J. Appl. Probab. **60,** 1386–1415 (2023)
*doi:*10.1017/jpr.2023.11

ON MARKOV CHAIN APPROXIMATIONS FOR COMPUTING BOUNDARY CROSSING PROBABILITIES OF DIFFUSION PROCESSES

VINCENT LIANG,* ** AND
KONSTANTIN BOROVKOV ,* *** *The University of Melbourne*

Abstract

We propose a discrete-time discrete-space Markov chain approximation with a Brownian bridge correction for computing curvilinear boundary crossing probabilities of a general diffusion process on a finite time interval. For broad classes of curvilinear boundaries and diffusion processes, we prove the convergence of the constructed approximations in the form of products of the respective substochastic matrices to the boundary crossing probabilities for the process as the time grid used to construct the Markov chains is getting finer. Numerical results indicate that the convergence rate for the proposed approximation with the Brownian bridge correction is $O(n^{-2})$ in the case of C^2 boundaries and a uniform time grid with n steps.

Keywords: boundary crossing probability; diffusion processes; Markov chains

2020 Mathematics Subject Classification: Primary 60J60
Secondary 60J70; 65C30

1. Introduction

We consider the problem of approximating the probability for a general diffusion process to stay between two curvilinear boundaries. Mathematically, the problem is solved: the non-crossing probability can be expressed as a solution to the respective boundary value problem for the backward Kolmogorov partial differential equation (this result goes back to the 1930s [34, 36, 37]). However, simple explicit analytic expressions are confined to the case of the Wiener process using the method of images [38], and most of the results for diffusion processes rely on verifying Cherkasov's conditions [12, 51, 53] and then transforming the problem to that for the Wiener process by using a monotone transformation. Outside this class of special cases, we should mostly rely on computing numerical approximations to the desired probabilities.

One popular approach to finding expressions for the first-passage-time density is through the use of Volterra integral equations. Much work was done on the method of integral equations for approximating the first-passage-time density for general diffusion processes [10, 11, 20, 32, 52, 53, 55], culminating with [24], which expressed the first-passage-time density of a general diffusion process in terms of the solution to a Volterra integral equation of the second kind. Volterra integral equations of the first kind for the first-passage time for Brownian motion were derived in [41, 47]. Although the method of integral equations is quite efficient

Received 9 December 2021; revision received 10 February 2023.
* Postal address: School of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia
** Email address: vliang@student.unimelb.edu.au
*** Email address: borovkov@unimelb.edu.au

for computational purposes, a drawback of the method is that the kernel of the integral equation is expressed in terms of the transition probabilities of the diffusion process. Therefore, for this method to work, we have to first compute these transition probabilities for all times on the time grid used. The method proposed in this paper only requires knowledge of the drift and diffusion coefficients, allowing it to be easily used in the general case. For further details on the connection between Volterra integral equations and the first-passage-time density, we refer the reader to [49]. Other computational techniques include Monte Carlo [21, 31], including exact simulation methods [4, 5, 27, 28] and continuity corrections employing boundary shifting to compensate for the 'discrete monitoring bias' [9, 22, 58], and numerical solving of partial differential equations (see, e.g., [48] and references therein).

The classical probabilistic approach to the boundary-crossing problem is the method of weak approximation, which involves proving that a sequence of simpler processes X_n weakly converges to the desired diffusion process in a suitable functional space, which entails the convergence of the corresponding non-crossing probabilities. Along with the already mentioned [34, 36, 37], one can say that this approach was effectively used in [17] in the case of 'flat boundaries' (see also [6, Chapter 2, 11]). More recently, the authors in [19] approximated the Wiener process with a sequence of discrete Markov chains with absorbing states and expressed the non-crossing probability as a product of transition matrices. The authors in [33] extended the results in [19] by approximating a general diffusion with a sequence of Markov chains and similarly expressed the non-crossing probability as a product of transition matrices. However, the convergence rates of these approximations were proved to be $O(n^{-1/2})$, which leaves much to be desired in practical applications.

Another standard approach is to approximate the true boundary with one for which the crossing probability is easier to compute. In the one-sided boundary case, [50, 61] used piecewise-linear approximations and the well-known formula for a one-sided linear boundary crossing probability for the Brownian bridge process to express the non-crossing probability in terms of a multiple Gaussian integral. This was generalised to the case of two-sided boundaries in [45] and to diffusion processes set up in that case in [62]. In the case of Brownian motion, a sharp explicit error bound for the approximation of the corresponding boundary crossing probabilities was obtained in [8] and extended to general diffusion processes in [15].

The present paper combines piecewise-linear boundary approximations, limit theorems on convergence of Markov chains to diffusions, and a modification of the matrix multiplication scheme from [19] to create an efficient and tractable numerical method for computing the boundary crossing probabilities for time-inhomogeneous diffusion processes in both one- and two-sided boundary cases. The approach in the paper consists of the following steps:

(i) Transform the original general diffusion process into one with unit diffusion coefficient, applying the same transformation to the boundaries.

(ii) Approximate the transformed diffusion process with a discrete-time Gaussian Markov process using a weak Taylor expansion.

(iii) Approximate the discrete-time process from step (ii) with a discrete Markov chain in discrete time, whose transition probabilities are given by the normalised values of the transition density of the process from step (ii). The state spaces of the discrete Markov chain are constructed in such a way that the Markov chain does not overshoot or undershoot the boundaries when hitting them.

(iv) Construct a continuous-time process that interpolates the Markov chain from step (iii) with a collection of Brownian bridges.

(v) Approximate the transformed boundaries with piecewise linear ones and compute the piecewise linear boundary crossing probability of the interpolated process constructed in step (iv) using matrix multiplication.

The paper is structured as follows. In Section 2 we describe in detail the above steps. Section 3 states the main results of the paper. These include convergence in distribution of our approximating discrete schemes with Brownian bridge interpolations to the original process and, as a corollary, convergence of the respective boundary crossing probabilities as well. That section also contains a sketch of a possible argument showing that the convergence rate of the boundary crossing probability approximations is $O(n^{-2})$ provided that the nth time interval partition has rank $O(n^{-1})$ as $n \to \infty$. In Section 4 we present the proofs of these results. Section 5 contains numerical examples and a brief discussion of the performance of our method compared to some existing alternative approaches.

2. Markov chain approximation of a diffusion

2.1. Step (i)

We are interested in the boundary crossing probability of a one-dimensional diffusion process Y, whose dynamics are governed by the following stochastic differential equation:

$$\begin{cases} \mathrm{d}Y(t) = \mu_Y(t, Y(t))\,\mathrm{d}t + \sigma_Y(t, Y(t))\,\mathrm{d}W(t), \quad t \in (0, 1], \\ Y(0) = y_0, \end{cases} \tag{1}$$

where $\{W_t\}_{t\ge 0}$ is a standard Brownian motion process defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ and y_0 is constant. We assume that $\mu_Y\colon [0, 1] \times \mathbb{R} \to \mathbb{R}$ and $\sigma_Y\colon [0, 1] \times \mathbb{R} \to (0, \infty)$ satisfy the following conditions sufficient for the uniqueness and existence of a strong solution to (1) [18, pp. 297–298]:

Condition 1. *The functions μ_Y and σ_Y are continuous in both variables and such that, for some $K < \infty$, $x\mu_Y(t, x) \le K(1 + x^2)$ and $\sigma_Y^2(t, x) \le K(1 + x^2)$ for all $(t, x) \in [0, 1] \times \mathbb{R}$ and, for any bounded open set $U \subset \mathbb{R}$, there exists a $K_U < \infty$ such that*

$$|\mu_Y(t, x) - \mu_Y(t, y)| + |\sigma_Y^2(t, x) - \sigma_Y^2(t, y)| \le K_U|x - y|$$

for all $t \in [0, 1]$, $x, y \in U$.

For the unit-diffusion transformation $\psi_t(y) := \int_0^y \sigma_Y(t, u)^{-1}\,\mathrm{d}u$ to be well defined, we will also assume the following.

Condition 2. *The function σ_Y is continuously differentiable with respect to (t, x) with bounded partial derivatives, and $\inf_{(t,x)\in[0,1]\times\mathbb{R}} \sigma_Y(t, x) > 0$.*

By Itô's lemma, the transformed process $X = \{X(t) := \psi_t(Y(t))\}_{t\in[0,1]}$ is a diffusion process with a unit diffusion coefficient (see, e.g., [35, Section 4.7] or [54]),

$$\begin{cases} \mathrm{d}X(t) = \mu_X(t, X(t))\,\mathrm{d}t + \mathrm{d}W(t), \quad t \in (0, 1], \\ X(0) = x_0 := \psi_0(y_0), \end{cases} \tag{2}$$

where $\mu_X(t, x) = (\partial_t\psi_t + \mu_Y/\sigma_Y - \frac{1}{2}\partial_x\sigma_Y) \circ \psi_t^{-1}(x)$ and $\psi_t^{-1}(z)$ is the inverse of $z = \psi_t(y)$ in y.

Denote by $C = C([0, 1])$ the space of continuous functions $x\colon [0, 1] \to \mathbb{R}$ equipped with the uniform norm $\|x\|_\infty := \sup_{t\in[0,1]}|x(t)|$. For a fixed $x_0 \in \mathbb{R}$, consider the class

$$\mathcal{G} := \{(f^-, f^+)\colon f^\pm \in C,\ f^-(0) < x_0 < f^+(0),\ \inf_{0\le t\le 1} (f^+(t) - f^-(t)) > 0\}$$

of pairs of functions from C and introduce the notation

$$S(f^-, f^+) := \{x \in C\colon f^-(t) < x(t) < f^+(t),\ t \in [0, 1]\}, \qquad (f^-, f^+) \in \mathcal{G}.$$

The problem we deal with in this paper is how to compute the probability of the form $\mathbb{P}(Y \in S(g_0^-, g_0^+))$ for $(g_0^-, g_0^+) \in \mathcal{G}$. Clearly, the desired probability coincides with $\mathbb{P}(X \in G)$, $G := S(g^-, g^+)$, where $g^\pm(t) := \psi_t(g_0^\pm(t))$, $t \in [0, 1]$. It is also clear that $(g^-, g^+) \in \mathcal{G}$ due to Condition 2. Henceforth, we work exclusively with the process X and the boundaries $g^\pm$.

2.2. Step (ii)

In the context of curvilinear boundary crossing probabilities, [19] approximated the Wiener process by discrete Markov chains with transition probabilities computed by first taking the values of the Wiener process transition densities on a lattice and then normalising these to obtain a probability distribution on that lattice (there is also a small adjustment of that distribution to perfectly match the first two moments of the original transition probabilities and the 'discretised' ones).

In the general diffusion case, due to the absence of closed-form expressions for the transition density of X, we use transition probabilities of the weak Taylor approximations of X to construct the approximating discrete process. As the second-order expansion is to be used, we further require that the following condition is met.

Condition 3. *For any fixed* $x \in \mathbb{R}$, $\mu_X(\,\cdot\,, x) \in C^1([0, 1])$, *and for any fixed* $t \in [0, 1]$, $\mu_X(t,\ \cdot\,) \in C^2(\mathbb{R})$. *Moreover, for any* $r > 0$, *there exists a* $K_r < \infty$ *such that*

$$|\mu_X(t, x)| + |\partial_t\mu_X(t, x)| + |\partial_x\mu_X(t, x)| + |\partial_{xx}\mu_X(t, x)| \le K_r, \qquad t \in [0, 1],\ |x| \le r.$$

Next, for any $n \ge 1$, let $t_{n,k} := k/n$, $k = 0, 1, \dots, n$, be the uniform partition of [0,1] of rank $\Delta_n := 1/n$. Introduce discrete scheme drift $\beta_{n,k}$ and diffusion $\alpha_{n,k}^{1/2}$ coefficients by setting, for $k = 1, 2, \dots, n$,

$$\beta_{n,k}(x) := \big(\mu_X + \tfrac{1}{2}\Delta_n(\partial_t\mu_X + \mu_X\partial_x\mu_X + \tfrac{1}{2}\partial_{xx}\mu_X)\big)(t_{n,k-1}, x), \tag{3}$$

$$\alpha_{n,k}^{1/2}(x) := 1 + \tfrac{1}{2}\Delta_n\partial_x\mu_X(t_{n,k-1}, x). \tag{4}$$

For each fixed $n \ge 1$, the nth second-order weak Taylor approximation of the diffusion (2) is defined as the discrete-time process $\zeta_n := \{\zeta_{n,k}\}_{k=0}^n$, specified by $\zeta_{n,0} = x_0$ and

$$\zeta_{n,k} = \zeta_{n,k-1} + \beta_{n,k}(\zeta_{n,k-1})\Delta_n + \alpha_{n,k}^{1/2}(\zeta_{n,k-1})\Delta_n^{1/2}Z_{n,k}, \qquad k = 1, \dots, n,$$

where $\{Z_{n,k}\}_{k=1}^n$ is a triangular array of independent standard normal random variables. For more detail on weak Taylor approximations to solutions of stochastic differential equations, see, e.g., [35, Chapter 14]. Clearly, the conditional distributions of the increments $\zeta_{n,k} - \zeta_{n,k-1}$ given $\zeta_{n,k-1}$ are Gaussian, and so the transition probabilities of the discrete-time process ζ_n can be easily obtained.

2.3. Step (iii)

Next, we further approximate the discrete-time continuous state space process ζ_n with a discrete-time discrete-state-space Markov chain $\xi_n := \{\xi_{n,k}\}_{k=0}^n$, whose transition probabilities are based on the normalised values of the transition density of ζ_n.

To improve the convergence rate for our approximations to $\mathbb{P}(X \in G)$, we construct our Markov chains ξ_n choosing, generally speaking, different state spaces for each time step. Namely, the state space $E_{n,k}$ for $\xi_{n,k}$, $k = 0, 1, \ldots, n$, is chosen to be a lattice such that $g^{\pm}(t_{n,k}) \in E_{n,k}$. This modification improves upon the Markov chain approximation suggested in [19], and is widely used for accelerating the convergence rate of numerical schemes in barrier option pricing [13]. More precisely, the spaces $E_{n,k}$ are constructed as follows. Let $g^{\pm}_{n,k} := g^{\pm}(t_{n,k})$, $k = 1, \ldots, n$, and, for fixed $\delta \in (0, \frac{1}{2}]$ and $\gamma > 0$, set

$$w_{n,k} := \begin{cases} \dfrac{(g^+_{n,k} - g^-_{n,k})/\Delta_n^{1/2+\delta}}{\lfloor \gamma (g^+_{n,k} - g^-_{n,k})/\Delta_n^{1/2+\delta} \rfloor}, & 1 \le k < n, \\ \dfrac{(g^+(1) - g^-(1))/\Delta_n}{\lfloor \gamma (g^+(1) - g^-(1))/\Delta_n \rfloor}, & k = n, \end{cases}$$

assuming that n is large enough that the integer parts in all the denominators are non-zero. We set the time-dependent space lattice step sizes to be

$$h_{n,k} := \begin{cases} w_{n,k}\Delta_n^{1/2+\delta}, & 1 \le k < n, \\ w_{n,n}\Delta_n, & k = n. \end{cases} \tag{5}$$

The state space for $\xi_{n,k}$ is the $h_{n,k}$-spaced lattice

$$E_{n,k} := \{g^+_{n,k} - jh_{n,k} \colon j \in \mathbb{Z}\}, \qquad k = 1, \ldots, n. \tag{6}$$

We also put $E_{n,0} := \{x_0\}$ and $E_n := E_{n,0} \times E_{n,1} \times \cdots \times E_{n,n}$. Note that

$$\max_{1 \le k \le n} |w_{n,k} - \gamma^{-1}| \to 0 \quad \text{as } n \to \infty, \tag{7}$$

and $\gamma^{-1} \le w_{n,k} \le 2\gamma^{-1}$ for all $1 \le k \le n$ (assuming, as above, that n is large enough).

Further, for $k = 1, \ldots, n$, recalling (3) and (4), we let

$$\mu_{n,k}(x) := \beta_{n,k}(x)\Delta_n, \qquad \sigma^2_{n,k}(x) := \alpha_{n,k}(x)\Delta_n, \tag{8}$$

and define $\xi_{n,k}$, $k = 1, \ldots, n$, $n \ge 1$, as a triangular array of random variables, where each row forms a Markov chain with one-step transition probabilities given by

$$p_{n,k}(x, y) := \mathbb{P}(\xi_{n,k} = y \mid \xi_{n,k-1} = x) = \varphi(y \mid x + \mu_{n,k}(x),\ \sigma^2_{n,k}(x)) \frac{h_{n,k}}{C_{n,k}(x)}$$

for $(x, y) \in E_{n,k-1} \times E_{n,k}$, where $\varphi(x \mid \mu, \sigma^2) := (2\pi\sigma^2)^{-1/2} \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$, $x \in \mathbb{R}$, denotes the Gaussian density with mean μ and variance σ^2, and

$$C_{n,k}(x) := \sum_{y \in E_{n,k}} \varphi(y \mid x + \mu_{n,k}(x), \sigma^2_{n,k}(x)) h_{n,k}, \qquad x \in \mathbb{R}. \tag{9}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

2.4. Step (iv)

As an initial approximation for $\mathbb{P}(X \in G)$ we could take $\mathbb{P}(g^-_{n,k} < X(t_{n,k}) < g^+_{n,k}, k = 1, \ldots, n)$, which can in turn be approximated using the Chapman–Kolmogorov equations by

$$\mathbb{P}(g^-_{n,k} < \xi_{n,k} < g^+_{n,k},\ k = 1, \ldots, n) = \sum_{\boldsymbol{x}\in E^G_n} \prod_{k=1}^{n} p_{n,k}(x_{k-1}, x_k),$$

where $\boldsymbol{x} = (x_0, x_1, \ldots, x_n)$ and $E^G_n := E_{n,0} \times E^G_{n,1} \times \cdots \times E^G_{n,n}$,

$$E^G_{n,k} := \{x \in E_{n,k} \colon g^-_{n,k} < x < g^+_{n,k}\}, \qquad k = 1, \ldots, n.$$

Without loss of generality, we can assume that n is large enough that none of $E^G_{n,k}$ is empty.

An approximation of this kind was used in [19] to approximate the boundary crossing probability of the Brownian motion. Such discrete-time approximations of boundary crossing probabilities are well known to converge slowly to the respective probability in the 'continuously monitored' case since they fail to account for the probability of boundary crossing by the continuous-time process at an epoch inside a time interval between consecutive points on the time grid used.

In order to correct for this, so-called 'continuity corrections' have been studied in the context of sequential analysis [58] and, more recently, in the context of barrier option pricing [9]. These types of corrections have also been used in [50] to correct for discrete-time monitoring bias in Monte Carlo estimates of the boundary crossing probabilities of the Brownian motion. Without such a correction, using the classical result from [43, 44], the convergence rate of the approximation from [19] was shown to be $O(n^{-1/2})$. However, as our numerical experiments demonstrate, using our more accurate approximations of the transition probabilities in conjunction with the Brownian bridge correction greatly improves it from $O(n^{-1/2})$ to $O(n^{-2})$.

In the case of standard Brownian motion, the correction consists of simply multiplying the one-step transition probabilities by the non-crossing probability of a suitably pinned Brownian bridge. Due to the local Brownian nature of a diffusion process, it was shown in [3] that the leading-order term of the diffusion bridge crossing probability is given by an expression close to that of the Brownian bridge. Thus, to account for the possibility of the process X crossing the boundary inside time intervals $[t_{n,k-1}, t_{n,k}]$, we define a process $\widetilde{X}_n := \{\widetilde{X}_n(t)\}_{t\in[0,1]}$ which interpolates between the subsequent nodes $(t_{n,k}, \xi_{n,k})$ with a collection of independent Brownian bridges:

$$\widetilde{X}_n(t) := B^{\xi_{n,k-1},\xi_{n,k}}_{n,k}(t), \qquad t \in [t_{n,k-1}, t_{n,k}],\ k = 1, \ldots, n, \tag{10}$$

where $B^{x,y}_{n,k}(t) := B^{\circ}_{n,k}(t) + x + n(t - t_{n,k-1})(y - x)$, $x, y \in \mathbb{R}$, and $B^{\circ}_{n,k}(t)$, $t \in [t_{n,k-1}, t_{n,k}]$, are independent Brownian motions 'pinned' at the time–space points $(t_{n,k-1}, 0)$ and $(t_{n,k}, 0)$, these Brownian bridges being independent of ξ_n. Analogous to [45, Theorem 1], using the Chapman–Kolmogorov equations, the non-crossing probability of the boundaries $g^{\pm}$ for $\widetilde{X}_n$ can be expressed as

$$\mathbb{P}(\widetilde{X}_n \in G) = \mathbb{E} \prod_{k=1}^{n} \left(1 - \pi_{n,k}(g^-, g^+ \mid \xi_{n,k-1}, \xi_{n,k})\right), \tag{11}$$

where

$$\pi_{n,k}(g^-, g^+ \mid x, y) := \mathbb{P}\left(\sup_{t\in[t_{n,k-1},t_{n,k}]} (B^{x,y}_{n,k}(t) - g^+(t))(B^{x,y}_{n,k}(t) - g^-(t)) \geq 0 \right)$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

is the probability that the trajectory of a Brownian motion process pinned at the points $(t_{n,k-1}, x)$ and $(t_{n,k}, y)$ will be outside of the 'corridor' between the boundaries $g^{\pm}(t)$ at some point during the time interval $[t_{n,k-1}, t_{n,k}]$.

2.5. Step (v)

The above representation can be equivalently written as a matrix product:

$$\mathbb{P}(\widetilde{X}_n \in G) = \mathbf{T}_{n,1}\mathbf{T}_{n,2} \cdots \mathbf{T}_{n,n}\mathbf{1}^{\top}, \tag{12}$$

where $\mathbf{1} = (1, \ldots, 1)$ is a row vector of length $|E^G_{n,n}|$, and the sub-stochastic matrices $\mathbf{T}_{n,k}$ of dimensions $|E^G_{n,k-1}| \times |E^G_{n,k}|$ have entries equal to the respective taboo transition probabilities

$$(1 - \pi_{n,k}(g^-, g^+ \mid x, y))p_{n,k}(x, y), \qquad (x, y) \in E^G_{n,k-1} \times E^G_{n,k}.$$

Unfortunately, closed-form expressions for curvilinear boundary Brownian bridge crossing probabilities $\pi_{n,k}(g^-, g^+ \mid x, y)$ are known in a few special cases only, so we approximate the original boundaries $g^{\pm}$ with piecewise linear functions $f_n^{\pm}$, which linearly interpolate between the subsequent nodes $(t_{n,k}, g^{\pm}_{n,k})$ for $k = 0, \ldots, n$. In the special case of a one-sided boundary (when $g^- = -\infty$), the expression for the linear boundary crossing probability of the Brownian bridge is well known [7, p. 63]:

$$\pi_{n,k}(-\infty, f_n^+ \mid x, y) = \exp\left\{\frac{-2}{\Delta_n}(g^+_{n,k-1} - x)(g^+_{n,k} - y)\right\}, \qquad x < g^+_{n,k-1}, y < g^+_{n,k}.$$

In the case when both the upper and lower boundaries $f_n^{\pm}$ are linear, if the time interval Δ_n is sufficiently small, we can approximate the Brownian bridge crossing probability with the sum of one-sided crossing probabilities:

$$\pi_{n,k}(f_n^-, f_n^+ \mid x, y) = \pi_{n,k}(f_n^-, \infty \mid x, y) + \pi_{n,k}(-\infty, f_n^+ \mid x, y) - \vartheta(x, y, \Delta_n),$$

where the positive error term ϑ admits the obvious upper bound

$$\begin{aligned}\vartheta(x, y, \Delta_n) \le \pi_{n,k}(f_n^-, \infty \mid x, y) \max_{t\in[t_{n,k-1},t_{n,k}]} P_{t,f_n^-(t);t_{n,k},y}\left(\sup_{s\in[t,t_{n,k}]}(W(s) - f_n^+(s)) \ge 0\right) \\ + \pi_{n,k}(-\infty, f_n^+ \mid x, y) \max_{t\in[t_{n,k-1},t_{n,k}]} P_{t,f_n^+(t);t_{n,k},y}\left(\inf_{s\in[t,t_{n,k}]}(W(s) - f_n^-(s)) \le 0\right),\end{aligned}$$

where $P_{s,a;t,b}(\,\cdot\,) := \mathbb{P}(\,\cdot \mid W(s) = a, W(t) = b)$. An infinite series expression for $\pi_{n,k}(f_n^-, f_n^+ \mid x, y)$ can be found, e.g., in [2, 25].

We can further apply our method to approximate probabilities of the form $\mathbb{P}(X \in G, X(1) \in [a, b])$ for some $[a, b] \subseteq [g^-(1), g^+(1)]$. We first replace the final space grid $E^G_{n,n}$ with

$$E^{[a,b]}_{n,n} := \left\{a \le x \le b \colon x = b - j\frac{b-a}{\lfloor\gamma(b-a)/\Delta_n\rfloor}, j \in \mathbb{Z}\right\}.$$

Then, instead of (12) we have $\mathbb{P}(\widetilde{X}_n \in G, \widetilde{X}_n(1) \in [a, b]) = \mathbf{T}_{n,1}\mathbf{T}_{n,2} \cdots \mathbf{T}_{n,n}\mathbf{1}^{\top}_{[a,b]}$, where $\mathbf{T}_{n,n}$ is now of dimension $|E^G_{n,n-1}| \times |E^{[a,b]}_{n,n}|$ and $\mathbf{1}_{[a,b]} = (1, \ldots, 1)$ is a row vector of length $|E^{[a,b]}_{n,n}|$.

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

2.6. Remarks

We conclude this section with several remarks aimed at clarifying the choices we made in our construction and the connections of our method with existing approaches in the literature.

Remark 1. Our main results actually hold in a more general setting where, instead of using *uniform partitions*, we employ general ones $0 = t_{n,0} < t_{n,1} < \cdots < t_{n,n} = 1$ subject to the condition that, for some constants $\eta_2 \geq \eta_1 > 0$ and all $n \geq 1$,

$$\frac{\eta_1}{n} \leq t_{n,k} - t_{n,k-1} \leq \frac{\eta_2}{n}, \qquad k = 1, 2, \ldots, n. \tag{13}$$

All the proofs extend to this case in a straightforward way.

Using a more general sequence of partitions satisfying (13) helps to improve convergence rates by choosing a higher frequency of partition nodes on time intervals where the second derivatives of the boundaries $g^{\pm}$ are large.

Remark 2. Note that the fact that the transition probabilities of the 'time-discretised' process ζ are Gaussian is important. High-order approximation of transition semigroups is achieved by matching sufficiently many moments for one-step transition probabilities (see, e.g., [29, Theorem 11.6.3] and [1]), and in the case of Gaussian distributions, the moments are available. When the transition kernels are Gaussian, for the discrete-space approximation used in our paper, convergence is extremely fast (cf. Lemma 3).

Observe also that directly using a weak Taylor expansion for a general diffusion process with a *space-dependent diffusion coefficient* results in a non-central χ^2 distribution for the transition probabilities [16], which artificially limits the domain of the approximating process.

For other approaches to approximating the transition density of a general diffusion process, see, e.g., [30, 39] and the references therein.

Remark 3. The choice of the lattice spans in (5) can be explained as follows. Our approximation scheme involves replacing the original boundaries with piecewise linear ones (in Step (v)), which introduces an error of the order $O(n^{-2})$ (under assumption (13) and given that the functions $g^{\pm}$ are twice continuously differentiable). Therefore, there is no point in using $h_{n,k}$ smaller than necessary to achieve the above precision. At the end of Section 3 we present a sketch of an argument indicating that the convergence rate for the boundary crossing probabilities is $O(n^{-2})$. It shows that it suffices to choose $h_{n,n} \asymp n^{-1}$, whereas $h_{n,k} \asymp n^{-1/2-\delta}$, $k < n$, can be much larger. This is so because our Markov chain computational algorithm can be viewed as repeated trapezoidal quadrature and the partial derivative in y of the taboo (on boundary crossing) joint density of $(X(t_{n,k-1}), X(t_{n,k+1}))$ given $X(t_{n,k}) = y$ at the boundary $y = g(t_{n,k})$ at all time steps before the terminal time is zero, ensuring a higher approximation rate at these steps compared to the terminal one (cf. the Euler–Maclaurin formula).

Remark 4. If we used $\delta = 0$ in (5), there would be no convergence of the sequence of processes $\{\xi_n\}_{n \geq 1}$ to the desired diffusion limit. This is because there would be no convergence of the moments of the increments (cf. Lemma 3). Note also that, instead of using the power function (5), we could take $h_{n,k} := w_{n,k}(\Delta_n)^{1/2}\psi(\Delta_n)$ for some $\psi(x) \to 0$ as $x \downarrow 0$, with $w_{n,k} := v_{n,k}/\lfloor v_{n,k} \rfloor$, where

$$v_{n,k} := \frac{g^{+}_{n,k} - g^{-}_{n,k}}{(\Delta_n)^{1/2}\psi(\Delta_n)},$$

(in our case, $\psi(x) = x^{\delta}/\gamma$). We chose the power function for simplicity's sake.

Remark 5. Choosing suitable $\delta < \frac{1}{2}$ and $\gamma > 0$ in the definition of $h_{n,k}$ for $k < n$ can be used to reduce the computational burden. In our numerical experiments, we found that reducing the value of $\delta \in \left(0, \frac{1}{2}\right)$ did not negatively affect the empirical convergence rates for boundary crossing probabilities if γ is sufficiently large ($\gamma \geq 1.5$). In the case $\delta = 0$, our proposed scheme no longer converges since the infinitesimal moments do not converge (cf. Lemma 3). To restore convergence, we could use the adjusted transition probabilities suggested in [19]. However, we would not be able to apply the shifting state space methodology above since the approximation in [19] relies on adjusting the transition probabilities on a state space that is not changing each time step.

Remark 6. In the standard Brownian motion case, the matrix multiplication scheme in (12) can be seen as recursive numerical integration using the trapezoidal rule (cf. [45, Remark 3]). We may be tempted to employ higher-order quadrature methods instead. However, constructing a Markov chain approximation based on numerical integration techniques that use variable node positioning (e.g. Gauss–Legendre quadrature) would require interpolation of the resulting evolved transition density, causing the matrix multiplication in (11) to lose its probabilistic meaning. An alternative approach using efficient numerical integration techniques based on analytic mappings (e.g. double-exponential quadrature [60]) that maintain the positions of nodes is numerically feasible. However, verifying theoretical weak convergence for the resulting sequence of Markov chains is more difficult. Furthermore, using the trapezoidal rule allows us to use the Euler–Maclaurin summation formula to improve the convergence rate. From our numerical experimentation, the one-dimensional trapezoidal rule that we propose in this paper strikes a good balance between simplicity, flexibility, and numerical efficiency.

3. Main results

Set $\nu_n(t) := \max\{k \geq 0\colon t \geq t_{n,k}\}$, $t \in [0, 1]$, and introduce an auxiliary pure jump process

$$X_n(t) := \xi_{n,\nu_n(t)}, \qquad t \in [0, 1]. \tag{14}$$

Clearly, the trajectories of the process X_n belong to the Skorokhod space $D = D([0, 1])$, which we will endow with the first Skorokhod metric,

$$d(x, y) = \inf_{\lambda\in\Lambda} \left\{\varepsilon \geq 0\colon \sup_{t\in[0,1]} |x(t) - y(\lambda(t))| \leq \varepsilon, \sup_{t\in[0,1]} |\lambda(t) - t| \leq \varepsilon\right\}, \qquad x, y \in D,$$

where Λ denotes the class of strictly increasing continuous mappings of [0, 1] onto itself [6, Chapter 3]. We will use $\Rightarrow$ to denote convergence in distribution of random elements of (D, d).

Theorem 1. *Under Condition* [3], $X_n \Rightarrow X$ *as* $n \to \infty$.

Due to the small amplitude of the interpolating Brownian bridges' oscillations, it is unsurprising that the sequence of processes $\{\widetilde{X}_n\}$ also converges weakly to X.

Corollary 1. $\widetilde{X}_n \Rightarrow X$ *as* $n \to \infty$.

The following result is a theoretical justification of the Markov chain approximation method proposed in this paper.

Corollary 2. *Let* (g^-, g^+), (g_n^-, g_n^+), $n \geq 1$, *be elements of* $\mathcal{G}$ *such that* $\|g_n^\pm - g^\pm\|_\infty \to 0$ *as* $n \to \infty$, $G_n := S(g_n^-, g_n^+)$. *Then, for any Borel set* B *with* ∂B *of zero Lebesgue measure,*

$$\lim_{n\to\infty} \mathbb{P}(\widetilde{X}_n \in G_n,\ \widetilde{X}_n(1) \in B) = \mathbb{P}(X \in G, X(1) \in B). \tag{15}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

It immediately follows from this corollary that, for any piecewise continuous $f\colon [g^-(1), g^+(1)] \to \mathbb{R}$, $\lim_{n\to\infty} \mathbb{E} f(\widetilde{X}_n(1))\mathbf{1}\{\widetilde{X}_n \in G_n\} = \mathbb{E} f(X(1))\mathbf{1}\{X \in G\}$. This relation can be used, for instance, for computing risk-neutral prices of barrier options [8].

The above results establish the validity of the proposed approximation scheme. However, numerical studies strongly suggest that, under suitable conditions (including $g^\pm \in C^2$), the convergence rate in (15) is $O(n^{-2})$ (cf. the classical much slower convergence rate $O(n^{-1/2})$ in the boundary crossing problem in the invariance principle from [43, 44], establishing which required a very technical lengthy argument). Unsurprisingly, proving such a sharp result turned out to be a very challenging task that requires making quite a few difficult steps and is still a work in progress even in the simplest standard Brownian motion process case. To provide some insight into why such a high convergence rate holds true for our scheme for diffusions (2), we now give a sketch of a possible argument leading to the desired bound. To simplify the argument, we assume that closed-form expressions for the 'taboo transition densities'

$$\phi_{n,k}(x, y) := \partial_y \mathbb{P}\big(g^-(t) < X(t) < g^+(t),\ t \in [t_{n,k-1}, t_{n,k}];\ X_{t_{n,k}} \le y \,|\, X_{t_{n,k-1}} = x\big)$$

are available (note that, say, $\phi_{n,k}(x, y) = 0$ if $x \notin (g^-(t_{n,k-1}), g^+(t_{n,k-1}))$). Therefore, we can and do skip approximation of these densities.

Recalling the definition (6) of the grids $E_{n,k}$, denote the 'taboo transition semigroup' for the discrete-time 'skeleton' of our process X and its discrete approximation by

$$(T_{n,k} f)(x) := \int_{g^-(t_{n,k})}^{g^+(t_{n,k})} \phi_{n,k}(x, y) f(y)\,\mathrm{d}y, \qquad (S_{n,k} f)(x) := \sum_{y \in E_{n,k}} \phi_{n,k}(x, y) f(y) h_{n,k},$$

respectively.

Conjecture 1. *Assume that* $g^\pm \in C^2([0, 1])$*. Then, for any bounded continuous function* $f\colon \mathbb{R} \to \mathbb{R}$ *and any* $x \in (g^-(0), g^+(0))$*, as* $n \to \infty$*,*

$$(S_{n,1} \cdots S_{n,n} - T_{n,1} \cdots T_{n,n}) f(x) = O(n^{-2}).$$

$f \equiv 1$ *yields the desired convergence rate for the boundary crossing probabilities.*

A possible approach to proving the above conjecture can be outlined as follows. First, without loss of generality, we can assume that $x = 0$ in this argument. Second, using the bound from [8], we replace the original boundaries $g^\pm$ with their polygonal approximations $\widehat{g}^\pm$ with nodes at the points $(t_{n,k}, g^\pm(t_{n,k}))$, $k = 0, 1, \dots, n$, which introduces an error of the order $O(n^{-2})$. Next, using the method of compositions approach [56], we get

$$S_{n,1} \cdots S_{n,n} f - T_{n,1} \cdots T_{n,n} f = \sum_{k=1}^{n} f_{n,k} + \varepsilon_n, \tag{16}$$

where

$$f_{n,k} := T_{n,1} \cdots T_{n,k-1} (S_{n,k} - T_{n,k}) T_{n,k+1} \cdots T_{n,n} f,$$

$$\varepsilon_n := \sum_{k=2}^{n} (S_{n,1} \cdots S_{n,k-1} - T_{n,1} \cdots T_{n,k-1})(S_{n,k} - T_{n,k}) T_{n,k+1} \cdots T_{n,n} f.$$

As ε_n is a sum of terms involving compositions of the 'small operator differences' $S_{n,1} \cdots S_{n,k-1} - T_{n,1} \cdots T_{n,k-1}$ and $S_{n,k} - T_{n,k}$, we can expect that its order of magnitude is

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

higher than that of $f_{n,k}$. Unfortunately, a formal proof of that claim appears to be a rather difficult technical task. For the rest of this semi-formal argument, we will ignore the term ε_n.

Now set $\phi_{n,k;z}(y) := \phi_{n,k}(y, z)$, and let $u_{n,k}(z) := (T_{n,1} \cdots T_{n,k-1}\phi_{n,k;z})(0)$, $v_{n,k}(z) := (T_{n,k+1} \cdots T_{n,n} f)(z)$. Using the Euler–Maclaurin summation formula [46, Section 8.1], we get, for $1 \le k \le n-1$,

$$((S_{n,k} - T_{n,k})v_{n,k})(x) \approx h_{n,k}^4 \frac{B_4}{4!} \partial_{zzz}(\phi_{n,k}(x, z)v_{n,k}(z))\Big|_{z=g^-(t_{n,k})}^{z=g^+(t_{n,k})},$$

where B_j denotes the jth Bernoulli number, and we used the boundary conditions

$$u_{n,k}(g^\pm(t_{n,k})) = v_{n,k}(g^\pm(t_{n,k})) = 0, \qquad k \le n-1, \tag{17}$$

assumed the existence of the limits of the third derivatives in the above formula, and retained the first term in the asymptotic expansion.

By the linearity of the operators $S_{n,k}$ and $T_{n,k}$, changing the order of integration and differentiation, we have, for $k \le n-1$,

$$(T_{n,1} \cdots T_{n,k-1}(S_{n,k} - T_{n,k})v_{n,k})(0) \approx h_{n,k}^4 \frac{B_4}{4!}(u_{n,k}(z)v_{n,k}(z))'''\Big|_{z=g^-(t_{n,k})}^{z=g^+(t_{n,k})}. \tag{18}$$

Using the product rule and the boundary conditions (17) yields, for $1 \le k \le n-1$,

$$(u_{n,k}(z)v_{n,k}(z))'''\Big|_{z=g^\pm(t_{n,k})} = 3u''_{n,k}(g^\pm(t_{n,k}))v'_{n,k}(g^\pm(t_{n,k})) + 3u'_{n,k}(g^\pm(t_{n,k}))v''_{n,k}(g^\pm(t_{n,k})). \tag{19}$$

From classical diffusion theory, the functions $u_{n,k}(z)$ and $v_{n,k}(z)$ coincide with $u(t_{n,k}, z)$ and $v(t_{n,k}, z)$, respectively, where $u(t,z)$ and $v(t,z)$ are solutions of the following boundary value problems for Kolmogorov equations in the domain $D := \{(t, z) \colon t \in (0, 1), z \in (\widehat{g}^-(t), \widehat{g}^+(t))\}$:

$$(\partial_t + L)v = 0, \qquad v(t, \widehat{g}^\pm(t)) = 0, \qquad v(1, z) = f(z), \tag{20}$$

$$(\partial_t - L^*)u = 0, \qquad u(t, \widehat{g}^\pm(t)) = 0, \qquad u(0, z) = \delta_0(z), \tag{21}$$

where

$$(Lw)(t, z) := \mu(t, z)\partial_z w(t, z) + \tfrac{1}{2}\partial_{zz}w(t, z),$$
$$(L^*w)(t, z) := -\partial_z(\mu(t, z)w(t, z)) + \tfrac{1}{2}\partial_{zz}w(t, z).$$

Using the boundary conditions on u and v on $\widehat{g} = \widehat{g}^\pm$ and the chain rule, we can show that, for both $\Psi(t) := v(t, \widehat{g}(t))$ and $\Psi^*(t) := u(t, \widehat{g}(t))$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Psi(t+) = \partial_t v(t, \widehat{g}(t)) + \widehat{g}'(t+)\partial_z v(t, \widehat{g}(t)),$$
$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Psi^*(t-) = \partial_t u(t, \widehat{g}(t)) + \widehat{g}'(t-)\partial_z u(t, \widehat{g}(t)).$$

Here and in what follows, we use notation conventions of the form $\partial_t v(t, \widehat{g}^+(t)) = \lim_{z\uparrow\widehat{g}^+(t)} \partial_t v(t, z)$, etc. With this in mind, we now get, from (20) and (21), the relations

$$-(Lv)(t, \widehat{g}(t)) + \widehat{g}'(t+)\partial_z v(t, \widehat{g}(t)) = 0, \tag{22}$$

$$(L^*u)(t, \widehat{g}(t)) + \widehat{g}'(t-)\partial_z u(t, \widehat{g}(t)) = 0. \tag{23}$$

Rearranging (23), using the product rule and the boundary condition $u(t, \widehat{g}(t)) = 0$,

$$\begin{aligned}\tfrac{1}{2}\partial_{zz}u(t, \widehat{g}(t)) &= \partial_z(\mu u)(t, \widehat{g}(t)) - \widehat{g}'(t-)\partial_z u(t, \widehat{g}(t)) \\ &= ((\partial_z\mu)u + \mu\partial_z u - \widehat{g}'(t-)\partial_z u)(t, \widehat{g}(t)) \\ &= (\mu(t, \widehat{g}(t)) - \widehat{g}'(t-))\partial_z u(t, \widehat{g}(t)).\end{aligned} \tag{24}$$

Similarly, using (22),

$$\tfrac{1}{2}\partial_{zz}v(t, \widehat{g}(t)) = -(\mu(t, \widehat{g}(t)) - \widehat{g}'(t+))\partial_z v(t, \widehat{g}(t)). \tag{25}$$

Substituting (24) and (25) into (19), since g is twice continuously differentiable,

$$\begin{aligned}(u_{n,k}(z)v_{n,k}(z))'''\big|_{z=g(t_{n,k})} &= 6\big[(\mu(t_{n,k}, z) - \widehat{g}'(t-))u'_{n,k}(z)v'_{n,k}(z) \\ &\quad - (\mu(t_{n,k}, z) - \widehat{g}'(t_{n,k}+))u'_{n,k}(z)v'_{n,k}(z)\big]_{z=g(t_{n,k})} \\ &= 6(\widehat{g}'(t_{n,k}+) - \widehat{g}'(t_{n,k}-))u'_{n,k}(g(t_{n,k}))v'_{n,k}(g(t_{n,k})) \\ &= 6g''(\theta_{n,k})u'_{n,k}(g(t_{n,k}))v'_{n,k}(g(t_{n,k}))\Delta_n\end{aligned}$$

for some $\theta_{n,k} \in [t_{n,k-1}, t_{n,k+1}]$. Substituting the above expression into (18) and letting, for convenience of summation, $g_1(t) := g^-(t)$ and $g_2(t) := g^+(t)$, we have, for $k \le n-1$,

$$f_{n,k} \approx \frac{B_4}{4}h_{n,k}^4 \sum_{i=1}^{2}(-1)^i g_i''(\theta_{n,k})u'_{n,k}(g_i(t_{n,k}))v'_{n,k}(g_i(t_{n,k}))\Delta_n,$$

where we can replace $h_{n,k}$ on the right-hand side with $h_{n,1}$ in view of (5) and (7). Substituting the resulting expression for $f_{n,k}$ into (16), and assuming that the emerging Riemann sums converge as $n \to \infty$ and that, when $t_{n,k} \to t$,

$$u'_{n,k}(g_i(t_{n,k})) = \partial_z\widetilde{u}(t_{n,k}, g_i(t_{n,k})) \to \partial_z\widetilde{u}(t, g_i(t)),$$

where $\widetilde{u}$ is the solution to the boundary value problem (21) with $\widehat{g}^{\pm}$ replaced with $g^{\pm}$, and that similar convergence holds for $v'_{n,k}$ and $\widetilde{v}$ (the latter solving (20) with $g^{\pm}$ instead of $\widehat{g}^{\pm}$), we have

$$\sum_{k=1}^{n-1} f_{n,k} \approx \frac{B_4}{4}h_{n,1}^4 \sum_{i=1}^{2}(-1)^i \int_0^1 g_i''(t)\partial_z\widetilde{u}(t, g_i(t))\partial_z\widetilde{v}(t, g_i(t))\,\mathrm{d}t.$$

We can show that the integrals on the right-hand side are finite using the observations that $\partial_z\widetilde{v}(t, g(t)) = O((1-t)^{-1/2})$ as $t \to 1$ and that $\partial_z\widetilde{u}(t, g(t))$ is uniformly bounded. Our numerical computations indicate that the above conjecture on the behaviour of the Riemann sums is correct.

For $k = n$, since $u_{n,n}(g(1)) = 0$, using the Euler–Maclaurin summation formula we have

$$T_{n,1}\cdots T_{n,n-1}(S_{n,n} - T_{n,n})f \approx \frac{B_2}{2!}h_{n,n}^2(u_{n,n}(z)f(z))'\big|_{z=g_1(1)}^{z=g_2(1)} = \frac{h_{n,n}^2}{12}u'_{n,n}(z)f(z)\big|_{z=g_1(1)}^{z=g_2(1)}.$$

It now follows from (16) that $(S_{n,1}\cdots S_{n,n} - T_{n,1}\cdots T_{n,n})f \approx C_1h_{n,1}^4 + C_2h_{n,n}^2$, with the constants

$$C_1 := \frac{-1}{120}\sum_{i=1}^{2}(-1)^i\int_0^1 g_i''(t)\partial_z\widetilde{u}(t, g_i(t))\partial_z\widetilde{v}(t, g_i(t))\,\mathrm{d}t, \qquad C_2 := \frac{1}{12}u'_{n,n}(z)f(z)\big|_{z=g_1(1)}^{z=g_2(1)}.$$

Since $h_{n,1} = o(n^{-1/2})$ and $h_{n,n} = O(n^{-1})$, we have obtained the claimed convergence rate.

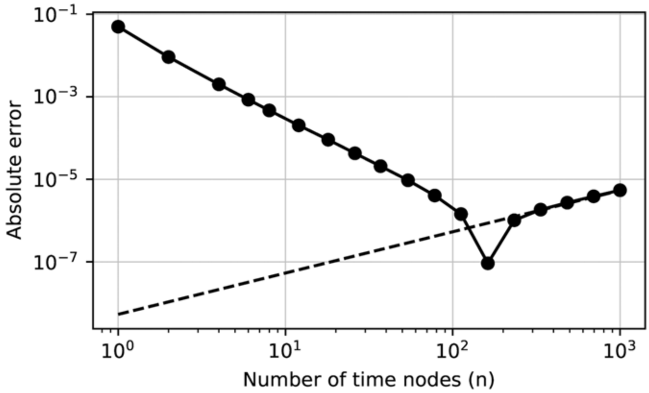

FIGURE 1. On this log–log plot, the bullets • show the absolute errors between Markov chain approximations with $\gamma = 1$, $\delta = 0$ and normalising factors $C_{n,k}(x)$ replaced with 1, and the true boundary crossing probability of the standard Brownian motion process of boundaries (28). The dashed line corresponds to the straight line $2n\mathrm{e}^{-2\pi^2\gamma^2}$.

Remark 7. To accelerate the numerical evaluation of $\mathbb{P}(\widetilde{X}_n \in G_n)$, we can ignore the normalising factors $C_{n,k}(x)$ as they are very close to 1. Indeed, let $\widehat{\mathbf{T}}_{n,k}$ be $|E^G_{n,k-1}| \times |E^G_{n,k}|$ matrices with entries

$$q_{n,k}(x, y) := (1 - \pi_{n,k}(f_n^-, f_n^+ \mid x, y))\varphi(y \mid x + \mu_{n,k}(x), \sigma^2_{n,k}(x))h_{n,k} \tag{26}$$

for $(x, y) \in E^G_{n,k-1} \times E^G_{n,k}$. These matrices differ from the $\mathbf{T}_{n,k}$ from (12) in that they do not involve the factors $C_{n,k}(x)$. For M and c_0 defined in Lemma 2, we show below that

$$\left|\mathbb{P}(\widetilde{X}_n \in G_n) - \widehat{\mathbf{T}}_{n,1}\widehat{\mathbf{T}}_{n,2}\cdots\widehat{\mathbf{T}}_{n,n}\mathbf{1}^\top\right| \le c_0 n \rho_n^{n-1} \exp\{-Mn^{2\delta}\}, \tag{27}$$

where $\rho_n := 1 \vee \max_{1\le j\le n} \sup_{|x|\le r} C_{n,j}(x)$. Note that some care must be exercised when choosing γ and δ. Using small values of γ and δ will result in the error from replacing the normalising factors with 1 becoming non-negligible. It will be seen from the derivation of (3.3) that when $\delta = 0$ the leading term of the error of the left-hand side is equal to $2n\mathrm{e}^{-2\pi^2\gamma^2}$. Figure 1 illustrates this observation numerically in the case of the standard Brownian motion process and boundaries

$$g^\pm(t) = \pm\frac{t}{3}\cosh^{-1}(2\mathrm{e}^{9/(2t)}), \qquad t \ge 0. \tag{28}$$

4. Proofs

The proof of Theorem 1 is based on several auxiliary results. For the reader's convenience, we will first state them as separate lemmata. Recall the quantity K_r that appeared in Condition 3, and also β and $\alpha^{1/2}$ from (3) and (4).

Lemma 1. *For* $r > |x_0|$, $\max_{1\le k\le n} \sup_{|x|\le r} \big(|\beta_{n,k}(x)| \vee \alpha_{n,k}(x)\big) \le K'_r$, *where* $K'_r := \big(K_r + \frac{1}{2}K_r\big(\frac{3}{2} + K_r\big)\big) \vee \big(1 + \frac{1}{2}K_r\big)^2$.

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

The result is directly obtained by substituting the upper bound on the partial derivatives of μ_X from Condition 3 into the definitions of β and α in (3) and (4), respectively.

Lemma 2. *For any $r > |x_0|$ and $n > 1 + 2K_r$, for the normalising factors $C_{n,k}(x)$ from (9) we have* $\max_{1\le k\le n} \sup_{x\in E_{n,k}(r)} |C_{n,k}(x) - 1| \le c_0 \mathrm{e}^{-Mn^{2\delta}}$, *where*

$$E_{n,k}(r) := E_{n,k-1} \cap [-r, r], \qquad r > 0, \tag{29}$$

δ is the quantity appearing in (5), $M = M(\gamma) := \gamma^2\pi^2/4$, and $c_0 := 2/(1 - \mathrm{e}^{-M})$.

Proof of Lemma 2. For brevity, let $\mu := \mu_{n,k}(x)$ and $\sigma^2 := \sigma^2_{n,k}(x)$. Set

$$\widehat{\varphi}_\sigma(s) := \int_{-\infty}^{+\infty} \varphi(x \,|\, 0, \sigma^2)\mathrm{e}^{-2\pi isx}\,\mathrm{d}x = \mathrm{e}^{-2\pi^2 s^2\sigma^2}, \qquad s \in \mathbb{R}. \tag{30}$$

Since both $\varphi(\,\cdot\,|\,0, \sigma^2)$ and $\widehat{\varphi}_\sigma(\,\cdot\,)$ decay at infinity faster than any power function, we can apply the Poisson summation formula [59, p. 252] to obtain

$$\begin{aligned} C_{n,k}(x) &= \sum_{j\in\mathbb{Z}} \varphi(jh_{n,k} + g^+(t_{n,k}) \,|\, x + \mu, \sigma^2)h_{n,k} \\ &= 1 + \sum_{\ell\in\mathbb{Z}\setminus\{0\}} \widehat{\varphi}_\sigma(\ell/h_{n,k})\mathrm{e}^{-2\pi i(x+\mu-g^+(t_{n,k}))\ell/h_{n,k}} \\ &= 1 + 2\sum_{\ell=1}^{\infty} \widehat{\varphi}_\sigma(\ell/h_{n,k}) \cos(2\pi(x + \mu - g^+(t_{n,k}))\ell/h_{n,k}), \qquad x \in E_{n,k-1}, \end{aligned}$$

since $\widehat{\varphi}_\sigma(0) = 1$. It follows that $|C_{n,k}(x) - 1| \le 2\sum_{\ell=1}^{\infty} |\widehat{\varphi}_\sigma(\ell/h_{n,k})|$. Using (30) and the elementary inequality $\mathrm{e}^{-a\ell^2} \le \mathrm{e}^{-a\ell}$, $\ell \ge 1$, $a > 0$, we obtain

$$|C_{n,k}(x) - 1| \le 2\sum_{\ell=1}^{\infty} \mathrm{e}^{-2\pi^2\ell\sigma^2/h^2_{n,k}} = \frac{2\mathrm{e}^{-2\pi^2\sigma^2/h^2_{n,k}}}{1 - \mathrm{e}^{-2\pi^2\sigma^2/h^2_{n,k}}}. \tag{31}$$

By Condition 3, for $n > 2K_r$ we have

$$\begin{aligned} \inf_{|x|\le r} \alpha_{n,k}(x) &= \inf_{|x|\le r} \left(1 + \tfrac{1}{2}\Delta_n \partial_x \mu_X(t_{n,k-1}, x)\right)^2 \\ &\ge 1 - \Delta_n \sup_{|x|\le r} |\partial_x \mu_X(t_{n,k-1}, x)| \ge 1 - K_r/n \ge \tfrac{1}{2}. \end{aligned} \tag{32}$$

Substituting $\sigma^2 = \sigma^2_{n,k}(x) = \alpha_{n,k}(x)\Delta_n \ge \Delta_n/2$ into (31) and using the inequality

$$\min_{1\le k\le n} \frac{\Delta_n}{h^2_{n,k}} = \frac{1}{w^2_{n,n}\Delta_n} \wedge \min_{1\le k\le n-1} \frac{1}{w^2_{n,k}\Delta_n^{2\delta}} \ge \frac{\gamma^2}{4}(n \wedge n^{2\delta}) = \frac{\gamma^2 n^{2\delta}}{4}, \tag{33}$$

which holds since $\max_{1\le k\le n} |w_{n,k}| \le 2\gamma^{-1}$, $\delta < \frac{1}{2}$, we obtain

$$\max_{1\le k\le n} \sup_{|x|\le r} |C_{n,k}(x) - 1| \le \frac{2\mathrm{e}^{-Mn^{2\delta}}}{1 - \mathrm{e}^{-Mn^{2\delta}}} \le c_0 \mathrm{e}^{-Mn^{2\delta}}.$$

□

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

Lemma 3. *Let Z be a standard normal random variable independent of $\{\xi_{n,k}\}_{k=1}^{n}$. Set*

$$Z_{n,k}(x) := \mu_{n,k}(x) + \sigma_{n,k}(x)Z, \qquad k = 1, \dots, n, \ x \in \mathbb{R}, \tag{34}$$

and let $\Delta\xi_{n,k} := \xi_{n,k} - \xi_{n,k-1}$. Denote by $\mathbb{E}_{k,x}$ the conditional expectation given $\xi_{n,k-1} = x$, and set, for $r > |x_0|$ and sub-grids $E_{n,k}(r)$ from (29),

$$e_{n,k}(m, r) := \sup_{x\in E_{n,k}(r)} |\mathbb{E}_{k,x}(\Delta\xi_{n,k})^m - \mathbb{E}Z_{n,k}^m(x)|, \qquad m = 1, 2, \dots \tag{35}$$

Then $\max_{1\le k\le n} e_{n,k}(m, r) \le c_m \mathrm{e}^{-Mn^{2\delta}}$ for $n > (1 + 2K_r) \vee M^{-1/(2\delta)}$, where $c_m \in (0, \infty)$ is a constant whose explicit value is given in (40).

Proof of Lemma 3. For brevity, we will often supress dependence on m, n, k, and x. This should cause no confusion. Set $C = C_{n,k}(x)$, $\lambda := \mathbb{E}_{k,x}(\Delta\xi_{n,k})^m$, $\widetilde{\lambda} := \mathbb{E}Z_{n,k}^m(x)$, for $m = 1, 2, \dots$ Using the triangle inequality,

$$|\lambda - \widetilde{\lambda}| = \left|\frac{1-C}{C}\widetilde{\lambda} + \frac{1}{C}(C\lambda - \widetilde{\lambda})\right| \le \left|\frac{C-1}{C}\right||\widetilde{\lambda}| + \frac{1}{C}|C\lambda - \widetilde{\lambda}|. \tag{36}$$

The term $C\lambda = C\mathbb{E}_{k,x}(\Delta\xi_{n,k})^m$ can be viewed as a trapezoidal approximation of $\widetilde{\lambda} = \mathbb{E}Z_{n,k}^m(x)$, so after rewriting $\widetilde{\lambda}$ as an integral, we can express $C\lambda - \widetilde{\lambda}$ as the quadrature error

$$\begin{aligned}\varepsilon_k(x) := C\lambda - \widetilde{\lambda} &= \sum_{j\in\mathbb{Z}} (jh_{n,k} + g^+(t_{n,k}) - x)^m \varphi(jh_{n,k} + g^+(t_{n,k}) - x \mid \mu, \sigma^2) h_{n,k} \\ &\quad - \int_{-\infty}^{+\infty} (u + g^+(t_{n,k}) - x)^m \varphi(u + g^+(t_{n,k}) - x \mid \mu, \sigma^2)\, \mathrm{d}u,\end{aligned}$$

where $\mu = \mu_{n,k}(x)$ and $\sigma^2 = \sigma_{n,k}^2(x)$. We further note that, due to (4), the condition $n \ge 2K_r$ ensures that $\sigma^2 > 0$. For $s \in \mathbb{R}$, set

$$\begin{aligned}\widehat{p}(s) := \int_{-\infty}^{+\infty} (v + \mu)^m \varphi(v \mid 0, \sigma^2) \mathrm{e}^{-2\pi isv}\, \mathrm{d}v &= \sum_{l=0}^{m} \binom{m}{l} \mu^{m-l} \int_{-\infty}^{+\infty} v^l \varphi(v \mid 0, \sigma^2) \mathrm{e}^{-2\pi isv}\, \mathrm{d}v \\ &= \sum_{l=0}^{m} \binom{m}{l} \mu^{m-l} (-i\sigma)^l H_l(2\pi s\sigma) \mathrm{e}^{-2\pi^2 s^2\sigma^2},\end{aligned} \tag{37}$$

where $H_l(x) := (-1)^l \mathrm{e}^{x^2/2}(\mathrm{d}^l/\mathrm{d}x^l)\mathrm{e}^{-x^2/2}$, $x \in \mathbb{R}$, $l \ge 1$, is the lth Chebyshev–Hermite polynomial. Since both $(\,\cdot\, - x)^m \varphi(\,\cdot\, - x \mid \mu, \sigma^2)$ and $\widehat{p}(\,\cdot\,)$ decay at infinity faster than any power function, using the Poisson summation formula [59, p. 252] and the change of variables $v = uh_{n,k} + g^+(t_{n,k}) - x$ in (37), we obtain

$$\begin{aligned}\varepsilon_k(x) &= \sum_{\ell\in\mathbb{Z}\setminus\{0\}} \widehat{p}(\ell/h_{n,k}) \exp\{-2\pi i(-g^+(t_{n,k}) + x + \mu)\ell/h_{n,k}\} \\ &= 2\sum_{\ell=1}^{\infty} \mathrm{Re}\big[\widehat{p}(\ell/h_{n,k}) \exp\{-2\pi i(-g^+(t_{n,k}) + x + \mu)\ell/h_{n,k}\}\big].\end{aligned}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

Since $|\mathrm{Re}z| \le |z|$, $z \in \mathbb{C}$, and $|\mathrm{e}^{is}| \le 1$, $s \in \mathbb{R}$, we obtain $|\varepsilon_k(x)| \le 2\sum_{\ell=1}^{\infty} |\widehat{p}(\ell/h_{n,k})|$. Note that $|H_l(u)| \le C_l'(|u|^{\ell} + 1)$, $u \in \mathbb{R}$, $l = 1, 2, \ldots$, where C_l' is a constant that depends on l only and we can assume without loss of generality that $\{C_l'\}_{l\ge1}$ is a non-decreasing sequence. Therefore, we have $|(-i\sigma)^l H_l(2\pi s\sigma)| \le C_l(s^l\sigma^{2l} + \sigma^l)$, $s \ge 1$, where $C_l := (2\pi)^l C_l'$. Using $\mu^{m-l}(\ell\sigma^2/h_{n,k})^l = \beta_{n,k}^{m-l}(x)\alpha_{n,k}^l(x)\ell^m {\Delta_n}^m/(h_{n,k}^l \ell^{m-l})$, and the fact that $h_{n,k}^{-l}\ell^{-(m-l)} \le h_{n,k}^{-m}$ for $l \le m$ as $h_{n,k} < 1$, we obtain from (37) that

$$
\begin{aligned}
|\varepsilon_k(x)| &\le 2\sum_{\ell=1}^{\infty}\sum_{l=0}^{m}\binom{m}{l} C_l|\mu|^{m-l}\left(\left(\frac{\ell\sigma^2}{h_{n,k}}\right)^l + \sigma^l\right)\mathrm{e}^{-2\pi^2\sigma^2\ell^2/h_{n,k}^2} \\
&\le 2\sum_{\ell=1}^{\infty}\sum_{l=0}^{m}\binom{m}{l} C_l\left(\left(\frac{\Delta_n}{h_{n,k}}\right)^m |\beta_{n,k}^{m-l}(x)|\alpha_{n,k}^l(x)\ell^m \right.\\
&\quad \left. + |\beta_{n,k}^{m-l}(x)|\alpha_{n,k}^{l/2}(x)\Delta_n^{m-l/2}\right)\mathrm{e}^{-2\pi^2\sigma^2\ell^2/h_{n,k}^2}.
\end{aligned}
$$

Since C_l is non-decreasing in l,

$$
\sup_{x\in E_{n,k}(r)} \sum_{l=0}^{m}\binom{m}{l} C_l\left(\frac{\Delta_n}{h_{n,k}}\right)^m |\beta_{n,k}^{m-l}(x)|\alpha_{n,k}^l(x) \le 2^m C_m(2K_r')^m =: L_{m,r},
$$

where we used Lemma 1 and the inequality $\max_{1\le k\le n}\Delta_n/h_{n,k} \le 2$, which follows from $\min_{1\le k\le n} h_{n,k} \ge \min\{(1/n)^{1/2+\delta}, 1/n\}/2 = 1/2n$. Again using Lemma 1 and the trivial bound $\Delta_n^{m-l/2} \le 1$, $0 \le l \le m$, we have

$$
\sup_{x\in E_{n,k}(r)} \sum_{l=0}^{m}\binom{m}{l} C_l|\beta_{n,k}^{m-l}(x)|\alpha_{n,k}^{l/2}(x)(\Delta_n)^{m-l/2} \le 2^m C_m(K_r')^m \le L_{m,r}.
$$

Hence,

$$
\sup_{x\in E_{n,k}(r)} |\varepsilon_k(x)| \le 2L_{m,r}\sum_{\ell=1}^{\infty}(\ell^m+1)\mathrm{e}^{-2\pi^2\sigma^2\ell^2/h_{n,k}^2} \le 4L_{m,r}\sum_{\ell=1}^{\infty}\ell^m\mathrm{e}^{-\ell Mn^{2\delta}},
$$

where we used (32), (33) and the bound $\mathrm{e}^{-a\ell^2} \le \mathrm{e}^{-a\ell}$, $\ell \ge 1$, $a > 0$, in the second inequality. Note that $\sum_{\ell=1}^{\infty}\ell^m z^\ell \le a_m z$, $z \in [0, \mathrm{e}^{-1}]$, where $a_m := \sum_{\ell=1}^{\infty}\ell^m\mathrm{e}^{-\ell+1} < \infty$. As $Mn^{2\delta} > 1$, we obtain from here that

$$
\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)} |\varepsilon_k(x)| \le 4a_m L_{m,r}\mathrm{e}^{-Mn^{2\delta}}. \tag{38}
$$

From Lemma 2,

$$
\sup_{x\in E_{n,k}(r)}\left|\frac{C_{n,k}(x)-1}{C_{n,k}(x)}\right| \le \frac{c_0\mathrm{e}^{-Mn^{2\delta}}}{1-c_0}. \tag{39}
$$

Using inequalities (38) and (39) in (36), we get

$$
\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)} |\mathrm{e}_{n,k}(m, r)| \le c_m\mathrm{e}^{-Mn^{2\delta}},
$$

where

$$
c_m := \frac{c_0}{1-c_0}\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}|\mathbb{E}Z_{n,k}^m(x)| + \frac{4a_m L_{m,r}}{1-c_0}. \tag{40}
$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

The boundedness of $\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}|\mathbb{E}Z^m_{n,k}(x)|$ can be proved by applying the inequality $|x+y|^m\le 2^{m-1}(|x|^m+|y|^m)$, $x, y\in\mathbb{R}$, $m\ge 1$, to obtain

$$\sup_{x\in E_{n,k}(r)}|\mathbb{E}Z^m_{n,k}(x)|\le 2^{m-1}\sup_{|x|\le r}(|\beta_{n,k}(x)\Delta_n|^m+|\alpha_{n,k}(x)\Delta_n|^{m/2}\mathbb{E}|Z|^m)$$
$$\le 2^{m-1}\left((K'_r)^m+(2K'_r)^{m/2}\pi^{-1/2}\Gamma\left(\frac{m+1}{2}\right)\right)<\infty,$$

where Γ is the gamma function. Lemma 3 is proved. □

To prove the convergence stated in Theorem 1 we will use the martingale characterisation method, verifying the sufficient conditions for convergence from [18, Theorem 4.1 in Chapter 7] (to be referred to as the EK theorem in what follows). For $x\in E_{n,k-1}$, let

$$b_{n,k}(x):=\frac{1}{\Delta_n}\mathbb{E}[\Delta\xi_{n,k}\mid \xi_{n,k-1}=x],\qquad a_{n,k}(x):=\frac{1}{\Delta_n}\mathrm{var}[\Delta\xi_{n,k}\mid \xi_{n,k-1}=x].$$

Using the standard semimartingale decomposition of X_n (see (14)), we set

$$B_n(t):=\sum_{k=1}^{\nu_n(t)}b_{n,k}(\xi_{n,k-1})\Delta_n,\qquad A_n(t):=\sum_{k=1}^{\nu_n(t)}a_{n,k}(\xi_{n,k-1})\Delta_n,$$

and let $M_n:=X_n-B_n$. With respect to the natural filtration

$$\mathbf{F}^n:=\{\sigma(X_n(s), B_n(s), A_n(s)\colon s\le t)\colon t\ge 0\}=\{\sigma(\xi_{n,k}\colon k\le\nu_n(t), t\ge 0)\},$$

our B_n, A_n, and M_n are the predictable drift, angle bracket, and martingale component, respectively, of the process X_n.

The EK theorem is stated for time-homogeneous processes. To use it in our case, we consider the vector-valued processes $\boldsymbol{X}_n(t):=(t, X_n(t))$ and, for a fixed $r>0$, let τ^r_n be localising $\mathbf{F}^n$-stopping times, $\tau^r_n:=\inf\{t\colon \|\boldsymbol{X}_n(t)\|\vee\|\boldsymbol{X}_n(t-)\|\ge r\}$, $\|\boldsymbol{u}\|=|u_1|\vee|u_2|$ being the maximum norm of $\boldsymbol{u}=(u_1, u_2)$.

Lemma 4. *For each fixed $r>|x_0|$, as $n\to\infty$,*

$$\sup_{t\le 1\wedge\tau^r_n}\left|B_n(t)-\int_0^t\mu_X(s, X_n(s))\,\mathrm{d}s\right|\xrightarrow{\text{a.s.}}0,\qquad \sup_{t\le 1\wedge\tau^r_n}|A_n(t)-t|\xrightarrow{\text{a.s.}}0.$$

Proof of Lemma 4. Since on each of the time intervals $[t_{n,k-1}, t_{n,k})$, $k=1,\dots,n$, the process X_n is equal to $\xi_{n,k-1}$, we have the following decomposition:

$$\begin{aligned}
&B_n(t)-\int_0^t\mu_X(s, X_n(s))\,\mathrm{d}s\\
&=\sum_{k=1}^{\nu_n(t)}\left[b_{n,k}(\xi_{n,k-1})\Delta_n-\int_{t_{n,k-1}}^{t_{n,k}}\mu_X(s, X_n(s))\,\mathrm{d}s\right]-\int_{t_{n,\nu_n(t)}}^t\mu_X(s, X_n(s))\,\mathrm{d}s\\
&=\sum_{k=1}^{\nu_n(t)}(bn, k(\xi_{n,k-1})-\beta_{n,k}(\xi_{n,k-1}))\Delta_n+\sum_{k=1}^{\nu_n(t)}(\beta_{n,k}(\xi_{n,k-1})-\mu_X(t_{n,k-1}, \xi_{n,k-1}))\Delta_n\\
&\quad+\sum_{k=1}^{\nu_n(t)}\int_{t_{n,k-1}}^{t_{n,k}}[\mu_X(t_{n,k-1}, X_n(s))-\mu_X(s, X_n(s))]\,\mathrm{d}s-\int_{t_{n,\nu_n(t)}}^t\mu_X(s, \xi_{n,\nu_n(t)})\,\mathrm{d}s.
\end{aligned}\tag{41}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

Due to the stopping-time localisation, the first term on the right-hand side of (41) has the following upper bound (see (8), (34), and (35)):

$$\begin{aligned}\sup_{t\le 1\wedge\tau_n^r}&\left|\sum_{k=1}^{\nu_n(t)}(b_{n,k}(\xi_{n,k-1})-\beta_{n,k}(\xi_{n,k-1}))\Delta_n\right|\\ &=\sup_{t\le 1\wedge\tau_n^r}\left|\sum_{k=1}^{\nu_n(t)}(\mathbb{E}[\Delta\xi_{n,k}\mid\xi_{n,k-1}]-\mathbb{E}[Z_{n,k}(\xi_{n,k-1})\mid\xi_{n,k-1}])\right|\\ &\le\sum_{k=1}^{n}\sup_{x\in E_{n,k}(r)}|\mathbb{E}_{k,x}\Delta\xi_{n,k}-\mathbb{E}Z_{n,k}(x)|\le\sum_{k=1}^{n}e_{n,k}(1,r).\end{aligned}\tag{42}$$

To bound the second term on the right-hand side of (41), we use the definition of $\beta_{n,k}$ in (3) and Condition 3 to get following inequality:

$$\begin{aligned}\sup_{x\in E_{n,k}(r)}|\beta_{n,k}(x)-\mu_X(t_{n,k-1},x)|&\le\tfrac{1}{2}\Delta_n\sup_{|x|\le r}\left|\left(\partial_t\mu_X+\mu_X\partial_x\mu_X+\tfrac{1}{2}\partial_{xx}\mu_X\right)(t_{n,k-1},x)\right|\\ &\le\tfrac{1}{2}\Delta_nK_r\left(K_r+\tfrac{3}{2}\right).\end{aligned}\tag{43}$$

The second-last term in (41) can be bounded from above by using the bound for $\partial_t\mu_X$ from Condition 3:

$$\sup_{t\le 1\wedge\tau_n^r}\sum_{k=1}^{\nu_n(t)}\left|\int_{t_{n,k-1}}^{t_{n,k}}[\mu_X(t_{n,k-1},X_n(s))-\mu_X(s,X_n(s))]\,\mathrm{d}s\right|\le K_r\Delta_n.\tag{44}$$

Again using Condition 3, the last term in (41) is bounded as follows:

$$\sup_{t\le 1\wedge\tau_n^r}\left|\int_{t_{n,\nu_n(t)}}^{t}\mu_X(s,X_{n,\nu_n(t)})\,\mathrm{d}s\right|\le K_r\Delta_n.\tag{45}$$

Using inequalities (42)–(45) in the decomposition (41), we obtain

$$\sup_{t\le 1\wedge\tau_n^r}\left|B_n(t)-\int_0^t\mu_X(s,X_n(s))\,\mathrm{d}s\right|\le\sum_{k=1}^{n}e_{n,k}(1,r)+K_r\left(\tfrac{1}{2}K_r+\tfrac{11}{4}\right)\Delta_n.$$

Since $\Delta_n\to 0$ as $n\to\infty$, the first half of the lemma follows after applying Lemma 3 with $m=1$.

Similarly,

$$A_n(t)-t=\sum_{k=1}^{\nu_n(t)}(a_{n,k}(\xi_{n,k-1})-\alpha_{n,k}(\xi_{n,k-1}))\Delta_n+\sum_{k=1}^{\nu_n(t)}(\alpha_{n,k}(\xi_{n,k-1})-1)\Delta_n-\int_{t_{n,\nu_n(t)}}^{t}\mathrm{d}s.\tag{46}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

To bound the first term on the right-hand side of (46), we use $\mathrm{Var}[Y \mid X] = \mathbb{E}[Y^2 \mid X] - (\mathbb{E}[Y \mid X])^2$ for a square-integrable random variable Y to obtain

$$\begin{aligned}
&\sup_{t\le 1\wedge\tau_n^r}\left|\sum_{k=1}^{\nu_n(t)}(a_{n,k}(\xi_{n,k-1})-\alpha_{n,k}(\xi_{n,k-1}))\Delta_n\right| \\
&\quad=\sup_{t\le 1\wedge\tau_n^r}\left|\sum_{k=1}^{\nu_n(t)}(\mathrm{Var}[\Delta\xi_{n,k}\mid\xi_{n,k-1}]-\mathrm{Var}[Z_{n,k}(\xi_{n,k-1})\mid\xi_{n,k-1}])\right| \\
&\quad\le \sup_{t\le 1\wedge\tau_n^r}\sum_{k=1}^{\nu_n(t)}\left|\mathbb{E}[(\Delta\xi_{n,k})^2\mid\xi_{n,k-1}]-\mathbb{E}[Z_{n,k}{}^2(\xi_{n,k-1})\mid\xi_{n,k-1}]\right| \\
&\qquad+\sup_{t\le 1\wedge\tau_n^r}\sum_{k=1}^{\nu_n(t)}\left|(\mathbb{E}[\Delta\xi_{n,k}\mid\xi_{n,k-1}])^2-(\mathbb{E}[Z_{n,k}(\xi_{n,k-1})\mid\xi_{n,k-1}])^2\right| \\
&\quad\le \sum_{k=1}^{n}e_{n,k}(2,r)+2\sum_{k=1}^{n}e_{n,k}(1,r)\left(e_{n,k}(1,r)+\sup_{x\in E_{n,k}(r)}|\mathbb{E}Z_{n,k}(x)|\right),
\end{aligned}\tag{47}$$

where we used the elementary bound

$$|x^2-y^2|\le|x-y|(|x-y|+2|y|)\tag{48}$$

in the final inequality. Furthermore,

$$\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}|\mathbb{E}Z_{n,k}(x)|\le\max_{1\le k\le n}\sup_{|x|\le r}|\beta_{n,k}(x)|\Delta_n\le K_r'.$$

Lemma 3 with $m=1$ and $m=2$ implies that the expression in the last line of (47) vanishes as $n\to\infty$. The second term on the right-hand side of (46) is bounded from above using (4), Lemma 1, and Condition 3:

$$\begin{aligned}
\sup_{t\le 1\wedge\tau_n^r}\sum_{k=1}^{\nu_n(t)}|\alpha_{n,k}(\xi_{n,k-1})-1|\Delta_n &\le \max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}\left|\alpha_{n,k}^{1/2}(x)-1\right|\left(\alpha_{n,k}^{1/2}(x)+1\right) \\
&\le \tfrac{1}{2}\Delta_n K_r(\sqrt{K_r'}+1).
\end{aligned}\tag{49}$$

For the last term in (46) we have

$$\sup_{t\le 1\wedge\tau_n^r}\left|\int_{t_{n,\nu_n(t)}}^{t}\mathrm{d}s\right|\le\Delta_n.\tag{50}$$

Applying inequalities (47)–(50) to (46), we complete the proof of Lemma 4.

Lemma 5. *For each fixed $r>|x_0|$,*

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|B_n(t)-B_n(t-)|^2=0,\tag{51}$$

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|A_n(t)-A_n(t-)|=0,\tag{52}$$

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|X_n(t)-X_n(t-)|^2=0.\tag{53}$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

Proof of Lemma 5. Set $\overline{\xi}_n := \max_{1\le k\le n}|\xi_{n,k}|$ and

$$\chi_n^r := \min\{k \le n\colon |\xi_{n,k}| \ge r\}\mathbf{1}\{\overline{\xi}_n \ge r\} + n\mathbf{1}\{\overline{\xi}_n < r\}.$$

The jumps of B_n are given by the conditional means of the increments, so

$$\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|B_n(t)-B_n(t-)|^2 = \mathbb{E}\max_{k\le\chi_n^r}(\mathbb{E}[\Delta\xi_{n,k}\mid\xi_{n,k-1}])^2 \le \mathbb{E}\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}(\mathbb{E}_{k,x}\Delta\xi_{n,k})^2.$$

By the triangle inequality, for $x\in E_{n,k}(r)$,

$$|\mathbb{E}_{k,x}\Delta\xi_{n,k}| \le |\mathbb{E}_{k,x}\Delta\xi_{n,k} - \mathbb{E}Z_{n,k}(x)| + |\mathbb{E}Z_{n,k}(x)| \le e_{n,k}(1, r) + |\beta_{n,k}(x)|\Delta_n.$$

By inequality (13) and Lemmata 1 and 3, we obtain (51). The jumps of A_n are given by the conditional variances of the increments, so

$$\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|A_n(t)-A_n(t-)| = \mathbb{E}\max_{k\le\chi_n^r}\mathrm{Var}[\Delta\xi_{n,k}\mid\xi_{n,k-1}] \le \mathbb{E}\max_{1\le k\le n}\sup_{x\in E_{n,k}(r)}\mathrm{Var}_{k,x}\Delta\xi_{n,k},$$

where $\mathrm{Var}_{k,x}[\,\cdot\,] := \mathrm{Var}[\,\cdot \mid \xi_{n,k-1}=x]$. Using (48), we obtain, for $x\in E_{n,k}(r)$,

$$\begin{aligned}|\mathrm{Var}_{k,x}\Delta\xi_{n,k}| &\le |\mathrm{Var}_{k,x}\Delta\xi_{n,k} - \mathrm{Var}Z_{n,k}(x)| + \mathrm{Var}Z_{n,k}(x)\\ &\le |\mathbb{E}_{k,x}(\Delta\xi_{n,k})^2 - \mathbb{E}Z^2_{n,k}(x)| + |(\mathbb{E}_{k,x}\Delta\xi_{n,k})^2 - (\mathbb{E}Z_{n,k}(x))^2| + \sigma^2_{n,k}(x)\\ &\le e_{n,k}(2, r) + e_{n,k}(1, r)(e_{n,k}(1, r) + 2|\beta_{n,k}(x)|\Delta_n) + \alpha_{n,k}(x)\Delta_n.\end{aligned}$$

By the local boundedness of $\beta_{n,k}$ and $\alpha_{n,k}$, from Lemma 1 we obtain (52). Further,

$$\mathbb{E}\sup_{t\le 1\wedge\tau_n^r}|X_n(t)-X_n(t-)|^2 = \mathbb{E}\max_{k\le\chi_n^r}(\Delta\xi_{n,k})^2.$$

Using Lyapunov's inequality, we obtain

$$\begin{aligned}\mathbb{E}\max_{k\le\chi_n^r}(\Delta\xi_{n,k})^2 \le \Big(\mathbb{E}\max_{k\le\chi_n^r}(\Delta\xi_{n,k})^4\Big)^{1/2} &\le \Bigg(\mathbb{E}\sum_{k\le\chi_n^r}(\Delta\xi_{n,k})^4\Bigg)^{1/2}\\ &\le \Bigg(\sum_{k=1}^n\sup_{x\in E_{n,k}(r)}\mathbb{E}_{k,x}(\Delta\xi_{n,k})^4\Bigg)^{1/2}.\end{aligned}$$

By the triangle inequality, we have

$$\begin{aligned}|\mathbb{E}_{k,x}(\Delta\xi_{n,k})^4| &\le |\mathbb{E}_{k,x}(\Delta\xi_{n,k})^4 - \mathbb{E}Z^4_{n,k}(x)| + \mathbb{E}Z^4_{n,k}(x)\\ &\le e_{n,k}(4, r) + (\beta_{n,k}(x)\Delta_n)^4 + 6(\beta_{n,k}(x)\Delta_n)^2\alpha_{n,k}(x)\Delta_n + 3(\alpha_{n,k}(x)\Delta_n)^2.\end{aligned}$$

Using Lemmata 1 and 3 we obtain (53). Lemma 5 is proved. □

Proof of Theorem 1. We verify the conditions of the EK theorem. Denote by L the generator of the bivariate process $\boldsymbol{X} := \{\boldsymbol{X}(t) = (t, X(t))\}_{t\in[0,1]}$,

$$Lf = \partial_t f + \mu\partial_x f + \tfrac{1}{2}\partial_{xx} f, \qquad f\in C_c^\infty(\mathbb{R}^2),$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

where $C_c^\infty(\mathbb{R}^2)$ is the space of infinitely many times differentiable functions with compact support. The distribution of $\boldsymbol{X}$ is the solution to the martingale problem for L, i.e., for $f \in C_c^\infty(\mathbb{R}^2)$,

$$f(\boldsymbol{X}(t)) - f(\boldsymbol{X}(0)) - \int_0^t Lf(\boldsymbol{X}(s))\,\mathrm{d}s, \qquad t \in [0,\,1],$$

is a martingale. Using Condition 3, by [18, Proposition 3.5 in Chapter 5] the martingale problem for L is well posed since the solution for the stochastic differential equation (2) exists and is unique. Therefore, the first condition in the EK theorem is met.

The martingale characteristics of $\boldsymbol{X}_n = (t,\,X_n(t))$ are given by

$$\boldsymbol{B}_n(t) := (t,\,B_n(t)), \qquad \boldsymbol{A}_n(t) := \begin{pmatrix} 0 & 0 \\ 0 & A_n(t) \end{pmatrix}, \qquad \boldsymbol{M}_n(t) := (0,\,M_n(t)).$$

A simple calculation shows that $\boldsymbol{M}_n$ and $\boldsymbol{M}_n^\top \boldsymbol{M}_n - \boldsymbol{A}_n$ are $\mathbf{F}^n$-martingales. It follows from Lemmata 4 and 5 that conditions (4.3)–(4.7) in the EK theorem are satisfied, which means that all the conditions of that theorem are met. Theorem 1 is proved. □

Proof of Corollary 1. Denote the process whose trajectories are polygons with nodes $(t_{n,k},\,\xi_{n,k})$ by $\widehat{X}_n$. By the triangle inequality,

$$\|\widetilde{X}_n - X_n\|_\infty \le \|\widetilde{X}_n - \widehat{X}_n\|_\infty + \|\widehat{X}_n - X_n\|_\infty,$$

where $\widetilde{X}_n$ was defined in (10). Using the distribution of the maximum of the standard Brownian bridge B° [7, p. 63], for any $\varepsilon > 0$ we obtain

$$\begin{aligned}\mathbb{P}(\|\widetilde{X}_n - \widehat{X}_n\|_\infty \ge \varepsilon) &= \mathbb{P}\bigg(\max_{1\le k\le n}\ \sup_{s\in[t_{n,k-1},t_{n,k}]} |{B^\circ}_{n,k}(s)| \ge \varepsilon\bigg) \\ &\le \sum_{k=1}^n \mathbb{P}\bigg(\sup_{t\in[0,1]} |B^\circ(t)| \ge \varepsilon/\sqrt{\Delta_n}\bigg) \le 2n \exp\big\{-2\varepsilon^2/\Delta_n\big\}.\end{aligned}$$

Hence, $\|\widetilde{X}_n - \widehat{X}_n\|_\infty \xrightarrow{\text{p}} 0$ since $\Delta_n = 1/n$. Further, $\|\widehat{X}_n - X_n\|_\infty = \sup_{t\in(0,1]}|X_n(t) - X_n(t-)| \xrightarrow{\text{p}} 0$ since $X_n \Rightarrow X$ and X is almost surely continuous. Therefore, $d(\widetilde{X}_n,\,X_n) \le \|\widetilde{X}_n - X_n\|_\infty \xrightarrow{\text{p}} 0$. It follows from [6, Theorem 4.1] that $\widetilde{X}_n \Rightarrow X$ as $n \to \infty$. □

Proof of Corollary 2. For sets $A \subset C$ and $B \subseteq \mathbb{R}$, set $A[B] := A \cap \{x \in C\colon x(1) \in B\}$. Recall that, for $(g^-,\,g^+) \in \mathcal{G}$, $\inf_{t\in[0,1]}(g^+(t) - g^-(t)) > 0$. For any $\varepsilon \in (0,\,\varepsilon')$, $\varepsilon' := [\inf_{t\in[0,1]}(g^+(t) - g^-(t))/2] \wedge (g^+(0) - x_0) \wedge (x_0 - g^-(0))$, and all sufficiently large n (such that $\|g_n^\pm - g^\pm\|_\infty < \varepsilon'$), the 'strips' $G^{\pm\varepsilon} := S(g^- \mp \varepsilon,\,g^+ \pm \varepsilon)$ are non-empty and

$$\mathbb{P}(\widetilde{X}_n \in G^{-\varepsilon}[B]) \le \mathbb{P}(\widetilde{X}_n \in G_n[B]) \le \mathbb{P}(\widetilde{X}_n \in G^{+\varepsilon}[B]).$$

Note that $\mathbb{P}(X \in \partial(G^{\pm\varepsilon})) = 0$ (the boundary is taken with respect to the uniform topology) due to the continuity of the distributions of $\sup_{0\le t\le 1}(X(t) - g^+(t))$ and $\inf_{0\le t\le 1}(X(t) - g^-(t))$ [6, p. 232]. Furthermore, due to $X(1)$ having a continuous density, it is clear that $\mathbb{P}(X(1) \in \partial B) = 0$ for any Borel set B with ∂B of Lebesgue measure zero. By subadditivity and the fact that $\partial(A_1 \cap A_2) \subseteq \partial A_1 \cup \partial A_2$ for arbitrary sets A_1 and A_2, it follows that

$$\mathbb{P}(X \in \partial(G[B])) \le \mathbb{P}(X \in \partial G) + \mathbb{P}(X(1) \in \partial B) = 0.$$

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

Hence, from Corollary 1,

$$\liminf_{n\to\infty} \mathbb{P}(\widetilde{X}_n \in G_n[B]) \geq \mathbb{P}(X \in G^{-\varepsilon}[B]), \qquad \limsup_{n\to\infty} \mathbb{P}(\widetilde{X}_n \in G_n[B]) \leq \mathbb{P}(X \in G^{+\varepsilon}[B]).$$

As $\mathbb{P}(X \in \partial(G[B])) = 0$, we also have

$$\mathbb{P}(X \in G^{+\varepsilon}[B]) - \mathbb{P}(X \in G^{-\varepsilon}[B]) = \mathbb{P}(X \in (\partial G)^{(\varepsilon)}[B]) \to 0 \quad \text{as } \varepsilon \downarrow 0,$$

where $(\partial G)^{(\varepsilon)}$ is the ε-neighbourhood of ∂G (also in the uniform norm). The result follows. □

Proof of relation (27). Let $r := \|g^-\|_\infty \vee \|g^+\|_\infty$. Using the elementary inequality

$$\left|1 - \prod_{j=1}^{n} a_j\right| \leq \left(\max_{i<n} \prod_{k=1}^{i} |a_k|\right) \sum_{j=1}^{n} |1 - a_j| \leq \left(1 \vee \max_{k\leq n} |a_k|\right)^{n-1} \sum_{j=1}^{n} |1 - a_j|,$$

we have

$$\begin{aligned}
|\mathbb{P}(\widetilde{X}_n \in G_n) - \widehat{\mathbf{T}}_{n,1}\widehat{\mathbf{T}}_{n,2}\cdots\widehat{\mathbf{T}}_{n,n}\mathbf{1}^\top| &= \left|\sum_{\boldsymbol{x}\in E_n^G} \prod_{k=1}^{n} \frac{q_{n,k}(x_{k-1}, x_k)}{C_{n,k}(x_{k-1})} - \sum_{\boldsymbol{x}\in E_n^G} \prod_{k=1}^{n} q_{n,k}(x_{k-1}, x_k)\right| \\
&= \left|\sum_{\boldsymbol{x}\in E_n^G} \prod_{k=1}^{n} \frac{q_{n,k}(x_{k-1}, x_k)}{C_{n,k}(x_{k-1})} \left(1 - \prod_{i=1}^{n} C_{n,i}(x_{i-1})\right)\right| \\
&\leq \mathbb{P}(\widetilde{X}_n \in G_n)\rho_n^{n-1} \sum_{i=1}^{n} \sup_{x\in E_{n,k}(r)} |1 - C_{n,i}(x)|,
\end{aligned}$$

where $\rho_n := 1 \vee \max_{1\leq j\leq n} \sup_{|x|\leq r} C_{n,j}(x)$. Using Lemma 2, we obtain (27). □

5. Numerical examples

To illustrate the efficiency of our approximation scheme (12), we implemented it in the programming language `Julia` run on a MacBook Pro 2020 laptop computer with an Intel Core i5 processor (2 GHz, 16 RAM). We used the package `HyperDualNumbers.jl` to evaluate the partial derivatives of μ_X in (3).

It is well known in the numerical analysis literature that trapezoidal quadrature is extremely accurate for analytic functions [23]. In light of (27), for numerical illustration purposes we drop the normalising constants $C_{n,k}(x)$ and use $\widehat{\mathbf{T}}_{n,1}\widehat{\mathbf{T}}_{n,2}\cdots\widehat{\mathbf{T}}_{n,n}\mathbf{1}^\top$ instead of $\mathbb{P}(\widetilde{X}_n \in G_n)$ to approximate $\mathbb{P}(X \in G)$.

5.1. The Wiener process with one-sided boundary

Using the method of images, [14] obtained a closed-form expression for the crossing probability of the boundary

$$g_{\mathrm{D}}(t) := \tfrac{1}{2} - t \ln\left(\tfrac{1}{4}(1 + \sqrt{1 + 8\mathrm{e}^{-1/t}})\right), \qquad t > 0,$$

for the standard Wiener process $W := \{W(t) : t \geq 0\}$.

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

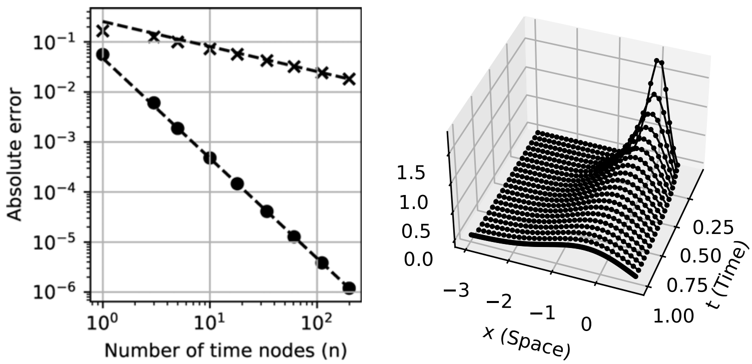

FIGURE 2. Approximation of the boundary g_D non-crossing probabilities for the Wiener process. The exact Daniels boundary crossing probability in this case is 0.479 749 35 . . . The left pane shows a log–log plot of the absolute approximation error as a function of n. The right pane shows the time evolution of the taboo transition density using the Markov chain approximation $\widetilde{X}_n$ with $n = 20$.

In order to make the state space E_n^G finite in the case of the one-sided boundary where $g^-(t) = -\infty$, we insert an absorbing lower boundary at a low enough fixed level $L < x_0$ and replace $E_{n,k}^G$ with

$$E_{n,k}^{G,L} := \{x \in E_{n,k} : L < x < g_{n,k}^+\}, \qquad k = 1, \dots, n,$$
$$E_n^{G,L} := E_{n,0} \times E_{n,1}^{G,L} \times \cdots \times E_{n,n}^{G,L}.$$

We approximate $\mathbb{P}(W(t) < g_D(t),\ t \in [0, 1])$ with $\big(\prod_{k=1}^n \mathbf{T}_{n,k}^L\big)\mathbf{1}^\top$, where $\mathbf{T}_{n,k}^L$ are sub-stochastic matrices of dimensions $\big(|E_{n,k-1}^{G,L}| + 1\big) \times \big(|E_{n,k}^{G,L}| + 1\big)$ with entries equal to the respective transition probabilities

$$\begin{cases} q_{n,k}(x, y), & (x, y) \in E_{n,k-1}^{G,L} \times E_{n,k}^{G,L}, \\ \sum_{z \in E_{n,k} \cap (-\infty, L]} q_{n,k}(x, z), & x \in E_{n,k-1}^{G,L},\ y = L, \\ 1, & x = L,\ y = L, \\ 0, & \text{otherwise}, \end{cases}$$

where we put $f_n^-(t) = -\infty$ in the definition of $q_{n,k}$ in (26). This approximation assumes that the lower auxiliary boundary L is sufficiently far away from the initial point x_0 and the upper boundary, such that after a sample path crosses the lower boundary it is highly unlikely that it will cross the upper boundary in the remaining time. In our example, we took $L = -3$. The probability of the Wiener process first hitting this level and then crossing g_D prior to time $t = 1$ is less than 1.26×10^{-6}. Further, we chose $x_0 = 0$, $\delta = 0$, and $\gamma = 2$. To guarantee convergence of the scheme, δ must be strictly positive; however, for the values of n we are interested in and the larger value of γ compared to the one in the example from Fig. 1, the error is negligible.

The left pane in Fig. 2 shows a log–log plot of the absolute approximation error as a function of n. The crosses $\times$ show the absolute error of the Markov chain approximation *without* the Brownian bridge correction, while the bullets $\bullet$ show the error when using the Brownian bridge correction. The upper and lower dashed lines correspond to $C_1 n^{-1/2}$ and $C_2 n^{-2}$ respectively, where C_1 and C_2 are fitted constants. We see that the convergence rate

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

$|\mathbb{P}(W \in G) - \mathbb{P}(\widetilde{X}_n \in G_n)|$ is of the order $O(n^{-2})$, which is the same as the boundary approximation order of the error $|\mathbb{P}(W \in G) - \mathbb{P}(W \in G_n)| = O(n^{-2})$ proved in [8] in the case of twice continuously differentiable boundaries. It appears that, due to the high accuracy of approximation of the increments' moments (Lemma 3) and the Brownian bridge correction applied to the transition probabilities, we achieve a much faster convergence rate compared to the convergence rate $O(n^{-1/2})$ achieved in [19].

The right pane in Fig. 2 shows the time evolution of the taboo transition density using the Markov chain approximation $\widetilde{X}_n$ with $n = 20$. The positions of the nodes on the surface correspond to the points from the respective $E^{G,-3}_{n,k}$. Note from (5) that the spacing between the nodes at the final time step $t = 1$ is finer compared to earlier time steps. This is crucial for the observed improved convergence rate.

5.2. The Ornstein–Uhlenbeck process

Let X be the Ornstein–Uhlenbeck (OU) process satisfying the stochastic differential equation

$$\begin{cases} \mathrm{d}X(t) = -X(t)\,\mathrm{d}t + \mathrm{d}W(t), & t \in (0, 1], \\ X(0) = 0. \end{cases}$$

The usual approach for computing the boundary crossing probability of the OU process is to express the process in terms of a time-changed Brownian motion. This is achieved by using the time substitution $\theta(t) := (\mathrm{e}^{2t} - 1)/2$, so that we can write $X(t) = \mathrm{e}^{-t} W(\theta(t))$.

To illustrate the effectiveness of our approximation, we consider the following two-sided boundary for which explicit boundary crossing probabilities are available for the OU process:

$$g^{\pm}_{\psi}(t) := \mathrm{e}^{-t}\psi_{\pm}(\theta(t)), \quad \text{where } \psi_{\pm}(t) = \pm\tfrac{1}{2} t \cosh^{-1}(\mathrm{e}^{4/t}), \qquad t > 0.$$

Letting $t := \theta(s)$, we obtain

$$\begin{aligned} P_{\psi}(T) &:= \mathbb{P}(\psi_-(t) < W(t) < \psi_+(t),\ 0 \le t \le T) \\ &= \mathbb{P}(\mathrm{e}^{-s}\psi_-(\theta(s)) < \mathrm{e}^{-s} W(\theta(s)) < \mathrm{e}^{-s}\psi_{\pm}(\theta(s)),\ 0 \le s \le \theta^{-1}(T)) \\ &= \mathbb{P}(g^-_{\psi}(s) < X(s) < g^+_{\psi}(s),\ s \in [0, \theta^{-1}(T)]), \end{aligned}$$

where we set $T := \theta(1)$ so that $s \in [0, 1]$. A closed-form expression for $P_{\psi}(T)$ can be found [38, p. 28]. The exact boundary crossing probability in this case is 0.750 502 88 . . .

Using (3) and (4), the approximate drift and diffusion coefficients of the weak second-order Itô–Taylor expansion for the OU process are given by $\beta_{n,k}(x) = -x + \frac{1}{2}\Delta_n x$, $\alpha^{1/2}_{n,k}(x) = 1 - \frac{1}{2}\Delta_n$. From the numerical results below, it appears that it is sufficient to use the weak second-order Itô–Taylor expansion of transition densities instead of the true transition density of the OU process for our Markov chain approximation to maintain a $O(n^{-2})$ convergence rate of the boundary crossing probabilities.

In the log–log plot in the left pane of Fig. 3, the crosses $\times$ show the absolute error of the Markov chain approximation *without* the Brownian bridge correction, while the bullets $\bullet$ show the error with the Brownian bridge correction. The markers $\triangle$ and $\circ$ show the absolute error of the Markov chain approximation using the exact transition density of the OU process and the Euler–Maruyama approximation instead of the transition density from the Itô–Taylor expansion, respectively. From this plot, we empirically observe that the convergence rate $|\mathbb{P}(X \in G) - \mathbb{P}(\widetilde{X}_n \in G_n)|$ is of the order $O(n^{-2})$, which is the same as the boundary

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

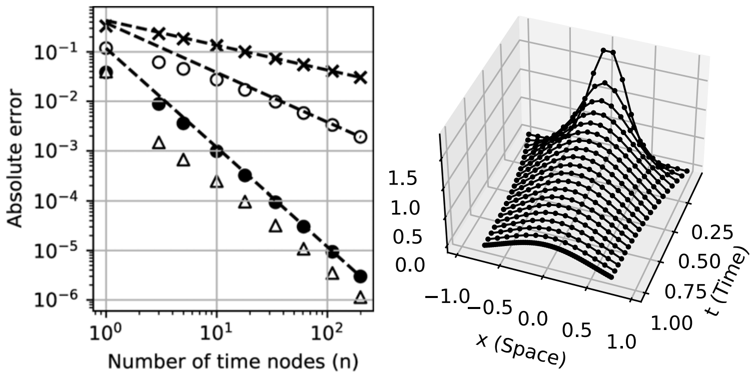

FIGURE 3. Approximation of non-crossing probabilities of the OU process with boundaries $g_{\psi}^{\pm}$. The left pane shows a log–log plot of the absolute approximation error as a function of n. The upper and lower dashed lines correspond to $C_1 n^{-1/2}$, $C_2 n^{-1}$, and $C_3 n^{-2}$ respectively, where C_1, C_2, and C_3 are fitted constants. The right pane depicts the time evolution of the taboo transition density using the Markov chain approximation $\widetilde{X}_n$ with $n = 20$.

approximation order of error $|\mathbb{P}(X \in G) - \mathbb{P}(X \in G_n)| = O(n^{-2})$ proved in [15] in the case of twice continuously differentiable boundaries. It appears that, at this level of accuracy, we might ignore the higher-order terms in the diffusion bridge crossing probability derived in [3].

5.3. The Bessel process

To test our method on a non–Gaussian diffusion process, we chose the Bessel process, due to the availability of a closed-form expression for its transition density. Note also that the Cox–Ingersoll–Ross (CIR) process popular in mathematical finance is a suitably time–space-transformed Bessel process, and hence boundary crossing probabilities for the CIR process can be immediately obtained from those for the Bessel case.

For $\nu \geq 0$, the Bessel process of order ν can be defined as the (strong) solution of the stochastic differential equation [7, p. 66]

$$\mathrm{d}X(t) = \frac{2\nu + 1}{2X(t)}\,\mathrm{d}t + \mathrm{d}W(t), \qquad t > 0; \quad X(0) = x_0 > 0.$$

The approximate drift and diffusion coefficients of the weak second-order Itô–Taylor expansion for the Bessel process are given by

$$\beta_{n,k}(x) = \frac{2\nu + 1}{2x} + \frac{1 - 4\nu^2}{8x^3}\Delta_n, \qquad \alpha_{n,k}^{1/2}(x) = 1 - \frac{2\nu + 1}{4x^2}\Delta_n.$$

For numerical illustration purposes, we tested our algorithm in the case when $\nu = \frac{1}{2}$, $x_0 = 1$, $g^-(t) = 0.7$ and $g^+(t) = +\infty$, for which a closed-form expression is available for benchmarking [26, Theorem 2.2]). The results of the numerical calculations are shown in Fig. 4. The left pane presents a log–log plot of the absolute approximation error. The $\times$ markers show the Markov chain approximation error *without* the Brownian bridge correction. The markers $\circ$, $\bullet$, and Δ show the error of the Markov chain approximation with the Brownian bridge correction when using the Euler–Maruyama approximation, the Itô–Taylor approximation, and the exact transition density respectively. We can see that the observed convergence rate is the

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

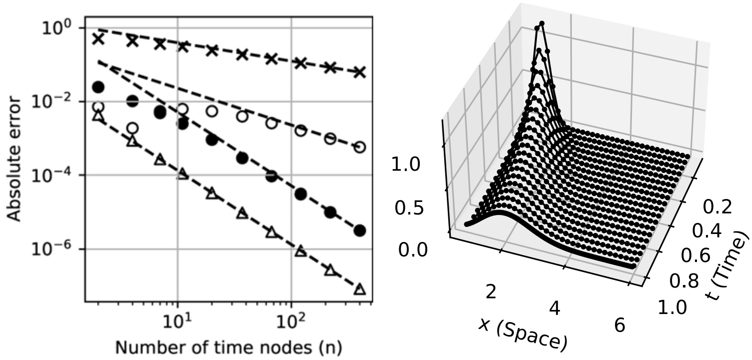

FIGURE 4. Approximation of the boundary crossing probabilities for the Bessel process with initial condition $X(0) = 1$, starting above the boundary $g^-(t) = 0.7$. The exact boundary crossing probability in this case is 0.534 924 . . . The left pane shows a log–log plot of the absolute approximation error as a function of n. The dashed lines correspond to $C_1 n^{-1/2}$, $C_2 n^{-1}$, $C_3 n^{-2}$, and $C_4 n^{-2}$ from top to bottom, where C_1, C_2, C_3, and C_4 are fitted constants. The right pane depicts the time evolution of the taboo transition density using the Markov chain approximation $\widetilde{X}_n$ with $n = 20$.

same, $O(n^{-2})$, whether one uses the Itô–Taylor approximation or the exact transition densities, but the rate drops to $O(n^{-1})$ when using the Euler–Maruyama approximation. The right pane shows the time evolution of the taboo transition density using the Markov chain approximation $\widetilde{X}_n$ with $n = 20$. An artificial absorbing boundary has been inserted at $x = 7$ to ensure that probability mass is not lost.

5.4. Comparison with other methods

In this subsection we will comment on the relative performance of our approach compared to the previously proposed ones in the special basic case of one-sided boundary g for the Brownian motion process.

Based on the results of our numerical experiments, one of the most efficient methods for numerical computation of the boundary crossing probability of the Brownian motion is the method of integral equations [49]. The method proposed in [41, 47] is based on numerically solving (using 'midpoint' quadratures) a Volterra integral equation of the first kind with the kernel $K(t, s) := \overline{\Phi}((g(t) - g(s))/\sqrt{t - s})$, $0 < s < t$, where $\overline{\Phi}$ is the standard normal 'tail'. Unfortunately, the convergence rate for that method was discussed in neither [47] nor [41]. If the above kernel K had no end-point singularity, it would follow from the proof of [40, Theorem 9.1] that the convergence rate is $O(n^{-2})$, n being the number of steps on the uniform time grid used, since that rate is essentially determined by the quadrature error. In view of the results of [42, Section 7], the proof of the above-mentioned theorem from [40] indicates that, in the case of our kernel K, the convergence rate will be $O(n^{-3/2})$. And indeed, our numerical experiments show that the convergence rate for that method is of that order and, moreover, that the error may have the form $C_0 n^{-3/2} + C_0 n^{-2} + C_0 n^{-5/2} + \cdots$. Hence, in our comparison, when comparing the results of different methods after applying Richardson's extrapolation [57, p. 27], we assumed that the error dependence on the small parameter n^{-1} for the above-mentioned integral equation method is of that form.

In [52] the first-passage-time density of the boundary g was shown to satisfy a Volterra integral equation of the second type. Its solution was approximated using the trapezoidal rule

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

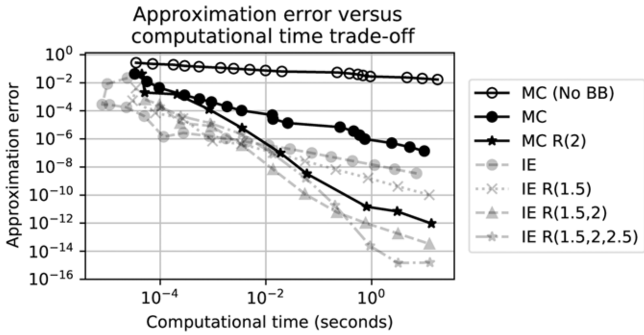

FIGURE 5. Log–log plot of the absolute approximation errors versus the computational time for our Markov chain approximation and the integral equation method applied to the standard Brownian motion crossing Daniels' boundary before time $T=1$. Labels ending with $\mathrm{R}(p_1, p_2, \ldots, p_k)$ indicate the use of Richardson's extrapolation applied repeatedly to the respective sequence $A(h)$ with an error of the form $a_1h^{p_1} + a_2h^{p_2} + \cdots + a_kh^{p_k} + o(h^{p_k})$, h being the appropriate small parameter in the scheme used.

in [11]; however, the convergence rates where not discussed in that paper. In [55], the authors obtained a series expansion for the solution to the integral equation using the method of successive approximations, with each term in the series containing an integral. Error analysis was done on the truncation of the series expansion. However, the integrals in the series expansion were obtained numerically using quadratures, and there was no mention of the convergence rate for the quadratures.

Note that both Volterra integral equations mentioned above are special cases of the spectrum of integral equations given in [49, Theorem 6.1]. Our numerical experiments showed that the approximation from [47] performs better, and hence we chose it to be the benchmark for the comparison in this subsection.

In Fig. 5 we show a comparison between the Markov chain approximation and the integral equation method proposed in [41], which is a version of the approach from [47] (labeled 'IE'). The label 'MC (No BB)' refers to the Markov chain approximation results obtained without using the Brownian bridge correction, whereas just 'MC' corresponds to the Markov chain method with that correction.

We see from Fig. 5 that applying the Brownian bridge correction dramatically improves the efficiency of our scheme. Pre-Richardson's extrapolation, the computational times for the integral equation method are generally lower, which is unsurprising since our method discretises both space and time variables, while the integral equation method only discretises time. However, our scheme becomes competitive once we apply Richardson's extrapolation. Furthermore, and most importantly, our method works for general diffusion processes whereas the method of integral equations requires an explicit expression for the transition density.

Another important advantage of our approach is that it can easily be modified to calculate expressions that involve the space variable, because our scheme essentially approximates taboo transition probabilities.

Acknowledgements

The authors are grateful to the anonymous referees for providing useful comments and drawing their attention to a number of relevant publications that helped to improve the exposition of this paper.

Funding information

Research supported by the University of Melbourne Faculty of Science Research Grant Support Scheme.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AIT-SAHALIA, Y. (2002). Maximum-likelihood estimation of discretely sampled diffusions: A closed-form approach. *Econometrica* **70**, 223–262.
[2] ANDERSON, T. W. (1960). A modification of the sequential probability ratio test to reduce the sample size. *Ann. Math. Statist.* **31**, 165–197.
[3] BALDI, P. AND CARAMELLINO, L. (2010). Asymptotics of hitting probabilities for general one-dimensional pinned diffusions. *Ann. Appl. Prob.* **12**, 1071–1095.
[4] BESKOS, A., PAPASPILIOPOULOS, O. AND ROBERTS, G. O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12**, 1077–1098.
[5] BESKOS, A. AND ROBERTS, G. O. (2005). Exact simulation of diffusions. *Ann. Appl. Prob.* **15**, 2422–2444.
[6] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
[7] BORODIN, A. N. AND SALMINEN, P. (1996). *Handbook of Brownian Motion – Facts and Formulae*. Birkhäuser, Basel.
[8] BOROVKOV, K. AND NOVIKOV, A. (2005). Explicit bounds for approximation rates of boundary crossing probabilities for the Wiener process. *J. Appl. Prob.* **42**, 82–92.
[9] BROADIE, M., GLASSERMAN, P. AND KOU, S. (1997). A continuity correction for discrete barrier options. *Math. Finance* **7**, 325–349.
[10] BUONOCORE, A., GIORNO, V., NOBILE, A. G. AND RICCIARDI, L. M. (1990). On the two-boundary first-crossing-time problem for diffusion processes. *J. Appl. Prob.* **27**, 102–114.
[11] BUONOCORE, A., NOBILE, A. G. AND RICCIARDI, L. M. (1987). A new integral equation for the evaluation of first-passage-time probability densities. *Adv. Appl. Prob.* **19**, 784–800.
[12] CHERKASOV, I. D. (1957). On the transformation of the diffusion process to a Wiener process. *Theor. Prob. Appl.* **2**, 373–377.
[13] CHEUK, T. H. F. AND VORST, T. C. F. (1996). Complex barrier options. *J. Derivatives* **4**, 8–22.
[14] DANIELS, H. E. (1969). The minimum of a stationary Markov process superimposed on a U-shaped trend. *J. Appl. Prob.* **6**, 399–408.
[15] DOWNES, A. N. AND BOROVKOV, K. A. (2008). First passage densities and boundary crossing probabilities for diffusion processes. *Methodology Comput. Appl. Prob.* **10**, 621–644.
[16] ELERIAN, O. (1998). *A Note on the Existence of a Closed-Form Conditional Transition Density for the Milstein Scheme*. Economics Discussion Paper 1998–W18, Nuffield College, Oxford.
[17] ERDÖS, P. AND KAC, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52**, 292–302.
[18] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. Wiley, New York.
[19] FU, J. C. AND WU, T.-L. (2010). Linear and nonlinear boundary crossing probabilities for Brownian motion and related processes. *J. Appl. Prob.* **47**, 1058–1071.
[20] GIORNO, V., NOBILE, A. G., RICCIARDI, L. M. AND SATO, S. (1989). On the evaluation of first-passage-time probability densities via non-singular integral equations. *Adv. Appl. Prob.* **21**, 20–36.
[21] GOBET, E. (2000). Weak approximation of killed diffusion using Euler schemes. *Stoch. Process. Appl.* **87**, 167–197.
[22] GOBET, E. AND MENOZZI, S. (2010). Stopped diffusion processes: Boundary corrections and overshoot. *Stoch. Process. Appl.* **120**, 130–162.
[23] GOODWIN, E. T. (1949). The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-x^2}\,\mathrm{d}x$. *Proc. Camb. Phil.. Soc.* 45, 241–245.
[24] GUTIÉRREZ, R., RICCIARDI, L. M., ROMÁN, P, and TORRES, F. (1997). First-passage-time densities for time-non-homogeneous diffusion processes. *J. Appl. Prob.* **34**, 623–631.
[25] HALL, W. J. (1997). The distribution of Brownian motion on linear stopping boundaries. *Sequent. Anal.* **16**, 345–352.

https://doi.org/10.1017/jpr.2023.11 Published online by Cambridge University Press

[26] HAMANA, Y. AND MATSUMOTO, H. (2013). The probability distributions of the first hitting times of Bessel processes. *Trans. Amer. Math. Soc.* **365**, 5237–5257.
[27] HERRMANN S. AND ZUCCA C. (2019). Exact simulation of the first-passage time of diffusions. *J. Sci. Comput.* **3**, 1477–1504.
[28] HERRMANN S. AND ZUCCA C. (2020). Exact simulation of first exit times for one-dimensional diffusion processes. *ESAIM: M2AN* **54**, 811–844.
[29] HILLE, E. AND PHILLIPS, R. S. (1996). *Functional Analysis and Semi-groups,* reprint, rev. (AMS Colloquium Publications **31**). AMS, Providence, RA.
[30] HURN, A. S., JEISMAN, J. I. AND LINDSAY, K. A. (2007). Seeing the wood for the trees: A critical evaluation of methods to estimate the parameters of stochastic differential equations. *J. Financial Econometrics* **5**, 390–455.
[31] ICHIBA, T. AND KARDARAS, C. (2011). Efficient estimation of one-dimensional diffusion first passage time densities via Monte Carlo simulation. *J. Appl. Prob.* **48**, 390–455.
[32] JÁIMEZ, R. G., GONZALEZ, A. J. AND ROMÁN, P. R. (1991). Construction of first-passage-time densities for a diffusion process which is not necessarily time-homogeneous. *J. Appl. Prob.* **28**, 903–909.
[33] JI, H. AND SHAO, J. (2015). First passage probabilities of one-dimensional diffusion processes. *Front. Math. China* **10**, 901–916.
[34] KHINTCHINE, A. (1933). *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung.* Springer, Berlin. [Reprinted by Chelsea Publishing Company, 1948.]
[35] KLOEDEN, P. E. AND PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin.
[36] KOLMOGOROV, A. (1931). Eine Verallgemeinerung des Laplace–Liapounoffschen Satzes. *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et naturelles*, 959–962.
[37] KOLMOGOROV, A. (1933). Über die Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et naturelles*, 363–372.
[38] LERCHE, H. R. (1986). *Boundary Crossing of Brownian Motion: Its Relation to the Law of the Iterated Logarithm and to Sequential Analysis*. Springer, Berlin.
[39] LI, C. (2013). Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.* **41**, 1350–1380.
[40] LINZ, P. (1987). *Analytical and Numerical Methods for Volterra Equations*. SIAM, Philadelphia, PA.
[41] LOADER, C. R. AND DEELY, J. J. (1987). Computations of boundary crossing probabilities for the Wiener process *J. Statist. Comput. Simul.* **27**, 95–105.
[42] LYNESS, J. N. AND NINHAM, B. W. (1967). Numerical quadrature and asymptotic expansions. *Math. Comput.* **21**, 162–178.
[43] NAGAEV, S. V. (1970). On the speed of convergence in a boundary problem I. *Theory Prob. Appl.* **15**, 163–186.
[44] NAGAEV, S. V. (1970). On the speed of convergence in a boundary problem II. *Theory Prob. Appl.* **15**, 403–429.
[45] NOVIKOV, A., FRISHLING, V. AND KORDZAKHIA, N. (1999). Approximations of boundary crossing probabilities for a Brownian motion. *J. Appl. Prob.* **36**, 1019–1030.
[46] OLVER, F. W. J. (1974). *Asymptotics and Special Functions*. Academic Press, New York.
[47] PARK, C. AND SCHUURMANN, F. J. (1976). Evaluations of barrier-crossing probabilities of Wiener paths. *J. Appl. Prob.* **13**, 267–275.
[48] PATIE, P. AND WINTER, C. (2008). First exit time probability for multidimensional diffusions: A PDE-based approach. *J. Comput. Appl. Math.* **222**, 42–53.
[49] PESKIR, G. (2002). On integral equations arising in the first-passage problem for Brownian motion. *J. Integral Equ. Appl.* **14**, 397–423.
[50] PÖTZELBERGER, K. AND WANG, L. (2001). Boundary crossing probability for Brownian motion. *J. Appl. Prob.* **38**, 152–164.
[51] RICCIARDI, L. M. (1976). On the transformation of diffusion processes into the Wiener process. *J. Math. Anal. Appl.* **54**, 185–199.
[52] RICCIARDI, L. M., SACERDOTE, L. AND SATO, S. (1984). On an integral equation for first-passage-time probability densities. *J. Appl. Prob.* **21**, 302–314.
[53] RICCIARDI, L. M. AND SATO, S. (1983). A note on the evaluation of first-passage-time probability densities. *J. Appl. Prob.* **20**, 197–201.
[54] ROGERS, L. C. G. (1985). Smooth transition densities for one-dimensional diffusions. *Bull. London Math. Soc.* **17**, 157–161.
[55] SACERDOTE, L. AND TOMASSETTI, F. (1996). Smooth transition densities for one-dimensional diffusions. *Adv. Appl. Prob.* **28**, 270–284.
[56] SAZONOV, V. V. (1981). *Normal Approximation – Some Recent Advances* (Lect. Notes Math. 879). Springer, Berlin.
[57] SIDI, A. (2003). *Practical Extrapolation Methods*. Cambridge University Press.

[58] SIEGMUND, D. AND YUH, Y. S. (1982). Brownian approximations to first passage probabilities. *Z. Wahrscheinlichkeitst.* **59**, 239–248.
[59] STEIN, E. M. AND WEISS, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press.
[60] TAKAHASI, H. AND MORI, M. (1973). Double exponential formulas for numerical integration. *Publ. Res. Inst. Math. Sci.* **9**, 721–741.
[61] WANG, L. AND PÖTZELBERGER, K. (1997). Boundary crossing probability for Brownian motion and general boundaries. *J. Appl. Prob.* **34**, 54–65.
[62] WANG, L. AND PÖTZELBERGER, K. (2007). Crossing probabilities for diffusion processes with piecewise continuous boundaries. *Methodology Comput. Appl. Prob.* **9**, 21–40.