



HAL
open science

General diffusion processes as the limit of time-space Markov chains

Alexis Anagnostakis, Antoine Lejay, Denis Villemonais

► **To cite this version:**

Alexis Anagnostakis, Antoine Lejay, Denis Villemonais. General diffusion processes as the limit of time-space Markov chains. 2020. hal-02897819v1

HAL Id: hal-02897819

<https://inria.hal.science/hal-02897819v1>

Preprint submitted on 12 Jul 2020 (v1), last revised 8 Jun 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

General diffusion processes as the limit of time-space Markov chains

Alexis Anagnostakis* Antoine Lejay[†] Denis Villemonais[‡]

July 12, 2020

Abstract

In this paper we prove the convergence of the law of grid-valued random walks, which can be seen as time-space Markov chains, to the law of a general diffusion process. This includes processes with sticky features, reflecting or absorbing boundaries and skew behavior. We show that for an arbitrary grid, the convergence occurs at a rate of order $1/4$ in terms of the maximum cell size of the grid for any p -Wasserstein distance. We also show that it is possible to achieve convergence rates of order $1/2$ if the grid is adapted to the speed measure of the diffusion, which is optimal. This result allows us to set up asymptotically optimal convergence schemes for general diffusion processes.

Finally, we give several examples where the quantities that determine the law of the random walk are non-tractable or semi-tractable and where the diffusion it approximates exhibits various singular features.

Keywords: diffusion process, Markov chain approximation, sticky diffusion, skew process, reflected process, Donsker theorem, slow reflection, Itô calculus, embeddable scheme, Wasserstein distance.

1 Introduction

In the diffusion process literature, the most well-studied and straightforward way to approximate diffusion processes is the Euler scheme. While this method works well when dealing with non-degenerate stochastic differential equations, it is not well defined for more general

*Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France, alexis.anagnostakis@univ-lorraine.fr

[†]Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France, antoine.lejay@univ-lorraine.fr

[‡]Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France, denis.villemonais@univ-lorraine.fr

diffusion processes. Examples of such processes are ones that exhibit sticky features, skew behavior or slowly reflecting boundaries.

First studied by Feller (see [14]), general diffusion processes are used to model a wide range of phenomena from semi-permeable layers in the case of skew diffusions (see [19]), principal agent problem dynamics in the case of sticky Brownian motion (see [24, 22]) to interest rates behavior close to 0 for slow reflection (see [21]). For more theoretical results and applications of general diffusions (see [20, 10, 6, 11, 17]). The resurgence of interest in these processes along with the need to have a universal method of simulation have motivated the search for new kind of approximation processes.

Several works aim at overcoming the shortcomings of the Euler scheme allowing us to approximate the law of more general diffusion processes. In [3], the author proposes to approximate the sticky Brownian motion with a simple random walk that stops for a fixed amount of time as it reaches 0. In [21], a continuous time Markov chain is used to approximate slowly reflected solutions of SDEs, where the jumping intensities are computed using a discretization of the infinitesimal generator of the diffusion. Another work where such a process is defined is [15], where the authors use a continuous time Markov chain to identify events in genomics evolution. An approximation scheme that deals with more general diffusions is given in [4], where a symmetric random walk with fixed time-step is proven to converge to the target process under a mild non-explosion condition.

In this paper, we prove the convergence in law of grid-valued random walks to any general diffusion process at an asymptotically optimal rate. This allows us to set up approximation schemes that, while make it straightforward to take into account for sticky points, can also be applied to any diffusion process that satisfy a mild non-explosion condition. This includes processes with boundary conditions like absorption, reflection or slow-reflection as well as the skew diffusions such as the Skew Brownian motion [20] and its generalizations. Thus, the values taken by the random walk correspond to values taken by the target process at random times, allowing us to classify it as an embeddable scheme along with [4] and [12]. We prove that for a grid adapted to the speed measure of the diffusion process, the law of the random walk converges at the rate of $1/2$ in terms of the maximum cell size. This convergence rate is optimal according the Donsker's invariance principle, as this is the rate a simple random walk converges to the standard Brownian motion (see [9]).

Besides the asymptotic optimal convergence rate, the usage of such an approximation process yield several advantages. Firstly, the static character of the grid makes the quantities that are involved in it good candidates for numerical approximation (see Sections 6 and 7). Moreover, this scheme makes it straightforward to take into account potential sticky points of the diffusion. At last, its universality is further validated by the fact that the Donsker's invariance principle [9, 3, 12] are all special cases of it.

We make heavy use of the speed measure/scale function characterization of diffusions (see [16, 23]). According to this formulation, the law of a one-dimensional diffusion process with state-space \mathbb{I} an open interval of \mathbb{R} is entirely determined by an increasing continuous function $s(x)$ and a positive locally finite measure $m(dx)$ defined on \mathbb{I} . The function $s(x)$ is called

scale function and is defined as the unique function¹ such that,

$$P_x(\tau_b < \tau_a) = \frac{s(x) - s(a)}{s(b) - s(a)},$$

for $a < x < b \in \mathbb{I}$ and with τ_a being the hitting time of a . The measure $m(dx)$ is called speed measure and is the unique positive locally finite measure such that,

$$E_x(\tau_{ab}) = \int_{(a,b)} G_{a,b}(x, y)m(dy),$$

for $a < x < b \in \mathbb{I}$ and where $G_{a,b}(x, y)$ is the Green function² defined as

$$G_{a,b}(x, y) = \begin{cases} \frac{(s(x) - s(a))(s(b) - s(y))}{s(b) - s(a)}, & \text{for } x \leq y, \\ \frac{(s(y) - s(a))(s(b) - s(x))}{s(b) - s(a)}, & \text{for } x > y. \end{cases} \quad (1)$$

We can also express the infinitesimal generator of the diffusion in terms of $m(dx)$ and $s(x)$ as,

$$\begin{aligned} L &= D^m D^s, \\ \text{dom}(L) &= \{f \in \mathcal{C}_b^0(\mathbb{I}, \mathbb{R}) : Lf \in \mathcal{C}_b^0(\mathbb{I}, \mathbb{R})\}, \end{aligned} \quad (2)$$

where D^m and D^s are defined as,

$$\begin{aligned} D^m f(x) &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{m(x, x+h]}, \\ D^s f(x) &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{s(x+h) - s(x)}. \end{aligned} \quad (3)$$

These results give us analytic formulations of quantities of the form $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$ for $k \in \mathbb{N}_0$ and to bound these quantities in terms of the size of the interval (a, b) . The latter will be particularly useful for proving the convergence results of the algorithm.

Outline. We begin the main part of the paper with Section 2 where we present the approximation scheme along with its properties. In Section 3 we give analytical characterizations of the quantities that determine the law of the random walk defined by the algorithm, allowing us to implement it. Section 4 is dedicated to proving the convergence of embedding times. In Section 5 we prove the main convergence result in terms of the maximum cell size of the grid. The case of solution of an SDE is studied in Section 6. Finally, in Section 7 we apply the scheme for several types of diffusions (reflection, stickiness, skew-behavior, singularity at 0 or ∞) and illustrate its convergence via histograms.

¹Unique up to an affine transformation.

²Alternative definition of Green function.

2 The Spatio-Temporal Markov Chain Approximation and its properties

2.1 The approximation scheme

In this section we define the approximation process for a one-dimensional diffusion process on natural scale³ with state space \mathbb{I} , an open interval of \mathbb{R} . The general case is obtained by a change of scale, as detailed in Section 2.2. Thus, let $(X_t)_{t \geq 0}$ be such a process defined on the filtered probability spaces $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (P_x)_{x \in I})$, where P_x is the law of $(X_t)_{t \geq 0}$ such that $P_x\{X_0 = x\} = 1$. Let $m(dx)$ be the speed measure of $(X_t)_{t \geq 0}$ and \mathbf{g} a grid on the state-space \mathbb{I} , which means that \mathbf{g} is an ordered subset of \mathbb{I} with no accumulation points within \mathbb{I} .

We define the time-space Markov chain $(\tilde{X}_t^{\mathbf{g}})_{t \geq 0}$ as the asymmetric random walk that has the following properties:

- its state-space is \mathbf{g} ,
- its initial value has the same distribution as the value of X_t the first time it touches the grid,
- it has the same transition probabilities and conditional expected transition times on \mathbf{g} as $(X_t)_{t \geq 0}$.

Thus, under P_x , if a and b are the closest lower and upper elements to x of \mathbf{g} ,

$$\tilde{X}_0 = \begin{cases} a, & \text{with probability } P_x(\tau_b < \tau_a), \\ b, & \text{with probability } P_x(\tau_a < \tau_b) = 1 - P_x(\tau_b < \tau_a). \end{cases} \quad (4)$$

For the rest of the trajectory, if τ_a is the hitting time of $a \in \mathbb{I}$ for the process $(X_t)_{t \geq 0}$, $\tau_{ab} := \tau_a \wedge \tau_b$ and $(T^{\mathbf{g}}(n))_{n \geq 0}$ are the consecutive jumping times of $(\tilde{X}_t^{\mathbf{g}})_{t \geq 0}$. Then, for all $k \in \mathbb{N}_0$ and $a < x < b$ adjacent points of \mathbf{g} ,

$$P(\tilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = b | \tilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x) = P_x(\tau_b < \tau_a), \quad (5)$$

and

$$T^{\mathbf{g}}(k+1) - T^{\mathbf{g}}(k) = \begin{cases} E_x(\tau_{ab} | \tau_b < \tau_a), & \text{on } \{\tilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = b\} \cap \{\tilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x\}, \\ E_x(\tau_{ab} | \tau_a < \tau_b), & \text{on } \{\tilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} = a\} \cap \{\tilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x\}. \end{cases} \quad (6)$$

As proved in Section 3, the quantities that appear on the right hand side of (5) and (6) are explicit functionals of the speed measure $m(dx)$.

Let \mathbf{c}_x be the cell of the grid \mathbf{g} containing x , i.e. the smallest open interval with endpoints in \mathbf{g} having x in its interior. We observe from (5), (6) and Bayes rule that for all cells of the grid $\mathbf{c}_x = (a, b)$ both $P(\tilde{X}_{T^{\mathbf{g}}(k+1)}^{\mathbf{g}} | \tilde{X}_{T^{\mathbf{g}}(k)}^{\mathbf{g}} = x)$ and $T^{\mathbf{g}}(k+1) - T^{\mathbf{g}}(k)$ depend only on x . Thus, if we define the following quantities on each cell of the grid,

$$\begin{aligned} p^+[x, (a, b)] &= P_x(\tau_b < \tau_a), & T^+[x, (a, b)] &= E_x[\tau_{ab} | \tau_b < \tau_a], \\ p^-[x, (a, b)] &= P_x(\tau_a < \tau_b), & T^-[x, (a, b)] &= E_x[\tau_{ab} | \tau_a < \tau_b], \end{aligned} \quad (7)$$

³Which means that $s(x) = x$.

we can simulate the random walk using Algorithm 1. We discuss in Section 3 on how to compute the quantities in (7). Practical examples are given in Sections 6 and 7.

Remark 1. In the case of sticky diffusions, where $m(dx) = m_c(dx) + \rho\delta_0(dx)$, the transition probabilities and transition times (7) can be directly inferred from the quantities of the diffusion without the sticky term. Indeed, Proposition 3.2 yields,

$$E_x[\tau_{ab}\mathbb{1}_{\tau_b < \tau_a}] = \int_{(a,b)} G_{a,b}(x, \zeta)v_0(\zeta)m_c(d\zeta) + \rho G_{a,b}(x, \theta)v_0(\theta).$$

Input : x initial value, T time horizon, \mathbf{g} grid on \mathbb{I}

Output: $(\widehat{X}[k])_k, (\widehat{t}[k])_k$

Initialization:

$\widehat{t}[0] = 0, n = 0$

$j = \arg \min\{|x_i - x|; x_i < x; i \in J\}$

$U \sim \text{Bernoulli}(p^+[x, (x_j, x_{j+1})])$

if $U == 1$ **then**

$j = j + 1$

end

$\widehat{X}[0] = x_j$

Algorithm:

while $\widehat{t}[n] < T$ **do**

$U \sim \text{Bernoulli}(p^+[x_j, (x_{j-1}, x_{j+1})])$

if $U == 1$ **then**

$j = j + 1$

$\widehat{t}[n + 1] = \widehat{t}[n] + T^+[x_j, (x_{j-1}, x_{j+1})]$

else

$j = j - 1$

$\widehat{t}[n + 1] = \widehat{t}[n] + T^-[x_j, (x_{j-1}, x_{j+1})]$

end

$\widehat{X}[n + 1] = x_j$

$n = n + 1$

end

STMCA^a Algorithm

^aSpace-Time Markov Chain Approximation

For the convergence, we make the further following assumption, that the diffusion process $(X_t)_{t \geq 0}$ satisfies the following non-explosion condition: There exist a constant $k_1 > 0$ such

that the speed measure of the diffusion process satisfies

$$m(dx) \geq k_1 \frac{1}{1+x^2} dx, \quad (8)$$

for all $x \in \mathbb{I}$. In Section 5, we prove that the law of the random walk $(\tilde{X}_t)_{t \geq 0}$ converges to the law of the target diffusion process for all p -Wasserstein distances, with $p \geq 1$. The p -Wasserstein distance \mathcal{W}_p between two laws μ and ν of processes with paths in a functional space $(\mathbb{X}, \|\cdot\|)$ is defined as,

$$\mathcal{W}_p[\mu, \nu] = \inf_{(\zeta, \xi) \sim \Gamma(\mu, \nu)} \left\| \|\zeta - \xi\| \right\|_{L^p}, \quad (9)$$

where by $\Gamma(\mu, \nu)$ we denote the collection of all measures with marginals μ and ν . Thus, if we take $(C^0([0, T], \mathbb{I}), \|\cdot\|_\infty)$ to be the functional space, we prove the following.

Theorem 2.1. *Let $(X_t)_{t \geq 0}$ be a diffusion process on natural scale whose speed measure satisfies Condition (8) for a constant $k_1 > 0$. Let also \mathbf{g} be a grid defined on the state-space \mathbb{I} of $(X_t)_{t \geq 0}$. Then, for all $p \geq 1$, $\delta \in (0, \frac{1}{4})$, $T > 0$ and $x \in \mathbb{I}$ there exists a positive constant C such that,*

$$\mathcal{W}_p \left[\text{Law} \left((\tilde{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \leq C |\mathbf{g}|_\star^\delta, \quad (10)$$

where $|\mathbf{g}|_\star = \sup_{\mathbf{c} \in \mathbf{g}} \{m(\mathbf{c})|\mathbf{c}|\}$.

Corollary 2.2. *The process $(\tilde{X}_t^{\mathbf{g}})_{t \in [0, T]}$ converges in law to $(X_t)_{t \in [0, T]}$ in $D([0, T], \mathbb{I})$ as $|\mathbf{g}|_\star \rightarrow 0$.*

Remark 2. In the case where $m(dx) \geq k_1 dx$, the constant⁴ $C > 0$ in Theorem 2.1 does not depend on the starting point of the diffusion.

2.2 Convergence rate for the general case

While the convergence results we established in the previous section were proven in the case of a diffusion process on natural scale, in this section we show how more general results can be inferred. Let $(X_t)_{t \geq 0}$ be a diffusion process with scale function $s(x)$ and speed measure $m(dx)$. We assume that s belongs to the Sobolev space $W^{1,1}(\mathbb{I})$. From Theorem 8.2 of [7] and since s is continuous, we have

$$s(x) - s(y) = \int_y^x s'(t) dt,$$

for all y, x in \mathbb{I} , and we assume that there exist $k_1 > 0$ and $k_2 \in \{0, 1\}$,

$$m(dx) \geq k_1 \frac{s'(x)}{1 + k_2(s(x))^2} dx. \quad (11)$$

⁴This results in the bound of Theorem 2.1 not depending on the starting point of the diffusion.

We also assume that the inverse of s is α -Hölder continuous, i.e. there exists a constant $C > 0$ such that for all $x \neq y \in \mathbb{I}$,

$$\frac{|s^{-1}(\bar{x}) - s^{-1}(\bar{y})|}{|\bar{x} - \bar{y}|^\alpha} \leq C. \quad (12)$$

Given a grid \mathbf{g} , we consider the random walk $\tilde{X}_t^{\mathbf{g}}$ defined by Algorithm 1, where the transition probabilities and transition times in (7) can be computed using the formulas derived in Section 3. We obtain the following Corollary of Theorem 2.1.

Corollary 2.3. *Let $(X_t)_{t \geq 0}$ be a diffusion process with scale function and speed measure satisfying the above conditions. Let also \mathbf{g} be a grid defined on the state-space \mathbb{I} of $(X_t)_{t \geq 0}$. Then, for all $p \geq 1$, $\delta \in (0, \frac{1}{4})$, $T > 0$ and $x \in \mathbb{I}$ there exists positive constants C_1 , and C_2 such that*

$$\mathcal{W}_p \left[\text{Law} \left((\tilde{X}_t^{\mathbf{g}})_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \leq C_1 (|\mathbf{g}|_\star^\delta + e^{-C_2/|\mathbf{g}|_\star}),$$

where $|\mathbf{g}|_\star = \sup_{\mathbf{c} \in \mathbf{g}} \{ |s(\mathbf{c})| m(\mathbf{c}) \}$.

Proof. We define the proxy process $(Y_t)_{t \geq 0}$ as $Y_t = s(X_t)$ which has scale function $s_Y(x) = x$ and speed measure $m_Y(dx) = m \circ s^{-1}(dx)$. From condition (11) and a change of variables, we get that $(Y_t)_{t \geq 0}$ satisfies condition (8) for the same constants k_1 and k_2 . We also define $\tilde{Y}_t^{s(\mathbf{g})}$ evolving according to Algorithm 1, with grid $s(\mathbf{g}) = \{s(x); x \in \mathbf{g}\}$. It can be defined on the canonical space of $(X_t)_{t \geq 0}$ so that $s(\tilde{X}_t^{\mathbf{g}}) = \tilde{Y}_t^{s(\mathbf{g})}$ almost surely. Thus, Condition (12) implies that,

$$|X_t - \tilde{X}_t^{\mathbf{g}}| \leq C |Y_t - \tilde{Y}_t^{s(\mathbf{g})}|^{\alpha'}. \quad (13)$$

Along with the fact that,

$$|\mathbf{g}|_\star = \sup_{\mathbf{c} \in \mathbf{g}} \{ |s(\mathbf{c})| m_Y(s(\mathbf{c})) \}, \quad (14)$$

Theorem 2.1 implies Corollary 2.3. □

2.3 Grid tuning

We observe that in the case of a Brownian motion this yields a convergence rate of $\mathcal{O}(|\mathbf{g}|^{1/2})$ which is optimal from Donsker's invariance principle (see [9]). While this is a result we would like to have for all diffusion processes, the following example illustrates that this is not the case. We show how we can remediate to this by using a custom grid and extrapolate this method to the general case via Corollary 2.5.

Example 2.4. Let X_t be a Brownian motion with a sticky point at 0, which is the process having scale function and speed measure

$$s(x) = x, \quad m(dx) = 2 dx + \rho \delta_0(dx),$$

with $\rho > 0$ being its stickiness parameter. From Theorem 2.1, the random walk defined by the approximation scheme converges at a rate of $\mathcal{O}(|\mathbf{g}|_\star^{1/4})$ which is $\mathcal{O}(|\mathbf{g}|^{1/4})$ as,

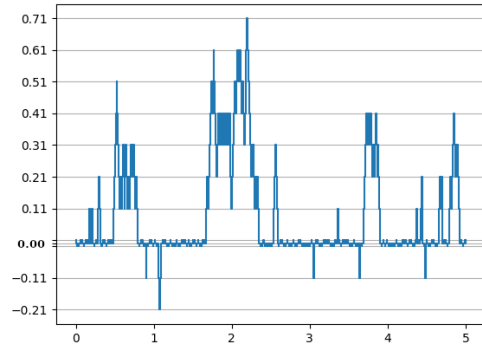
$$\rho|\mathbf{c}_0| \leq \sup_{x \in \mathbf{g}} \{m(c_x)|\mathbf{c}_x|\},$$

so for a uniform grid of step size h ,

$$2h\rho \leq |\mathbf{g}|_\star.$$



(a) Approximation process of a Brownian motion on the time interval $[0, 5]$ with a sticky point at 0 of stickiness $\rho = 4.0$ using a uniform grid \mathbf{g} with step size $h = 0.1$



(b) Approximation process of a Brownian motion on the time interval $[0, 5]$ with a sticky point at 0 of stickiness $\rho = 4.0$ using the tuned grid described in (15) with $h = 0.1$.

Figure 1: Sticky Brownian motion trajectory approximation.

This means that there are functionals of the trajectory for which the convergence rate is much slower for this process in comparison with a standard Brownian motion. In order to remediate to this, we propose a preliminary step to the approximation scheme that involves finding a grid that is “adapted” to the speed measure of the process. In the case of the Brownian motion with a sticky point at 0, such a grid can be defined as one that has uniform non-adjacent cells to 0 of size h and with $\mathbf{c}_0 = (-h^2/2\rho, h^2/2\rho)$, i.e.

$$\mathbf{g} = \left\{ \bigcup_{k \in \mathbb{N}_0} \left\{ -\frac{h^2}{2\rho} - k\frac{h}{2} \right\} \right\} \cup \{0\} \cup \left\{ \bigcup_{k \in \mathbb{N}_0} \left\{ \frac{h^2}{2\rho} + k\frac{h}{2} \right\} \right\}. \quad (15)$$

As the approximation process is a random walk, for every k steps it makes, it spends $\mathcal{O}(\sqrt{k})$ steps on the cell containing 0. Thus, running the algorithm on either grid yields the same algorithmic complexity.

The general case is covered by the following Corollary.

Corollary 2.5. *Let \mathbf{g} be a grid such that $|\mathbf{g}|_\star \leq C|\mathbf{g}|^2$, then we can bound the p -Wasserstein distance between the laws of $(\tilde{X}_t^\mathbf{g})_{t \in [0, T]}$ and $((X_t)_{t \in [0, T]})$ in Theorem 2.1 by $|\mathbf{g}|^{2\delta}$ instead of $|\mathbf{g}|_\star^\delta$. Thus, the law of the random walk converges at a rate of $\mathcal{O}(|\mathbf{g}|^{1/2})$ instead of $\mathcal{O}(|\mathbf{g}|^{1/4})$.*

2.4 The case of diffusions with boundary conditions

When presenting the results and the structure of the scheme, we considered only processes where $s(I)$ is an open set, thus considering diffusion with unreachable boundaries. Our results also adapt to the situation where either $s(\ell)$ and/or $s(r)$ are reachable, and in this case some adjustments are needed, depending on the nature of finite boundaries and on the condition at regular boundaries. In order to keep the presentation simple, we assume that the process is on natural scale and that $\mathbb{I} = [0, +\infty)$ (the adaptation to $\mathbb{I} = (l, r]$ or $\mathbb{I} = [l, r]$ or $I = [l, r)$ with $\ell = \mathbb{R} \cup \{-\infty\}$ and $r \in \mathbb{R} \cup \{+\infty\}$ is straightforward).

It is well known (see for instance Section 5.11 of Itô's book [18]) that the finite boundary 0 can be of four types. Setting, for some fixed $c > 0$,

$$\mathcal{I} = \int \int_{0 < y < x < c} m(dx) dy, \quad \mathcal{II} = \int \int_{0 < y < x < c} m(dy) dx,$$

then

- 0 is an exit boundary if $\mathcal{I} < \infty$ and $\mathcal{II} = \infty$,
- 0 is a regular boundary if $\mathcal{I} < \infty$ and $\mathcal{II} < \infty$,
- 0 is a natural boundary if $\mathcal{I} = \infty$ and $\mathcal{II} = \infty$,
- 0 is an entrance boundary if $\mathcal{I} = \infty$ and $\mathcal{II} < \infty$.

The entrance type been excluded for a finite boundary of a diffusion process on natural scale, and the natural type been considered in the settings of Theorem 2.1, this leaves us with two possible types for the boundary 0: exit or regular. If 0 is an exit boundary, then the diffusion process X is absorbed at the boundary 0. If 0 is a reflecting boundary, then the diffusion process can be either absorbed or reflected at 0. In these cases, the convergence result of Theorem 2.1 can be extended by considering a grid \mathbf{g} on \mathbb{I} containing 0 and by adapting the dynamics of $\tilde{X}^\mathbf{g}$ as follows. The dynamic of $\tilde{X}^\mathbf{g}$ is the same as in Algorithm 1, up to the time when it reaches 0, then:

- if 0 is an absorbing boundary (exit or regular), then the result can be immediately extended by stopping $\tilde{X}^\mathbf{g}$ when it reaches 0;
- if 0 is a reflecting regular boundary, then the process $\tilde{X}^\mathbf{g}$ jumps from 0 to $b := \min \mathbf{g} \setminus \{0\}$ with probability 1 and after a time $\int_{[0, b)} (b - \zeta) m(d\zeta)$. We emphasize that in this configuration, 0 may be a sticky boundary (i.e. with $m(0) \in (0, +\infty)$).

The proof of the convergence in these situations is omitted here, since it is a straightforward adaptation of the proof of Theorem 2.1, using in particular the fact that, in the case of a reflecting boundary,

$$E_0[\tau_b] = \int_{[0,b)} (b - \zeta) m(d\zeta).$$

The case of killing boundaries, and in general of a process with non-zero killing measure, leads to additional non-trivial difficulties. Devising an algorithm and a similar result as Theorem 2.1 remains an active area of research.

2.5 Markovian embedding

The consecutive values of the process $(\tilde{X}_t)_{t \geq 0}$ defined in form a Markov chain with, by construction, the same transition probabilities as $(X_t)_{t \geq 0}$ on \mathbf{g} . We define the embedding times of $(X_t)_{t \geq 0}$ in \mathbf{g} as,

$$\begin{aligned} \tau_0^{\mathbf{g}} &= 0, \\ \tau_k^{\mathbf{g}} &= \inf \{t > \tau_{k-1}^{\mathbf{g}} : X_t \in \mathbf{g} \setminus \{X_{\tau_{k-1}^{\mathbf{g}}}\}\}, \quad \forall k \geq 1. \end{aligned} \tag{16}$$

As both $\tilde{X}_{T^{\mathbf{g}}(n)}$ and $X_{\tau_n^{\mathbf{g}}}$ are both Markov chains with the same transition probabilities with \tilde{X}_0 forced to be equal in law to $X_{\tau_1^{\mathbf{g}}}$ (see Section 2.1), the following equality in law holds,

$$\text{Law}(\tilde{X}_{T^{\mathbf{g}}(n)}; n \geq 0) = \text{Law}(X_{\tau_n^{\mathbf{g}}}; n \geq 1). \tag{17}$$

We define $K^{\mathbf{g}}(t)$ as the inverse of $T^{\mathbf{g}}(n)$, i.e.,

$$K^{\mathbf{g}}(t) = \inf \left\{ n \in \mathbb{N} : \sum_{k=1}^n \mathbb{E} [\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}] > t \right\}. \tag{18}$$

Thus, we get the following Proposition.

Proposition 2.6. *Let X_t be a diffusion process, \mathbf{g} a grid defined over its state space \mathbb{I} and $\tilde{X}_t^{\mathbf{g}}$ be the approximation process defined in (5) and (6). Then, if $\tau_n^{\mathbf{g}}$ are the embedding times of X_t in \mathbf{g} , the following equality in law holds,*

$$\text{Law}(\tilde{X}_t; t \geq 0) = \text{Law}(X_{\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}}; t \geq 0),$$

where $K^{\mathbf{g}}(t)$ is a random index defined defined in (18).

3 Moment characterization of conditional exit times

The law of the approximation process defined in the previous section was shown to be determined by its transition probabilities $P_x(\tau_b < \tau_a)$ and conditional transition times

$\mathbb{E}_x(\tau_{ab}|\tau_b < \tau_a)$. In this section we show that quantities of the form $v_k(x) = \mathbb{E}_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$ yield an integral formulation with respect to the speed measure of the diffusion and involving the scale function (we do not assume that the diffusion is on natural scale in the present section). We also show that this results in them being solutions to Dirichlet problems where the differential operator is the infinitesimal generator L of the diffusion. This allows us to simulate such processes via Algorithm 1 and thus to approximate the law of the target diffusion process $(X_t)_{t \geq 0}$.

In terms of Algorithm 1, we need to compute for three adjacent points a, x, b of the grid the quantities

$$v_0(x) = \mathbb{P}_x(\tau_b < \tau_a), \quad v_1(x) = \mathbb{E}_x(\tau_{ab} \mathbb{1}_{\tau_b < \tau_a}) \quad \text{and} \quad \bar{v}_1(x) = \mathbb{E}_x(\tau_{ab} \mathbb{1}_{\tau_a < \tau_b}).$$

The quantities of (7) are then

$$\begin{aligned} p^+[x, (a, b)] &= v_0(x), & T^+[x, (a, b)] &= \frac{v_1(x)}{v_0(x)}, \\ p^-[x, (a, b)] &= 1 - v_0(x), & T^-[x, (a, b)] &= \frac{\bar{v}_1(x)}{1 - v_0(x)}. \end{aligned} \tag{19}$$

Proposition 3.1. *The function $v_0(x) = \mathbb{P}_x(\tau_b < \tau_a)$ is solution to the problem with Dirichlet boundary conditions*

$$\begin{cases} L u = 0, & x \in (a, b), \\ u(a) = 0, \\ u(b) = 1, \end{cases} \tag{20}$$

where L is the infinitesimal generator of $(X_t)_{t \geq 0}$, which also implies that $v_0 \in \text{dom}(L)$.

Proof. Let $x \in (a, b)$, from the definition of the scale function and the factorization of the infinitesimal generator $L = D^m D^s$

$$L v_0 = D^m D^s \frac{s(\cdot) - s(a)}{s(b) - s(a)} = D^m \frac{1}{s(b) - s(a)}.$$

which equals 0 as $m(dx)$ is a positive measure. As v_0 and $L v_0 = 0$ are both functions in C_b^0 , we deduce that $v_0 \in \text{dom}(L)$ and $L v_0 = 0$. Under P_b the stopping time $\tau_b = 0$ a.s. and as the process has a.s. continuous trajectories, $\tau_a > 0$ a.s., i.e.

$$v_0(b) = \mathbb{P}_b(\tau_b < \tau_a) = \mathbb{P}_b(0 < \tau_a) = 1.$$

This, along with the symmetrical argument, allow us to retrieve the boundary conditions of (20). \square

Proposition 3.2. *Let $v_n(x) = \mathbb{E}_x(\tau_{ab}^n \mathbb{1}_{\tau_b < \tau_a})$ for $n \in \mathbb{N}$. Then,*

$$v_n(x) = n \int_{(a,b)} G_{a,b}(x, \zeta) v_{n-1}(\zeta) m(d\zeta). \tag{21}$$

Proof. Since $\int_0^{\tau_{ab}} (\tau_{ab} - t)^{n-1} dt = \tau_{ab}^n/n$,

$$v_n(x) = n \mathbb{E}_x \left[\mathbb{1}_{\tau_b < \tau_a} \int_0^{\tau_{ab}} (\tau_{ab} - t)^{n-1} dt \right] = n \mathbb{E}_x \left[\mathbb{1}_{\tau_b < \tau_a} \int_0^\infty \mathbb{1}_{t \leq \tau_{ab}} (\tau_{ab} - t)^{n-1} dt \right].$$

By conditioning on \mathcal{F}_t , as $\mathbb{1}_{t \leq \tau_{ab}}$ is \mathcal{F}_t -measurable and from Markov property,

$$v_n(x) = n \mathbb{E}_x \left[\int_0^\infty \mathbb{1}_{t \leq \tau_{ab}} \mathbb{E} \left[\mathbb{1}_{\tau_b < \tau_a} (\tau_{ab} - t)^{n-1} | \mathcal{F}_t \right] dt \right] = n \mathbb{E}_x \left[\int_0^{\tau_{ab}} E_{X_t} \left[\mathbb{1}_{\tau_b < \tau_a} \tau_{ab}^{n-1} \right] dt \right].$$

The equality (21) results by applying directly Green's formula. \square

Proposition 3.3. *The function $v_k(x) = \mathbb{E}_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$ is solution to the problem with Dirichlet boundary conditions*

$$\begin{cases} \mathbb{L} u = -k v_{k-1}, & x \in (a, b), \\ u(a) = 0, \\ u(b) = 0. \end{cases} \quad (22)$$

Lemma 3.4. *Let $g(x) = \int_{(a,b)} G_{a,b}(x, y) f(y) m(dy)$, where $f \in C_b^0(a, b)$ and $G_{a,b}(x, y)$ is the Green function defined in (1). Then $g \in \text{dom}(\mathbb{L})$ and*

$$\mathbb{L} g(x) = -f(x), \quad \forall x \in (a, b).$$

Proof. Let $x \in (a, b)$. Using the $D^m D^s$ factorization of \mathbb{L} and the dominated convergence theorem we get

$$\begin{aligned} \mathbb{L} g(x) &= D^m D^s \int_{(a,b)} \left[\mathbb{1}_{y \leq x} \frac{(s(y) - s(a))(s(b) - s(x))}{s(b) - s(a)} + \mathbb{1}_{y > x} \frac{(s(x) - s(a))(s(b) - s(y))}{s(b) - s(a)} \right] f(y) m(dy) \\ &= -D^m \int_{(a,x]} v_0(y) f(y) m(dy) + D^m \int_{(x,b)} (1 - v_0(y)) f(y) m(dy) \\ &= -v_0(x) f(x) - (1 - v_0(x)) f(x) = -f(x). \end{aligned}$$

The continuity of g is a consequence of monotone convergence theorem for integrals. Moreover as $G_{a,b}(x, y)$ is bounded by $s(b) - s(a)$, $m(dx)$ is locally finite and f is bounded, g is also bounded. So, we deduce that $f \in \text{dom}(\mathbb{L})$ and that on (a, b) we have $\mathbb{L} g = -f$. \square

Proof of Proposition 3.3. From Proposition 3.2,

$$v_k(x) = \int_{(a,b)} G_{a,b}(x, y) k v_{k-1}(y) m(dy).$$

As $v_0 \in C_b^0(a, b)$, from Lemma 3.4 we deduce iteratively that $v_k \in C_b^0(a, b)$ for all $k \in \mathbb{N}$ and that $\mathbb{L} v_k = -k v_{k-1}$ on (a, b) . For the boundary conditions we have that $\tau_{ab} = \tau_a \wedge \tau_b = 0$ a.s. for P_a and P_b , so for $k \geq 1$

$$v_k(a) = \mathbb{E}_a(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a}) \leq \mathbb{E}_a(\tau_{ab}^k) = 0.$$

With the same argument we show that $v_k(b) \leq 0$ and as they are obviously positive quantities $v_k(a) = v_k(b) = 0$. \square

4 Convergence of the embedding times

In order to prove the convergence of the process $(\tilde{X}_t^g)_{t \geq 0}$ we need to control quantities of the form $E_x[\sup_{t \leq T} |t - \tau_{K(t, \mathbf{g})}^g|^p]$, where τ_k^g are the embedding times of the process $(X_t)_{t \geq 0}$ in the grid \mathbf{g} . In this section we show the existence of such bounds in terms of $|\mathbf{g}|_\star$ which is a measure on the grid defined as

$$|\mathbf{g}|_\star = \sup_{\mathbf{c} \in \mathbf{g}} \{m(\mathbf{c})|\mathbf{c}|\},$$

where by $|\mathbf{c}|$ we denote the Lebesgue measure of the cell \mathbf{c} and by $\mathbf{c} \in \mathbf{g}$ we denote the cells \mathbf{c} of the grid \mathbf{g} which are the intervals of endpoints on the grid that have exactly one element of the grid in their interior. For example if $\{x_j\}_{j \in J}$ is an ordered grid, a typical cell will have the form (x_{j-1}, x_{j+1}) . We could consider a bound on $m(\mathbf{c})$ for $\mathbf{c} \in \mathbf{g}$ and bound the quantities of interest in terms of $|\mathbf{g}| = \sup_{\mathbf{c} \in \mathbf{g}} \{|\mathbf{c}|\}$, but in doing so we do not track correctly the convergence rates of processes like the standard Brownian motion where $|\mathbf{g}|_\star = |\mathbf{g}|^2$. Moreover as shown in Section 2.3, having such bounds give us a direct way to adapt the grid to the speed measure in order to accelerate the convergence rate of the scheme.

4.1 Bounds on the conditional moments of the exit times

Lemma 4.1. *Let $v_k(x) = E_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$, then for all $k \in \mathbb{N}$*

$$\left\| \frac{v_k}{v_{k-1}} \right\|_\infty \leq k|\mathbf{g}|_\star.$$

Proof. We first observe that

$$G_{a,b}(x, y) \frac{v_0(y)}{v_0(x)} = \begin{cases} \frac{(y-a)(b-x)}{(b-a)} \frac{y-a}{x-a}, & x > y, \\ \frac{(x-a)(b-y)}{(b-a)} \frac{y-a}{x-a}, & x \leq y. \end{cases}$$

As for $x > y$ the ratio $\frac{y-a}{x-a} < 1$,

$$G_{a,b}(x, y) \frac{v_0(y)}{v_0(x)} \leq \begin{cases} \frac{(y-a)(b-x)}{(b-a)}, & x > y, \\ \frac{(y-a)(b-y)}{(b-a)}, & x \leq y, \end{cases}$$

which is bounded by $(b-a)$ in both cases as $x, y \in (a, b)$.

From Proposition 3.2, $v_k(x) = k \int_{(a,b)} G_{a,b}(x, y) v_{k-1}(y) m(dy)$. So, by induction,

$$\begin{aligned}
v_k(x) &= k! \int_{(a,b)} G_{a,b}(x, x_k) \int_{(a,b)} G_{a,b}(x_k, x_{k-1}) \dots \\
&\quad \dots \int_{(a,b)} G_{a,b}(x_2, x_1) v_0(x_1) m(dx_1) \dots m(dx_k) \\
&= k! \int_{(a,b)} G_{a,b}(x, x_k) \int_{(a,b)} G_{a,b}(x_k, x_{k-1}) \dots \\
&\quad \dots \int_{(a,b)} G_{a,b}(x_3, x_2) v_0(x_2) \left(\int_{(a,b)} G_{a,b}(x_2, x_1) \frac{v_0(x_1)}{v_0(x_2)} m(dx_1) \right) m(dx_2) \dots m(dx_k) \\
&\leq k!(b-a)m((a,b)) \int_{(a,b)} G_{a,b}(x, x_k) \int_{(a,b)} G_{a,b}(x_k, x_{k-1}) \dots \\
&\quad \dots \int_{(a,b)} G_{a,b}(x_3, x_2) v_0(x_2) m(dx_2) \dots m(dx_k) \\
&= k(b-a)m((a,b))v_{k-1}(x).
\end{aligned}$$

Since $(b-a)m((a,b)) \leq |\mathbf{g}|_\star$, we get the desired result on v_k/v_{k-1} . \square

Corollary 4.2. *Let $k \in \mathbb{N}$, then we have the following bound for $v_k(x) = \mathbb{E}_x(\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a})$,*

$$\|v_k\|_\infty \leq k! |\mathbf{g}|_\star^k.$$

Lemma 4.3. *Let X_t be a diffusion process with speed measure $m(dx)$, $a < x < b$ elements of the state space \mathbb{I} , $\mathbf{c}_x = (a, b)$, $|\mathbf{c}_x|_\star = m(\mathbf{c}_x)|b-a|$ and $\lambda > 0$ such that $\lambda|\mathbf{c}_x|_\star \in (0, 1)$. Then,*

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) \leq \exp\left(\frac{\lambda}{1-\lambda|\mathbf{c}_x|_\star} \mathbb{E}_x(\tau_{ab})\right). \quad (23)$$

Proof. Developing the exponential series,

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) = 1 + \sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \mathbb{E}_x(\tau_{ab}^N) = 1 + \mathbb{E}_x(\tau_{ab}) \sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \frac{\mathbb{E}_x(\tau_{ab}^N)}{\mathbb{E}_x(\tau_{ab})}.$$

From Corollary 4.2 we can bound the ratio of expected values by $(N-1)!|\mathbf{c}_x|_\star^{N-1}$ and as $\lambda|\mathbf{c}_x|_\star \in (0, 1)$,

$$\mathbb{E}_x(e^{\lambda\tau_{ab}}) \leq 1 + \mathbb{E}_x(\tau_{ab}) \sum_{N=1}^{\infty} \frac{\lambda}{N} (\lambda|\mathbf{c}_x|_\star)^{N-1} \leq 1 + \mathbb{E}_x(\tau_{ab}) \frac{\lambda}{1-\lambda|\mathbf{c}_x|_\star}.$$

Thus, we need only to apply the inequality $1+x \leq e^x$ to get (23). \square

Lemma 4.4. *Let $t, M > 0$ and $\lambda > 0$ such that $\lambda|\mathbf{g}|_\star \in (0, 1)$. Then,*

$$\mathbb{P}_x(\tau_{K\mathbf{g}(t)}^\mathbf{g} > M) \leq e^{-\lambda M} e^{\lambda \frac{t+|\mathbf{g}|_\star}{1-\lambda|\mathbf{g}|_\star}}. \quad (24)$$

Proof. From Markov's inequality,

$$\mathbb{P}_x(\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} > M) \leq e^{-\lambda M} \mathbb{E}_x \left[e^{\lambda \tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}}} \right] = e^{-\lambda M} \mathbb{E}_x \left[e^{\lambda \sum_{k=1}^{K^{\mathbf{g}}(t)} (\tau_k - \tau_{k-1})} \right]. \quad (25)$$

Conditioning on the σ -algebra \mathcal{B} generated by the trajectory of X_t on the grid \mathbf{g} , i.e. $\mathcal{B} = \sigma\{X_{\tau_k}; k \in \mathbb{N}_0\}$, and as $K^{\mathbf{g}}(t)$ is \mathcal{B} -measurable,

$$\mathbb{E}_x \left[e^{\lambda \sum_{k=1}^{K^{\mathbf{g}}(t)} (\tau_k - \tau_{k-1})} \right] = \mathbb{E}_x \left[\prod_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E} \left[e^{\lambda(\tau_k - \tau_{k-1})} \middle| \mathcal{B} \right] \right]. \quad (26)$$

From the definition of $|\mathbf{g}|_*$, $\lambda|\mathbf{c}|_* \leq \lambda|\mathbf{g}|_* \in (0, 1)$ for each cell \mathbf{c} of the grid \mathbf{g} . Thus, applying Lemma 4.3 on each term of the product in (26),

$$\mathbb{P}_x(\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} > M) \leq e^{-\lambda M} \mathbb{E}_x \left[\exp \left(\frac{\lambda}{1 - \lambda|\mathbf{g}|_*} \sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E}(\tau_k - \tau_{k-1} | \mathcal{B}) \right) \right]$$

From the definition of $K^{\mathbf{g}}(t)$, we have that $\sum_{k=1}^{K^{\mathbf{g}}(t)} \mathbb{E}(\tau_k - \tau_{k-1} | \mathcal{B}) \leq t + \|v_1/v_0\|_\infty \leq t + |\mathbf{g}|_*$, thus proving (24). \square

4.2 Convergence of the embedding times

Lemma 4.5. *Let $(X_t)_{t \geq 0}$ be a diffusion process, $\tau_k^{\mathbf{g}}$ its embedding times on a grid \mathbf{g} and $K^{\mathbf{g}}(t)$ as defined in (18). Then,*

$$\sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \leq 2|\mathbf{g}|_*(T + |\mathbf{g}|_*).$$

Proof. For all $x \in \mathbf{g}$, let \mathbf{c}_x be the cell of \mathbf{g} containing x . Then,

$$\begin{aligned} & \sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var} \left(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right) \\ &= \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}})^2 \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] - \left(\mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \right)^2 \\ &\leq \sup_{k \leq K^{\mathbf{g}}(t)} \left\{ \frac{\mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}})^2 \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]}{\mathbb{E} \left[\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]} \right\} \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}) \middle| X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right]. \end{aligned}$$

So from Lemma 4.1 and the definition of $K^{\mathbf{g}}(t)$,

$$\begin{aligned} \sum_{k=1}^{K^{\mathbf{g}}(T)} \text{Var}(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}) \\ \leq \left\| \frac{v_2}{v_1} \right\|_{\infty} \sum_{k=1}^{K^{\mathbf{g}}(T)} \mathbb{E} \left[(\tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}) | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}} \right] \\ \leq \left\| \frac{v_2}{v_1} \right\|_{\infty} \left(T + \left\| \frac{v_1}{v_0} \right\|_{\infty} \right) \leq 2|\mathbf{g}|_{\star} (T + |\mathbf{g}|_{\star}), \end{aligned}$$

which is the desired inequality. \square

Proposition 4.6. *Let \mathcal{F}_t be the canonical filtration of the diffusion process $(X_t)_{t \geq 0}$ and $\tau_k^{\mathbf{g}}$ the embedding times of $(X_t)_{t \geq 0}$. Let also $\mathcal{A}_n = \mathcal{F}_{\tau_n^{\mathbf{g}}}$, $\mathcal{B} = \sigma((X_{\tau_k^{\mathbf{g}}})_{k \in \mathbb{N}_0})$ and $\Delta\tau_k^{\mathbf{g}} = \tau_k^{\mathbf{g}} - \tau_{k-1}^{\mathbf{g}}$. If we define the augmented filtration $\mathcal{G}_n = \mathcal{A}_n \vee \mathcal{B}$, then the process*

$$M_n = \sum_{k=1}^n \Delta\tau_k^{\mathbf{g}} - \mathbb{E} [\Delta\tau_k^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}],$$

is a \mathcal{G}_n -martingale.

Proof. Let $m \leq n$, then

$$\begin{aligned} \mathbb{E} [M_n | \mathcal{G}_m] &= \mathbb{E} [M_n | \mathcal{A}_m, \mathcal{B}] \\ &= \mathbb{E} \left[\sum_{k=1}^m \Delta\tau_k^{\mathbf{g}} - \mathbb{E} [\Delta\tau_k^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}] \right. \\ &\quad \left. + \sum_{k=m+1}^n \Delta\tau_k^{\mathbf{g}} - \mathbb{E} [\Delta\tau_k^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}] \middle| \mathcal{A}_m, \mathcal{B} \right]. \end{aligned}$$

From Markov's property and as $\tau_k^{\mathbf{g}}$ is \mathcal{A}_k measurable, we have

$$\mathbb{E} [M_n | \mathcal{G}_m] = M_m + \sum_{k=m+1}^n \mathbb{E}_{X_{\tau_m^{\mathbf{g}}}} \left[\Delta\tau_k^{\mathbf{g}} - \mathbb{E} [\Delta\tau_k^{\mathbf{g}} | X_{\tau_{k-1}^{\mathbf{g}}}, X_{\tau_k^{\mathbf{g}}}] \middle| \mathcal{B} \right] = M_m.$$

This proves the result. \square

Theorem 4.7. *Let X_t be a diffusion process, \mathbf{g} a grid on the state space \mathbb{I} and $\tau_k^{\mathbf{g}}$ the embedding times of X_t in \mathbf{g} . Then, for a time horizon $T > 0$, the following inequality holds*

$$\mathbb{E} \left[\sup_{t \in [0, T]} |\tau_{K^{\mathbf{g}}(t)}^{\mathbf{g}} - t|^2 \right] \leq 2|\mathbf{g}|_{\star} (4(T + |\mathbf{g}|_{\star}) + 1), \quad (27)$$

where $K^{\mathbf{g}}(t)$ is defined in (18).

Proof. From Minkowski's inequality, we have $(a + b)^p \leq 2^{p-1}(a^p + b^p)$, which yields for $p = 2$

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} |\tau_{K^g(t)}^g - t|^2 \right] &\leq 2 \mathbb{E} \left[\sup_{t \in [0, T]} \left| \tau_{K^g(t)}^g - \sum_{k=1}^{K^g(t)} \mathbb{E} [\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}] \right|^2 \right] \\ &\quad + 2 \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{k=1}^{K^g(t)} \mathbb{E} [\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}] - t \right|^2 \right]. \end{aligned} \quad (28)$$

The term $\sum_{k=1}^n \Delta \tau_k^g - \mathbb{E} [\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}]$ is a \mathcal{G}_n -martingale from Proposition 4.6, where $\mathcal{G}_n = \mathcal{F}_{\tau_n^g} \vee \mathcal{B}$ and $\mathcal{B} = \sigma((X_{\tau_k^g})_{k \in \mathbb{N}_0})$. Thus, from Doob's L^p inequality,

$$\mathbb{E} \left[\sup_{t \in [0, T]} |M_{K^g(t)}|^2 \right] = \mathbb{E} \left[\sup_{k \leq K^g(T)} |M_k|^2 \right] \leq 2 \mathbb{E} [|M_{K^g(T)}|^2]. \quad (29)$$

By conditioning on \mathcal{B} ,

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{k=1}^{K^g(t)} \Delta \tau_k^g - \mathbb{E} [\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}] \right|^2 \right] \\ \leq 2 \mathbb{E} \left[\left(\sum_{k=1}^{K^g(T)} \Delta \tau_k^g - \mathbb{E} [\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}] \right)^2 \right] \\ = 2 \mathbb{E} \left[\sum_{k=1}^{K^g(T)} \text{Var} (\Delta \tau_k^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}) \right], \end{aligned} \quad (30)$$

which from Lemma 4.5 is bounded by $4|\mathbf{g}|_\star(T + |\mathbf{g}|_\star)$. For the second term on the right hand side of (28) we have from the definition of $K^g(t)$,

$$\begin{aligned} \sum_{k=1}^{K^g(t)} \mathbb{E} [\tau_k^g - \tau_{k-1}^g | X_{\tau_{k-1}^g}, X_{\tau_k^g}] - t &\leq \mathbb{E} [\tau_{K^g(t)+1}^g - \tau_{K^g(t)}^g | X_{\tau_{K^g(t)-1}^g}, X_{\tau_{K^g(t)}^g}] \\ &\leq \left\| \frac{v_1}{v_0} \right\|_\infty \leq |\mathbf{g}|_\star. \end{aligned} \quad (31)$$

So having bounded both additive parts of the right hand side of (28), we get (27). \square

5 Convergence rate of the Markov chain

5.1 Moment bounds

Lemma 5.1. *Let X_t be a diffusion process on natural scale with a speed measure $m_X(dx)$ that satisfies Condition (8). Then, for all $p \geq 2$, there exist two constants $C, C' > 0$ such that*

$$\mathbb{E}_x \left[\sup_{t \in [0, T]} |X_t - x|^p \right] \leq C' [1 + |x|^p] e^{CT}. \quad (32)$$

Moreover, $C \leq 2p(p-1)/k_1$, where k_1 is such that Condition (8) is satisfied and $C' > 0$ is a constant that depends only on p .

Proof. Let $(Z_t)_{t \geq 0}$ be the diffusion process on natural scale with speed measure,

$$m_Z(dx) = \mathbb{1}_{|x| < 1} \frac{k_1}{2} dx + \mathbb{1}_{|x| \geq 1} \frac{k_2}{2x^2} dx.$$

We note that $\frac{k_1}{2x^2} \leq \frac{k_1}{1+x^2}$ for all $x \geq 1$. The dynamic of the process Z_t can be shown to be

$$dZ_t = \begin{cases} \frac{2}{\sqrt{k_1}} dB_t, & \text{for } |Z_t| < 1, \\ \frac{2}{\sqrt{k_1}} Z_t dB_t, & \text{for } |Z_t| \geq 1, \end{cases}$$

where B_t is a standard Brownian motion. We also suppose that $Z_0 = x$ a.s. for P_x . As X_t and Z_t are on natural scale, they can be expressed as time-changed Brownian motion, i.e. $X_t = B_{\gamma_X(t)}$ and $Z_t = W_{\gamma_Z(t)}$, where B_t and W_t are two standard Brownian motions⁵ with local times $L_t^x(B)$ and $L_t^x(W)$ and with $\gamma_X(t)$ and $\gamma_Z(t)$ being the generalized inverses⁶ of $A_X(t) = \frac{1}{2} \int_{\mathbb{I}} L_t^x(B) m_X(dx)$ and $A_Z(t) = \frac{1}{2} \int_{\mathbb{I}} L_t^x(W) m_Z(dx)$ respectively. It is straightforward that, using the same underlying Brownian motion in these definitions, we have $A_Z(t) \leq A_X(t)$, and hence $\gamma_X(t) \leq \gamma_Z(t)$. Thus,

$$\begin{aligned} \mathbb{E}_x \left[\sup_{t \in [0, T]} |X_t - x|^p \right] &= \mathbb{E}_x \left[\sup_{t \in [0, \gamma_X(T)]} |B_t - x|^p \right] \\ &\leq \mathbb{E}_x \left[\sup_{t \in [0, \gamma_Z(T)]} |B_t - x|^p \right] = \mathbb{E}_x \left[\sup_{t \in [0, T]} |Z_t - x|^p \right], \end{aligned}$$

which is itself bounded by $\frac{p}{p-1} \mathbb{E}_x [|Z_T - x|^p]$ from Doob's L^p inequality. From Minkowski's inequality,

$$\mathbb{E}_x [|Z_T - x|^p] \leq 2^{p-1} \left(\mathbb{E}_x [|Z_T|^p] + |x|^p \right). \quad (33)$$

Using Itô's formula, followed by a classical localization argument, Fatou's Lemma along with the standard dominated convergence theorem, one obtains that for all $q > 2/3$,

$$\begin{aligned} &\mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{\{|Z_t| \geq 1\}} + 1 \right] \\ &\leq (|x| - 1)^{3q} \mathbb{1}_{\{|x| \geq 1\}} + 1 + \int_0^t \frac{3q}{2} (3q - 1) \mathbb{E}_x \left[(|Z_s| - 1)^{3q-2} \mathbb{1}_{\{|Z_s| \geq 1\}} \frac{4Z_s^2}{k_1} \right] ds. \end{aligned}$$

⁵We will suppose that $P_x(B_0 = x) = P_x(W_0 = x) = 1$.

⁶The generalized inverse of a function f is given by,

$$f^{-1}(x) = \inf\{\zeta \geq 0 : f(\zeta) > x\}.$$

Using the inequality $|x - 1|^{3q-2}x^2 \leq |x - 1|^{3q} + 1$ for all $x \geq 1$,

$$\begin{aligned} & \mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{\{|Z_t| \geq 1\}} + 1 \right] \\ & \leq (|x| - 1)^{3q} \mathbb{1}_{\{|x| \geq 1\}} + 1 + \int_0^t \frac{6q}{c} (3q - 1) \mathbb{E}_x \left[(|Z_s| - 1)^{3q} \mathbb{1}_{\{|Z_s| \geq 1\}} + 1 \right] ds. \end{aligned}$$

Using Gronwall's Lemma, we deduce that, for all $t \geq 0$,

$$\mathbb{E}_x \left[(|Z_t| - 1)^{3q} \mathbb{1}_{\{|Z_t| \geq 1\}} + 1 \right] \leq [(|x| - 1)^{3q} \mathbb{1}_{\{|x| \geq 1\}} + 1] e^{6q(3q-1)t/k_1}.$$

Let $C'_q > 0$ be a constant such that $(|x| - 1)^{3q} \mathbb{1}_{\{|x| \geq 1\}} + 1 \geq C'_q |x|^{3q}$ for all $x \in \mathbb{R}$. Thus,

$$\mathbb{E}_x \left[|Z_t|^{3q} \right] \leq C'_q [(|x| - 1)^{3q} \mathbb{1}_{\{|x| \geq 1\}} + 1] e^{6q(3q-1)t/k_1}, \quad \forall t \geq 0. \quad (34)$$

Hence, for all⁷ $p > 2$, we have

$$\mathbb{E}_x \left[|Z_t|^p \right] \leq C'_{p/3} [1 + |x|^p] e^{2p(p-1)t/k_1}, \quad \forall t \geq 0.$$

Replacing (34) in (33), we get the bound (32) for $C' = \frac{p}{p-1} 2^{p-1} (C'_{p/3} + 1)$ and $C = 2p(p-1)/k_1$. This proves the result. \square

From Lemma 5.1 and using the same arguments as in the proof of Theorem 3.1 in [5], we derive the same result with a sharper constant.

Proposition 5.2. *If X_t is a diffusion process with a speed measure $m_X(dx)$ that satisfies Condition (8). Then, for each $T > 0$ and $\gamma \in (0, \frac{1}{2})$, there exists a constant $C > 0$ such that*

$$\left\| \sup_{s \neq t \leq T} \frac{|X_t - X_s|}{|t - s|^\gamma} \right\|_{L^p(\mathbb{P}_x)} \leq C(1 + |x|).$$

5.2 Proof of the convergence rate for a process on natural scale

Proof of Theorem 2.1. From Proposition 2.6,

$$\begin{aligned} & \mathcal{W}_p \left[\text{Law} \left((\tilde{X}_t^g)_{t \in [0, T]} \right), \text{Law} \left((X_t)_{t \in [0, T]} \right) \right] \\ & = \inf_{\zeta \sim \text{Law} \left((\tilde{X}_t^g)_{t \in [0, T]} \right), \xi \sim \text{Law} \left((X_t)_{t \in [0, T]} \right)} \left\| \|\zeta - \xi\|_{L^p} \right\| \\ & \leq \left\| \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t| \right\|_{L^p(\mathbb{P}_x)}. \end{aligned}$$

⁷The extension to $p = 2$ is straightforward.

Let $M > T$, then from Minkowski inequality,

$$\begin{aligned} & \left\| \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \quad + \left\| \mathbb{1}_{\tau_{K^g}^g(T) > M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \end{aligned} \quad (35)$$

For the first additive term of (35), by multiplying and dividing by $|\tau_{K^g}^g(t) - t|^\gamma$,

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & = \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} \left\{ \frac{|X_{\tau_{K^g}^g(t)} - X_t|}{|\tau_{K^g}^g(t) - t|^\gamma} |\tau_{K^g}^g(t) - t|^\gamma \right\} \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} \left\{ \frac{|X_{\tau_{K^g}^g(t)} - X_t|}{|\tau_{K^g}^g(t) - t|^\gamma} \right\} \sup_{t \in [0, T]} |\tau_{K^g}^g(t) - t|^\gamma \right\|_{L^p(\mathbb{P}_x)} \end{aligned}$$

As $\tau_{K^g}^g$ is increasing with respect to t ,

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left\| \sup_{s \neq t \leq M} \left\{ \frac{|X_s - X_t|}{|s - t|^\gamma} \right\} \sup_{t \in [0, T]} |\tau_{K^g}^g(t) - t|^\gamma \right\|_{L^p(\mathbb{P}_x)} \\ & = \left(\mathbb{E}_x \left[\sup_{s \neq t \leq M} \frac{|X_s - X_t|}{|s - t|^\gamma} \sup_{t \in [0, T]} |\tau_{K^g}^g(t) - t|^\gamma \right]^p \right)^{1/p}. \end{aligned}$$

Using Hölder's inequality for $q \geq 1$ and $q/(q-1)$ conjugates exponents and then applying Kolmogorov's continuity theorem, there exists a constant $C_1 = C_1(M, \gamma, p(q-1)/q, x) > 0$ such that,

$$\begin{aligned} & \left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \\ & \leq \left(\mathbb{E}_x \left[\sup_{s \neq t \leq M} \frac{|X_s - X_t|}{|s - t|^\gamma} \right]^{p(q-1)/q} \right)^{q/p(q-1)} \left(\mathbb{E}_x \left[\sup_{t \in [0, T]} |\tau_{K^g}^g(t) - t|^\gamma \right]^{pq} \right)^{1/pq} \\ & \leq C_1^{q/p(q-1)} \left(\mathbb{E}_x \left[\sup_{t \in [0, T]} |\tau_{K^g}^g(t) - t|^\gamma \right]^{pq} \right)^{1/pq}. \end{aligned}$$

Choosing q so that $\gamma pq = 2$ and from Theorem 4.7,

$$\left\| \mathbb{1}_{\tau_{K^g}^g(T) \leq M} \sup_{t \in [0, T]} |X_{\tau_{K^g}^g(t)} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C_1^{1/p(1-\gamma p)} (2|\mathbf{g}|_\star (4T + |\mathbf{g}|_\star))^{\gamma/2}. \quad (36)$$

For the second additive term of (35),

$$\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g > M} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] = \sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right], \quad (37)$$

where for each term of the sum from Hölder's inequality,

$$\begin{aligned} & \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \\ & \leq \left[\mathbb{P}_x(\tau_{K^g(T)}^g \in [m, m+1)) \right]^{1/q'} \left[\mathbb{E}_x \left[\sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \right]^{q'/(q'-1)} \\ & \leq \left[\mathbb{P}_x(\tau_{K^g(T)}^g > m) \right]^{1/q'} \left[\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g < m+1} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \right]^{q'/(q'-1)}. \end{aligned}$$

As each m in the sum in (37) satisfies $m+1 > M > T$, from Minkowski's inequality,

$$\begin{aligned} & \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g < m+1} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \\ & \leq \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g < m+1} 2^{p-1} \sup_{t \in [0, T]} \left\{ |X_{\tau_{K^g(t)}^g} - x|^p + |X_t - x|^p \right\} \right] \\ & \leq 2^p \mathbb{E}_x \left[\sup_{t \in [0, m+1]} |X_t - x|^p \right], \end{aligned}$$

which from Lemma 5.1 is bounded by $C_2 [1 + |x|^p] e^{C_3(m+1)}$, where C_2 and C_3 are positive constants depending on p . This, along with Lemma 4.4 and Hölder's inequality gives us for a $\lambda > 0$ chosen such that $\alpha = \lambda |\mathbf{g}|_* < 1$,

$$\mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \leq \left[C_2 [1 + |x|^{pq'}] e^{C_3(m+1)} \right]^{1/q'} \left[e^{-\lambda m} e^{\lambda \frac{T + |\mathbf{g}|_*}{1 - \lambda |\mathbf{g}|_*}} \right]^{(q'-1)/q'}$$

where $C_2(pq') \leq 2pq'(pq' - 1)/c$ and $C_2(pq') > 0$ a positive constant depending only on pq' .

Thus, setting $C_4(pq') := \left[C_2(pq') [1 + |x|^{pq'}] \right]^{1/q'} e^{C_3/q'} > 0$,

$$\begin{aligned} & \mathbb{E}_x \left[\mathbb{1}_{\tau_{K^g(T)}^g \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K^g(t)}^g} - X_t|^p \right] \\ & \leq C_4 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \left(\lambda \frac{T + |\mathbf{g}|_*}{1 - \lambda |\mathbf{g}|_*} - \lambda m \right) \right] \right\} \\ & = C_4 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \frac{\alpha}{|\mathbf{g}|_*} \left(\frac{T + |\mathbf{g}|_*}{1 - \alpha} - m \right) \right] \right\} \\ & = C_5 \exp \left\{ \frac{1}{q'} \left[C_3 m + (q' - 1) \frac{\alpha}{|\mathbf{g}|_*} \left(\frac{T}{1 - \alpha} - m \right) \right] \right\}, \end{aligned}$$

where $C_5 := C_4 \exp(\frac{q'-1}{q'} \frac{\alpha}{1-\alpha})$. If we choose $q' > 1$ such that⁸ $A := C_3 - (q' - 1)\alpha/|\mathbf{g}|_\star < 0$,

$$\begin{aligned} \sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \\ \leq C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_\star} \right\} \sum_{m=M}^{\infty} (e^A)^m \\ = C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_\star} \right\} \frac{e^{AM}}{1 - e^A} \\ = C_5 \exp \left\{ \frac{(q' - 1)\alpha T}{(1 - \alpha)|\mathbf{g}|_\star} \right\} \frac{\exp \left\{ \left(C_3 - \frac{(q'-1)\alpha}{|\mathbf{g}|_\star} \right) M \right\}}{1 - \exp \left\{ C_3 - \frac{(q'-1)\alpha}{|\mathbf{g}|_\star} \right\}}. \end{aligned}$$

If M is chosen such that $M > T/(1 - \alpha)$, then

$$\sum_{m=M}^{\infty} \mathbb{E}_x \left[\mathbb{1}_{\tau_{K\mathbf{g}(T)}^{\mathbf{g}} \in [m, m+1)} \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t|^p \right] \leq \frac{C_5 e^{C_3 M}}{1 - e^A} \exp \left\{ \frac{(q' - 1)\alpha}{|\mathbf{g}|_\star} \left(\frac{T}{1 - \alpha} - M \right) \right\}, \quad (38)$$

where the bound is $O(e^{-1/|\mathbf{g}|_\star})$, and can be rewritten as $C^{(1)} e^{-C^{(2)}/|\mathbf{g}|_\star}$. Thus, from (35), (36) and (38),

$$\left\| \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C_1^{1/p(1-\gamma p)} (2|\mathbf{g}|_\star (4T + |\mathbf{g}|_\star))^{\gamma/2} + C^{(1)} e^{-C^{(2)}/|\mathbf{g}|_\star}. \quad (39)$$

As $|\mathbf{g}|_\star^\gamma$ and $e^{-1/|\mathbf{g}|_\star}$ are both $\mathcal{O}(|\mathbf{g}|_\star^{\gamma/2})$, there exists a constant $C > 0$, such that,

$$\left\| \sup_{t \in [0, T]} |X_{\tau_{K\mathbf{g}(t)}^{\mathbf{g}}} - X_t| \right\|_{L^p(\mathbb{P}_x)} \leq C |\mathbf{g}|_\star^{\gamma/2}, \quad (40)$$

which is (10). □

6 Computations for the classical SDE case and beyond

While Theorem 2.1 and Corollary 2.3 guarantee the convergence for all diffusion processes, most such processes encountered in practice are solutions to SDE with sticky points and/or boundary conditions. In this section we show how to compute the quantities (7) in the pure SDE case, allowing us via Algorithm 1 to simulate all aforementioned processes. Thus, let X_t be the solution of an SDE with time-homogeneous coefficients,

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t,$$

⁸We remark that if this is satisfied for a grid \mathbf{g} , then it is satisfied for all grids \mathbf{g}' such that $|\mathbf{g}'|_\star \leq |\mathbf{g}|_\star$.

where $\mu(x)$ and $\sigma(x)$ are real measurable function and B_t is a standard Brownian motion. A straightforward computation using Itô's formula gives us the infinitesimal generator of X_t ,

$$(\mathbf{L}, \text{dom}(\mathbf{L})) = \begin{cases} \mathbf{L} f(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x), & \forall f \in \text{dom}(\mathbf{L}), \\ \text{dom}(\mathbf{L}) = \mathcal{C}^2(\mathbb{R}). \end{cases} \quad (41)$$

From the analytical expression of the infinitesimal generator (41) we can infer,

$$s'(x) = e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} \text{ and } m(dx) = \frac{1}{s'(x)} \frac{2}{\sigma^2(x)} dx. \quad (42)$$

Thus, from Proposition 3.1,

$$v_0(x) = \frac{\int_a^x e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}{\int_a^b e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}. \quad (43)$$

Replacing the expression of $m(dx)$ in (21),

$$\begin{aligned} v_1(x) &= \int_a^b G_{a,b}(x, \zeta) v_0(\zeta) \frac{1}{s'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta, \\ \bar{v}_1(x) &= \int_a^b G_{a,b}(x, \zeta) (1 - v_0(\zeta)) \frac{1}{s'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta. \end{aligned} \quad (44)$$

The scale functions and the speed measures are defined up to additive and multiplicative constant: for $\alpha \in \mathbb{R}$ and $\lambda > 0$, the pairs (s, m) and $(\alpha + \lambda S, \lambda^{-1}m)$ are associated to the same diffusion. In particular, as we are only concerned with points $x \in [a, b]$, we could use v_0 for the scale function. The speed measure shall be adapted accordingly. For $x, \zeta \in [a, b]$, the Green function in (1) takes the simpler form,

$$G_{a,b}(x, \zeta) = v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)).$$

Thus, expressions (44) become

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) \frac{1 - v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2(\zeta)} d\zeta. \end{aligned} \quad (45)$$

The quantities (43) and (44) or (45) can be computed analytically as we do for the Ornstein-Uhlenbeck process in Section 7.4 or through a numerical approximation procedure as for the Cox-Ingersoll-Ross process in Section 7.4.

7 Examples

7.1 Standard Brownian motion

The standard Brownian motion can be defined as the diffusion process with scale function and speed measure

$$s(x) = x \quad \text{and} \quad m(dx) = 2 dx.$$

Let $v_k(x) := E_x[\tau_{ab}^k \mathbb{1}_{\tau_b < \tau_a}]$ and $\bar{v}_k(x) := E_x[\tau_{ab}^k \mathbb{1}_{\tau_a < \tau_b}]$ for all $k \in \mathbb{N}_0$, where $\mathbf{c}_x = (a, b)$. Then, from the definition of the scale function,

$$v_0(x) = \frac{x - a}{b - a}.$$

From Proposition 3.2,

$$v_1^{BM}(x) = (x - a)(b - x) \left(\frac{2(x - a)^2}{3(b - a)^2} + \frac{b - x}{b - a} - \frac{2(b - x)^2}{3(b - a)^2} \right),$$

$$\bar{v}_1^{BM}(x) = (x - a)(b - x) \left(\frac{2(b - x)^2}{3(b - a)^2} + \frac{x - a}{b - a} - \frac{2(x - a)^2}{3(b - a)^2} \right),$$

where by *BM* we mean these quantities are associated with the Brownian motion. Thus from (7), we have all the necessary quantities we need to implement the algorithm.

7.2 Sticky Brownian motion

The Brownian motion with a sticky point at 0 where the stickiness parameter is $\rho > 0$ can be defined as the diffusion process with scale function and speed measure

$$s(x) = x \quad \text{and} \quad m(dx) = 2 dx + \rho \delta_0(dx).$$

From the definition of the scale function,

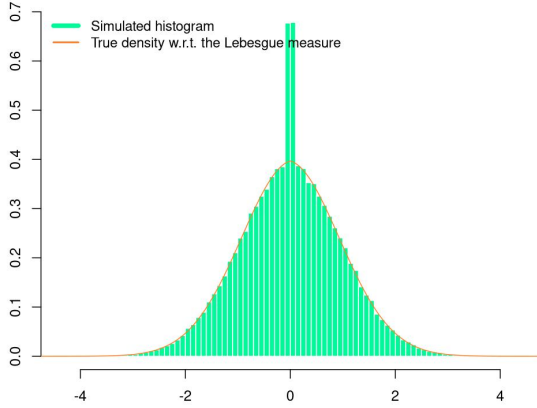
$$v_0(x) = \frac{x - a}{b - a}.$$

From Proposition 3.2 we may deduce the following expressions for the conditional exit times of $\mathbf{c}_x = (a, b)$,

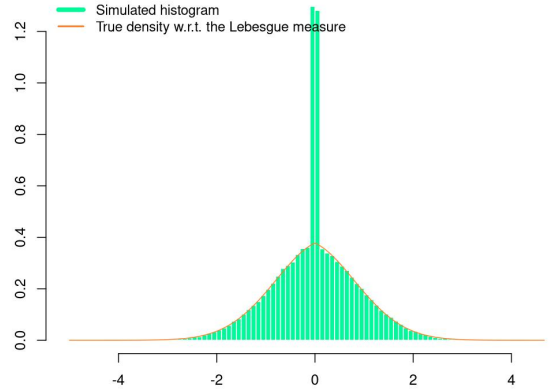
$$v_1(x) = v_1^{BM}(x) + \rho \mathbb{1}_{\{0 \in (a, b)\}} G_{(a, b)}(x, 0) v_0(0),$$

$$\bar{v}_1(x) = \bar{v}_1^{BM}(x) + \rho \mathbb{1}_{\{0 \in (a, b)\}} G_{(a, b)}(x, 0) (1 - v_0(0)),$$

where $v_1^{BM}(x)$ and $\bar{v}_1^{BM}(x)$ are the analogous quantities for the standard Brownian motion.

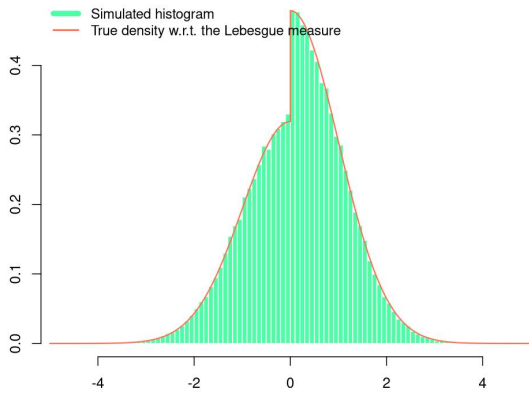


(a) Histogram of a sticky Brownian motion with stickiness parameter $\rho = 0.3$ at $T = 1$ using Algorithm 1 with a tuned grid with parameter $h = 0.01$.

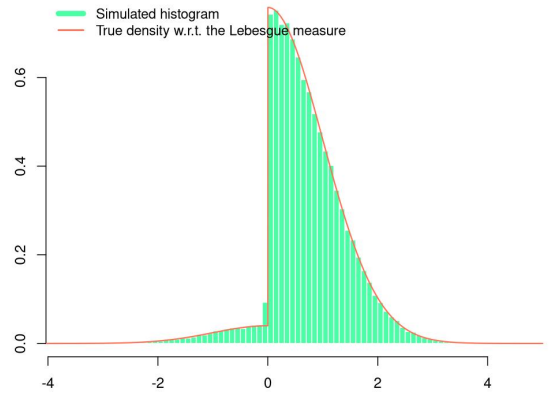


(b) Histogram of a sticky Brownian motion with stickiness parameter $\rho = 1.0$ at $T = 1$ using Algorithm 1 with a tuned grid with parameter $h = 0.01$.

Figure 2: Sticky Brownian motion approximation.



(a) Histogram of a skew Brownian motion with skew parameter $\beta = 0.6$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.



(b) Histogram of a skew Brownian motion with skew parameter $\beta = 0.95$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

Figure 3: Skew Brownian motion approximation.

7.3 Skew Brownian motion

The skew Brownian motion at 0 with parameter $\beta \in (0, 1)$ can be defined as the diffusion process with scale function and speed measure

$$s(x) = \begin{cases} \frac{x}{\beta}, & x \geq 0, \\ \frac{x}{1-\beta}, & x \leq 0 \end{cases} \quad \text{and} \quad m(dx) = \begin{cases} 2\beta dx, & x > 0, \\ 2(1-\beta) dx, & x < 0. \end{cases}$$

From the definition of the scale function,

$$v_0(x) = \frac{s(x) - s(a)}{s(b) - s(a)}.$$

As the β and $(1 - \beta)$ terms between the speed measure and the scale function compensate themselves in the expressions giving v_k in Proposition 3.2,

$$v_1(x) = v_1^{BM}(x) \text{ and } \bar{v}_1(x) = \bar{v}_1^{BM}(x).$$

7.4 Ornstein-Uhlenbeck process

Let X_t be an Ornstein-Uhlenbeck with mean reversion force $\theta > 0$, long-term mean μ and diffusion parameter $\sigma > 0$, i.e.

$$dX_t = \theta(\mu - X_t) dt + \sigma dB_t,$$

where B_t is a standard Brownian motion. We will see that v_0 can be expressed in terms of the Gaussian imaginary error function as $s(x) = \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - x)\right)$,

$$v_0(x) = \frac{\int_a^x e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy}{\int_a^b e^{-\int_a^y \frac{2\mu(\zeta)}{\sigma^2(\zeta)} d\zeta} dy} = \frac{\operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - x)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)}{\operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - b)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)},$$

where $\operatorname{erfi}(x) = \frac{2}{\pi} \int_0^x e^{t^2} dt = \frac{2}{\pi} e^{x^2} D_+(x)$, with $D_+(x)$ being Dawson function⁹. Thus, for an Ornstein-Uhlenbeck process (44) becomes,

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) v_0(\zeta) c \exp\left(\frac{\theta(\zeta - \mu)^2}{\sigma^2}\right) \frac{2}{\sigma^2} dx, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta) (1 - v_0(x \vee \zeta)) (1 - v_0(\zeta)) c \exp\left(\frac{\theta(\zeta - \mu)^2}{\sigma^2}\right) \frac{2}{\sigma^2} dx, \end{aligned} \tag{46}$$

where $c = \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - b)\right) - \operatorname{erfi}\left(\sqrt{\frac{\theta}{\sigma^2}}(\mu - a)\right)$.

7.5 Cox-Ingersol-Ross process

The Cox-Ingersol-Ross process or CIR process [8], introduced first by W. Feller [13], is the diffusion that solves the SDE

$$dX_t = \theta(\mu - X_t) dt + \sigma\sqrt{X_t} dB_t, \tag{47}$$

⁹We use the latter representation in our numerical results.

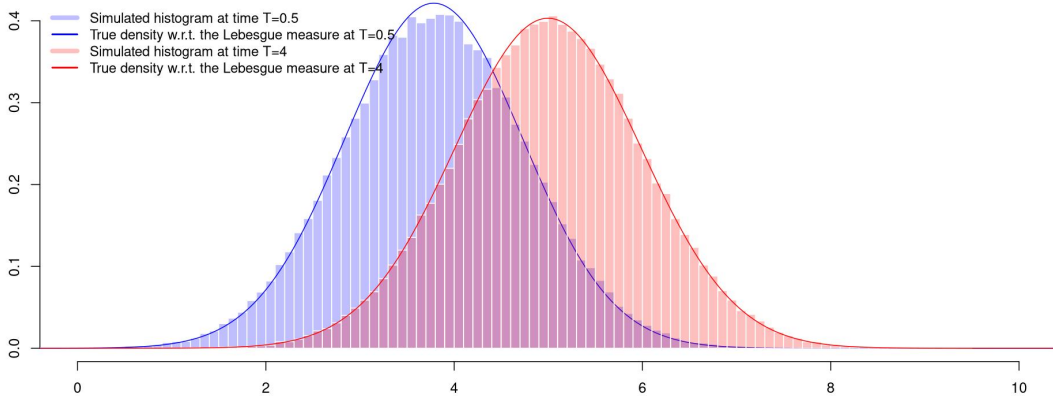


Figure 4: Histogram of an Ornstein-Uhlenbeck process with parameters $(\theta, \mu, \sigma) = (1, 8, 1)$ with an initial value of $x_0 = 0$ at times $T = 0.5$ and $T = 4$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

where B_t is a standard Brownian motion. The parameter $\theta > 0$ expresses its mean reversion speed, μ its long term speed and $\sigma > 0$ is its diffusivity parameter. The parameter $\theta > 0$ expresses its mean reversion speed, μ its long term speed and $\sigma > 0$ is its diffusivity parameter. This equation has a diffusion coefficients that degenerates at 0. It however remains non-negative given $X_0 \geq 0$ and almost surely never hit 0 when $2\theta\mu > \sigma^2$. A large body of work have been devoted to the simulation of the CIR and related process, see e.g. [1].

From (43) and (45),

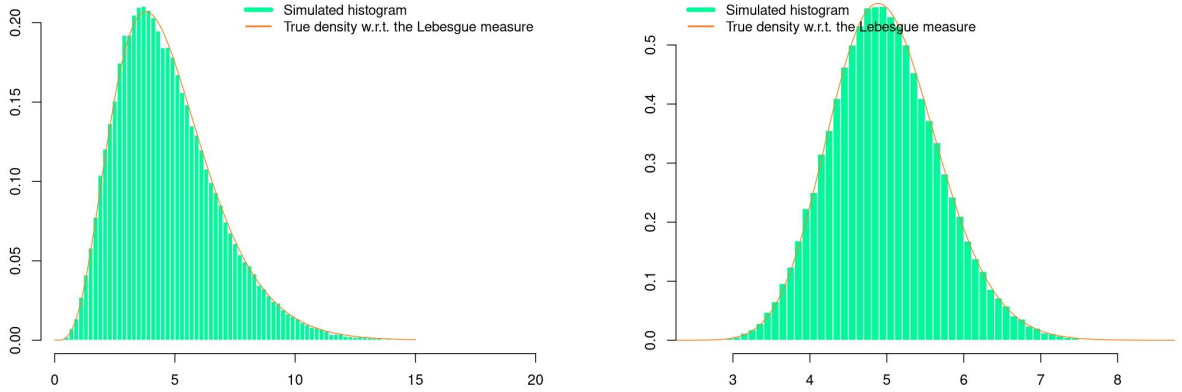
$$v_0(x) = \frac{\int_a^x y^{-\frac{2\theta\mu}{\sigma^2}} e^{\frac{2\theta}{\sigma^2}y} dy}{\int_a^b y^{-\frac{2\theta\mu}{\sigma^2}} e^{\frac{2\theta}{\sigma^2}y} dy} \quad (48)$$

and

$$\begin{aligned} v_1(x) &= \int_a^b v_0(x \wedge \zeta)(1 - v_0(x \vee \zeta)) \frac{v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2 \zeta} d\zeta, \\ \bar{v}_1(x) &= \int_a^b v_0(x \wedge \zeta)(1 - v_0(x \vee \zeta)) \frac{1 - v_0(\zeta)}{v_0'(\zeta)} \frac{2}{\sigma^2 \zeta} d\zeta. \end{aligned} \quad (49)$$

As the scale function yield no satisfactory closed formula, we perform a numerical approximation of both v_0 and the couple (v_1, \bar{v}_1) .

These functions may be computed numerically. One may choose a suitable grid when the process is close to 0 and it is noteworthy that the process may not cross 0, a problem which arise when using Euler type schemes.



(a) Histogram of a CIR process with parameters $(\theta, \mu, \sigma) = (2, 5, 2)$ with an initial value of $x_0 = 3$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

(b) Histogram of a CIR process with parameters $(\theta, \mu, \sigma) = (5, 5, 1)$ with an initial value of $x_0 = 3$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

Figure 5: CIR process approximation.

7.6 Skew and reflected Bessel process

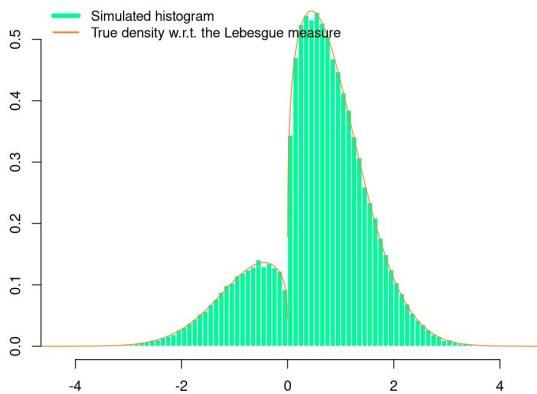
The Skew-Bessel process (see [2]) is the diffusion process with the following scale function and speed measure,

$$s(x) = \begin{cases} \frac{1}{\beta} \frac{x^{2-\delta}}{2-\delta}, & x > 0, \\ -\frac{1}{1-\beta} \frac{|x|^{2-\delta}}{2-\delta}, & x \leq 0, \end{cases} \quad m(dx) = \begin{cases} 2\beta x^{\delta-1} dx, & x > 0, \\ 2(1-\beta)|x|^{\delta-1} dx, & x \leq 0. \end{cases} \quad (50)$$

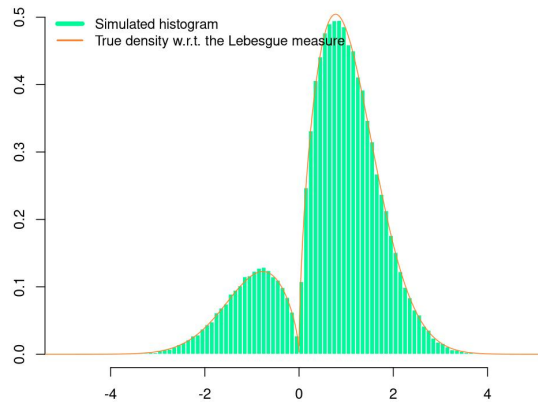
This yield the following expressions for the quantities we need to compute in order to implement the algorithm,

$$\begin{aligned} v_0(x) &= \frac{s(x) - s(a)}{s(b) - s(a)}, \\ v_1(x) &= \int_a^b G_{a,b}(x, \zeta) v_0(\zeta) 2|\zeta|^{\delta-1} d\zeta, \\ \bar{v}_1(x) &= \int_a^b G_{a,b}(x, \zeta) (1 - v_0(\zeta)) 2|\zeta|^{\delta-1} d\zeta, \end{aligned}$$

where $G_{a,b}(x, \zeta)$ is defined in (1). The probability transition kernels plotted in Figures 6a, 6b, 7a and 7b are computed in [2].

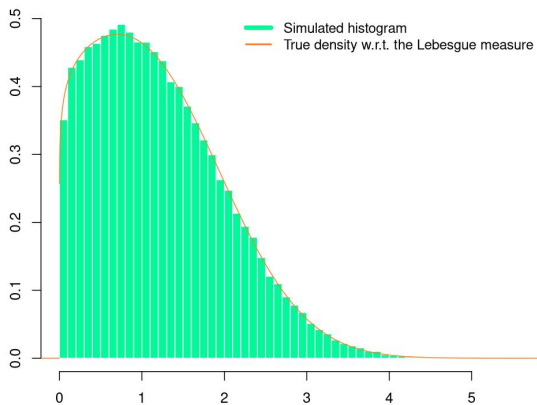


(a) Histogram of a skew Bessel process of dimension 1.2 with skew parameter $\beta = 0.8$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

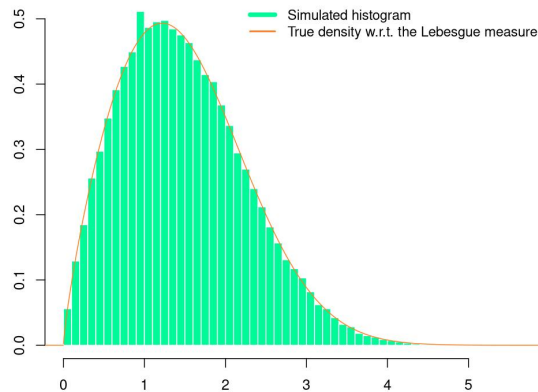


(b) Histogram of a skew Bessel process of dimension 1.6 with skew parameter $\beta = 0.8$ at $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

Figure 6: Skew Bessel process approximation.



(a) Histogram of Bessel process of dimension 1.1 reflected at 0 with a starting point of $x_0 = 1$ and time horizon $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.



(b) Histogram of Bessel process of dimension 1.8 reflected at 0 with a starting point of $x_0 = 1$ and time horizon $T = 1$ simulated using Algorithm 1 with a uniform grid of cell size $h = 0.01$.

Figure 7: Reflected Bessel process approximation.

Acknowledgement. The PhD thesis of A. Anagostakis is supported by a scholarship from the Grand-Est Region (France).

References

- [1] Aurélien Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384, 2005. doi: 10.1163/156939605777438569.
- [2] Larbi Alili and Andrew Aylwin. On the semi-group of a scaled skew Bessel process. *Statist. Probab. Lett.*, 145:96–102, 2019. doi: 10.1016/j.spl.2018.08.014.
- [3] Madjid Amir. Sticky Brownian motion as the strong limit of a sequence of random walks. *Stochastic Process. Appl.*, 39(2):221–237, 1991. doi: 10.1016/0304-4149(91)90080-V.
- [4] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. A functional limit theorem for irregular SDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1438–1457, 2017. doi: 10.1214/16-AIHP760.
- [5] Stefan Ankirchner, Thomas Kruse, and Mikhail Urusov. Wasserstein convergence rates for coin tossing approximations of continuous markov processes, 03 2019. Preprint arXiv:1903.07880.
- [6] Richard F. Bass. A stochastic differential equation with a sticky point. *Electron. J. Probab.*, 19:no. 32, 22, 2014. doi: 10.1214/EJP.v19-2350.
- [7] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.
- [8] John C. Cox, Jonathan E. Ingersoll, Jr., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985. doi: 10.2307/1911242.
- [9] Monroe D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6:12, 1951.
- [10] Andreas Eberle and Raphael Zimmer. Sticky couplings of multidimensional diffusions with different drifts. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(4):2370–2394, 2019. doi: 10.1214/18-AIHP951.
- [11] Hans-Jürgen Engelbert and Goran Peskir. Stochastic differential equations for sticky Brownian motion. *Stochastics*, 86(6):993–1021, 2014. doi: 10.1080/17442508.2014.899600.
- [12] Pierre Étoré and Antoine Lejay. A Donsker theorem to simulate one-dimensional processes with measurable coefficients. *ESAIM Probab. Stat.*, 11:301–326, 2007. doi: 10.1051/ps:2007021.
- [13] William Feller. Two singular diffusion problems. *Ann. of Math. (2)*, 54:173–182, 1951. doi: 10.2307/1969318.
- [14] William Feller. The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math. (2)*, 55:468–519, 1952. doi: 10.2307/1969644.
- [15] Anna Ferrer-Admetlla, Christoph Leuenberger, Jeffrey Jensen, and Daniel Wegmann. An approximate markov model for the wright-fisher diffusion and its application to time series data. *Genetics*, 203, 04 2016. doi: 10.1534/genetics.115.184598.
- [16] David Freedman. *Brownian motion and diffusion*. Springer-Verlag, New York-Berlin, second edition, 1983.
- [17] Hatem Hajri, Mine Caglar, and Marc Arnaudon. Application of stochastic flows to the sticky Brownian motion equation. *Electron. Commun. Probab.*, 22:Paper No. 3, 10, 2017. doi: 10.1214/16-ECP37.
- [18] Kiyosi Itô. *Essentials of stochastic processes*, volume 231. American Mathematical Soc., 2006.

- [19] Antoine Lejay. Monte Carlo methods for fissured porous media: a gridless approach. *Monte Carlo Methods Appl.*, 10(3-4):385–392, 2004. doi: 10.1515/mcma.2004.10.3-4.385.
- [20] Antoine Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.*, 3:413–466, 2006. doi: 10.1214/154957807000000013.
- [21] Yutian Nie and Vadim Linetsky. Sticky reflecting ornstein-uhlenbeck diffusions and the vasicek interest rate model with the sticky zero lower bound. *Stochastic Models*, 0(0):1–19, 2019. doi: 10.1080/15326349.2019.1630287.
- [22] Tomasz Piskorski and Mark M. Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *J. Econom. Theory*, 166:242–281, 2016. doi: 10.1016/j.jet.2016.08.003.
- [23] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. doi: 10.1007/978-3-662-06400-9.
- [24] John Y. Zhu. Optimal contracts with shirking. *Rev. Econ. Stud.*, 80(2):812–839, 2013. doi: 10.1093/restud/rds038.