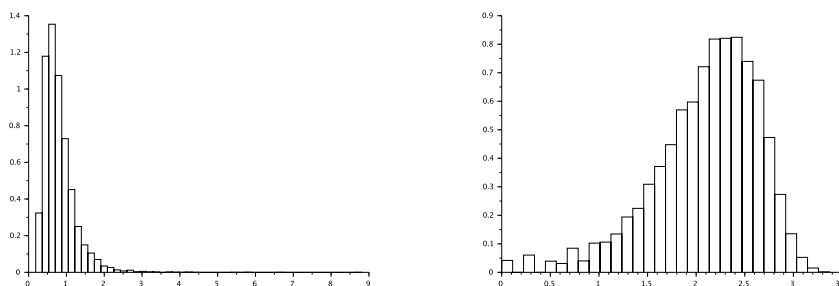


Figure 1: Histogram of the hitting time distribution for 10 000 simulations corresponding to the level  $L = 2$  and starting position  $X_0 = 0$  (left), histogram of the number of iterations in Algorithm (A1) in the  $\log_{10}$ -scale (right).

used in this algorithm is of course random and the associated histogram (Figure 2) illustrates its distribution.

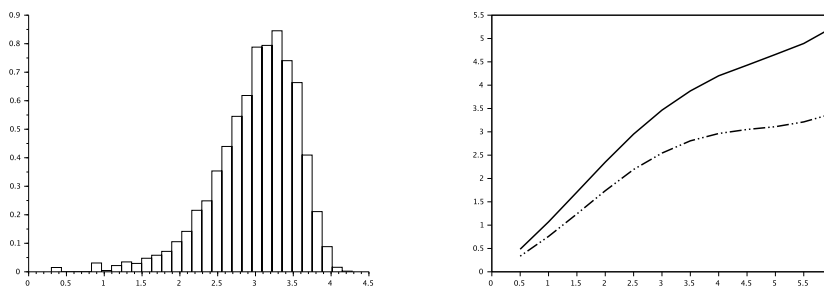


Figure 2: Number of random variables used in Algorithm (A1) for 10 000 simulations with  $L = 2$ ,  $X_0 = 0$  in the  $\log_{10}$ -scale (left) and mean number of iterations versus the level height  $L$  for Algorithm (A1)<sub>shift</sub> respectively (A1) (dashed line resp. solid line), both curves are in the  $\log_{10}$ -scale (10 000 simulations have been used for the average estimation).

Let us just note that Algorithm (A1) is quite time consuming since simulating 10 000 hitting times (corresponding to  $L = 2$  and  $X_0 = 0$ ) requires about 300 cpu. This time increases very fast as  $L$  becomes large: the figure (right)

presents the number of iterations versus  $L$ . In order to increase the algorithm efficiency and consequently reduce the computation time, one way is to shift the function  $\gamma$  as described in Proposition 2.2. The function  $\gamma$  introduced in (5.1) satisfies  $\gamma(x) \geq \gamma_0 := 1/4$ . The comparison of the efficiency of Algorithm (A1) (solid curve) and Algorithm (A1)<sub>shift</sub> (dashed curve) is presented in Figure 2.

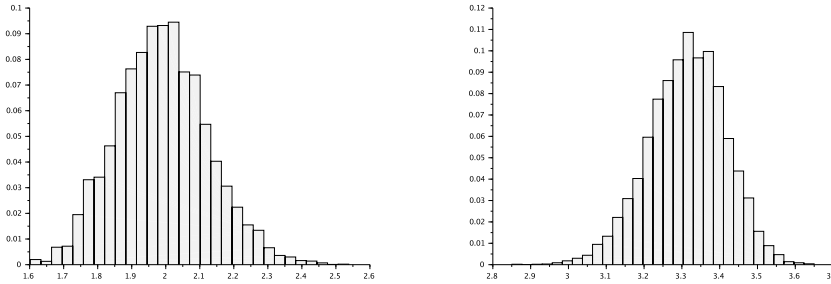


Figure 3: Histogram of the number of random variables in Algorithm (A1) using space splitting for 10000 with  $L = 2$ ,  $X_0 = 0$ ,  $k = 20$  (left),  $L = 20$ ,  $k = 20$  (right) both in the  $\log_{10}$ -scale.

The example (5.1) also permits to illustrate the second modification of the algorithm developed in Section 2.1: the *space splitting* method. Let us recall that we can replace the simulation of  $\tau_L$  for a diffusion starting in  $x$  (denoted here by  $\tau(x \rightarrow L)$ ) by the simulation of  $k$  independent first-passage times  $\tau(x + (i - 1)(L - x)/k \rightarrow x + i(L - x)/k)$  for  $i = 1, \dots, k$ . It suffices then to take the sum of all these quantities. Figure 3 (left) represents the number of random variables for the particular case:  $L = 2$ ,  $X_0 = 0$  and  $k = 20$ . We can notice that such a procedure is slightly more efficient for this particular case: the average number of random variables (estimation with 10000 simulations) namely goes from about 1791 down to 102. It permits also to deal with large level values (Figure 3 – right) which is a huge task with Algorithm (A1). For each level  $L$  there is an optimal splitting (optimal value  $k$ ) illustrates by Figure 4.

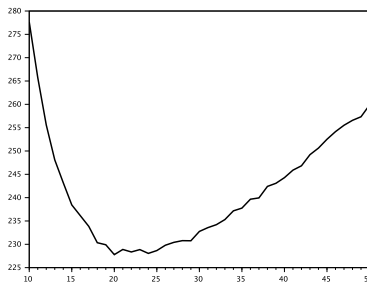


Figure 4: Averaged number of random variables used in Algorithm (A1) versus the number of slices  $k$  with  $X_0 = 0$  and  $L = 5$ . The averaging uses 10000 simulations.

Up to this point, all simulations use Algorithm (A1). Of course, it is also



possible to deal with Algorithm (A2) in such a situation. We obtain the same hitting time distribution and can therefore compare the number of iterations and the number of random variables used. Let us notice that both Algorithms (A1) and (A2) are based on the simulation of a Poisson point process defined on the rectangle  $[0, T] \times [0, \kappa]$ : for one algorithm, the simulation starts with the point with smallest abscissa and then goes from one point to the next one by increasing the abscissa, for the other one, the order of simulation depends on the ordinate. It is possible of course to have a one-to-one correspondance between these two simulation technics: each point  $(t, x) \in [0, T] \times [0, \kappa]$  of the Poisson process is then transformed as follows

$$(t, x) \mapsto (xT/\kappa, t\kappa/T).$$

Consequently the same sequence of random points can be used for both Algorithms and it becomes easy to compare the total number of random points. Let us introduce  $\mathcal{N}_1$  the total number of random variables used by Algorithm (A1) and respectively  $\mathcal{N}_2$  by Algorithm (A2), then we define the difference

$$\Delta := \mathcal{N}_2 - \mathcal{N}_1. \quad (5.2)$$

In the diffusion case (5.1), the estimated mean  $\bar{\Delta}$  is equal to  $-196.0$  and its standard deviation equals  $1853.2$  when observing  $10\,000$  simulations with  $x = 0$  and  $L = 2$ . We claim therefore that Algorithm (A2) is more efficient than Algorithm (A1) for Example 1 since  $\bar{\Delta}$  is significantly negative (p-value:  $1.9E - 26$ ). However this difference is small in comparison to the mean number of random variables:  $\bar{\Delta}/\bar{\mathcal{N}}_1 \approx -0.11$ . In other situations, this difference can be much stronger or positive. It essentially depends on the drift term  $b$  of the diffusion. Let us consider a second situation.

**Example 2.** Let us consider the following SDE:

$$dX_t = (1 + \arctan(1 - X_t)) dt + dB_t, \quad t \geq 0 \quad X_0 = 0, \quad (5.3)$$

associated to the first-passage time for  $L = 1$ . Let us note that  $\tau_1 < \infty$  a.s.

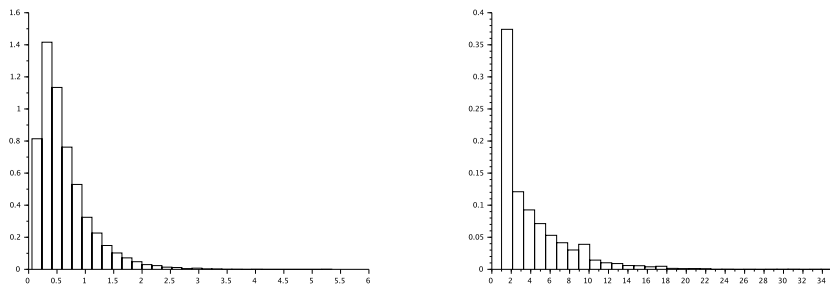


Figure 5: Histogram of the hitting time distribution using Algorithm (A1) (left) and histogram of the number of iterations (right) for  $X_0 = 0$ ,  $L = 2$  and  $10\,000$  simulations.

We observe that  $\gamma(x) = \frac{1}{2} \left( (1 + \arctan(1 - x))^2 - \frac{1}{1 + (1 - x)^2} \right)$  satisfies  $0 \leq \gamma \leq \frac{1}{2}(1 + \pi/2)^2$ . Both Algorithms (A1) and (A2) can be used in order to

simulate the first-passage time  $\tau_L$  and the same kind of study is achieved: here, the estimated difference  $\overline{\Delta} = -169.7$ , its standard deviation 235.2 and finally the reduction rate  $\overline{\Delta}/\overline{\mathcal{N}}_1 \approx -0.67$  for  $x = 0$ ,  $L = 1$  and 10 000 simulations. The choice between (A1) and (A2), in such a situation, becomes an important question (the hitting time distribution is illustrated by Figure 5).

## 5.2 Simulation of $\tau_L$ given $\tau_L \leq t_0$ .

In this section, we illustrate Algorithm (A3) presented in Section 3 which permits to overcome the condition  $\gamma \geq 0$ . Indeed this condition can be weakened (Assumption 3.1) provided the study only concerns the distribution of the first-passage time  $\tau_L$  conditioned on the event  $\tau_L \leq t_0$ . Here  $t_0$  is any fixed time.

We therefore introduce the following SDE:

$$dX_t = -\arctan(X_t) dt + dB_t, \quad 0 \leq t \leq t_0. \quad (5.4)$$

We focus our attention to the case  $X_0 = 0$ ,  $L = 1$  and different values for  $t_0$ . Let us observe that  $\gamma(x) = (\arctan(x)^2 - 1/(1+x^2))/2$  satisfies  $-1/2 \leq \gamma(x) \leq \pi^2/8$  and that the first-passage time is almost surely finite (Assumption 1.2). As described in the statement of Proposition 3.1, the simulation of  $\tau_L$  given  $\tau_L \leq t_0$  is based on a two-steps acceptance/rejection method:

- In order to simulate  $G^2$  given  $G^2 \geq (L-x)^2/t_0$  (here  $G$  stands for a standard gaussian variable), we use shifted exponentially distributed random variables with an optimal (classical) rejection rule.
- Using the acceptance/rejection procedure introduced in Algorithm (A3) we transform the distribution of  $(L-x)^2/G^2$  given  $(L-x)^2/G^2 \leq t_0$  into  $\tau_L$  given  $\tau_L \leq t_0$ .

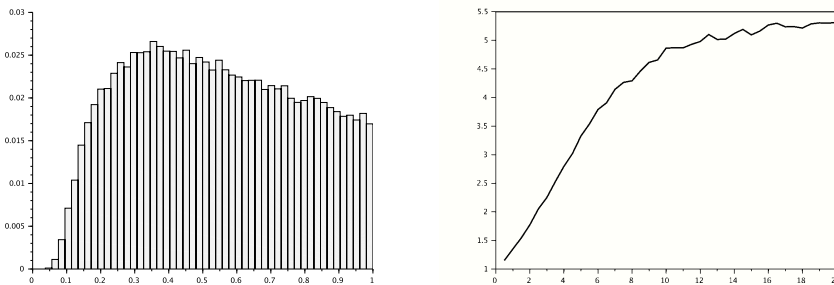


Figure 6: Histogram of the hitting time distribution using Algorithm (A3) for  $t_0 = 1$  and 100 000 simulations (left) and averaged number of iterations in Algorithm (A3) versus  $t_0$  (right) for  $X_0 = 0$ ,  $L = 1$  and 10 000 simulations.

In this illustration, we choose for the Poisson point process appearing in Algorithm (A3) the same kind of simulation as Algorithm (A1). Of course, it is also possible to adapt the procedure in order to follow the technics of Algorithm (A2) and finally to compare the efficiency of both Algorithms (we let this task to the reader).

### 5.3 Modified Algorithm (A1)<sup>ρ</sup>

In this section we consider diffusion processes  $(X_t, t \geq 0)$  satisfying (1.1) with unbounded drift terms  $b$ . We apply the modified algorithm described in Section 4 in order to approximate the first-passage time through the level  $L$ . Let us consider as example an Ornstein-Uhlenbeck process with  $b(x) = -\alpha x + \beta$  where  $\alpha$  and  $\beta$  are positive constants. We compute easily

$$\gamma(x) = \frac{1}{2}(-\alpha x + \beta)^2 - \frac{\alpha}{2}, \quad (5.5)$$

and observe that in general  $\gamma$  is neither a bounded function on the whole interval  $] -\infty, L]$  nor a positive one. The following choice of parameters  $\alpha = 0.3$ ,  $\beta = 1$  with starting position  $X_0 = 0$  and boundary  $L = 1$  ensures that  $\gamma$  is a positive function but  $b$  remains unbounded, that's why we replace the original drift term by its modified version (4.2):

$$b_\rho(x) = \begin{cases} -\alpha x + \beta & \text{if } -\rho \leq x \leq L, \\ \alpha\rho + \beta - \alpha(x + \rho)e^{x+\rho} & \text{if } x < -\rho. \end{cases} \quad (5.6)$$

The modified  $\gamma$  satisfies  $\gamma_\rho(x) = \gamma(x)$  for  $x \in [-\rho, L]$  and

$$\gamma_\rho(x) = \frac{1}{2}(\alpha\rho + \beta - \alpha(x + \rho)e^{x+\rho})^2 - \frac{\alpha}{2}(1 + x + \rho)e^{x+\rho} \quad \text{for } x < -\rho. \quad (5.7)$$

The function  $\gamma_\rho$  is now positive on the whole interval  $] -\infty, L]$  and admits the following upper-bound:

$$\kappa = \frac{1}{2} \left( \alpha\rho + \beta + \frac{\alpha}{e} \right)^2 + \frac{\alpha}{2e^2}.$$

To obtain such an expression, it suffices to upper-bound separately both main terms of (5.7). We can therefore apply Algorithm (A1) or Algorithm (A2) in order to simulate the approximated first-passage time  $\tau_L^\rho$ .

Figure 7 presents a comparison between the histograms of the hitting time distribution (10 000 independent simulations and  $\rho = 5$ ) with the approximation of the hitting time density obtained via the numerical algorithm proposed in [8]. We can notice that the histogram and the numerical *pdf* perfectly fit. Moreover, the choice of different larger values for  $\rho$  does not significantly affect the results.

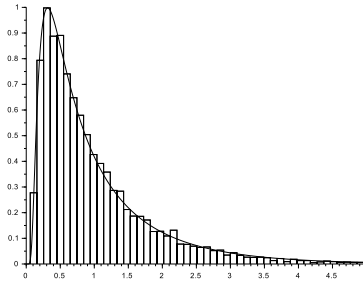


Figure 7: Histogram of the hitting time distribution for 10000 simulations corresponding to the level  $L = 1$ , starting position  $X_0 = 0$  and parameters  $a = 0.3$ ,  $b = 1$  for  $\rho = 5$  using Algorithm (A1) with modified drift, compared with the approximation of the hitting time density obtained via a numerical algorithm.

## References

- [1] *A Guide to First-Passage Processes*. Cambridge University Press, 2001.
- [2] P. Baldi and L. Caramellino. Asymptotics of hitting probabilities for general one-dimensional pinned diffusions. *Ann. Appl. Probab.*, 12(3):1071–1095, 2002.
- [3] E. Benedetto, L. Sacerdote, and C. Zucca. A first passage problem for a bivariate diffusion process: Numerical solution with an application to neuroscience when the process is gauss-markov. 242:41–52, April 2013.
- [4] A. Beskos, O. Papaspiliopoulos, and G.O. Roberts. Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, 12(6):1077–1098, 2006.
- [5] A. Beskos, O. Papaspiliopoulos, and G.O. Roberts. A factorisation of diffusion measure and finite sample path constructions. *Methodol. Comput. Appl. Probab.*, 10(1):85–104, 2008.
- [6] A. Beskos and G.O. Roberts. Exact simulation of diffusions. *The Annals of Applied Probability*, 15(4):2422–2444, 2005.
- [7] M. Broadie, P. Glasserman, and S. Kou. A continuity correction for discrete barrier options. *Math. Finance*, 7(4):325–349, 1997.
- [8] A. Buonocore, A.G. Nobile, and L.M. Ricciardi. A new integral equation for the evaluation of first-passage-time probability densities. *Advances in Applied Probability*, pages 784–800, 1987.
- [9] A.N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological cybernetics*, 95(1):1–19, 2006.
- [10] L. Devroye. *Nonuniform random variate generation*. Springer-Verlag, New York, 1986.

- [11] J. Durbin. The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Probab.*, 22(1):99–122, 1985.
- [12] J. Durbin. The first-passage density of the Brownian motion process to a curved boundary. *J. Appl. Probab.*, 29(2):291–304, 1992. With an appendix by D. Williams.
- [13] B. Ferebee. An asymptotic expansion for one-sided Brownian exit densities. *Z. Wahrsch. Verw. Gebiete*, 63(1):1–15, 1983.
- [14] V. Giorno, A. G. Nobile, L. M. Ricciardi, and S. Sato. On the evaluation of first-passage-time probability densities via nonsingular integral equations. *Adv. in Appl. Probab.*, 21(1):20–36, 1989.
- [15] M.T. Giraudo and L. Sacerdote. Simulation methods in neuronal modeling. *BioSystems*, 48:77–83, 1998.
- [16] M.T. Giraudo and L. Sacerdote. An improved technique for the simulation of first passage times for diffusion processes. *Comm. Statist. Simulation Comput.*, 28(4):1135–1163, 1999.
- [17] M.T. Giraudo, L. Sacerdote, and C. Zucca. A Monte Carlo method for the simulation of first passage times of diffusion processes. *Methodol. Comput. Appl. Probab.*, 3(2):215–231, 2001.
- [18] E. Gobet. Weak approximation of killed diffusion using Euler schemes. *Stochastic Process. Appl.*, 87(2):167–197, 2000.
- [19] E. Gobet and S. Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Process. Appl.*, 120(2):130–162, 2010.
- [20] S. Herrmann and E. Tanré. The first-passage time of the Brownian motion to a curved boundary: an algorithmic approach. *SIAM J. Sci. Comput.*, 38(1):A196–A215, 2016.
- [21] Q. Hu, Y. Wang, and X. Yang. The Hitting Time Density for a Reflected Brownian Motion. *Computational Economics*, 40(1):1–18, 2012.
- [22] T. Ichiba and C. Kardaras. Efficient estimation of one-dimensional diffusion first passage time densities via monte carlo simulation. *Journal of Applied Probability*, pages 699–712, 2011.
- [23] J. Janssen, O. Manca, and R. Manca. *Applied Diffusion Processes from Engineering to Finance*. Wiley, 2013.
- [24] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets*. Springer Finance. Springer-Verlag London, Ltd., London, 2009.
- [25] P. A. Jenkins. Exact simulation of the sample paths of a diffusion with a finite entrance boundary. *ArXiv e-prints*, November 2013.
- [26] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

- [27] J. Kent. Some probabilistic properties of Bessel functions. *Ann. Probab.*, 6(5):760–770, 1978.
- [28] V. Linetsky. Lookback options and diffusion hitting times: A spectral expansion approach. *Finance Stoch.*, 8:373398, 2004.
- [29] D.J. Navarro and I.G. Fuss. Fast and accurate calculations for first-passage times in wiener diffusion models. 53:222–230, 2009.
- [30] V. Pieper, M. Domin, and P. Kurth. Level crossing problems and drift reliability. *Math. Methods Oper. Res.*, 45:347–354, 1997.
- [31] J. Pitman and M. Yor. The law of the maximum of a Bessel bridge. *Electron. J. Probab.*, 4:no. 15, 35 pp. 1999.
- [32] K. Pötzelberger and L. Wang. Boundary crossing probability for Brownian motion. *J. Appl. Probab.*, 38(1):152–164, 2001.
- [33] L. M. Ricciardi, L. Sacerdote, and S. Sato. On an integral equation for first-passage-time probability densities. *J. Appl. Probab.*, 21(2):302–314, 1984.
- [34] L. Sacerdote and M.T. Giraudo. Stochastic integrate and fire models: a review on mathematical methods and their applications. In *Stochastic Biomathematical Models*, pages 99–148. Springer, 2013.
- [35] L. Sacerdote, M. Tamborrino, and C. Zucca. First passage times of two-dimensional correlated processes: Analytical results for the wiener process and a numerical method for diffusion processes. *Journal of Computation and Applied Mathematics*, 2016.
- [36] L. Sacerdote and F. Tomassetti. On evaluations and asymptotic approximations of first-passage-time probabilities. *Adv. in Appl. Probab.*, 28(1):270–284, 1996.
- [37] L. Wang and K. Pötzelberger. Boundary crossing probability for Brownian motion and general boundaries. *J. Appl. Probab.*, 34(1):54–65, 1997.
- [38] L. Wang and K. Pötzelberger. Crossing probabilities for diffusion processes with piecewise continuous boundaries. *Methodol. Comput. Appl. Probab.*, 9(1):21–40, 2007.
- [39] D. Williams. Path decomposition and continuity of local time for one-dimensional diffusions. I. *Proc. London Math. Soc. (3)*, 28:738–768, 1974.
- [40] C. Zucca and P. Tavella. The clock model and its relationship with the allan and related variances. 52(2):289–296, 2005.
- [41] C. Zucca, P. Tavella, and G. Peskir. Detecting atomic clock frequency trends using an optimal stopping method. 53(3):S89–S95, 2016.