# EXACT SAMPLING OF DIFFUSIONS WITH A DISCONTINUITY IN THE DRIFT

BY OMIROS PAPASPILIOPOULOS, GARETH O. ROBERTS AND KASIA B. TAYLOR

**Abstract**

We introduce exact methods for the simulation of sample paths of one-dimensional diffusions with a discontinuity in the drift function. Our procedures require the simulation of finite-dimensional candidate draws from probability laws related to those of Brownian motion and its local time, and are based on the principle of retrospective rejection sampling. A simple illustration is provided.

*Keywords:* Exact algorithm; sampling of diffusions; diffusion with discontinuous drift; retrospective rejection sampling

2010 Mathematics Subject Classification: Primary 65C30
Secondary 60H35; 65C05

## 1. Introduction

Beskos *et al.* [2], [3] introduced a collection of efficient 'exact algorithms' for the simulation of skeletons of diffusion processes. While the methodology is intrinsically limited in the multivariate case to processes that can be transformed into unit-volatility diffusions with drifts which can be written as the gradient of a potential, for one-dimensional nonexplosive diffusions the algorithm's applicability depends, more or less, only on smoothness conditions on the diffusion and drift coefficients.

Being able to simulate one-dimensional diffusions with discontinuous drifts is however of considerable interest, not least because this then allows us to tackle reflecting boundaries by suitable unfolding procedures. Therefore, in this paper we focus on extending these exact algorithms to the case of discontinuous drift. Namely, we consider solutions to one-dimensional stochastic differential equations (SDEs) of the form

$$\mathrm{d}X_t = \alpha(X_t)\,\mathrm{d}t + \mathrm{d}B_t, \qquad X_0 = x, \qquad t \in [0, T],\ T < \infty, \tag{1}$$

where $B$ denotes one-dimensional Brownian motion and $\alpha$ is a discontinuous function which satisfies assumptions as specified later.

These exact algorithms carry out rejection sampling in the diffusion trajectory space. The difficulty for discontinuous drift lies in the choice of a suitable candidate measure for proposals in rejection sampling and in performing the acceptance/rejection step, which reduces to sampling a random variable with a Bernoulli distribution whose probability of success is not explicitly computable. We address these problems by suggesting a new methodology for sampling certain conditional laws of Brownian motion and its local time. We construct a stochastic process to be used as a proposal within rejection sampling in this context, which we call the 'local-time-tilted biased Wiener process'; this is to be contrasted with the simpler

*doi*:10.1017/apr.2016.54 

'endpoint-tilted Wiener process' which has been used when the drift is continuous. Overall, our approach is a natural evolution of the research program put forward in [2] for simulation of diffusion sample paths using the Wiener–Poisson factorisation of the diffusion measure together with retrospective rejection sampling principles. The present work forms part of the doctoral thesis [6].

Concurrent to our work is that of Étoré and Martinez [5], who addressed the same fundamental problem as we do here. They followed a different approach to that which we take in this article, one based on the limiting argument (with $n \to \infty$) applied to their earlier results for exact sampling of solutions to SDEs of the type $\mathrm{d}X_t = \alpha(X_t)\,\mathrm{d}t + \mathrm{d}B_t + n^{-1}\,\mathrm{d}L_t$. They then proved by indirect analytic arguments that their limiting algorithm does successfully simulate exactly from the SDE in (1). In contrast, our paper offers a much simpler and more direct algorithm with a direct probabilistic interpretation as rejection sampling on path space. It is difficult to be precise about the computational cost comparison between the two methods, although we estimate that our algorithms can be anything from 2–20 times quicker than those in [5].

## 2. Derivation of algorithms

Our aim here is to sample from $\mathbb{Q}$, the measure induced by the diffusion $X$ on $(C, \mathcal{C})$, the space of continuous functions on $[0, T]$ with the supremum norm and cylinder $\sigma$-algebra. Note that the strong solution to (1) exists under mild conditions, namely, if $\alpha$ is bounded and measurable (see Theorem 4 of [7]). Denote by $\mathbb{W}$ the measure induced by Brownian motion started at $x$ on $(C, \mathcal{C})$. We need the following Assumption 1 about the Radon–Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{W}$.

**Assumption 1.** *The Cameron–Martin–Girsanov theorem applies and the Radon–Nikodym derivative* $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{W}$ *is a true martingale:*

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{W}} = \exp\left\{\int_0^T \alpha(X_t)\,\mathrm{d}X_t - \frac{1}{2}\int_0^T \alpha^2(X_t)\,\mathrm{d}t\right\}.$$

The first step towards performing an acceptance/rejection step is the substitution of the stochastic integral $\int_0^T \alpha(X_t)\,\mathrm{d}X_t$. In existing exact algorithms for diffusions with sufficiently smooth coefficients (see, e.g. [2] and [1]) this is done by application of Itô's lemma to $A(X_t)$, where $A := \int_0^u \alpha(y)\,\mathrm{d}y$. The discontinuity of $\alpha$ precludes proceeding in the same way, although we can generalise the approach by appealing to the generalised Itô formula.

**Assumption 2.** *Let* $r_1 < r_2 < \cdots < r_n$ *be real numbers, and define* $D = \{r_1, r_2, \ldots, r_n\}$. *Assume that the drift function* $\alpha\colon \mathbb{R} \to \mathbb{R}$ *is continuous on* $\mathbb{R} \setminus D$, *that* $\alpha'$ *exists and is continuous on* $\mathbb{R} \setminus D$, *and that the limits*

$$\alpha'(r_k+) := \lim_{x \downarrow r_k} \alpha'(x) \quad \textit{and} \quad \alpha'(r_k-) := \lim_{x \uparrow r_k} \alpha'(x), \qquad k \in \{1, \ldots, n\},$$

*exist and are finite.*

Under Assumption 2, $A$ is the difference of two convex functions and, almost surely (a.s.) for $t \in [0, T]$,

$$A(X_t) = A(x) + \int_0^t \alpha(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t \alpha'(X_s)\,\mathrm{d}s + \frac{1}{2}\sum_{k=1}^n L_t^{r_k}[\alpha(r_k+) - \alpha(r_k-)],$$

where $L_t^{r_k}$ denotes local time at level $r_k$ and time $t$; $r_k$ is omitted when $r_k = 0$.

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

The algorithms that we present address the problem in the case where the drift function $\alpha$ has one point of discontinuity which without loss of generality can be assumed to be at 0, and we denote the discontinuity jump by $\theta = \frac{1}{2}(\alpha(0+) - \alpha(0-))$. Then the Radon–Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{W}$ is

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{W}} = \exp\left\{A(X_T) - A(x) - \frac{1}{2}\int_0^T (\alpha'(X_t) + \alpha^2(X_t))\,\mathrm{d}t - \theta L_T^0(X)\right\} \quad \text{a.s.} \tag{2}$$

The main question here is how to use (2) for simulation. If (2) is used in rejection sampling, it needs to be uniformly bounded above.

**Assumption 3.** *There exist constants $\kappa$ and $M$ with*

$$-\infty < \kappa \le \tfrac{1}{2}(\alpha^2(u) + \alpha'(u))\mathbf{1}_{\{u\neq 0\}} \le \kappa + M.$$

On the basis of this assumption, we define $\varphi(u) = \frac{1}{2}(\alpha^2(u) + \alpha'(u))\mathbf{1}_{\{u\neq 0\}} - \kappa$.

The main problem here is that the multiplicative term $\exp\{A(X_T) - \theta L_T^0(X)\}$ may be unbounded. We address this issue by biasing—using exponential tilting—the dominating measure $\mathbb{W}$, by these terms. Thus we define $\mathbb{S}$ on $(C, \mathcal{C})$ with paths starting at $x$ and satisfying

$$\frac{\mathrm{d}\mathbb{S}}{\mathrm{d}\mathbb{W}} \propto \mathbf{1}_{\{L_T(X)>0\}}(X)\frac{g(X_T, L_T(X))}{f_T^x(X_T, L_T(X))} + \mathbf{1}_{\{L_T(X)=0\}}(X)\frac{g_*(X_T)}{f_{*,T}^x(X_T)},$$

with

$$g(b, l) := c_g f_T^x(b, l)\mathrm{e}^{A(b)-l\theta} \quad \text{for } l > 0, \tag{3}$$

$$g_*(b) := c_g f_{*,T}^x(b)\mathrm{e}^{A(b)}, \tag{4}$$

such that $\int_{\mathbb{R}}\int_{(0,\infty)} g(b,l)\,\mathrm{d}l\,\mathrm{d}b + \int_{\mathbb{R}} g_*(b)\,\mathrm{d}b = 1$. Above, $f$ and $f_*$ describe the joint law of Brownian motion and its local time at level zero (see, e.g. [4]); thus,

$$\begin{aligned} f_s^x(b,l)\,\mathrm{d}b\,\mathrm{d}l &:= \mathbb{P}(B_s \in \mathrm{d}b,\ L_s \in \mathrm{d}l \mid B_0 = x) \\ &= \frac{l + |b| + |x|}{s\sqrt{2\pi s}}\exp\left\{-\frac{(l+|b|+|x|)^2}{2s}\right\}\mathrm{d}b\,\mathrm{d}l \quad \text{for } l > 0. \end{aligned}$$

If $x > 0$ and $b > 0$ (or $x < 0$ and $b < 0$) then

$$f_{*,s}^x(b)\,\mathrm{d}b := \mathbb{P}(B_s \in \mathrm{d}b,\ L_s = 0 \mid B_0 = x) = \phi_{x,s}(b)(1 - \mathrm{e}^{-2bx/s})\,\mathrm{d}b,$$

where $\phi_{\mu,s}$ denotes the density ($\Phi_{\mu,s}$ denotes the cumulative distribution function) of the normal distribution with mean $\mu$ and variance $s$.

The above definitions of $g$ and $g_*$ rely on Assumption 4 below.

**Assumption 4.** *It holds that*

$$\int_{\mathbb{R}}\int_{(0,\infty)} f_T^x(b,l)\mathrm{e}^{A(b)-l\theta}\,\mathrm{d}l\,\mathrm{d}b < \infty \quad \textit{and} \quad \int_{\mathbb{R}} f_{*,T}^x(b)\mathrm{e}^{A(b)}\,\mathrm{d}b < \infty.$$

A fairly direct calculation then yields

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{S}} \propto \exp\left\{-\int_0^T \varphi(X_t)\,\mathrm{d}t\right\} \quad \text{a.s.}$$

Beskos *et al.* [2] showed that, for Radon–Nikodym derivatives of this form for bounded $\varphi > 0$, provided finite-dimensional distributions of $\mathbb{S}$ can be simulated exactly, exact simulation of sample paths from $\mathbb{Q}$ is feasible using retrospective rejection sampling employing auxiliary

Poisson processes. We now present an algorithm which requires Assumptions 1, 2 (with $n = 1$), 3, and 4. We denote by $0 = t^0, t^1, \ldots, t^n, t^{n+1} = T$ the time instances at which we want to sample the diffusion. Recall that $X_0 = x$. The following algorithm returns values of the diffusion $X$ together with its local time at a collection of chosen and random time points.

**Algorithm 1.** (Exact algorithm for drift with discontinuity.)

**I** *Generate* $(X_T, L_T)$ *according to the law given by (3) and (4).*

**II** *Sample an auxiliary Poisson process* $\Psi$ *on* $[0, T]$ *with a rate parameter* $M$ *to get* $\tau_1, \ldots, \tau_k$, *and then* $\psi_i \sim U(0, M)$ *independent and identically distributed for* $i = 1, \ldots, k$.

**III** *Generate* $X$ *and* $L$ *at times* $\tau_1, \tau_2, \ldots, \tau_k$ *conditioned on values at* $0$ *and* $T$ *(see Sections 3.2 and 3.3).*

**IV** *Use* $\Psi$ *to perform rejection sampling:*

(i) *compute* $\varphi(X_{\tau_i})$ *for* $i \in \{1, \ldots, k\}$;

(ii) *if* $\varphi(X_{\tau_i}) < \psi_{\tau_i}$ *for each* $i \in \{1, \ldots, k\}$, *proceed to* **V**; *otherwise, start again at* **I**.

**V** *Generate X jointly with L at times* $t^1, \ldots, t^n$ *conditioned on values at* $\{0, \tau_1, \tau_2, \ldots, \tau_k, T\}$ *(see Sections 3.2 and 3.3).*

**VI** *Return* $(0, x, 0), (t_1, X_{t_1}, L_{t_1}), \ldots, (t_{k+n}, X_{t_{k+n}}, L_{t_{k+n}}), (T, X_T, L_T)$.

The practical and probabilistic challenges with this algorithm are contained in step **I**, which requires simulation from a bivariate distribution for the endpoint of the path and its accumulated local time at 0, and step **III**, which requires finite-dimensional distributions of Brownian motion and its local time. In the rest of this section we address the first challenge and in Section 3 the second challenge. However, as we show below, solving the second problem can also provide alternative and more efficient solutions to the first, given an additional assumption.

### 2.1. Simulating the endpoint distribution

We first provide some further insight on the exponential tilting employed in the definition of $\mathbb{S}$. Note that we are biasing the joint distribution of $(X_T, L_T(X))$ under the Wiener measure, with the terms $\mathrm{e}^{A(X_T)}$ and $\mathrm{e}^{-\theta L_T(X)}$. This observation generates our first method for simulating from $\mathbb{S}$, when the discontinuity jump is positive, i.e. if $\theta > 0$. We decompose the law of $(X_T, L_T(X))$ as the marginal law of $X_T$ and the conditional law of $L_T(X) \mid X_T$. The simulation of $X_T$ in this case is done according to

$$h(u) \propto \mathrm{e}^{A(u)} \phi_{x,T}(u).$$

This step is in fact common to all exact algorithms for diffusions. Conditionally on $X_T$ drawn according to this scheme, $L_T \mid X_T$ is proposed according to the conditional law of local time given point evaluation under the Wiener measure, as described in Section 3.1. Any value produced in this way is then accepted with probability $\mathrm{e}^{-\theta L_T(X)}$. An accepted pair $(X_T, L_T(X))$ produced by this procedure is drawn from $\mathbb{S}$.

When the jump is negative, $\theta < 0$, the above procedure cannot work and the biasing due to the local time has to be dealt with in a slightly more involved way, albeit still by rejection sampling. We will simulate $(X_T, L_T(X))$ jointly in this case, and find it useful to use a mixture distribution consisting of six mixture components appropriately chosen to dominate either the tails or the central part of the probability distribution given by $g$ and $g_*$ in the case $(b \geq 0,\ l > 0)$, $(b < 0,\ l > 0)$, or $(b > 0,\ l = 0)$ (or $(b < 0,\ l = 0)$ if $x < 0$). Suppose that $\xi_1, \xi_3 > 0$ and $\xi_2, \xi_4 < 0$. Let $h_i \colon \mathbb{R} \times (0, \infty) \to \mathbb{R}$ $(i = 1, \ldots, 4)$ and $h_i \colon \mathbb{R} \to \mathbb{R}$ $(i = 5, \ldots, 8)$ be

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

probability distributions such that

$$\begin{aligned}
h_1 > 0 &\quad \text{if and only if} \quad (b, l) \in [0, \xi_1] \times (0, \infty),\\
h_2 > 0 &\quad \text{if and only if} \quad (b, l) \in (\xi_1, \infty) \times (0, \infty),\\
h_3 > 0 &\quad \text{if and only if} \quad (b, l) \in [\xi_2, 0) \times (0, \infty),\\
h_4 > 0 &\quad \text{if and only if} \quad (b, l) \in (-\infty, \xi_2) \times (0, \infty),\\
h_5 > 0 &\quad \text{if and only if} \quad b \in [0, \xi_3],\\
h_6 > 0 &\quad \text{if and only if} \quad b \in (\xi_3, \infty),\\
h_7 > 0 &\quad \text{if and only if} \quad b \in [\xi_4, 0],\\
h_8 > 0 &\quad \text{if and only if} \quad b \in (-\infty, \xi_4).
\end{aligned}$$

We use the $h_i$ with $i \in \{1, \dots, 6\}$ when $x > 0$, with $i \in \{1, 2, 3, 4, 7, 8\}$ when $x < 0$ and with $i \in \{1, \dots, 4\}$ when $x = 0$. Assume for now that $x > 0$. Choose $K$ such that $6g(b, l)/h_i(b, l) < K$ whenever $h_i > 0$ for $i = 1, \dots, 4$, and $6g_*(b)/h_i(b) < K$ whenever $h_i > 0$ for $i = 5, 6$; then the procedure for sampling the candidate end-pair $(b, l)$ is as described below.

**Procedure 1.** (Endpoint rejection sampling.)

**repeat**
  $U_1 \sim$ DiscreteUniform on $\{1, 2, \dots, 6\}$
  $U_2 \sim$ U(0, 1)
  **for** $i = 1$ to 4 **do**
    **if** $u_1 = i$ **then**
      $(b, l) \sim h_i$
      **if** $u_2 \leq 6g(b, l)/(Kh_i(b, l))$ **then**
        accept $(b, l)$
  **for** $i = 5$ to 6 **do**
    **if** $u_1 = i$ **then**
      $b \sim h_i$
      **if** $u_2 \leq 6g_*(b)/(Kh_i(b))$ **then**
        accept $b$, set $l = 0$
**until** $b$ accepted
**return** $(b, l)$

## 3. Sampling procedures for Brownian motion and its local time

In this section we discuss sampling from the conditional laws $\mathcal{L}(L_{s_2} \mid B_{s_1}, B_{s_2}, L_{s_1})$ and $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$, where $0 \leq s_1 < s_2 < s_3$. Note that $(B, L)$ is a Markov process, with increments whose distribution depends only on the first coordinate of the process, which facilitates simulation of finite-dimensional distributions of the process as well as finite-dimensional distributions of the process conditionally on its endpoints. Furthermore, we observe that as far as simulating skeletons of the process conditionally on a starting point, it suffices to describe how to generate observations from $\mathcal{L}(L_{s_2} \mid B_{s_1}, B_{s_2}, L_{s_1})$, which we do in Section 3.1. Simulating finite-dimensional distributions conditionally on past and future endpoints requires simulation from $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$; we tackle this problem in two steps, first in the simpler case where $L_{s_1} = L_{s_3}$, in which case $L_{s_2}$ is also equal to

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

those values, and then for the general case where the local time has changed between the two endpoints.

### 3.1. Sampling from $\mathcal{L}(L_{s_2} \mid B_{s_1}, B_{s_2}, L_{s_1})$

We compute the probability that the local time does not increase on $[s_1, s_2]$. Assuming that $b_1 < 0$ and $b_2 < 0$ (or $b_1 > 0$ and $b_2 > 0$),

$$\mathbb{P}(L_{s_2} = l_1 \mid B_{s_1} = b_1,\ B_{s_2} = b_2,\ L_{s_1} = l_1) = \frac{f^{b_1}_{*,s_2-s_1}(b_2)}{\phi_{b_1,s_2-s_1}(b_2)} = 1 - \mathrm{e}^{-2b_1b_2/(s_2-s_1)}.$$

Now suppose that $l_2 > l_1$; then

$$\begin{aligned}\mathbb{P}(L_{s_2} &\in \mathrm{d}l_2 \mid B_{s_1} = b_1,\ B_{s_2} = b_2,\ L_{s_1} = l_1)\\ &= \frac{f^{b_1}_{s_2-s_1}(b_2, l_2 - l_1)}{\phi_{b_1,s_2-s_1}(b_2)}\,\mathrm{d}l_2\\ &= \frac{l_2 - l_1 + |b_2| + |b_1|}{s_2 - s_1}\exp\bigg\{\frac{-(l_2 - l_1 + |b_2| + |b_1|)^2}{2(s_2 - s_1)}\bigg\}\exp\bigg\{\frac{(b_2 - b_1)^2}{2(s_2 - s_1)}\bigg\}\,\mathrm{d}l_2,\end{aligned}$$

and further, by substituting $l := l_2 - l_1 + |b_2| + |b_1|$,

$$\propto \frac{l}{s_2 - s_1}\exp\bigg\{-\frac{l^2}{2(s_2 - s_1)}\bigg\}\,\mathrm{d}l \quad \text{for } l > |b_2| + |b_1|.$$

**Procedure 2.** (Sampling from $\mathcal{L}(L_{s_2} \mid B_{s_1}, B_{s_2}, L_{s_1})$ with $B_{s_1} = b_1$, $B_{s_2} = b_2$, and $L_{s_1} = l_1$.)

**I** *If* $b_1b_2 \le 0$, *proceed to* **II**.

*Otherwise,*

*sample* $U \sim U(0, 1)$;

*if* $u \le 1 - \exp\{-2b_1b_2/(s_2 - s_1)\}$, *set* $l_2 = l_1$ *and finish here;*

*otherwise, proceed to* **II**.

**II** *Sample* $Z \sim U(1 - \exp\{-(|b_2| + |b_1|)^2/(2(s_2 - s_1))\}, 1)$;

*set* $y = \sqrt{-2(s_2 - s_1)\ln(1 - z)}$;

*set* $l_2 = y + l_1 - |b_2| - |b_1|$.

### 3.2. Sampling from $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$, where $L_{s_1} = L_{s_3}$

In this and the next subsection we are interested in conditioning Brownian motion and its local time on both past and future values. Here we consider $L_{s_1} = L_{s_3}$, which implies that $B_s \ne 0$ for all $s \in [s_1, s_3]$ a.s. Also, here $L_{s_2}$ is trivially equal to $L_{s_1}$. Throughout this subsection, we suppose that $b_1 > 0$, $b_2 > 0$, and $b_3 > 0$; the case $b_1 < 0$, $b_2 < 0$ and $b_3 < 0$ can be treated completely symmetrically.

Using Bayes' theorem, we obtain

$$\begin{aligned}\nu_1(\mathrm{d}b_2) &:= \mathbb{P}(B_{s_2} \in \mathrm{d}b_2 \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_1)\\ &= \frac{f^{b_1}_{*,s_2-s_1}(b_2)\,f^{b_2}_{*,s_3-s_2}(b_3)}{f^{b_1}_{*,s_3-s_1}(b_3)}\,\mathrm{d}b_2\\ &= \phi_{\mu,\sigma^2}(b_2)\frac{(1 - \exp\{-2b_2b_1/(s_2 - s_1)\})(1 - \exp\{-2b_3b_2/(s_3 - s_2)\})}{1 - \exp\{-2b_3b_1/(s_3 - s_1)\}}\,\mathrm{d}b_2,\end{aligned}$$

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

where

$$\mu := \frac{b_1(s_3 - s_2) + b_3(s_2 - s_1)}{s_3 - s_1}, \qquad \sigma^2 := \frac{(s_2 - s_1)(s_3 - s_2)}{s_3 - s_1}. \tag{5}$$

Next we introduce a measure $\nu_2$ which will be useful for auxiliary rejection sampling:

$$\nu_2(\mathrm{d}b_2) := \frac{\mathbb{P}(B_{s_2} \in \mathrm{d}b_2 \mid B_{s_1} = b_1,\ B_{s_3} = b_3)}{\mathbb{P}(B_{s_2} > 0 \mid B_{s_1} = b_1,\ B_{s_3} = b_3)};$$

so

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2} \propto (1 - \mathrm{e}^{-2b_2b_1/(s_2-s_1)})(1 - \mathrm{e}^{-2b_3b_2/(s_3-s_2)}).$$

**Procedure 3.** (Sampling from $\mathcal{L}(B_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$, where $L_{s_1} = L_{s_3}$.)

**I** *Sample $Z \sim$ truncated normal distribution $N(\mu, \sigma^2)$ on $(0, \infty)$, where $\mu$ and $\sigma^2$ are as in (5).*

**II** *Sample $U \sim U(0, 1)$.*

**II** *If $u < (1 - \mathrm{e}^{-2b_2b_1/(s_2-s_1)})(1 - \mathrm{e}^{-2b_3b_2/(s_3-s_2)})$, set $b_2 = z$; otherwise, start again at* **I**.

### 3.3. Sampling from $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$, where $L_{s_1} \neq L_{s_3}$

In this subsection we concentrate on sampling Brownian motion and its local time at $s_2$ when $L_{s_1} = l_1 \neq L_{s_3} = l_3$. We need to consider three cases, namely when local time stays constant over $[s_1, s_2]$, over $[s_2, s_3]$, and when not constant over either of these intervals.

1. Suppose that $l_1 = l_2$, but $l_1 \neq l_3$. Note that here we consider only $b_1 > 0$ and $b_2 > 0$ (or $b_1 < 0$ and $b_2 < 0$). Define $\xi_1(b_2, l_1)$ and $p_1$ as follows:

$$\mathbb{P}(B_{s_2} \in \mathrm{d}b_2,\ L_{s_2} = l_1 \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_3) =: \xi_1(b_2, l_1)\,\mathrm{d}b_2,$$

$$p_1 := \int_{-\infty}^{\infty} \xi_1(b_2, l_1)\,\mathrm{d}b_2 = \mathbb{P}(L_{s_2} = l_1 \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_3).$$

The upper limit of integration, if $b_1 < 0$, or the lower, if $b_1 > 0$, can be changed to 0.

2. Suppose that $l_2 = l_3$ but $l_1 \neq l_3$. Note that here we consider only $b_3 > 0$ and $b_2 > 0$ (or $b_3 < 0$ and $b_2 < 0$). Define $\xi_3(b_2, l_3)$ and $p_3$ as follows:

$$\mathbb{P}(B_{s_2} \in \mathrm{d}b_2,\ L_{s_2} = l_3 \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_3) =: \xi_3(b_2, l_3)\,\mathrm{d}b_2,$$

$$p_3 := \int_{-\infty}^{\infty} \xi_3(b_2, l_3)\,\mathrm{d}b_2 = \mathbb{P}(L_{s_2} = l_3 \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_3).$$

The upper limit of integration, if $b_3 < 0$, or the lower, if $b_3 > 0$, can be changed to 0.

3. Suppose that $l_2 \in (l_1, l_3)$ and $l_1 \neq l_3$. Define $\xi_2(b_2, l_2)$ and $p_2$ as follows:

$$\mathbb{P}(B_{s_2} \in \mathrm{d}b_2, L_{s_2} \in \mathrm{d}l_2 \mid B_{s_1} = b_1, B_{s_3} = b_3, L_{s_1} = l_1, L_{s_3} = l_3) =: \xi_2(b_2, l_2)\,\mathrm{d}b_2\,\mathrm{d}l_2,$$

$$\begin{aligned} p_2 &:= \int_{(l_1,l_3)} \int_{\mathbb{R}} \xi_2(b_2, l_2)\,\mathrm{d}b_2\,\mathrm{d}l_2 \\ &= \mathbb{P}(L_{s_2} \in (l_1, l_3) \mid B_{s_1} = b_1,\ B_{s_3} = b_3,\ L_{s_1} = l_1,\ L_{s_3} = l_3). \end{aligned}$$

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

We introduced $p_1$, $p_2$, and $p_3$ (where $p_1 + p_2 + p_3 = 1$) so that we can use them to split the simulation into two steps. First we determine the case and then conditioned on the case we sample the value of $B_{s_2}$ (or in case 3 of both $B_{s_2}$ and $L_{s_2}$). Observe that

$$\begin{aligned}\xi_2(b_2, l_2) &= \frac{f^{b_1}_{s_2-s_1}(b_2, l_2 - l_1) f^{b_2}_{s_3-s_2}(b_3, l_3 - l_2)}{f^{b_1}_{s_3-s_1}(b_3, l_3 - l_1)} \\ &= c(l_2 - l_1 + |b_2| + |b_1|) \exp\biggl\{-\frac{(l_2 - l_1 + |b_2| + |b_1|)^2}{2(s_2 - s_1)}\biggr\} \\ &\quad \times (l_3 - l_2 + |b_3| + |b_2|) \exp\biggl\{-\frac{(l_3 - l_2 + |b_3| + |b_2|)^2}{2(s_3 - s_2)}\biggr\},\end{aligned}$$

where

$$c := \frac{1}{2\pi(s_2 - s_1)^{3/2}(s_3 - s_2)^{3/2} f^{b_1}_{s_3-s_1}(b_3, l_3 - l_1)}.$$

Note in the above formula for $\xi_2(b_2, l_2)$ the symmetry in $b_2$ about 0; that is, $\xi_2(y, l_2) = \xi_2(-y, l_2)$ for $y \in \mathbb{R}$ and $l_2 \in (l_1, l_3)$. Assume that $b_2 > 0$. Then, by substituting

$$u := l_2 - l_1 + |b_2| + |b_1|, \qquad v := l_3 - l_2 + |b_3| + |b_2|, \tag{6}$$

we further have

$$\xi_2(b_2, l_2) \propto u \exp\biggl\{-\frac{u^2}{2(s_2 - s_1)}\biggr\} v \exp\biggl\{-\frac{v^2}{2(s_3 - s_2)}\biggr\}. \tag{7}$$

Recall that $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$ is a.s. equal to 0 outside $\mathbb{R} \times [l_1, l_3]$. Under the linear transformation $(b_2, l_2) \to (u, v)$ given by (6), the region $R_1 := [0, \infty) \times [l_1, l_3]$ is mapped onto the region $R_2$ bounded by the lines

$$v = u + l_1 - l_3 - |b_1| + |b_3|, \qquad v = u - (l_1 - l_3) - |b_1| + |b_3|,$$
$$v = -u - (l_1 - l_3) + |b_1| + |b_3|.$$

Observe that the form of (7) allows sampling of two independent random variables with Rayleigh distributions, but with different scale parameters. However, it is important to remember that this distribution needs to be truncated to region $R_2$.

**Procedure 4.** (Sampling from $\mathcal{L}(B_{s_2}, L_{s_2} \mid B_{s_1}, B_{s_3}, L_{s_1}, L_{s_3})$, where $L_{s_1} \neq L_{s_3}$.)

**I** *Sample* $U \sim U(0, 1)$.

**II** *Compute* $p_1$. *If* $u > p_1$, *proceed to* **III**; *otherwise,*

*set* $l_2 = l_1$,

*sample* $Z \sim h(z) \propto \xi_1(z, l_1)$,

*set* $b_2 = z$ *and finish here.*

**III** *Compute* $p_3$. *If* $u > p_1 + p_3$, *proceed to* **IV**; *otherwise,*

*set* $l_2 = l_3$,

*sample* $Z \sim h(z) \propto \xi_3(z, l_3)$,

*set* $b_2 = z$ *and finish here.*

**IV** *Sample* $(U, V) \sim h(u, v) \propto \mathbf{1}_{(u,v)\in R_2}(u, v) u \mathrm{e}^{-u^2/(2(s_2-s_1))} v \mathrm{e}^{-v^2/(2(s_3-s_2))}$.

*Set* $l_2 = \frac{1}{2}(u - v + l_3 + l_1 - |b_1| + |b_3|)$.

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

*Sample* $Y \sim \text{Bernoulli}(0.5)$.

*If* $y = 1$, *set* $b_2 = \frac{1}{2}(u + v - l_3 + l_1 - |b_1| - |b_3|)$*;*

*otherwise, set* $b_2 = -\frac{1}{2}(u + v - l_3 + l_1 - |b_1| - |b_3|)$.

Explicit formulæ for $p_1$ and $p_3$ can be found in Appendix A.

## 4. Examples

In this section we present simple examples of numerical simulation of diffusions with discontinuous drift, namely satisfying SDEs (i) $\mathrm{d}X_t = a_i\,\mathrm{d}t + \mathrm{d}B_t$ and (ii) $\mathrm{d}X_t = \sin(X_t - \theta_i)\,\mathrm{d}t + \mathrm{d}B_t$ with $a_1, \theta_1$ if $X_t \geq 0$ and $a_2, \theta_2$ if $X_t < 0$.

We produce 100 000 observations of diffusions at time $T = 1$ applying our exact methods, and use them for kernel density estimation. Our method is substantially quicker than using the Euler–Maruyama scheme with $\Delta t = 0.0001$ (as used for the plots in Figure 1) or even $\Delta t = 0.001$. Moreover, coarser discretisation leaves the Euler–Maruyama scheme competitive on running time, but with appreciable bias. In each example we set $X_0 = 0$ to observe the effect of the discontinuity in the drift.

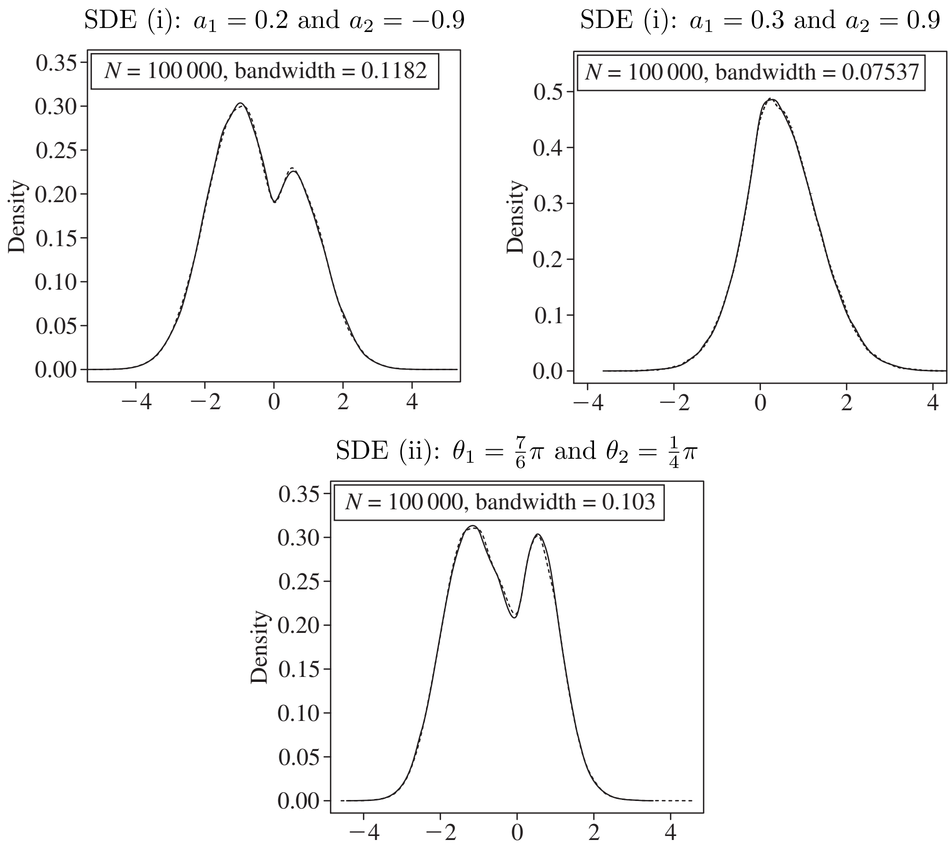


FIGURE 1: Kernel density estimation for $X$ at time $T = 1$ using observations obtained by exact methods (*solid line*) and the Euler–Maruyama method (*dashed line*).

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

The computing time for implementation of these algorithms on an Apple MacBook Air computer with 1.86GHz Intel® Core™ 2 Duo CPU was around 500s when using a mixture distribution (see Section 2.1) to produce candidate $(X_T, L_T)$, and 50s when the two-step procedure (first sampling $X_T$ then $L_T \mid X_T$) was applied, figures which compare favourably with Étoré and Martinez [5], who reported running times of over 1000s on a similar example. All code was implemented in *R*, and it is considered that the algorithm using the mixture distribution could be much more efficient with optimised code.

## Appendix A. Explicit formulæ for $p_1(l_1)$ and $p_3(l_3)$ used in Section 3.3

$$\begin{aligned} p_1(l_1) &= \mathbf{1}_{\{b_1>0\}} c_1 \\ &\quad \times \bigg[\exp\bigg\{-\frac{(b_1+k_1)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\mu_1^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\mu_1+k_1)(1-\Phi_{\mu_1,\sigma^2}(0))\bigg) \\ &\qquad - \exp\bigg\{-\frac{(b_1-k_1)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\mu_2^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\mu_2+k_1)(1-\Phi_{\mu_2,\sigma^2}(0))\bigg)\bigg] \\ &\quad + \mathbf{1}_{\{b_1<0\}} c_1 \\ &\qquad \times \bigg[\exp\bigg\{-\frac{(b_1-k_1)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\mu_3^2}{2\sigma^2}\bigg\} - \sqrt{2\pi}\sigma(\mu_3-k_1)\Phi_{\mu_3,\sigma^2}(0)\bigg) \\ &\qquad\quad + \exp\bigg\{-\frac{(b_1+k_1)^2}{2(s_3-s_1)}\bigg\}\bigg(-\sigma^2 \exp\bigg\{-\frac{\mu_4^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\mu_4-k_1)\Phi_{\mu_4,\sigma^2}(0)\bigg)\bigg], \\ p_3(l_3) &= \mathbf{1}_{\{b_3>0\}} c_2 \\ &\quad \times \bigg[\exp\bigg\{-\frac{(b_3+k_2)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\nu_1^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\nu_1+k_2)(1-\Phi_{\nu_1,\sigma^2}(0))\bigg) \\ &\qquad - \exp\bigg\{-\frac{(b_3-k_2)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\nu_2^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\nu_2+k_2)(1-\Phi_{\nu_2,\sigma^2}(0))\bigg)\bigg] \\ &\quad + \mathbf{1}_{\{b_3<0\}} c_2 \\ &\qquad \times \bigg[\exp\bigg\{-\frac{(b_3-k_2)^2}{2(s_3-s_1)}\bigg\}\bigg(\sigma^2 \exp\bigg\{-\frac{\nu_3^2}{2\sigma^2}\bigg\} - \sqrt{2\pi}\sigma(\nu_3-k_2)\Phi_{\nu_3,\sigma^2}(0)\bigg) \\ &\qquad\quad + \exp\bigg\{-\frac{(b_3+k_2)^2}{2(s_3-s_1)}\bigg\}\bigg(-\sigma^2 \exp\bigg\{-\frac{\nu_4^2}{2\sigma^2}\bigg\} + \sqrt{2\pi}\sigma(\nu_4-k_2)\Phi_{\nu_4,\sigma^2}(0)\bigg)\bigg], \end{aligned}$$

where

$$c_1 = \frac{1}{f^{b_1}_{s_3-s_1}(b_3, l_3-l_1)2\pi(s_2-s_1)^{1/2}(s_3-s_2)^{3/2}},$$

$$c_2 = \frac{1}{f^{b_1}_{s_3-s_1}(b_3, l_3-l_1)2\pi(s_2-s_1)^{3/2}(s_3-s_2)^{1/2}},$$

$$k_1 = l_3 - l_1 + |b_3|, \qquad k_2 = l_3 - l_1 + |b_1|,$$

$$\mu_1 = \frac{b_1(s_3-s_2) - k_1(s_2-s_1)}{s_3-s_1}, \qquad \mu_2 = \frac{-b_1(s_3-s_2) - k_1(s_2-s_1)}{s_3-s_1},$$

$$\mu_3 = -\mu_2, \qquad \mu_4 = -\mu_1,$$

$$\nu_1 = \frac{b_3(s_2-s_1) - k_2(s_3-s_2)}{s_3-s_1}, \qquad \nu_2 = \frac{-b_3(s_2-s_1) - k_2(s_3-s_2)}{s_3-s_1},$$

$$\nu_3 = -\nu_2, \qquad \nu_4 = -\nu_1,$$

$$\sigma^2 = \frac{(s_2-s_1)(s_3-s_2)}{s_3-s_1}.$$

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press

## Acknowledgements

The work of OP was supported by grant MTM2012-37195. GOR was supported by EPSRC grants EP/D002060/1 and EP/K014463/1. KBT was supported by EPSRC EP/D002060/1, ERC 226488, and the Leverhulme Trust RPG-2013-270 grants. KBT would like to thank Dan Crisan for valuable comments related to the article. The authors thank an anonymous referee for helpful comments on a first draft of this paper. The authors are also indebted to the editors of this volume for encouragement, patience, and careful proof reading.

## References

[1] Beskos, A. and Roberts, G. (2005). Exact simulation of diffusions. *Ann. Appl. Prob.* **15,** 2422–2444.
[2] Beskos, A., Papaspiliopoulos, O. and Roberts, G. O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12,** 1077–1098.
[3] Beskos, A., Papaspiliopoulos, O. and Roberts, G. O. (2008). A factorisation of diffusion measure and finite sample path constructions. *Methodology Comput. Appl. Prob.* **10,** 85–104.
[4] Borodin, A. N. and Salminen, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd edn. Birkhäuser, Basel.
[5] Étoré, P. and Martinez, M. (2014). Exact simulation for solutions of one-dimensional stochastic differential equations with discontinuous drift. *ESAIM Prob. Statist.* **18,** 686–702.
[6] Taylor, K. B. (2015). Exact algorithms for simulation of diffusions with discontinuous drift and robust curvature Metropolis-adjusted Langevin algorithms. Doctoral Thesis, University of Warwick.
[7] Zvonkin, A. K. (1974). A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb., Nov. Ser.* **93,** 129–149 (in Russian). English translation: *Math. USSR—Sbornik* **22,** 129–149.

OMIROS PAPASPILIOPOULOS, *ICREA and Universitat Pompeu Fabra*

Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain. Email address: omiros.papaspiliopoulos@upf.edu

GARETH O. ROBERTS, *University of Warwick*

Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK.
Email address: gareth.o.roberts@warwick.ac.uk

KASIA B. TAYLOR, *University of Warwick*

Department of Statistics and Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK.
Email address: kasia.taylor@warwick.ac.uk

https://doi.org/10.1017/apr.2016.54 Published online by Cambridge University Press