Estimating Boundary Crossing Probabilities by Adaptive Control Variables

Pötzelberger Klaus
Vienna University of Economics, Institute for Statistics and Mathematics
Augasse 2-6
A-1090 Vienna, Austria
E-mail: Klaus.Poetzelberger@wu.ac.at

1 Introduction

Let $(W_t)_{t\geq 0}$ be a standard Brownian motion starting from 0, $T > 0$ a fixed time-horizon, $a(t)$ and $b(t)$ real functions (boundaries) satisfying $a(t) \leq b(t)$ for all $0 < t \leq T$ and $a(0) < 0 < b(0)$. In this paper, we are concerned with Monte Carlo methods for computing the following boundary crossing probability (BCP)

(1) $$P\left(a,b\right) = P\left(a(t) < W_t < b(t), \forall t \in [0,T]\right),$$

or, in the one-sided case,

(2) $$P\left(b\right) = P\left(W_t < b(t), \forall t \in [0,T]\right).$$

The BCP for a diffusion process(X_t), i.e. the solution of the stochastic differential equation

(3) $$dX_t = \mu(t,X_t)dt + \sigma(t,X_t)dW_t, \qquad X_0 = x_0,$$

is in its general form

(4) $$P\left(a,b;h\right) = \mathbb{E}[h((X_t)_{0\leq t\leq T})I(a(t) < X_t < b(t),\ \forall t \in [0,T])],$$

where h is an integrable function of $(X_t)_{0\leq t\leq T}$.

Computing boundary crossing probabilities is of considerable importance in various fields, such as sequential methods in statistics (Siegmund 1985). A prominent example in financial mathematics is the pricing of barrier options. The derivative pays a payoff at maturity if the underlying never crosses a barrier (knock-out option) or only if it crosses the barrier (knock-in option). In credit risk analysis an interesting approach is to assume that a firm issuing a bond defaults in case its value crosses below a boundary (Roberts and Shortland (1997), Beibel and Lerche (1997), Lin (1998)).

Unfortunately, there are very few boundaries for which the BCP can be computed in closed form. For the linear boundary Doob (1949) derived the one-sided BCP, for quadratic boundaries and square-root boundaries series expansions or results on the Laplace transform of the density of the hitting time exist. For Daniel's boundaries the BCP may be computed explicitly, see Daniels (1969). Lerche (1986) has more on the method of images and a list of boundaries for which the BCP can be computed analytically.

Therefore in most cases one has to resort to numerical methods to compute a BCP approximately. Popular approaches derive the BCP as a solution of a differential or an integral equation (Ricciardi et al. (1984), Loader and Deely (1987), Sacerdote and Tomassetti (1996)) or use numerical quadrature, see Novikov et al. (1999).

The methods we propose and analyze are based on approximating the boundary b (by a piecewise linear boundary b_m and estimating the BCP for b_m by a MC (Monte Carlo) method. m is the number of intervals on which b_m is linear. Steps that lead to a highly efficient procedure are presented in the subsequent sections.

The central topic of the paper is the MSE (mean squared error) of the MC procedure in terms of N, the number of univariate random variables used. Typically for infinite-dimensional methods, the rate of the MSE is $O(1/N^\beta)$ with $\beta < 1$. We present the adaptive control variables approach to reduce the variance of the MC estimators. Convergence rates of the MSE are derived. Approximating the control variables themselves by control variables leads to iterated adaptive control variables, leading to a MSE of order $O(1/N)$. A numerical example emphazises the usefulness of the proposed approach. Proofs and technical details are skipped, but may be obtained from the author

2 Monte Carlo Estimation

2.1 Piecewise Linear Boundaries

In this and the next section we propose MC estimators for the boundary crossing probability of a standard Brownian motion $(W_t)_{0\le t\le T}$. Let b and a denote functions $[0,T]\to\mathbb{R}$ with $a(0)<0<b(0)$ and $a(t)\le b(t)$ for $t\in[0,T]$. Furthermore, assume that b and a are primitives, i.e. that measurable functions b' and a' exist, such that for all $t\in[0,T]$, $b(t)=b(0)+\int_0^t b'(s)ds$ and $a(t)=a(0)+\int_0^t a'(s)ds$.

For a linear boundary $b(t)=\alpha+\beta t$ with $\alpha>0$, we have for the one-sided BCP (Φ and ϕ denote the cdf and the pdf of the standard normal distribution),

$$P(b)=\Phi\left(\frac{\alpha+\beta T}{\sqrt{T}}\right)-\Phi\left(\frac{\beta T-\alpha}{\sqrt{T}}\right)e^{-2\alpha\beta}.$$

This result follows from the BCP of the Brownian bridge

$$P(W_t\le b(t),\,\forall t\in[t_0,t_1]\mid W_{t_0}=w_{t_0},W_{t_1}=w_{t_1})=1-\exp\left\{-\frac{2}{t_1-t_0}(b(t_0)-w_{t_0})(b(t_1)-w_{t_1})\right\}.$$

Pötzelberger and Wang (2001) used this result to construct a MC procedure for piecewise linear boundaries. Let b be linear on $[t_{i-1},t_i]$, $i=1,\dots,m$, where $0=t_0<t_1<\cdots<t_m=T$ is a partition of the interval $[0,T]$. Let $\Delta t_i=t_i-t_{i-1}$. Then

$$P(b)=\mathbb{E}[g_m(W_{t_1},\dots,W_{t_m})],$$

(5) $$g_m(w_{t_1},\dots,w_{t_m}) = \prod_{i=1}^m\left(1-\exp\left\{-\frac{2}{\Delta t_i}(b(t_{i-1})-w_{t_{i-1}})(b(t_i)-w_{t_i})\right\}\right)I(w_{t_i}<b(t_{i-1})).$$

2.2 Conditional Estimator

A possible approach to estimating the BCP by a MC procedure is as follows. Approximate the boundary b by a boundary b_m, which is linear on m intervals. Generate n discrete paths $(W_{t_i}^k)_{i=1,m}$, $k=1,\dots,n$, of the Brownian motion, estimate $P(b)$ by

(6) $$\hat{P}_n(b_m)=\frac{1}{n}\sum_{k=1}^n g_m((W_{t_i}^k)_{i=1,m}).$$

This procedure leads to an error depending on

(7) $$\Delta_m=|P(b)-P(b_m)|$$

and n, the number of discrete paths generated. Let $N=mn$ denote the number of univariate Gaussian random variables generated. σ_m^2 is the variance of g_m. The mean squared error (MSE) is then

$$\Delta_m^2+\mathbb{E}[(P(b_m)-\hat{P}_n(b_m))^2]=\Delta_m^2+\sigma_m^2\frac{m}{N}.$$

In finite-dimensional problems the MSE is $O(1/N)$. Here, as a typical infinite-dimensional estimation problem, the rate $O(1/N)$ is achieved due to the bias $P(b) - P(b_m)$. Note that since the MSE depends not only on the number of paths generated, but also on the bias, we have to increase m and N to get a consistent estimate. The behavior of Δ_m for $m \to \infty$ is well- known. Pötzelberger and Wang (2001) and later Borovkov and Novikov (2003) have shown that if b_m is the secant of b on the intervals $[t_{i-1}, t_i]$, then

$$\Delta_m = O(1/m^2)$$

with $\Delta_m = o(1/m^2)$ only if b is itself piecewise linear. Therefore the MSE is $O(1/m^4) + O(m/N) = O(1/N^{4/5})$ when $m \propto N^{1/5}$.

There exists an unconditional version of the estimator. Generate additionally to each discrete path $W^k = (W^k_{t_i})_{i=1,m}$ m variables $(Y^k_1, \ldots, Y^k_m)$, which are, given W^k, independent, with $Y^k_i \in \{0, 1\}$, $Y^k_i = 0$ if $W^k_{t_{i-1}} > b(t_{i-1})$ or $W^k_{t_i} > b(t_i)$ or else

$$P(Y^k_i = 1) = 1 - \exp\left\{ -\frac{2}{\Delta t_i}(b(t_{i-1}) - W^k_{t_{i-1}})(b(t_i) - W^k_{t_i}) \right\}.$$

Let

$$\hat{P}_n(b_m) = \sum_{k=1}^{n} \frac{\prod_{i=1}^{m} Y^k_i}{n}. \tag{8}$$

This estimator is interesting only from a theoretical point of view, see the proof of Theorem 2. It uses twice the number of random variables as the conditional estimator and its variance is greater.

2.3 Assumptions and Results on the Bias

Theorem 1 gives for Brownian motions the relevant results for the approximation of a boundary b by boundaries $\tilde{b}$ in the norm $\|b - \tilde{b}\|_\infty = \sup\{|b(t) - \tilde{b}(t)| \mid 0 \le t \le T\}$. We call a boundary b_m the approximating polygon of b of order m, if for $t_i = iT/m$, b_m is linear on $[t_{i-1}, t_i]$ and $b_m(t_i) = b(t_i)$.

Theorem 1 *Let (W_t) denote a Brownian motion and b a boundary with $b(0) > 0$. Let*

$$\tau = \inf\{t < T \mid W_t \ge b(t)\}$$

with $\tau = \infty$ on $\{W_t < b(t)$ for all $t < T\}$.

1. *Let b be Lipschitz continuous on $[0, T]$ with Lipschitz constant*

$$b'_\infty := \sup\{|b(t) - b(u)|/|t - u| \mid 0 \le u < t \le T\} < \infty.$$

Then, for $\epsilon > 0$,

$$\Delta(b, b + \epsilon) \le \epsilon \left(2b'_\infty(1 - P(b)) + \frac{2}{\sqrt{T}}\phi\left(\frac{b(T)}{\sqrt{T}}\right) \right). \tag{9}$$

2. *Let b be additionally twice differentiable with $|b''|$ bounded by b''_∞ on $]0, T[$ and let $b(0) > 0$. Then for the approximating polygon b_m of order m,*

$$\|b - b_m\|_\infty \le \frac{T^2 b''_\infty}{8m^2}. \tag{10}$$

Thus

$$\Delta(b, b_m) \le \frac{1}{m^2}\left(2b'_\infty(1 - P(b)) + \frac{2}{\sqrt{T}}\phi\left(\frac{b(T)}{\sqrt{T}}\right) \right) \frac{T^2 b''_\infty}{8}. \tag{11}$$

Remark. Uniform partitions, i.e. equally spaced (t_i), are not optimal. Pötzelberger and Wang (2001) derived properties of asymptotically optimal partitions and proved that $\Delta(b, b_m) = O(1/m^2)$, when the curvature of b is taken into account. For uniform partitions, $\Delta(b, b_m) = O(1/m^2)$ has been proved by Borovkov and Novikov (2005). Their proof and that of Theorem 1 are essentially the same. Our result is an improvement also for Lipschitz continuous boundaries. Moreover, it is new for boundaries with unbounded derivative for which the density of τ is bounded on neighborhoods of the poles of b', like $b(t) = b(0) + \sqrt{t}$.

2.4 Naive MC Estimator

The convergence rate $O(1/N^{4/5})$ of the MSE for MC estimators that are based on approximating the boundary by a piecewise linear boundary my be improved considerably, for instance by the methods presented in the next section. To show that it is not too bad either, i.e. that one could do much worse, let us briefly discuss the naive MC approach to estimating a BCP. Here $P(b)$ is approximated by the discrete BCP

$$P_m^E(b) = P(W_{t_i} \le b(t_i), i = 0, \ldots, m). \tag{12}$$

Let $t_i = Ti/m, i = 0, \ldots, m$, generate $n = N/m$ paths $W^k = (W^k_{t_1}, \ldots, W^k_{t_m})$, estimate $P(b)$ by

$$\hat{P}(W_{t_i} \le b(t_i), i = 0, \ldots, m) = \hat{P}_n. \tag{13}$$

It is known that we have

$$\Delta_m^E = P_m^E(b) - P(b) = O(m^{-1/2}). \tag{14}$$

To see this, consider a constant boundary. The maximum of a Brownian motion in an interval of length $1/m$ has a size proportional to $m^{-1/2}$. More precisely, define

$$\begin{aligned} W^* &= \max\{W_t \mid t \in [0, 1]\} \\ W_m^* &= \max\{W_{i/m} \mid i = 0, \ldots, m\} \end{aligned}$$

Chernoff (1965) showed

$$W^* - W_m^* = \frac{c}{\sqrt{m}} + o\left(\frac{1}{\sqrt{m}}\right), \quad c = \frac{\zeta(1/2)}{\sqrt{2\pi}} \approx 0.5826,$$

with ζ the Riemann zeta function. Thus

$$\begin{aligned} P(b) - P_m^E(b) &= P(b) - P(b - \frac{c}{\sqrt{m}}) + o\left(\frac{1}{\sqrt{m}}\right) \\ &= 2\Phi(b) - 2\Phi(b - \frac{c}{\sqrt{m}}) + o\left(\frac{1}{\sqrt{m}}\right) \\ &= \frac{2\phi(b)c}{\sqrt{m}} + o\left(\frac{1}{\sqrt{m}}\right) \end{aligned}$$

From (14) we conclude that the MSE is of the form

$$(\Delta_m^E)^2 + c_2\frac{m}{N} = c_1\frac{1}{m} + c_2\frac{m}{N} = c_3\frac{1}{N^{1/2}} \tag{15}$$

with $m \propto N^{1/2}$.

The result gets worse when the naive method is applied for estimating boundary crossing probabilities of general diffusion processes. In the case of the Brownian motion, at least the distribution of the discrete path $(W_{i/m})_{i=0,\ldots,m}$ is exact. In general, a discrete path is generated by some method like the Euler approximation, increasing errors considerably.

3 Adaptive Control Variables and Iterated Adaptive Control Variables

3.1 Adaptive Control Variables

Needless to say, there is no best or optimal MC method. There exists no lower bound to the variance except 0, so that every single method may be improved (unless the expectation is computed exactly, which is not regarded as a MC approach). The field of variance reducing methods is rich, especially for infinite-dimensional problems. Here adaptive methods can lead to essentially optimal rates (but only rates) of convergence of estimators.

The idea behind the method of control variable (to reduce the variance) is most simple. If $\mathbb{E}[g^*]$ has to be estimated, choose g with $\mathbb{E}[g^*] = \mathbb{E}[g]$ and $\mathbb{V}ar[g] < \mathbb{V}ar[g^*]$. Of course, since $\mathbb{E}[g^*]$ is unknown, the choice of g is the crucial part of the method. A control variable h is a variable, for which the expectation, $\mu_h = \mathbb{E}[h]$ is known, i.e. can be computed analytically. Then $g = g^* - (h - \mu_h)$. h should approximate g^*, since an upper bound of $\mathbb{V}ar[g]$ is $\mathbb{E}[(g^* - h)^2]$. In finite-dimensional problems h is usually specified up to a multiplicative constant that can be estimated from the generated sample. The method is equivalent to regressing g^* onto h with the variance of g^* replaced by the residual variance.

In infinite-dimensional problems the method of control variable may even be adaptive if a sequence (h_m), with $h_m \to g^*$ in the quadratic mean, is chosen. Here a hierarchy of difficulties for estimating expectations exists. The interesting and crucial fact is that the method works even if $\mathbb{E}[h_k]$ cannot be computed analytically, but if it can be estimated with “less difficulty”than $\mathbb{E}[g^*]$.

Let us describe the method of adaptive control variables (AC) for one-sided boundary crossing probabilities and the conditional estimator. Let $k < m$ be integers. The boundary b is approximated by piecewise linear boundaries b_m and b_k. $P(b_m)$ is the expectation of g_m given by (5). As control variable we choose g_k. Its expectation, $P(b_k)$, has to be estimated. Let $N = N_1 + N_2$, where in a first step N_2 univariate random variables are used to generate $n_2 = N_2/k$ discrete paths to estimate $P(b_k)$ by $\hat{P}_{n_2}(b_k)$ and then N_1 univariate random variables are used to estimate $P(b_m)$ by the method of control variable as the expectation of $g_m - (g_k - \hat{P}_{n_2}(b_k))$. To simplify the exposition of the procedure, assume that m is a multiple of k, $m = kd$. For the second, the control variable step, generate $n_1 = N_1/m$ discrete paths $W^j = (W^j_{t_i})_{i=1,\ldots,m}, j = 1, \ldots, n_1$, and estimate $P(b_m)$ by

$$\text{(16)} \qquad \hat{P}^C_{n_1}(b_m) = \sum_{j=1}^{n_1} \frac{g_m(W^j) - (g_k(W^j) - \hat{P}_{n_2}(b_k))}{n_1}.$$

Note that $g_k(W^j)$ is $g_k((W^j_{t_{id}})_{i=1,\ldots,k})$ and uses only the path W^j at the time-points t_{id}, $i = 1, \ldots, k$. Essentially, the conditional estimator (without control variable) counts the proportion of discrete paths that do not cross the boundary b_m, with a correction term that takes the probability of crossing the boundary between the discrete time-points into consideration. The conditional estimator with control variable counts only the paths that cross one boundary, but not the other. Now if both b_m and b_k converge to b, the proportion of these paths goes to 0, increasing the convergence rate of the estimator.

The error is

$$\text{(17)} \qquad P(b) - \hat{P}^C_{n_1}(b_m) = P(b) - P(b_m) + P(b_m) - \hat{P}^C_{n_1}(b_m),$$

the MSE is

$$\text{(18)} \qquad (\Delta_m)^2 + \mathbb{V}ar[g_m(W) - P(b_m) - g_k(W) + P(b_k)]\frac{m}{N_1} + \mathbb{V}ar[g_k(W) - P(b_k)]\frac{k}{N_2}.$$

It depends on $\mathbb{E}[(g_m(W) - P(b_m) - g_k(W) + P(b_k))^2]$, which can be shown to be $O(1/k^2)$ for the conditional estimator. We get

$$\text{MSE} \approx c_1 \frac{1}{m^4} + c_2 \frac{m}{k^2 N_1} + c_3 \frac{k}{N_2} \approx c_4 \frac{1}{N^{12/13}} \tag{19}$$

with $m \propto N^{3/13}$ and $k \propto N^{1/13}$.

Theorem 2 *Assume the BCP for the boundary b is estimated by an AC approach corresponding to the MC estimator defined by g_m.*

Let $\Delta_m = O(1/m^\alpha)$ for $m \to \infty$ and $\sup_{m>k} \mathbb{V}ar[g_m(W) - g_k(W)] = O(1/k^\beta)$ for $k \to \infty$. If $m = m_N$ and $k = k_N$ with

$$m \approx N^{\frac{1+\beta}{1+2\alpha(1+\beta)}}, \tag{20}$$

$$k \approx N^{\frac{1}{1+2\alpha(1+\beta)}}, \tag{21}$$

then the MSE of the AC estimator is $O(1/N^\gamma)$ with

$$\gamma = \frac{2\alpha(1+\beta)}{1+2\alpha(1+\beta)}. \tag{22}$$

Remark. 1. For sequences (m_N) and (m'_N) the notation $m_N \approx m'_N$ is short for: (m_N/m'_N) is bounded from above and below.
2. With the choices (20) and (21) and $N_1 = \pi_1 N$, $N_2 = (1-\pi_1)N$, the rate $O(1/N^\gamma)$ holds. Constants in the definition of m and k and π_1 are chosen in order to minimize the MSE.

3.2 Iterated Adaptive Control Variables

An adaptive control variable for the control variable itself can further improve the performance of the MC estimator. This leads to the method of iterated adaptive control variables (IAC).

Let $k_r < k_{r-1} < \cdots < k_1 < k_0$ be integers and let b_{k_i}, $i = 0, \ldots, r$, denote the approximating polygons of the boundary b. For each of the $r+1$ boundaries the corresponding BCP $P(b_{k_i})$ is the expectation of a function g_{k_i}, the choice of the functions g_{k_i} depend on the underlying MC method.

Proceed as follows. Let b_{k_i} be linear on $[t_j(i), t_{j+1}(i)]$, with $t_j(i) = Tj/k_i$ and $j = 0, \ldots, k_i$. Let $\tau_i = \{t_j(i) \mid j = 0, \ldots, k_i\}$. In the simplest case the $\tau_{i+1} \subseteq \tau_i$. This is the case if k_{i+1}/k_i is an integer. If the sets τ_i are not decreasing, define $\tau_i^* = \tau_i + \tau_{i+1}$ for $i < r$ and $\tau_r^* = \tau_r$.

Choose $N_0, \ldots, N_r$ with $N_0 + \cdots + N_r = N$.

Generate $n_r = N_r/k_r$ copies of the discrete Brownian motion $(W_t)_{t \in \tau_r^*}$ and estimate $P(b_{k_r})$ by $\hat{P}_{n_r}(b_{k_r}) = \overline{(g_{k_r})}_{n_r}$. The remaining BCP's are estimated with AC, starting with $i = r-1$. Generate $n_i = N_i/k_i$ copies of the discrete Brownian motion $(W_t)_{t \in \tau_i^*}$ and estimate $P(b_{k_i})$ by $\hat{P}_{n_i}(b_{k_i}) = \overline{(g_{k_i} - (g_{k_{i+1}} - \hat{P}_{n_{i+1}}(b_{k_{i+1}})))}_{n_i}$.

Notation. Let $\Delta_m = O(1/m^\alpha)$ and $\sup_{m>k} \mathbb{V}ar[g_m(W) - g_k(W)] = O(k^{-\beta})$. Let ν_Δ, ν^* and ν denote constants, such that ν_Δ is an upper bound for $\Delta_m m^\alpha$, ν^* for $\mathbb{V}ar[g_k(W)]$ and ν for $\sup_{m>k} \mathbb{V}ar[g_m(W) - g_k(W)] k^\beta$. Define

$$\text{MSE}^* = \nu_\Delta^2 \frac{1}{k_0^{2\alpha}} + \nu \frac{k_0}{k_1^\beta N_0} + \nu \frac{k_1}{k_2^\beta N_1} + \frac{k_{r-1}}{k_r^\beta N_{r-1}} + \cdots + \nu^* \frac{k_r}{N_r}. \tag{23}$$

Then $\text{MSE} \le \text{MSE}^*$.

Theorem 3 *Assume the BCP for the boundary b is estimated by an IAC variables approach corresponding to the MC estimator defined by g_m.*

1. *Let* $\Delta_m = O(1/m^\alpha)$ *for* $m \to \infty$ *and* $\sup_{m>k} \mathbb{V}ar[g_m(W) - g_k(W)] = O(1/k^\beta)$, *for* $k \to \infty$. *If* $k_i \propto N^{\omega_i}$ *with*

$$(24) \qquad \omega_i = \frac{\beta^{r+1-i} - 1}{2\alpha(\beta^{r+1} - 1) + \beta - 1} \qquad \text{for } \beta \neq 1,$$

$$(25) \qquad \omega_i = \frac{1 - i/(r+1)}{2\alpha + 1/(r+1)} \qquad \text{for } \beta = 1,$$

then the MSE of the AC estimator is $O(1/N^{\gamma_r})$ *with*

$$(26) \qquad \gamma_r = \frac{2\alpha}{2\alpha + \frac{1-\beta}{1-\beta^{(r+1)}}} \qquad \text{for } \beta \neq 1,$$

$$(27) \qquad \gamma_r = \frac{2\alpha}{2\alpha + \frac{1}{r+1}} \qquad \text{for } \beta = 1.$$

2. *If* $\Delta_m = O(1/m^2)$ *for* $m \to \infty$, *we have for the conditional estimator estimator*

$$\gamma_r = \frac{2^{r+3} - 4}{2^{r+3} - 3}.$$

For $\beta < 1$, γ_r is bounded by $2\alpha/(2\alpha + 1 - \beta) < 1$. However, if $\beta \geq 1$, then $\lim_{r\to\infty} \gamma_r = 1$, suggesting that choosing $r = r_N$ with $r_N \to \infty$ and suitable $k_0, \ldots, k_r$ and $N_0, \ldots, N_r$, leads to a procedure with MSE of order $O(1/N)$. In fact, this is true for $\beta > 1$, the upper bound (23) of the MSE is $O(1/N)$.

Let $\beta > 1$. Let ρ be a constant, its choice will be discussed later. Let $\tilde{r}$ be defined by

$$(28) \qquad \frac{\log N}{2\alpha(\beta^{\tilde{r}} - 1)/(\beta - 1) + 1} = \rho$$

and set $r = \lfloor \tilde{r} \rfloor$. Define

$$(29) \qquad \varphi_j = \frac{\beta^{j+1} - 1}{\beta - 1},$$

$$(30) \qquad c = \frac{\nu^*}{\nu_\Delta^2} \left(\frac{1 - \beta^{-(r+1)/2}}{1 - \beta^{-1/2}} \right)^2,$$

$$(31) \qquad \mu = \frac{\nu^*}{\nu \beta^{\beta/(\beta-1)}},$$

$$(32) \qquad \pi_i = \beta^{(i-r)/2} \frac{1 - \beta^{-1/2}}{1 - \beta^{-(r+1)/2}},$$

$$(33) \qquad \tilde{N}_i = \pi_i N,$$

$$(34) \qquad a_j = \mu^{\varphi_{j-1}} \beta^{j/(\beta-1)},$$

$$(35) \qquad \tilde{k}_r = \left(\frac{N 2\alpha \varphi_r}{c a_r^{2\alpha}} \right)^{1/(2\alpha\varphi_r + 1)},$$

$$(36) \qquad \tilde{k}_i = \tilde{k}_r^{\varphi_{r-i}} a_{r-i}, i = 0, \ldots, r-1.$$

Note that the quantities $\tilde{k}_i, \tilde{N}_i$ are not necessarily integers. Let us postpone this problem. With the choice $\tilde{k}_i$ for k_i, $\tilde{N}_i$ for N_i, MSE* is

$$\text{MSE}^* = \nu_\Delta^2 \left(\frac{1}{\tilde{k}_0^{2\alpha}} + \frac{c\tilde{k}_r}{N} \right) = \frac{\nu_\Delta^2 c \mu^{1/\beta} e^\rho}{N} (1 + o(1))$$

for $N \to \infty$. If $k_i = \lceil \tilde{k}_i \rceil$ and $N_i = \lfloor \tilde{N}_i \rfloor$, we get

$$\text{MSE}^* = \frac{\nu_\Delta^2 c \mu^{1/\beta} e^\rho k_r}{N(k_r - 1)}(1 + o(1)). \tag{37}$$

For $N \to \infty$, we have

$$k_r = e^\rho \mu^{1/\beta}(1 + o(1)).$$

Fix some lower bound k_r^* for k_r, such that $k_i > k_{i-1}$ for $i = 0, \ldots, r-1$. This is the case if $k_r^* \geq \mu^{-1/\beta}$. k_r^* is controlled by the choice of the parameter ρ, since

$$k_r = e^\rho \mu^{1/\beta}(1 + o(1)).$$

Theorem 4 *Assume the BCP for the boundary b is estimated by an IAC variables approach corresponding to the MC estimator defined by g_m. Let $\Delta_m = O(1/m^\alpha)$ for $m \to \infty$ and $\sup_{m>k} \mathbb{V}ar[g_m(W) - g_k(W)] = O(1/k^\beta)$ for $k \to \infty$. If $\beta > 1$, then, with the choices of r, k_i and N_i described above, we have*

$$MSE \leq \frac{\nu^* \mu^{1/\beta} e^\rho k_r^*}{N(1 - \beta^{-1/2})^2 (k_r^* - 1))} + o\left(\frac{1}{N}\right) \tag{38}$$

4 Diffusion Processes

To estimate boundary crossing probabilities $P(b)$ for diffusion processes (X_t) defined by (3), it is convenient to proceed as follows. The feasibility of the proposed steps depend on regularity conditions, which we assume to hold and which have to be checked for the considered processes and boundaries. Again, we consider the one-sided problem only, but this is done only to facilitate the exposition.

1. Check whether the process may be transformed by $Y_t = F(t, X_t)$ to a Gaussian process $dY_t = \sigma(t)dW_t$ with deterministic $\sigma(t)$. In this case, (Y_t) has the same distribution as $(W_{a(t)})$ with $a(t)$ deterministic. The BCP can then be transformed to a BCP for the Brownian motion. Necessary and sufficient conditions on μ and σ for the existence of such a transformation are given in Wang and Pötzelberger (2007).

2. If Step 1 is not operable, apply the Lamberti transform $Y_t = F(t, X_t)$ with $\partial F/\partial x = 1/\sigma(t, x)$ to get a constant diffusion coefficient. If $F(t, .)$ is invertible, (Y_t) is of the form $dY_t = \eta(t, Y_t)dt + dW_t$.

3. Apply Girsanov's transform. Let $Q << P$ with

$$\frac{dQ}{dP} = \exp\left(-\int_0^T \eta(t, Y_t)^2 dt/2 + \int_0^T \eta(t, Y_t) dW_t\right). \tag{39}$$

   W.l.g. assume that $Y_0 = 0$. Note that under Q, (Y_t) is a Brownian motion and for $h = dP/dQ$,

$$P(Y_t \leq b(t)\ \forall t \leq T) = \mathbb{E}^Q[I(Y_t \leq b(t)\ \forall t \leq T)h] = P(b; h).$$

4. Apply a method such as the Euler or Milshtein scheme (see Kloeden and Platen (1992)) to approximate h by $\hat{h}$ which is a function of $(W_{t_i})_{i=0}^m$. Replace $\Delta_m = \Delta(b, b_m)$ by $\Delta_m = \Delta(b, b_m) + |P(b_m; h) - P(b_m, \hat{h})|$. Estimate $P(b_m, \hat{h})$ by the conditional estimator with control variables.

The following table gives the rates of convergence of the MSE for the thus described procedure. The MSE is $O(1/N^\gamma)$:

Scheme	# of control variables	γ
Euler	1	3/4
Euler	2	7/8
Milshtein	1	6/7
Milshtein	2	14/15
Order 1.5	1	9/10
Order 1.5	2	21/22
Order 2	1	12/13
Order 2	2	28/29

5 Numerical Example

To verify the results on the convergence rates of the proposed methods, we estimated the BCP for the boundary

(40) $$b(t) = 1 - t\log\left(1/20 + \left(1/400 + 50\,e^{-4/t}\right)^{1/2}\right)$$

and $T = 1$. This BCP can be computed by the method of images, it is

(41) $$P(b) = \Phi(b(1)) - 0.1\,\Phi(b(1) - 2) - 50\,\Phi(b(1) - 4) = 0.7579922.$$

A boundary, for which the BCP may be computed analytically, allows to assess the accuracy of the estimators. We estimated the $P(b)$ by the naive approach and by the conditional estimator, with 0, 1 and 2 adaptive control variables. N, the number of univariate random variables used, ranged from $N = 10^4$ to $N = 10^7$. Note that Theorems 2 and 3 give the rates for $k_0, \ldots, k_r$. Proportionality factors are chosen to minimize the MSE* (23), the upper bound of the MSE. The parameters ν_Δ, ν and ν^* can be estimated.

We used the following parameters:

1. For the naive estimator, we chose in each step $m = \sqrt{N}$ intervals and therefore also $\sqrt{N}$ discrete paths.

2. For the conditional estimator, the parameters $k_0 (= m), k_1, k_2$ were (AC denotes the number of adaptive control variables):

   0 AC: $k_0 = 5, 10, 15, 25$ for $N = 10^4, 10^5, 10^6, 10^7$.

   1 AC: $(k_0, k_1) = (8, 2), (15, 3), (24, 3), (40, 4)$ for $N = 10^4, 10^5, 10^6, 10^7$.

   2 AC: $(k_0, k_1, k_2) = (9, 3, 1), (15, 3, 1), (24, 4, 2), (36, 4, 2)$ for $N = 10^4, 10^5, 10^6, 10^7$.

The MSE was estimated by computing 1000 repetitions for $N = 10^4, 10^5, 10^6$ and 200 repetitions for $N = 10^7$. The following table summarizes the results. It gives the estimated MSE, standard errors are in brackets.

Method	$N = 10^4$	$N = 10^5$	$N = 10^6$	$N = 10^7$
Naive	3.89×10^{-3}	1.21×10^{-3}	4.06×10^{-4}	1.26×10^{-4}
	(2.8×10^{-5})	(8.9×10^{-6})	(1.0×10^{-5})	(3.2×10^{-6})
0 AC	1.61×10^{-4}	3.25×10^{-5}	5.09×10^{-6}	8.77×10^{-7}
	(1.1×10^{-6})	(7.2×10^{-7})	(1.1×10^{-7})	(2.3×10^{-8})
1 AC	1.21×10^{-4}	1.87×10^{-5}	2.86×10^{-6}	3.85×10^{-7}
	(4.2×10^{-6})	(6.3×10^{-7})	(9.6×10^{-8})	(2.5×10^{-8})
2 AC	1.24×10^{-4}	2.08×10^{-5}	2.69×10^{-6}	3.44×10^{-7}
	(4.1×10^{-6})	(5.7×10^{-7})	(1.7×10^{-7})	(3.8×10^{-8})

It is quite evident that adaptive control variables improve the estimator. For small sample sizes 1 AC variable outperforms 2 AC variables, but for moderate to large sample sizes 2 AC variables are better. For the sample sizes considered, 3 AC variables were inferior to 2, however further experiments suggest that for $N \geq 10^8$, or greater, 3 AC are better.

REFERENCES

Beibel M. and Lerche H. R. (1997). A New Look at Optimal Stopping Problems related to Mathematical Finance. *Statistica Sinica* 7, 93-108.
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.
Daniels, H. E. (1969). The minimum of stationary Markov process superimposed on a U-shaped trend. *J. Appl. Prob.* 6, 399-408.
Doob, J. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* 20, 393-403.
Kloeden, P. E. and Platen, E. (1992). *The Numerical Solution to Stochastic Differential Equations*, Springer-Verlag, New York.
Lerche, H. R. (1986). *Boundary Crossing of Brownian Motion*, Lecture Notes in Statistics 40. Springer, Berlin.
Lin, X. S. (1998). Double barrier hitting time distributions with applications to exotic options. *Insurance: Mathematics and Economics* 23, 45-58.
Loader, C. R. and Deely, J. J. (1987). Computations of boundary crossing probabilities for the Wiener process. *J. Statist. Comput. Simul.* 27, 95-105.
Novikov, A., Frishling, V. and Kordzakhia, N. (1999). Approximations of boundary crossing probabilities for a Brownian motion. *J. Appl. Prob.* 99, 1019-1030.
Pötzelberger, K. and Wang, L. (2001). Boundary crossing probability for Brownian Motion. *J. Appl. Prob.* 38, 152-164.
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.
Roberts, G. O. and Shortland, C. F. (1997). Pricing barrier options with time-dependent coefficients. *Mathematical Finance* 7, 83-93.
Sacerdote, L. and Tomassetti, F. (1996). On evaluations and asymptotic approximations of first-passage-time probabilities. *Adv. Appl. Prob.* 28, 270–284.
Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer, New York.
Wang, L. and Pötzelberger, K. (1997). Boundary crossing probability for Brownian Motion and general boundaries. *J. Appl. Prob.* 34, 54-65.
Wang, L. and Pötzelberger, K. (2007).Crossing probabilities for diffusions with piecewise continuous boundaries. *Methodol. Comput. Appl. Probab.*, 9, 21-40.

ABSTRACT

The method of adaptive control variables is an efficient Monte Carlo approach to compute boundary crossing probabilities (BCP) for Brownian motion and a large class of diffusion processes. Let N *denote the number of (univariate) Gaussian variables used for the MC estimation. Typically for infinite-dimensional MC methods, the convergence rate is less than the finite-dimensional* $O(1/N)$.

The boundary b *is approximated by a piecewise linear boundary* b_m*, which is linear on* m *intervals. The mean squared error for the boundary* b_m *is of order* $O(m/N)$*, leading to a mean squared error for the boundary* b *order* $O(1/N^{\beta})$ *with* $\beta = 2\alpha/(2\alpha + 1)$*, if the difference of the (exact) BCP's for* b *and* b_m *is* $O(1/m^{\alpha})$*. Let* b_k *be a further approximating boundary which is linear on* k *intervals. If* k *is small compared to* m*, the corresponding BCP may be estimated with high accuracy. The BCP for* b_k *is the control variable. Iterated it improves the convergence rate of the MC estimator to* $O(1/N^{1-\epsilon})$ *for all* $\epsilon > 0$*, reducing the problem of estimating the BCP to an essentially finite-dimensional problem.*