EFFICIENT ESTIMATION OF ONE-DIMENSIONAL DIFFUSION FIRST PASSAGE TIME DENSITIES VIA MONTE CARLO SIMULATION

TOMOYUKI ICHIBA,* *University of California Santa Barbara*
CONSTANTINOS KARDARAS,** *Boston University*

Abstract

We propose a method for estimating first passage time densities of one-dimensional diffusions via Monte Carlo simulation. Our approach involves a representation of the first passage time density as the expectation of a functional of the three-dimensional Brownian bridge. As the latter process can be simulated exactly, our method leads to almost unbiased estimators. Furthermore, since the density is estimated directly, a convergence of order $1/\sqrt{N}$, where N is the sample size, is achieved, which is in sharp contrast to the slower nonparametric rates achieved by kernel smoothing of cumulative distribution functions.

Keywords: First passage time; Monte Carlo density estimation; one-dimensional diffusion; three-dimensional Brownian bridge; rate function

2010 Mathematics Subject Classification: Primary 65C05; 60G44

1. Introduction

The problem of computing the distribution of the first time that a diffusion crosses a certain level naturally arises in many different contexts. As probably the most prevalent, we mention quantitative finance, where first passage times are used in credit risk (times of default) as well as in defining exotic contingent claims (so-called barrier options). In this paper, we focus on the numerical computation of the probability density function of first passage times associated with a general one-dimensional diffusion.

Analytic expressions of densities of first passage times are known only in very particular cases. The primary example is Brownian motion with certain (constant) drift and diffusion rates, for which a combination of Girsanov's theorem and the special case of standard (driftless) Brownian motion is used. The first passage time density for the latter case can be obtained by the reflection principle; see, for example, [11, Section 2.6A]. An explicit form of the first passage time density can also be obtained in the case of the radial Ornstein–Uhlenbeck process; see [6] and [9]. Smoothness of the first passage time and related transition densities has been studied in [14], [17], and [22], amongst others.

In the absence of general analytic expressions for first passage time distributions, computational methods are indispensable and, in fact, widely used. One approach is to use Volterra integral equations—we mention [1], [4], and [18] as representative papers dealing with this approach. Alternatively, we can use Monte Carlo simulation. The simplest scheme uses the so-called Euler scheme to approximate the solution of the stochastic differential equation

governing the diffusion at predetermined grid time points ih, $i = 0, 1, \ldots$, where h is the step size, stops at the first time that the diffusion crosses the level of interest, and continues this way to obtain an estimator for the cumulative distribution function of the first passage time. As the Euler scheme is only approximate, it causes bias in the estimation of the probability distribution function. (To account for the fact that the passage can potentially happen in-between the sampled points, a Brownian bridge interpolation may be used—this means that, given the sampled points, the conditional probability that the first passage time occurs in each of the consecutive intervals is approximated by the probability of first passage time of a corresponding Brownian bridge This could potentially reduce the bias, but the whole scheme is still only approximate and some bias remains.) See [7, Chapters 6–7] for a general discussion, [8] for the evaluation of error via partial differential equations, and [2] for a sharp large deviation principle approach. The issue with the bias becomes significantly more severe when numerically computing the density, as some kind of numerical differentiation of the (nonsmooth) empirical distribution function is necessary. Even if we use an exact simulation approach for the diffusion in question (which is, of course, available only in special cases), the estimator for the density will have huge variance. To top it all, even if the aforementioned problems can be eliminated, we can never hope for convergence of the estimators to the true density to be of order $1/\sqrt{N}$, where N is the 'path sample' size, as the problem is nonparametric.

In this work we offer an alternative approach which has clear advantages. First, we arrive at a representation of the density function in terms of the expectation of a functional of a three-dimensional Brownian bridge. This makes it possible to estimate *directly* the first passage time density without having to rely on estimators of the cumulative distribution function, thus achieving the 'parametric' rate of convergence $1/\sqrt{N}$, where N is the sample size. Furthermore, only the three-dimensional Brownian bridge is involved in the simulation, which can be carried out exactly. There is an integral involving the previous three-dimensional Brownian bridge, which can be approximated via a Riemann sum; therefore, the error of the approximation can be estimated efficiently. By construction, our method significantly improves both the bias and variance of the density estimation obtained via the empirical distribution function. The only potential problem of our approach is large-time density estimation, since the thin grid that has to be used in the simulation of the Brownian bridge will result in high computational effort. To circumvent this issue, we note that the tails of the first passage distribution usually decrease exponentially with a rate that can be expressed as the principal eigenvalue of a certain Dirichlet boundary problem involving a second-order ordinary differential equation. This implies that a mixture of Monte Carlo and ordinary differential equation techniques can be efficiently utilized to improve the quality of our estimator.

The structure of the paper is as follows. In Section 2, the problem is formulated and the key representation formula is obtained. In Section 3 we discuss the Monte Carlo estimator of the first passage time density function, and study its large sample properties. In Section 4, the relation between the exponential tail decay of the probability density and the eigenvalues of a Dirichlet boundary problem is discussed. The proofs of all the results are deferred to Appendix A in order to keep the presentation smooth in the main body of the paper.

2. A representation of first passage time densities

2.1. The setup

Consider a one-dimensional diffusion X with dynamics

$$\mathrm{d}X_t = a(X_t)\,\mathrm{d}t + \mathrm{d}W_t, \qquad t \in \mathbb{R}_+, \tag{2.1}$$

where W is a standard one-dimensional Brownian motion. The restrictions on the drift function a that we will impose later (see Assumptions 2.1) ensure that (2.1) has a weak solution, unique in the sense of the probability law, for any initial condition $x \in (0, \infty)$. Let P_x^a denote the law on the canonical path space $C(\mathbb{R}_+; \mathbb{R})$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$ that makes the coordinate processes behave according to (2.1) and is such that $P_x^a[X_0 = x] = 1$.

Define $\tau_0 := \inf\{t \in \mathbb{R}_+ \mid X_t = 0\}$ to be the first passage time of X at level 0. We will consider the problem of finding convenient, in terms of the numerical approximation using the Monte Carlo simulation technique, representations of the quantity

$$p_x^a(t) := \frac{\partial}{\partial t} P_x^a[\tau_0 \le t], \qquad x \in (0, \infty), \ t \in \mathbb{R}_+,$$

i.e. the density of the first passage time of the diffusion at level 0.

Remark 2.1. The fact that we are using a unit diffusion coefficient in (2.1) by no means entails loss of generality in our discussion. Indeed, consider a general one-dimensional diffusion Y with dynamics

$$dY_t = b(Y_t) \, dt + \sigma(Y_t) \, dW_t, \qquad t \in \mathbb{R}_+, \tag{2.2}$$

where W is a standard one-dimensional Brownian motion, such that $Y_0 = y \in \mathbb{R}$. If (2.2) has a weak solution unique in the sense of the probability law, we may assume without loss of generality that $\sigma \ge 0$. (Indeed, otherwise we replace σ by $|\sigma|$ in (2.2) and we obtain the same law for the process Y.) Consider a level $\ell < y$. Under the mild assumption that $1/\sigma$ is locally integrable, the transformation $X = \int_\ell^Y (1/\sigma(z)) \, dz$ defines a diffusion with dynamics $dX_t = a(X_t) \, dt + dW_t$ for a function a that is easily computable from b and σ. With $x := \int_\ell^y (1/\sigma(z)) \, dz$, the first passage time of Y with $Y_0 = y$ at level ℓ is equal to the first passage time of X with $X_0 = x$ at level 0.

2.2. The representation

The following assumption on the drift function in (2.1) will allow us to arrive at a very convenient representation for the density function p_x^a.

Assumption 2.1. *The function a restricted on $[0, \infty)$ is continuously differentiable, and satisfies*

$$\int_0^\infty \exp\left(-2 \int_0^w a(z) \, dz\right) dw = \infty.$$

In particular, under Assumption 2.1, a is locally square integrable on $[0, \infty)$ and the function

$$\gamma := \frac{a^2 + a'}{2} \tag{2.3}$$

is continuous and locally integrable. The theory of one-dimensional diffusions ensures that, for all $x \in (0, \infty)$, there exists a probability P_x^a on $C(\mathbb{R}_+, \mathbb{R})$ such that the coordinate process X has dynamics given by (2.1). Assumption 2.1 also ensures that X^{τ_0}, which is X stopped at level 0, does not explode to ∞—see, for example, [11, Chapter 5, Proposition 5.32(iii)].

Proposition 2.1. *Suppose that Assumption 2.1 holds. On $C([0, 1]; \mathbb{R}^3)$, consider the probability P_{BB^3} under which the coordinate process β is a standard three-dimensional Brownian bridge. Then,*

$$p_x^a(t) = q_x(t) \exp\left(-\int_0^x a(v) \, dv\right) E_{\mathrm{BB}^3}\left[\exp\left(-t \int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|) \, du\right)\right]$$

holds for all $t \in \mathbb{R}_+$, where $e_1 := (1, 0, 0)^\top$, $|\cdot|$ stands for the usual Euclidean norm, and q_x is the density given for all $t \in \mathbb{R}_+$ by

$$q_x(t) \equiv p_x^0(t) = \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right), \tag{2.4}$$

corresponding to the first passage time to 0 of a standard Brownian motion starting from x.

3. Monte Carlo density estimation

We now discuss issues related to the estimation of the density p_x^a. For the purposes of this and the next section, we *fix* a drift function a satisfying Assumption 2.1 and we write p_x for p_x^a in order to simplify the notation.

3.1. Convergence

It is clear how to get an estimate of the density $p_x(t)$ for a given $t \in \mathbb{R}_+$, at least in theory. One simulates N independent paths of the three-dimensional Brownian bridge, $\hat{\beta}^1, \ldots, \hat{\beta}^N$, and then defines the estimator $\hat{p}_x^N(t)$ for $p_x(t)$ via

$$\hat{p}_x^N(t) := q_x(t) \exp\left(-\int_0^x a(v) \, dv\right) \frac{1}{N} \sum_{i=1}^N \exp\left(-t \int_0^1 \gamma(|uxe_1 + \sqrt{t}\hat{\beta}_u^i|) \, du\right), \tag{3.1}$$

where recall that q_x is given in (2.4). By the strong law of large numbers, the estimator $\hat{p}_x^N(t)$ converges almost surely to the true density $p_x(t)$ as N goes to ∞ for each fixed $t \in \mathbb{R}_+$. Moreover, the estimator $\hat{p}_x^N(t)$ is unbiased and the variance of the estimator $\hat{p}_x^N(t)$ decreases in the order of $1/N$, for every fixed $t \in \mathbb{R}_+$, as a direct consequence of (3.1). (Of course, $E_{\mathsf{BB}^3}[\hat{p}_x^N(t)] = p_x(t)$ holds only if we assume that we actually have the whole path of each Brownian bridge simulated exactly, which is not possible in practice. However, we can simulate exactly discretized paths of the Brownian bridge, and then we can easily estimate the order of bias from the Riemann approximation of the integral. In this respect, see also Section 4.2.) In order to get weak convergence of the whole empirical densities $(\hat{p}_x^N(t))_{t \in \mathbb{R}_+}$, as well as the uniform rate of convergence over compact time intervals, we introduce an additional assumption.

Assumption 3.1. *Together with Assumption 2.1, we suppose that the function γ in (2.3) is such that $\inf_{z \in \mathbb{R}_+} \gamma(z) > -\infty$, and that a' is locally Lipschitz continuous on $\mathbb{R}_+$ with Lipschitz constant growing at most polynomially, that is, there exist constants $c_1 > 0$, $c_2 > 0$, and $n \in \mathbb{N}$ such that*

$$\sup_{0 \le v_1 < v_2 \le \kappa} \left| \frac{a'(v_2) - a'(v_1)}{v_2 - v_1} \right| \le c_1 + c_2 \kappa^n \quad \text{holds for all } \kappa > 0. \tag{3.2}$$

The next two results are concerned with a central limit theorem for the whole density function estimator as well as the uniform rate of convergence on compact intervals of $\mathbb{R}_+$.

Proposition 3.1. *For all $N \in \mathbb{N}$, define $\eta^N := \sqrt{N}(\hat{p}_x^N - p_x)/q_x$. Suppose that Assumption 3.1 holds. Then, the family of stochastic processes $\{\eta^N \mid N \in \mathbb{N}\}$ is tight. As $N \to \infty$, η^N converges weakly to a centered Gaussian process with continuous covariance function Γ, where, with $I(t) = \int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|) \, du$ for $t \in \mathbb{R}_+$,*

$$\Gamma(s, t) = \exp\left(-2\int_0^x a(v) \, dv\right) \operatorname{cov}_{\mathsf{BB}^3}[\exp(-sI(s)), \exp(-tI(t))], \quad (s, t) \in \mathbb{R}_+^2.$$

Proposition 3.2. *Under Assumption 3.1, for any fixed $T \in \mathbb{R}_+$, the sequence*

$$\left(\sqrt{N} \max_{t \in [0,T]} |\hat{p}_x^N(t) - p_x(t)| \right)_{N \in \mathbb{N}}$$

is bounded in probability.

Remark 3.1. In a similar manner, for fixed x_1 and x_2 in $(0, \infty)$ with $x_1 < x_2$, we may show that

$$\left(\sqrt{N} \max_{x \in [x_1, x_2]} |\hat{p}_x^N(t) - p_x(t)| \right)_{N \in \mathbb{N}}$$

is bounded in probability. Moreover, under some additional conditions on the differentiability of a, we may estimate the partial derivatives of $p_x(t)$ with respect to (x, t) by differentiating the estimator with respect to the variable of interest.

4. The rate function

Recall that we are dropping the qualifying 'a' from 'p_x^a' in order to simplify the notation. Define implicitly the function λ_x via

$$p_x(t) = q_x(t) \exp\left(-\int_0^x a(v)\, dv \right) \exp(-t\lambda_x(t)), \qquad t \in \mathbb{R}_+.$$

In other words, and in view of Proposition 2.1, we have

$$\lambda_x(t) := -\frac{1}{t} \log\left(\mathrm{E}_{\mathsf{BB}^3}\left[\exp\left(-t \int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\, du \right) \right] \right) \quad \text{for } t \in \mathbb{R}_+. \qquad (4.1)$$

4.1. Theoretical results

Proposition 3.2 ensures the uniform convergence of the estimator on finite intervals $[0, T]$ for fixed $T \in (0, \infty)$. Of course, it is almost always the case that $\lim_{t \to \infty} p_x(t) = 0$. For large $t \in \mathbb{R}_+$, $\lambda_x(t)$ gives a better understanding of the behavior of the density function, as it represents in a certain sense the exponential decrease of $p_x(t)$; therefore, it makes more sense to focus on λ_x rather than p_x. In fact, the following result implies that the function λ_x is frequently bounded on $\mathbb{R}_+$—this is the case, for example, when γ is bounded from below.

Proposition 4.1. *Let Assumption 2.1 hold. Then*

$$\inf_{z \in \mathbb{R}_+} \gamma(z) \leq \inf_{t \in \mathbb{R}_+} \lambda_x(t) \leq \limsup_{t \to \infty} \lambda_x(t) \leq \inf_{\kappa > 0} \left\{ m(\kappa + x) + \frac{\pi^2}{2\kappa^2} \right\}, \qquad (4.2)$$

where $m(w) = \max_{0 \leq z \leq w} \gamma(z)$ for $w > 0$. Furthermore, if γ is bounded from below then

$$\lim_{t \downarrow 0} \lambda_x(t) = \frac{1}{x} \int_0^x \gamma(u)\, du = \int_0^1 \gamma(ux)\, du. \qquad (4.3)$$

Remark 4.1. Inequalities (4.2) only imply bounds for the inferior and superior limits of $\lambda_x(t)$ as $t \to \infty$. In fact, it is expected that $\lim_{t \to \infty} \lambda_x(t)$ exists, possibly except in pathological cases. Let us now argue for this point on a rather loose and intuitive level.

The stopping time τ_0 can be approximated by the sequence $(\tau_0^n)_{n \in \mathbb{N}}$, where, for all $n \in \mathbb{N}$, $\tau_0^n := \inf\{t \geq 0 \mid X_t \notin (0, n)\}$ is the first exit time of X from the interval $(0, n)$. Therefore,

the rate function λ_x may be approximated by the rate functions corresponding to τ_0^n, $n \in \mathbb{N}$. It follows from [5] that the density function p_x^n of τ_0^n has the eigenvalue expansion (see also [12] and [19, Chapter 5] for similar problems)

$$p_x^n(t) = \frac{\partial}{\partial t} P_x^a[\tau_0^n \le t] = \sum_{k=1}^{\infty} e^{-\mu_k^n t} \psi_k^n(x) \quad \text{for } (t, x) \in \mathbb{R}_+ \times (0, n), \ n \in \mathbb{N},$$

with the functions $\{\psi_k^n \mid k \in \mathbb{N}\}$ computed from the eigenfunctions $\{\varphi_k^n \mid k \in \mathbb{N}\}$ and the corresponding eigenvalues $0 < \mu_1^n < \mu_2^n < \cdots$ of the Dirichlet problem

$$\tfrac{1}{2}\varphi''(z) + a(z)\varphi'(z) = -\mu\varphi(z) \quad \text{for } z \in (0, n), \tag{4.4}$$

where $\varphi \in C^{2,\alpha}([0, n], \mathbb{R}_+)$ with $\lim_{z\to 0}\varphi(z) = 0 = \lim_{z\to n}\varphi(z)$. Thus, the limit as $t \to \infty$ for the rate function of τ_0^n is exactly the principal eigenvalue μ_1^n, i.e.

$$\lim_{t\to\infty} \lambda_x^n(t) = -\lim_{t\to\infty} \frac{1}{t} \log\left(\frac{p_x^n(t)}{q_x(t)}\right) = \mu_1^n \quad \text{for all } n \in \mathbb{N},$$

which does not depend on the initial value $x > 0$. Since $\tau_0 = \lim_{n\to\infty} \tau_{0,n}$, it is conjectured that the limit of $\lambda_x(t)$ as $t \to \infty$ actually exists and is equal to $\lim_{n\to\infty} \mu_1^n$. A thorough study aimed at finding reasonable sufficient conditions for

$$\lim_{t\to\infty} \lambda_x(t) = \lim_{n\to\infty} \mu_1^n = \lim_{n\to\infty} \lim_{t\to\infty} \lambda_x^n(t)$$

to hold for $x > 0$ lies beyond the scope of this paper.

4.2. Practical issues

In view of Proposition 3.1 and Proposition 3.2, the estimator $(\hat{p}_x^N(t))_{t\in\mathbb{R}_+}$ of (3.1) convergences uniformly with rate $1/\sqrt{N}$ over compact time intervals. In practice, the computation of (3.1) is implemented by generating a standard three-dimensional Brownian bridge, which is simulated in an exact way over a thin enough grid. The approximation error for the Riemann integral over the *finite* interval [0, 1] in (3.1) can be controlled very efficiently. More precisely, the numerical computation of the exponential functional of the Brownian bridge in (3.1) can be carried out using the fourth-order Runge–Kutta scheme which is proposed and analyzed in [15]. Under appropriate mild regularity conditions on the function γ, it is shown that this numerical scheme has weak order four. For this numerical issue, consult the original paper [15], and the related monographs of [7], [13], and [16].

A potential problem with our estimator $(\hat{p}_x^N(t))_{t\in\mathbb{R}_+}$ can arise for *large t*, that is, the density function at the tail. Note that what is meant here is that the *relative* error of the estimator of $p_x(t)$ tends to be large; the absolute error tends to be extremely small, as $p_x(t)$ is very close to being 0 for large $t \in \mathbb{R}_+$. To visualize the issue, it is helpful to study by experiment the large-time behavior of the rate function (4.1) when X is an Ornstein–Uhlenbeck (OU) process starting with $x = 1$. Here $a(z) = -z$ and $\gamma(z) = (z^2 - 1)/2$ for $z \in \mathbb{R}_+$. The first passage time density of this OU process is known analytically (see, for example, [9, Equation (8)]) and reads

$$p_1(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sinh^{3/2}(t)} \exp\left(\frac{1 + t - \coth(t)}{2}\right) \quad \text{for } t \in \mathbb{R}_+. \tag{4.5}$$

The true density (4.5) and the estimated density (3.1) with $N = 100$ simulations are shown over the interval [0, 10] in Figure 1(a). Note that even with this small number of simulations,

(a) Density functions p_x and p_x^N
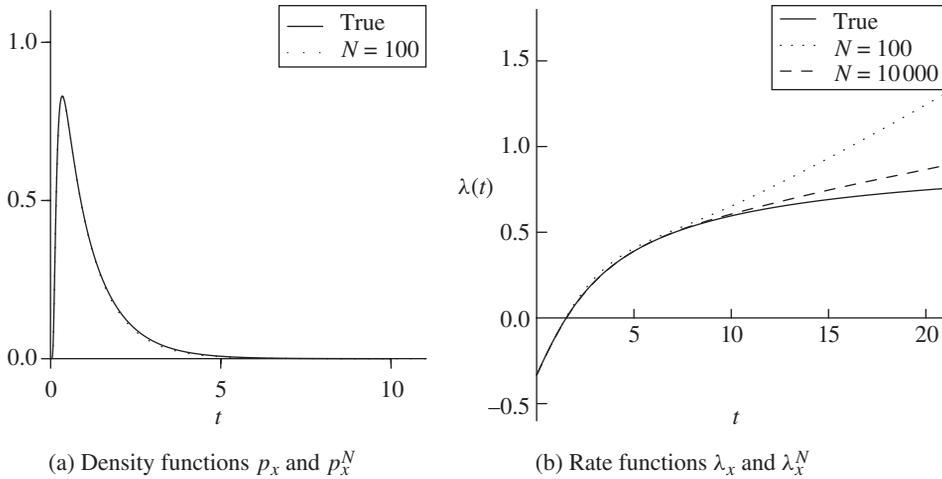
(b) Rate functions λ_x and λ_x^N

FIGURE 1: Comparisons of the simulated density and rate functions in (3.1) and (4.6), with theoretical values for the OU process with $a(z) = -z$ ($z \in \mathbb{R}_+$) and $x = 1$.

the two curves are almost indistinguishable. Figure 1(b) contains graphs of the more refined rate functions. In this scale, we can see that the estimation of the exponential rate of decay of the density for large t is not as good. (As can be seen from Figure 1(b), on the interval $[0, 7]$, the true and estimated rate functions almost coincide. However, on the interval $[7, 20]$, the estimated rate functions $\hat{\lambda}_x^N$ are much larger than the true rate function λ_x, which implies that the estimator (3.1) underestimates the tail probability.) In fact, the estimated asymptotic rate seems to increase linearly instead of converging to a finite limit. This becomes clear once we note that, in this OU example,

$$-t \int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|) \, du = t^2 \frac{1}{2} \int_0^1 |\beta_u|^2 \, du + t^{3/2} x \int_0^1 \langle e_1, \beta_u \rangle \, du + t \frac{x^2}{6};$$

therefore, the estimator for λ_x for sample size $N \in \mathbb{N}$ becomes

$$\hat{\lambda}_x^N(t) = \frac{x^2}{6} - \frac{1}{t} \log\left(\frac{1}{N} \sum_{i=1}^N \exp\left(t^2 \frac{1}{2} \int_0^1 |\hat{\beta}_u^i|^2 \, du + t^{3/2} x \int_0^1 \langle e_1, \hat{\beta}_u^i \rangle \, du \right) \right). \quad (4.6)$$

Observe that the leading term in the estimator of λ_x will be increasing linearly in t.

In order to overcome this poor situation in the tail, we can use a mixture method that combines the estimator (3.1) on the finite interval $[0, T]$ and an estimator for the tail probability of the form

$$c_* q_x(t) \exp(-\lambda t) \quad \text{for } t \in [T, \infty),$$

for some choice of large threshold T, where λ is the principal eigenvalue of the Dirichlet problem (4.4) for some choice of large threshold n, and c_* is chosen so that the density estimator is continuous. The principal eigenvalue can be numerically computed from the Sturm–Liouville problem

$$-\left[\exp\left(2 \int_0^z a(u) \, du \right) \varphi'(z) \right]' = 2\lambda \exp\left(2 \int_0^z a(u) \, du \right) \varphi(z) \quad \text{for } z \in [0, n]$$

(see [23, Chapter 23]), with the same Dirichlet boundary conditions, by use of either the variational method or the Liouville transform method. In the variational method, the principal eigenvalue is obtained by minimizing numerically the corresponding Rayleigh quotient. The Liouville transform method turns the Sturm–Liouville equation into a Schrödinger-type equation, which is then numerically solvable with a discrete approximation. Both numerical methods are well studied—see [10, pp. 333–376] and the references within.

Appendix A. Proofs

A.1. Proof of Proposition 2.1

Consider the nonnegative P_x^a-supermartingale

$$Z := \exp\left(-\int_0^{\cdot \wedge \tau_0} a(X_u)\,\mathrm{d}W_u - \frac{1}{2}\int_0^{\cdot \wedge \tau_0} a^2(X_u)\,\mathrm{d}u\right). \tag{A.1}$$

As follows from a modification of [11, Exercise 5.5.38] for the restricted state space $[0, \infty)$, Z is a P_x^a-martingale. With Q_x being the probability on $C(\mathbb{R}_+; \mathbb{R})$ that makes the coordinate process X behave like a Brownian motion starting from x and stopping when it reaches level 0, Girsanov's theorem implies that

$$\left.\frac{\mathrm{d}Q_x}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_t \tag{A.2}$$

for all $t \in \mathbb{R}_+$. Moreover, since $X_{\tau_0} = 0$, Itô's formula (under Q_x) implies that, on the set $\{\tau_0 < \infty\}$,

$$-\int_0^x a(v)\,\mathrm{d}v = \int_x^{X_{\tau_0}} a(v)\,\mathrm{d}v = \int_0^{\tau_0} a(X_u)\,\mathrm{d}X_u + \frac{1}{2}\int_0^{\tau_0} a'(X_u)\,\mathrm{d}u.$$

Combining this with (2.1) and (2.3), the stochastic exponential defined in (A.1) under Q_x satisfies

$$\begin{aligned}
\frac{1}{Z_{\tau_0}} &= \exp\left(\int_0^{\tau_0} a(X_u)[\mathrm{d}X_u - a(X_u)\,\mathrm{d}u] + \frac{1}{2}\int_0^{\tau_0} a^2(X_u)\,\mathrm{d}u\right) \\
&= \exp\left(\int_0^{\tau_0} a(X_u)\,\mathrm{d}W_u^{Q_x} - \frac{1}{2}\int_0^{\tau_0} a^2(X_u)\,\mathrm{d}u\right) \\
&= \exp\left(-\int_0^x a(v)\,\mathrm{d}v - \frac{1}{2}\int_0^{\tau_0} a'(X_u)\,\mathrm{d}u - \frac{1}{2}\int_0^{\tau_0} a^2(X_u)\,\mathrm{d}u\right) \\
&= \exp\left(-\int_0^x a(v)\,\mathrm{d}v - \int_0^{\tau_0} \gamma(X_u)\,\mathrm{d}u\right),
\end{aligned} \tag{A.3}$$

on $\{\tau_0 < \infty\}$, where W^{Q_x} is a standard Brownian motion under Q_x. Therefore, (A.2) and (A.3) imply that

$$P_x^a[\tau_0 \leq t] = E_{Q_x}\left[\frac{1}{Z_{t \wedge \tau_0}}\mathbf{1}_{\{\tau_0 \leq t\}}\right] = E_{Q_x}\left[\exp\left(-\int_0^x a(v)\,\mathrm{d}v - \int_0^{\tau_0} \gamma(X_u)\,\mathrm{d}u\right)\mathbf{1}_{\{\tau_0 \leq t\}}\right]$$

for $(t, x) \in (\mathbb{R}_+)^2$.

Note that the density function of τ_0 under Q_x is given by q_x in (2.4). Using the regular conditional Q_x-expectation of $\exp(\int_0^{\tau_0} \gamma(X_u)\,\mathrm{d}u)$, given $\tau_0 = t$, we can write

$$p_x^a(t) = P_x^a[\tau_0 \in \mathrm{d}t] = q_x(t)\exp\left(-\int_0^x a(v)\,\mathrm{d}v\right)E_{Q_x}\left[\exp\left(-\int_0^{\tau_0} \gamma(X_u)\,\mathrm{d}u\right)\,\bigg|\,\tau_0 = t\right]. \tag{A.4}$$

Given $\tau_0 = t$, the regular conditional Q_x-distribution of $(X_{t-s}, 0 \le s \le t)$ is that of a three-dimensional Bessel bridge from 0 to x over $[0, t]$, owing to [21, Proposition VI.3.10 and Proposition VII.4.8]. On the canonical space $(C([0, 1], \mathbb{R}^3), P_{\mathsf{BB}^3})$ with coordinate process β, the process $\{|(s/t)xe_1 + \sqrt{t}\beta_{s/t}|, 0 \le s \le t\}$ has the exact law of the aforementioned Bessel bridge. Therefore,

$$\mathrm{E}_{Q_x}\left[\exp\left(-\int_0^{\tau_0} \gamma(X_u)\,\mathrm{d}u\right) \,\middle|\, \tau_0 = t\right]$$

$$= \mathrm{E}_{\mathsf{BB}^3}\left[\exp\left(-\int_0^t \gamma\left(\left|\frac{s}{t}xe_1 + \sqrt{t}\beta_{s/t}\right|\right)\mathrm{d}s\right)\right]$$

$$= \mathrm{E}_{\mathsf{BB}^3}\left[\exp\left(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,\mathrm{d}u\right)\right], \qquad (t, x) \in (\mathbb{R}_+)^2.$$

Combining this with (A.4) completes the proof of Proposition 2.1.

A.2. Proof of Proposition 3.1

The following technical result is the backbone of the proof.

Lemma A.1. *Suppose that Assumption 3.1 holds, and define*

$$\xi(t) := \exp(-tI(t)) := \exp\left(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,\mathrm{d}u\right), \qquad t \in \mathbb{R}_+. \qquad (\text{A.5})$$

Then we have

$$|\xi(t) - \xi(s)| \le \Phi_T |t - s| \quad \text{for all } s \in [0, T] \text{ and } t \in [0, T],$$

where $\mathrm{E}_{\mathsf{BB}^3}[|\Phi_T|^m] < \infty$ *for all* $T \in \mathbb{R}_+$ *and* $m \in \mathbb{N}$.

Proof. First, note that (3.2) in Assumption 3.1 implies that, for every $\kappa > 0$,

$$|a'(v_1)| \le |a'(0)| + |a'(v_1) - a'(0)| \le |a'(0)| + c_1 v_1 + c_2 \kappa^n v_1, \qquad 0 \le v_1 \le \kappa,$$

$$|a(\kappa)| = \left|a(0) + \int_0^\kappa a'(u)\,\mathrm{d}u\right|$$

$$\le |a(0)| + \int_0^\kappa |a'(u)|\,\mathrm{d}u$$

$$\le |a(0)| + |a'(0)|\kappa + c_1 \kappa^2 + c_2 \kappa^{n+2},$$

$$|a(v_1) - a(v_2)| \le \left|\int_{v_1}^{v_2} a'(u)\,\mathrm{d}u\right|$$

$$\le (|a'(0)| + c_1 \kappa + c_2 \kappa^{n+1})|v_2 - v_1|, \qquad 0 \le v_1 \le v_2 \le \kappa.$$

Using these inequalities, we obtain estimates for γ for every $\kappa > 0$ and every $0 \le v_1 \le v_2 \le \kappa$:

$$|\gamma(v_1)| = \left|\frac{a^2(v_1) + a'(v_1)}{2}\right|$$

$$\le \tfrac{1}{2}(|a(0)| + |a'(0)|\kappa + c_1 \kappa^2 + c_2 \kappa^{n+2})^2 + \tfrac{1}{2}(|a'(0)| + c_1 \kappa + c_2 \kappa^{n+1})$$

$$=: \varphi_1(\kappa),$$

$$\begin{aligned}
|\gamma(v_1) - \gamma(v_2)| &\leq \tfrac{1}{2}|a^2(v_1) - a^2(v_2)| + \tfrac{1}{2}|a'(v_1) - a'(v_2)| \\
&\leq \tfrac{1}{2}|a(v_1) + a(v_2)||a(v_1) - a(v_2)| + \tfrac{1}{2}(c_1 + c_2\kappa^n)|v_1 - v_2| \\
&\leq \big[(|a(0)| + |a'(0)|\kappa + c_1\kappa^2 + c_2\kappa^{n+2})(|a'(0)| + c_1\kappa + c_2\kappa^{n+1}) \\
&\quad + \tfrac{1}{2}(c_1 + c_2\kappa^n)\big]|v_1 - v_2| \\
&=: \varphi_2(\kappa)|v_1 - v_2|.
\end{aligned} \tag{A.6}$$

Here $\mathbb{R}_+ \ni \kappa \mapsto \varphi_j(\kappa)$, $j = 1, 2$, are polynomial functions of $\kappa \in \mathbb{R}_+$ and do not depend on v_1 and v_2.

Fix $T \in \mathbb{R}_+$. For $s \in [0, T]$, $t \in [0, T]$, and $u \in [0, 1]$, consider the random variables $\kappa = \sqrt{T} \max_{0 \leq u \leq 1} |\beta_u| + x$, $v(s, u) = |uxe_1 + \sqrt{s}\beta_u|$, and $v(t, u) = |uxe_1 + \sqrt{t}\beta_u|$. Using the estimates established before, we obtain estimates for $I(t)$ in (A.5):

$$\begin{aligned}
|I(t)| &= \left| \int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\, du \right| \\
&\leq \int_0^1 |\gamma(|uxe_1 + \sqrt{t}\beta_u|)|\, du \\
&= \int_0^1 |\gamma(v(t, u))|\, du \\
&\leq \varphi_1(\kappa),
\end{aligned}$$

$$\begin{aligned}
s|I(t) - I(s)| &\leq s \int_0^1 |\gamma(v(t, u)) - \gamma(v(s, u))|\, du \\
&\leq s \int_0^1 |v(t, u) - v(s, u)|\varphi_2(\kappa)\, du \\
&\leq s(\sqrt{t} - \sqrt{s}) \int_0^1 |\beta_u|\varphi_2(\kappa)\, du \\
&\leq \frac{s(t - s)}{\sqrt{t} + \sqrt{s}} \frac{\kappa}{\sqrt{T}} \varphi_2(\kappa) \\
&\leq \frac{s\kappa\varphi_2(\kappa)}{2\sqrt{sT}}(t - s) \\
&\leq \frac{\kappa\varphi_2(\kappa)}{2}(t - s).
\end{aligned} \tag{A.7}$$

Here we have used (A.6) in the second inequality, since $0 \leq v(s, u) \leq \kappa$ and $0 \leq v(t, u) \leq \kappa$, the fact that $|v(t, u) - v(s, u)| \leq (\sqrt{t} - \sqrt{s})|\beta_u|$ in the third inequality, and the fact that $\max_{0 \leq u \leq 1} |\beta_u| \leq \kappa/\sqrt{T}$ in the fourth inequality for $0 < s < t \leq T$.

Finally, since γ is bounded from below by Assumption 3.1, so is I in (A.5), that is, $I(t) \geq \inf_{z \in \mathbb{R}_+} \gamma(z) > -\infty$ for every $t \geq 0$. With this observation, because of monotonicity and the differentiability of the exponential function, we obtain, for $0 \leq s < t \leq T$,

$$\begin{aligned}
\left| \xi(t) - \xi(s) \right| &= |e^{-tI(t)} - e^{-sI(s)}| \\
&\leq (e^{-T \inf \gamma} \vee 1)|tI(t) - sI(s)| \\
&= c_3|(t - s)I(t) + s(I(t) - I(s))| \\
&\leq c_3|t - s||I(t)| + c_3 s|I(t) - I(s)|,
\end{aligned}$$

where $c_3 := \exp(-T \inf \gamma) \vee 1 < \infty$. Combining this with the estimates for $|I(t)|$ and $s|I(t) - I(s)|$ in (A.7), we obtain, for $0 \le s < t \le T$,

$$|\xi(t) - \xi(s)| \le c_3\left(\varphi_1(\kappa) + \frac{\kappa\varphi_2(\kappa)}{2}\right)|t - s| =: \varphi_3(\kappa)|t - s|,$$

where φ_1 and φ_2 are defined in (A.6), and, hence, φ_3 can be written as a polynomial function of κ whose coefficients do not depend on s or t, but do depend on T. Letting Φ_T be $\varphi_3(\kappa)$, and noting that all positive integer moments of the maximum of a standard three-dimensional Bessel Bridge are finite (see [20, Corollary 7]) completes the proof of Lemma A.1.

Let us define $\nu(t) := \mathrm{E}_{\mathsf{BB}^3}[\xi(t)]$ for $0 \le t \le T$. It follows from Lemma A.1 that ξ is locally Lipschitz continuous and, moreover,

$$\mathrm{E}_{\mathsf{BB}^3}[|\xi(t) - \xi(s) - (\nu(t) - \nu(s))|^2] \le c_4|t - s|^2, \qquad 0 \le s < t \le T.$$

Since the random paths $\{\hat{\beta}^i, \ i = 1, \dots, N\}$ are independent and identically distributed, for any $s, t \le T$, we obtain

$$
\begin{aligned}
\mathrm{E}_{\mathsf{BB}^3}&|\eta^N(t) - \eta^N(s)|^2 \\
&= \exp\left(-2\int_0^x a(v)\,\mathrm{d}v\right)\mathrm{E}_{\mathsf{BB}^3}[|\xi(t) - \xi(s) - (\nu(t) - \nu(s))|^2] \\
&\le \exp\left(-2\int_0^x a(v)\,\mathrm{d}v\right)c_4|t - s|^2 \\
&=: c_5|t - s|^2,
\end{aligned}
\tag{A.8}
$$

where the constant c_5 depends on T and x, but not on N, s, and t. This inequality is a sufficient condition for the tightness of the sequence $\{\eta^N \mid N \in \mathbb{N}\}$ of continuous stochastic processes starting at 0 in $C(\mathbb{R}_+, \mathbb{R})$—see [11, Problem 2.4.12]. By the usual multi-dimensional central limit theorem, for each $n \ge 1$ and $0 \le t_1 < \cdots < t_n < \infty$, the sequence $\{(\eta^N(t_1), \eta^N(t_2), \dots, \eta^N(t_n)) \mid N \in \mathbb{N}\}$ of random vectors converges in distribution to a Gaussian random vector with mean 0 and variance-covariance matrix $(\Gamma(t_i, t_j))_{1 \le i, j \le n}$, where

$$\Gamma(s, t) = \exp\left(-2\int_0^x a(v)\,\mathrm{d}v\right)\mathrm{cov}_{\mathsf{BB}^3}[\exp(-sI(s)), \exp(-tI(t))], \qquad (s, t) \in \mathbb{R}_+^2.$$

Therefore, we conclude that the tight sequence $\{\eta^N \mid N \in \mathbb{N}\}$ converges weakly to a continuous Gaussian process with mean 0 and continuous covariance function Γ.

A.3. Proof of Proposition 3.2

Define the Gaussian tail function $\bar{\Phi}$ via

$$\bar{\Phi}(z) = \int_z^\infty \frac{\mathrm{e}^{-y^2/2}}{\sqrt{2\pi}}\,\mathrm{d}y \quad \text{for } z \in \mathbb{R}.$$

Furthermore, for fixed $T \in \mathbb{R}_+$, define the modulus of continuity in $\mathbb{L}^2$:

$$\psi_T(h) := \max_{(s,t)\in[0,T]^2, \ |t-s|\le h}(\mathrm{E}_{\mathsf{BB}^3}[\eta(t) - \eta(s)]^2)^{1/2} \quad \text{for } h \in [0, T]. \tag{A.9}$$

It follows from (A.8) that $\psi_T(h) \le \sqrt{c_5}h$ for $h \in \mathbb{R}_+$; therefore, $\int_1^\infty \psi_T(\mathrm{e}^{-y^2})\,\mathrm{d}y < \infty$. We now recall Fernique's inequality for Gaussian processes, which we will use.

Lemma A.2. (Fernique's inequality—see Equation (2.2) of [3].) *If the function ψ_T in (A.9) satisfies $\int_1^\infty \psi_T(e^{-y^2})\,dy$ then, for that fixed $T > 0$ and any integer $m \geq 2$,*

$$P_{\mathsf{BB}^3}\Big[\max_{0 \leq t \leq T}|\eta(t)| > C_1(T,m)z\Big] \leq C_2(m)\bar{\Phi}(z) \quad \text{for all } z > (1 + 4\log m)^{1/2},$$

where $C_1(T,m) := \max_{(s,t)\in[0,T]^2} \Gamma(s,t)^{1/2} + (2+\sqrt{2})\int_1^\infty \psi_T(Tm^{-y^2})\,dy$ and $C_2 := 5m^2\sqrt{2\pi}/2$.

The weak convergence of $\{\eta^N \mid N \in \mathbb{N}\}$ to η and the invariance principle for the maximum function imply that the sequence $(\sqrt{N}\max_{t\in[0,T]}(|\hat{p}_x^N(t) - p_x(t)|/q_x(t)))_{N\in\mathbb{N}}$ converges weakly to $\max_{0\leq t\leq T}\eta(t)$, where η is the limiting Gaussian process, as N goes to ∞. Since the law of the last random variable does not charge ∞, and since $\max_{t\in[0,T]}q_x(t) < \infty$, we conclude that the family $\{\sqrt{N}\max_{t\in[0,T]}|\hat{p}_x^N(t) - p_x(t)| \mid N \in \mathbb{N}\}$ is bounded in probability.

A.4. Proof of Proposition 4.1

For the lower bound in (4.2), observe that

$$\log E_{\mathsf{BB}^3}\Big[\exp\Big(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,du\Big)\Big] \leq -t\inf_{z\in\mathbb{R}_+}\gamma(z), \qquad t \in \mathbb{R}_+,$$

by Assumption 3.1. Therefore,

$$\inf_{z\in\mathbb{R}_+}\gamma(z) \leq \inf_{t\in\mathbb{R}_+}\Big[-\frac{1}{t}\log E_{\mathsf{BB}^3}\Big[\exp\Big(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,du\Big)\Big]\Big] = \inf_{t\in\mathbb{R}_+}\lambda_x(t).$$

For the upper bound in (4.2), for a fixed $\kappa > 0$, consider $A_\kappa := \{\sqrt{t}\max_{0\leq u\leq 1}|\beta_u| \leq \kappa\}$. On A_κ, $|uxe_1 + \sqrt{t}\beta_u| \leq \kappa + x$ holds for $0 \leq u \leq 1$; hence,

$$\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,du \leq \max_{0\leq z\leq\kappa+x}\gamma(z) = m(\kappa + x),$$

where $m(w) := \max_{0\leq z\leq w}\gamma(z)$ for $w > 0$. It follows that, for $t \in \mathbb{R}_+$,

$$E_{\mathsf{BB}^3}\Big[\exp\Big(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,du\Big)\Big]$$

$$\geq E_{\mathsf{BB}^3}\Big[\exp\Big(-t\int_0^1 \gamma(|uxe_1 + \sqrt{t}\beta_u|)\,du\Big)\mathbf{1}_{A_\kappa}\Big]$$

$$\geq \exp(-tm(\kappa + x))P_{\mathsf{BB}^3}\Big[\sqrt{t}\max_{0\leq u\leq 1}|\beta_u| \leq \kappa\Big],$$

and, hence,

$$\lambda_x(t) \leq m(\kappa + x) - \frac{1}{t}\log P_{\mathsf{BB}^3}\Big[\max_{0\leq u\leq 1}|\beta_u| \leq \kappa t^{-1/2}\Big]. \tag{A.10}$$

The distribution for the maximum of the absolute value of the standard three-dimensional Bessel bridge $|\beta|$ is known—see, for example, [20, Equation (5)]. More precisely, we have

$$P_{\mathsf{BB}^3}\Big[\max_{0\leq u\leq 1}|\beta_u| \leq \kappa t^{-1/2}\Big]$$

$$= \frac{2}{\kappa^3}\sqrt{\frac{2t}{\pi}}\sum_{n=1}^\infty \frac{n\pi}{J_{3/2}^2(n\pi)}\exp\Big(-\frac{\pi^2 n^2}{2\kappa^2}t\Big) \quad \text{for } \kappa > 0 \text{ and } t \in \mathbb{R}_+, \tag{A.11}$$

where $J_{3/2}$ is the Bessel function of index $\frac{3}{2}$. In particular,

$$\lim_{t\to\infty}\left(-\frac{1}{t}\log \mathrm{P}_{\mathsf{BB}^3}\left[\max_{0\le u\le 1}|\beta_u|\le \kappa t^{-1/2}\right]\right)=\frac{\pi^2}{2\kappa^2},$$

as only the first term in the summand in the series in (A.11) will play a role in the limit. Combining the last limiting relationship with inequality (A.10), we obtain

$$\limsup_{t\to\infty}\lambda_x(t)\le m(\kappa+x)+\frac{\pi^2}{2\kappa^2}.$$

Upon minimizing the right-hand side of the above inequality, the upper bound in (4.2) is obtained.

Finally, to verify (4.3), observe that γ being bounded from below implies that the random variables $\exp(-t\int_0^1\gamma(|uxe_1+\sqrt{t}\beta_u|)\,\mathrm{d}u)$ are uniformly bounded for small $t\in\mathbb{R}_+$. Then, l'Hôpital's rule and the bounded convergence theorem give

$$\begin{aligned}
\lim_{t\downarrow 0}\lambda_x(t) &= (-1)\lim_{t\downarrow 0}\frac{(\partial/\partial t)\,\mathrm{E}_{\mathsf{BB}^3}[\exp(-t\int_0^1\gamma(|uxe_1+\sqrt{t}\beta_u|)\,\mathrm{d}u)]}{\mathrm{E}_{\mathsf{BB}^3}[\exp(-t\int_0^1\gamma(|uxe_1+\sqrt{t}\beta_u|)\,\mathrm{d}u)]}\\
&= \int_0^1\gamma(ux)\,\mathrm{d}u\\
&= \frac{1}{x}\int_0^x\gamma(u)\,\mathrm{d}u.
\end{aligned}$$

Acknowledgements

References

[1] ALILI, L. AND PATIE, P. (2010). Boundary-crossing identities for diffusions having the time-inversion property. *J. Theoret. Prob.* **23**, 65–84.

[2] BALDI, P., CARAMELLINO, L. AND IOVINO, M. G. (1999). Pricing general barrier options: a numerical approach using sharp large deviations. *Math. Finance* **9**, 293–322.

[3] BERMAN, S. M. (1985). An asymptotic bound for the tail of the distribution of the maximum of a Gaussian process. *Ann. Inst. H. Poincaré Prob. Statist.* **21**, 47–57.

[4] DURBIN, J. (1985). The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Prob.* **22**, 99–122.

[5] ELLIOTT, J. (1955). Eigenfunction expansions associated with singular differential operators. *Trans. Amer. Math. Soc.* **78**, 406–425.

[6] ELWORTHY, K. D., LI, X.-M. AND YOR, M. (1999). The importance of strictly local martingales; applications to radial Ornstein–Uhlenbeck processes. *Prob. Theory Relat. Fields* **115**, 325–355.

[7] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering* (Appl. Math. **53**). Springer, New York.

[8] GOBET, E. (2000). Weak approximation of killed diffusion using Euler schemes. *Stoch. Process. Appl.* **87**, 167–197.

[9] GÖING-JAESCHKE, A. AND YOR, M. (2003). A clarification note about hitting times densities for Ornstein–Uhlenbeck processes. *Finance Stoch.* **7**, 413–415.

[10] HINTON, D. AND SCHAEFER, P. W. (eds) (1997). *Spectral Theory and Computational Methods of Sturm–Liouville Problems* (Lecture Notes Pure Appl. Math. **191**). Marcel Dekker, New York.

[11] KARATZAS, I. AND SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus* (Graduate Texts Math. **113**), 2nd edn. Springer, New York.

[12] KENT, J. T. (1980). Eigenvalue expansions for diffusion hitting times. *Z. Wahrscheinlichkeitsth.* **52,** 309–319.

[13] KLOEDEN, P. E. AND PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations* (Appl. Math. **23**). Springer, Berlin.

[14] LEHMANN, A. (2002). Smoothness of first passage time distributions and a new integral equation for the first passage time density of continuous Markov processes. *Adv. Appl. Prob.* **34,** 869–887.

[15] MILSTEIN, G. N. AND TRETYAKOV, M. V. (2004). Evaluation of conditional Wiener integrals by numerical integration of stochastic diferential equations. *J. Comput. Phys.* **197,** 275–298.

[16] MILSTEIN, G. N. AND TRETYAKOV, M. V. (2004). *Stochastic Numerics for Mathematical Physics*. Springer, Berlin.

[17] PAUWELS, E. J. (1987). Smooth first-passage densities for one-dimensional diffusions. *J. Appl. Prob.* **24,** 370–377.

[18] PESKIR, G. (2002). On integral equations arising in the first-passage problem for Brownian motion. *J. Integral Equations Appl.* **14,** 397–423.

[19] PINSKY, R. G. (1995). *Positive Harmonic Functions and Diffusion*. (Camb. Stud. Adv. Math. **45**). Cambridge University Press.

[20] PITMAN, J. AND YOR, M. (1999). The law of the maximum of a Bessel bridge. *Electron. J. Prob.* **4,** 35pp.

[21] REVUZ, D. AND YOR, M. (1999). *Continuous Martingales and Brownian Motion*. (Fundamental Principles Math. Sci. **293**), 3rd edn. Springer, Berlin.

[22] ROGERS, L. C. G. (1985). Smooth transition densities for one-dimensional diffusions. *Bull. London Math. Soc.* **17,** 157–161.

[23] ZETTL, A. (2005). *Sturm-Liouville Theory* (Math. Surveys Monogr. **121**). American Mathematical Society, Providence, RI.