The Annals of Applied Probability
1995, Vol. 5, No. 4, 875–896

DISCRETIZATION ERROR IN SIMULATION OF ONE-DIMENSIONAL REFLECTING BROWNIAN MOTION

BY SØREN ASMUSSEN, PETER GLYNN[1] AND JIM PITMAN[2]

Aalborg University, Stanford University and University of California at Berkeley

This paper is concerned with various aspects of the simulation of one-dimensional reflected (or regulated) Brownian motion. The main result shows that the discretization error associated with the Euler scheme for simulation of such a process has both a strong and weak order of convergence of precisely 1/2. This contrasts with the faster order 1 achievable for simulations of SDE's without reflecting boundaries. The asymptotic distribution of the discretization error is described using Williams' decomposition of a Brownian path at the time of a minimum. Improved methods for simulation of reflected Brownian motion are discussed.

1. Introduction. In this paper, we study the asymptotic discretization error associated with the simulation of one-dimensional reflected Brownian motion (RBM) $\overline{B} = \{\overline{B}(t)\}_{t \geq 0}$ with drift μ and variance σ^2.

Our interest in this problem stems from the fact that one of us recently needed to compute the distribution of the r.v.

$$T_0^{-1} \int_0^{T_0} \overline{B}(t)\, dt, \quad \text{where } T_0 = \sup\{t < T: \overline{B}(t) = 0\}. \tag{1.1}$$

This r.v. arises in the limit distribution needed to produce confidence intervals for estimation of steady-state quantities associated with queues in heavy traffic; see Asmussen [3, 4] for details. In [3], the distribution (1.1) was numerically evaluated by simulating a discretized version of the BM. However, it turned out that the discretized time increment h had to be made surprisingly fine in order to obtain an accurate assessment of the limit distribution ($h = 1/2000$ for $T = 8$ and $\mu = -1$, $\sigma^2 = 1$). The question thus arises as to whether one can quantify the discretization error associated with such simulations, and whether more efficient schemes can be developed.

There is a large literature that deals with discretization errors associated with the simulation of solutions to stochastic differential equations (SDE's).

Received February 1994; revised May 1995.

[1]Research supported in part by U.S. Army Research Office Grant DAAL03-91-G0101 and by NSF Grant DDM-91-01580.

[2]Research supported in part by NSF Grant DMS-91-07531.

AMS 1991 *subject classifications*. 65C05, 60J65, 60H10.

Key words and phrases. Bessel bridge, Bessel process, bias, excursion, Euler scheme, path decomposition, Riemann zeta function, Spitzer's identity, stochastic differential equation.

In particular, given functions $b: \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma: \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$ and a standard m-dimensional Brownian motion $\mathbf{B}^* = \{\mathbf{B}^*(t)\}_{t \geq 0}$, consider the $\mathbb{R}^d$-valued solution $\mathbf{Y} = \{\mathbf{Y}(t)\}_{t \geq 0}$ of the SDE

$$d\mathbf{Y}(t) = b(\mathbf{Y}(t))\,dt + \sigma(\mathbf{Y}(t))\,d\mathbf{B}^*(t)$$

subject to $\mathbf{Y}(0) = \mathbf{y}$. In general, the above SDE cannot be solved analytically, and numerical simulation must be used to provide information about the solution $\mathbf{Y}$. The most straightforward way to simulate an approximation to $\mathbf{Y}$ is to use the Euler approximation $\mathbf{Y}_h = \{\mathbf{Y}_h(t)\}_{t \geq 0}$ defined by

$$\begin{aligned} \mathbf{Y}_h((k+1)h) &= \mathbf{Y}_h(kh) + b(\mathbf{Y}_h(kh))h \\ &\quad + \sigma(\mathbf{Y}_h(kh))(\mathbf{B}^*((k+1)h) - \mathbf{B}^*(kh)) \end{aligned} \tag{1.2}$$

[subject to $\mathbf{Y}_h(0) = y$] on the lattice $h\mathbb{N}$, and $\mathbf{Y}_h(t) = \mathbf{Y}_h(\lfloor t/h \rfloor h)$ off the lattice. One measure of the quality of the approximation at time t is the absolute error criterion

$$\mathbb{E}\|\mathbf{Y}_h(t) - \mathbf{Y}(t)\|_2, \tag{1.3}$$

where $\| \ \|_2$ is the Euclidean distance on $\mathbb{R}^d$. An approximation is said to converge strongly with order $\gamma > 0$ at time t if (1.3) is $O(h^\gamma)$ as $h \downarrow 0$. This measure is of particular importance when good pathwise approximations to $\mathbf{Y}$ are needed; see Section 9.3 of Kloeden and Platen [19] for further discussion. On the other hand, an approximation $\mathbf{Y}_h$ is said to converge weakly with order $\beta > 0$ at time t if, for each $g \in \mathscr{C}_p^{2(\beta+1)}(\mathbb{R}^d, \mathbb{R})$ (the space of functions $g: \mathbb{R}^d \to \mathbb{R}$ such that all partial derivatives up to and including $\lceil 2(\beta+1) \rceil$ exist, are continuous and have polynomial growth),

$$|\mathbb{E}g(\mathbf{Y}_h(t)) - \mathbb{E}g(\mathbf{Y}(t))| = O(h^\beta)$$

as $h \downarrow 0$. Weak approximation is important in those situations in which only an approximation of the probability distribution of $\mathbf{Y}(t)$ is required, as occurs, for example, when moments must be calculated. See Chapter 9 of Kloeden and Platen [19] for additional discussion of these convergence concepts.

Under appropriate conditions on b and σ, it has been established that the Euler approximation typically converges strongly with order 1/2, whereas it converges weakly with order 1. However, when the function σ is identically constant, the Euler scheme then enjoys both strong and weak convergence of order 1. Extensive effort has been expended in developing discretization schemes that enjoy strong and/or weak convergence orders strictly greater than 1. See, for example, Mil'stein [26], Rumelin [32], Pardoux and Talay [27], Talay and Tubaro [36] and Kloeden and Platen [19] for further results of the above type.

Duffie and Glynn [13] consider the impact of these convergence rates on the issue of how to trade off the desire to decrease h against the increased computational effort required to generate a replicate of $\mathbf{Y}_h(t)$. Given c units of computer time, it is shown that the Euler estimator for $\mathbb{E}g(\mathbf{Y}(t))$ [$g \in \mathscr{C}_p^4(\mathbb{R}^d, \mathbb{R})$] converges at rate $c^{-1/3}$ in the budget level c.

This paper is concerned with the discretization error associated with the Euler scheme $\overline{B}_h = \Gamma B_h$ for approximating one-dimensional RBM $\overline{B} = \Gamma B$, where B is Brownian motion with drift μ and variance σ^2, B_h is defined by (1.2) with $d = m = 1$ and Γ is the reflection mapping defined by

$$(\Gamma z)(t) = z(t) - \Big(\inf_{0 \le s \le t} z(s) \wedge 0\Big).$$

As we shall see, the presence of the reflecting barrier has a substantial impact on the quality of the approximation. In particular, we will show that both the strong and weak convergence orders for the Euler scheme are $1/2$ (see Theorem 2 and Proposition 3). Thus the presence of the reflecting boundary reduces the order from 1 to $1/2$ and serves partially to explain the computational results obtained in the study of the r.v. (1.1). [The explanation is only "partial" because (1.1) is a functional depending on the entire path, which neither the strong nor weak convergence orders previously described directly addresses.]

It is well known that a wide variety of queueing networks in heavy traffic can be weakly approximated by a d-dimensional RBM, in which d typically corresponds to the number of queueing stations in the network; see, for example, Reiman [29]. While much progress has been made in developing good numerical solvers for the steady-state distribution of such RBM's (see Dai and Harrison [11]), current algorithms become inefficient when d is moderately large. Monte Carlo techniques are often viewed as the method of choice in dealing with numerical solution of high-dimensional differential equations. However, our results suggest that Monte Carlo simulation of high-dimensional RBM's is likely to be particularly challenging from a computational viewpoint, due to the presence of the reflecting boundaries. This puts a premium both on the study of non–Monte Carlo solvers for high-dimensional RBM's on the one hand, and on the development of improved simulation methods for dealing with reflecting barriers on the other.

There has been some previous work on the discretization errors associated with the simulation of diffusion processes subjected to reflection in a domain $D \subseteq \mathbb{R}^d$. In particular, Chitashvili and Lazrieva [9], Kinkladze [18] and Lépingle [22] have studied the mean square rate of convergence for the case in which $D = \mathbb{R}_+ \times \mathbb{R}^{d-1}$ and the first component of the diffusion undergoes reflection at the origin. Slominski [35] extends those results to convex domains with normal reflection and establishes that the numerical analogue $\mathbf{Y}_h$ of the Euler scheme (1.2) converges a.s. on compact time intervals to $\mathbf{Y}$ at rate $O(h^{1/2-\varepsilon})$ for any $\varepsilon > 0$. Petterson [28] derives an expected value version of [35], with a somewhat improved error rate. Liu [24] obtains similar results for the special case in which the diffusion exhibits normal reflection at the boundary, and develops (for the case of normal reflection) a variant that exhibits weak convergence of the order 1.

Our results provide a sharper description of the discretization error associated with numerical simulation of RBM in the one-dimensional case. We show that the scaled discretization error $h^{-1/2}(\overline{B}(t) - \overline{B}_h(t))$ converges weakly to a proper nonzero limit, and we identify this limit precisely. This type of limit theorem is in the spirit of limit results obtained by Rootzén [31] and Kurtz and Protter [21] for SDE's without boundaries. It turns out that our limit r.v. can be constructed from two independent Bessel processes, and our proof uses various facts from the theory of path decompositions and excursions of Brownian motion. Our analysis shows, in a precise sense, that the discretization error is of exact order $1/2$, whereas the more general theorems described above show only that the rate is greater than $1/2 - \varepsilon$. We even precisely identify the form of the discretization error, both in terms of approximating expected values (Proposition 3) and approximating distributions.

The paper is organized as follows. Section 2 is devoted to computing the asymptotic distribution of the discretization error associated with the simple Euler approximation $\overline{B}_h = \Gamma B_h$ to one-dimensional RBM $\overline{B} = \Gamma B$. In Section 3, we obtain the precise form of both the strong and weak error of the approximation, and we offer some additional refinements. Finally, in Section 4, we provide a number of improved approximations to the simulation of RBM, some of which generalize to higher-dimensional diffusions.

2. The asymptotic distribution of the discretization error. We first briefly look into the interpolation error in the Euler scheme (1.2) in the special case of one-dimensional BM $B(t) = x + \sigma B^*(t) + \mu t$, where B^* is a standard one-dimensional BM. Then

$$B_h(t) = x + \sigma B^*(h\lfloor t/h \rfloor) + \mu h \lfloor t/h \rfloor. \tag{2.1}$$

Because of the constant drift and variance coefficients associated with the diffusion B, the Euler approximation B_h coincides exactly with B on the grid $h\mathbb{N}$. However, off the grid, the approximation tends to be poor because we are approximating the process B that has highly irregular paths by a process B_h that is piecewise constant. In Appendix A, we show the following result:

PROPOSITION 1. *Suppose that $g \in \mathscr{C}_p^2(\mathbb{R}, \mathbb{R})$. Then as $h \downarrow 0$, the following hold*:

(i) $$\frac{1}{\sqrt{h}} \max_{0 \le s \le t} (B(s) - B_h(s)) - \sqrt{2\sigma^2 \log\left(\frac{t}{h}\right)} \to_{\mathscr{D}} 0;$$

(ii) $$\mathbb{E} g(B_h(t)) = \mathbb{E} g(B(t)) - \frac{1}{2} \mathbb{E} g''(B(t)) \sigma^2 (t - h\lfloor t/h \rfloor) + o(h).$$

This shows that the distance of B_h from B in the uniform metric on $[0, t]$ is roughly $\sqrt{2h\sigma^2 \log(t/h)}$, whereas (ii) shows that the error in approximating expectations is of order h. Of course, any smooth approximation B_h will have poor pathwise approximation properties off the grid (in the sense of distance in the uniform metric). On the other hand, because $\mathbb{E} g(B(t))$ tends to be a smooth function of t, the weak error is easier to control between grid points. Better interpolation schemes could be expected to improve the weak error beyond the order h asserted in (ii). However, the weak error at time t will clearly depend on how close t is to a grid point [see, e.g., (ii) above].

To avoid uninteresting complications arising from the above interpolation issues, we shall only consider the discretization error for a fixed t which corresponds to a grid point. Thus we let h tend to 0 through the sequence $h = t/n$, $n = 1, 2, \dots$. Let $\varepsilon_n(t) = \overline{B}(t) - \overline{B}_h(t)$ for $h = t/n$, and note that

$$\varepsilon_n(t) = \left(\min_{0 \le k \le n} B\left(\frac{kt}{n}\right) \wedge 0 \right) - \left(\min_{0 \le s \le t} B(s) \wedge 0 \right).$$

Since the discretized minimum is taken over a smaller set, it will be greater than the continuous minimum, and hence $\varepsilon_n(t) \ge 0$. Furthermore, it is clear that $\varepsilon_n(t) = 0$ for $t < T_1$, where $T_1 = \inf\{s \ge 0: B(s) \le 0\}$. By applying the strong Markov property at time T_1 and noting that $B(T_1) = 0$, we can reduce the computation of the distribution of $\varepsilon_n(t)$ for initial condition $B(0) = x$ to one involving $B(0) = 0$. In particular, this implies that

$$\mathbb{P}_x(\varepsilon_n(t) \in \cdot) = \mathbb{E}_x(\mathbb{P}_0(\varepsilon_n(t - T_1) \in \cdot); T_1 \le t).$$

We shall therefore henceforth assume that $B(0) = \overline{B}(0) = 0$.

The main result of this section provides an approximate distribution for $\varepsilon_n(t)$ for fixed t when n is large. The approximating distribution involves the three-dimensional Bessel process BES(3), denoted by $R = \{R(t)\}_{t \ge 0}$. The process R is defined as the radial part

$$R(t) = \|\mathbf{B}^*(t)\| = \sqrt{B_1^*(t)^2 + B_2^*(t)^2 + B_3^*(t)^2}$$

of three-dimensional standard Brownian motion

$$\{\mathbf{B}^*(t)\}_{t \ge 0} = \{(B_1^*(t), B_3^*(t), B_3^*(t))\}_{t \ge 0},$$

where the B_i^*, $i = 1, 2, 3$, are independent copies of standard BM B^*.

THEOREM 1. *Let* $\check{R} = \{\check{R}(t)\}_{-\infty < t < \infty}$ *be a two-sided version of* R,

$$\check{R}(t) = \begin{cases} R_1(t), & t \ge 0, \\ R_2(-t), & t \le 0, \end{cases}$$

where R_1, R_2 *are independent copies of* R. *Then as* $n \to \infty$,

$$\sqrt{n}\, \varepsilon_n(t) \to_{\mathscr{D}} \sqrt{\sigma^2 t}\, W,$$

where $W = \min_{n = 0, \pm 1, \pm 2, \dots} \check{R}(U + n)$ *with* U *uniform on* $(0, 1)$ *and independent of* $\check{R}$.

Before proceeding to the proof, we note that this result establishes that $\varepsilon_n(t)$ is of order $n^{-1/2}$ at the (grid) point t. Thus, even though the Euler approximation B_h to B incurs zero error at the grid points, the introduction of the reflecting barrier at the origin sends the error to order $h^{1/2}$. In fact, in

Section 3, we will show that $\overline{B}_h$ converges to $\overline{B}$ at order 1/2, both weakly and strongly, at time t.

In the proof, we will assume that $\sigma^2 = 1$, $t = 1$, which can be achieved by standard transformations. It is also convenient to switch from minima to maxima by sign reversion. Recall that we have already restricted the discussion to the initial value $B(0) = \overline{B}(0) = 0$. The objective is then to show that $\sqrt{n}\,\varepsilon_n \to_{\mathscr{D}} W$, where

$$\varepsilon_n = \max_{0 \le t \le 1} B(t) - \max_{k=0,1,\dots,n} B\left(\frac{k}{n}\right) \tag{2.2}$$

and B has unit variance constant and a general drift μ.

It is intuitively clear that the r.v. ε_n, for n large, is determined by the behavior of the BM B in a neighborhood of its maximizer τ (the a.s. unique random time $\tau \in [0,1]$ at which B attains its maximum value $M = \max_{0 \le t \le 1} B(t)$ over $[0,1]$). Because the asymptotics of ε_n solely depend on the local structure of the BM around τ, the distribution of the limit r.v. W does not depend on the drift μ.

The proof of Theorem 1, provided later in this section, uses a variation of a path decomposition of Williams [38] for a one-dimensional diffusion process at its global maximum into components related to BES(3). For BM with zero drift, Denisov [12] found a similar decomposition at τ: the processes

$$\left\{\frac{1}{\sqrt{\tau}} B(\tau(1-u))\right\}_{0 \le u \le 1} \quad \text{and} \quad \left\{\frac{1}{\sqrt{1-\tau}} B(\tau + (1-\tau)u)\right\}_{0 \le u \le 1}$$

are two independent Brownian meanders (see, e.g., Revuz and Yor [30], page 455, for definition of the meander). As noted by Imhof [16], the laws of the Brownian meander and the law of a BES(3) process $\{R(t)\}_{0 \le t \le 1}$ are mutually absolutely continuous, with Radon–Nikodym derivative which is a function of $R(1)$. Consequently, the two processes share a common family of conditional distributions given $R(1) = y$, $y \ge 0$. This is the family of laws $\{\mathrm{BB}(3,1,y)\}_{y \ge 0}$, where for $t > 0$, $y > 0$, a $\mathrm{BB}(3,t,y)$ is a *three-dimensional Bessel bridge from* $(0,0)$ to (t,y), that is, a process identical in law to

$$(R(s), 0 \le s \le t \,|\, R(t) = y),$$

where the BES(3) process R is started at $R(0) = 0$ (see, e.g., Salminen [33] and Fitzsimmons, Pitman and Yor [15] for rigorous treatment of the conditioning). The $\mathrm{BB}(3,t,y)$ is an inhomogeneous Markov process whose transition function can be obtained from the well-known one for BES(3) by Bayes' rule. Combination of the above results, Brownian scaling and the well-known density relation between the laws of B for general drift μ and for zero drift yields the following proposition, which is a special case of an unpublished result of Fitzsimmons [14] for one-dimensional diffusion.

PROPOSITION 2. *Let B be a BM with $B(0) = 0$, unit variance coefficient and constant drift. Let τ be the (a.s. unique) time in $[0,1]$ at which B attains its maximum $M = \max_{0 \le t \le 1} B(t)$. Conditionally on $\tau = s$, $M = m$ and $B(1) =$*

$m - y$, the process $\{m - B(s-u)\}_{0 \le u \le s}$ is a BB(3, s, m), *independent of $\{m - B(s+v)\}_{0 \le v \le 1-s}$, which is a* BB(3, $1-s$, y).

REMARK 1. The joint density of $(\tau, B(\tau), B(1))$ was obtained by Shepp [34]. See also Fitzsimmons [14] or Csáki, Földes and Salminen [10] for more general diffusions.

To proceed from Proposition 2 to Theorem 1, we need the following lemmas.

LEMMA 1. *Let $A, T, z > 0$ be fixed. Then conditionally on $R(T) = z$, $\{\sqrt{n}\, R(t/n)\}_{0 \le t \le A}$ converges in distribution to $\{R(t)\}_{0 \le t \le A}$ in $C[0, A]$ as $n \to \infty$.*

PROOF. The result is well known (see, e.g., Lemma 11 of [6]) but since the proof is short, we include it for the convenience of the reader. Let $\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$, $\|\mathbf{w}\| = z$. Then conditionally on $\mathbf{B}(T) = \mathbf{w}$, $\mathbf{B}(A/n)$ is approximately normal with mean $A\mathbf{w}/n$ and covariance matrix $\boldsymbol{\Delta}_{A/n} - \boldsymbol{\Delta}_{A/n}\boldsymbol{\Delta}_T^{-1}\boldsymbol{\Delta}_{A/n}$, where $\boldsymbol{\Delta}_{\sigma^2} = \sigma^2 \mathbf{I}$ and $\mathbf{I}$ is the 3×3 identity matrix. Thus, in the conditional limit $\sqrt{n}\,\mathbf{B}(A/n)$ is normal $(0, A\mathbf{I})$, that is, has the marginal unconditional distribution of $\mathbf{B}_A$. Involving the Markov property of $\{\mathbf{B}(t)\}$, it follows that, in the conditional limit,

$$\left\{\sqrt{n}\,\mathbf{B}\left(\frac{t}{n}\right)\right\}_{0 \le t \le A} \to_{\mathscr{D}} \{\mathbf{B}(t)\}_{0 \le t \le A}.$$

From this, the lemma easily follows. □

Combining Proposition 2 and Lemma 1, we obtain the following lemma.

LEMMA 2. *Let $A > 0$ be fixed. Then conditionally on $M = m$, $\tau = s$, $B(1) = m - y$,*

$$\left\{\left(\sqrt{n}\left(M - B\left(\tau - \frac{t}{n}\right)\right), \sqrt{n}\left(M - B\left(\tau + \frac{t}{n}\right)\right)\right)\right\}_{0 \le t \le A}$$
$$\to_{\mathscr{D}} \{(R_1(t), R_2(t))\}_{0 \le t \le A}$$

in $C[0, A] \times C[0, A]$ as $n \to \infty$.

LEMMA 3. (a) *Let $a, b > 0$. The probability that a Brownian bridge from $(0,0)$ to $(T, 0)$ crosses the line from $(0, a)$ to (T, b) is $e^{-2ab/T}$.*

(b) *Let $L_{T,b}$ be the minimum of a Brownian bridge from $(0,0)$ to (T, b). Then for $x \in [0, \varepsilon]$,*

$$\mathbb{P}(L_{T,b} \le -x \mid L_{T,b} > -\varepsilon) = \frac{e^{-2(b+x)/T} - e^{-2\varepsilon(b+\varepsilon)/T}}{1 - e^{-2\varepsilon(b+\varepsilon)/T}}. \tag{2.3}$$

PROOF. Part (a) is formula (20) of Lévy [23]. Noting that a Brownian bridge from $(0,0)$ to (T,b) is identical in law to the process obtained by adding drift b/T to a Brownian bridge from $(0,0)$ to $(T,0)$, we observe that (b) is then a simple consequence of (a). □

Now write $\varepsilon_n = \min(\varepsilon_{n,A}, \delta_{n,A})$, where

$$\varepsilon_{n,A} = \min_{k=0,\dots,n,\, |k/n-\tau|\le A/n} \left(M - B\left(\frac{k}{n}\right)\right),$$

$$\delta_{n,A} = \min_{k=0,\dots,n,\, |k/n-\tau|\ge A/n} \left(M - B\left(\frac{k}{n}\right)\right).$$

LEMMA 4. *For any fixed* $a < \infty$,

$$\lim_{A\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\sqrt{n}\,\delta_{n,A} \le a\right) = 0.$$

PROOF. Clearly, $\delta_{n,A} \ge \min(\delta_{n,A}^{(-)}, \delta_{n,A}^{(+)})$, where

$$\delta_{n,A}^{(-)} = \min_{0\le t\le \tau - A/n} (M - B(t)), \qquad \delta_{n,A}^{(+)} = \min_{\tau + A/n \le t \le 1} (M - B(t)).$$

Thus, it suffices to prove the lemma with $\delta_{n,A}$ replaced by (say) $\delta_{n,A}^{(+)}$, and it is enough to do this with the basic probability $\mathbb{P}$ replaced by the conditional probability

$$\mathbb{P}_{m,s,y} = \mathbb{P}(\cdot \,|\, M = m,\, \tau = s,\, B(1) = m - y)$$

considered in Lemma 2. Define $Z = \sqrt{n}\,(M - B(s + A/n))$,

$$\mathbb{P}_{m,s,y,z,n} = \mathbb{P}_{m,s,y}(\cdot \,|\, Z = z).$$

Then, under $\mathbb{P}_{m,s,y,z,n}$,

$$\left\{ m - \frac{z}{\sqrt{n}} - B\left(s + \frac{A}{n} + t\right)\right\}_{0\le t\le 1-s-A/n}$$

is a Brownian bridge from $(0,0)$ to $(1 - s - A/n,\, y - z/\sqrt{n})$ conditioned on having minimum greater than or equal to $-z/\sqrt{n}$. Thus, for $0 \le a \le z$, $\mathbb{P}_{m,s,y,z,n}(\sqrt{n}\,\delta_{n,A}^{(+)} \le a)$ is given by (2.3) with

$$T = T_n = 1 - s - \frac{A}{n}, \qquad \varepsilon = \varepsilon_n = \frac{z}{\sqrt{n}}, \qquad b = b_n = y - \frac{z}{\sqrt{n}},$$

$$x = x_n = \frac{z - a}{\sqrt{n}}.$$

Since $\sup_n T_n < \infty$ and $\inf_n b_n > 0$, the asymptotic form of (2.3) as $n \to \infty$ is just a/z. That is, given $Z = z$, $\sqrt{n}\,\delta_{n,A}$ is uniform on $(0,z)$. Now by Lemma 2, Z is approximately distributed as $\sqrt{AV}$ where V is χ^2 with 3 degrees of freedom. Thus, under $\mathbb{P}_{m,s,y}$ the r.v. $\sqrt{n}\,\delta_{n,A}^{(+)}$ is approximately distributed as $U\sqrt{AV}$ for large n where U is uniform$(0,1)$ and independent of V. Noting that the limit is independent of m, s, y and letting $A \to \infty$, the lemma follows. □

PROOF OF THEOREM 1. Easily,

$$\left|\mathbb{P}\left(\sqrt{n}\,\varepsilon_n \le b\right) - \mathbb{P}\left(\sqrt{n}\,\varepsilon_{n;A} \le b\right)\right| \le \mathbb{P}\left(\sqrt{n}\,\delta_{n,A} \le b\right).$$

Let

$$W_A = \min_{n=0,\pm 1,\pm 2,\dots,\ |U+n|\le A} R^*(U+n).$$

Then by Lemma 2, $\sqrt{n}\,\varepsilon_{n;A} \to_{\mathscr{D}} W_A$, and hence

$$\limsup_{n\to\infty}\left|\mathbb{P}\left(\sqrt{n}\,\varepsilon_n \le b\right) - \mathbb{P}(W_A \le b)\right| \le \limsup_{n\to\infty}\mathbb{P}\left(\sqrt{n}\,\delta_{n,A} \le b\right).$$

Letting $A \to \infty$ yields $W_A \to_{\mathscr{D}} W$. Thus, by Lemma 4,

$$\mathbb{P}\left(\sqrt{n}\,\varepsilon_n \le b\right) \to \mathbb{P}(W \le b). \qquad \square$$

REMARK 2. The proof of Theorem 1 shows that $\sqrt{n}\,\varepsilon_n$ is asymptotically independent of B. That is to say, the $\mathbb{R} \times C[0,1]$-valued random pair $(\sqrt{n}\,\varepsilon_n, B)$ converges in distribution to (W, B), where W is independent of B.

REMARK 3. In view of Theorem 1, it seems reasonable to ask for a functional limit theorem for $\{\sqrt{n}\,\varepsilon_n(s)\}_{s\ge 0}$. It is obvious from the above analysis what the limit should be: a process which is constant on each excursion interval I of BM, the values on different intervals being i.i.d. and distributed as W. In fact, it is easy to show that if one restricts $n = 2^k$ to powers of 2 and $s/t = i/2^l$ (t fixed) to dyadic rationals, then finite-dimensional distributions converge. However, in the functional setting we are faced with the interpolation issue and, more important, the difficulty that the candidate for the limit does not have D-paths (right-continuity fails at time points where the BM is at a maximum but no excursion occurs), so that one cannot work in the standard setup of weak convergence in D.

It is of interest to derive properties of the distribution of the limiting r.v. W in Theorem 1. It is easy to see that the distribution has a density, but we do not have any useful expression for the density, or any transform or moments, apart from the mean $\mathbb{E}W$ which is found in the next section. Another partial result is the following sandwich between simpler distributions.

COROLLARY 1. *Let U, U_1, U_2, V, V_1, V_2 be independent r.v.'s such that U, U_1, U_2 are uniform on $(0,1)$ and V, V_1, V_2 are χ^2 with 3 degrees of freedom. Then*

$$\min\left(U_1\sqrt{UV_1}, U_2\sqrt{(1-U)V_2}\right) \le W \le \min\left(\sqrt{UV_1}, \sqrt{(1-U)V_2}\right)$$

in the sense of stochastical ordering.

PROOF. The upper bound follows immediately by restricting the minimum in the definition of W to the time points $U, U-1$ and noting that $R(t)$ is distributed as $\sqrt{tV}$, where V is χ^2 with 3 degrees of freedom. For the lower

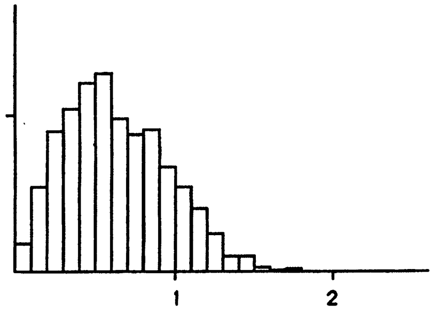

FIG. 1.

bound, we use the fact that the minimum of R starting from $R(0) = r$ is uniform on $(0, r)$ (see [39]) to conclude that $\min_{n=0,1,\dots} R(U+n)$ is stochastically larger than $U_1 R(U) =_{\mathscr{D}} U_1\sqrt{UV_1}$, together with a similar bound for $\min_{n=0,1,\dots} R(U-1-n)$. □

A histogram of 1000 simulated values of W is given in Figure 1; due to the infinite horizon of the minimum, it is nontrivial to generate W exactly, and we explain the algorithm in Appendix B. Similar histograms of the upper and lower bounds, which we omit, show that the upper bound is somewhat better. This is also confirmed by the simulated means, which came out as 0.66 for the upper bound and 0.26 for the lower one. This is to be compared with $\mathbb{E}W \approx 0.58$; see the next section.

3. Further results on the asymptotic discretization error. Both the weak and strong convergence orders of an approximation scheme are defined in terms of expected values. Our first order of business is therefore to study the mean discretization error. Let $\zeta(x)$ be the Riemann zeta function.

THEOREM 2.

$$\mathbb{E}\left|\overline{B}(t) - \overline{B}_{t/n}(t)\right| \approx -\sqrt{\frac{\sigma^2 t}{2\pi n}}\,\zeta\left(\frac{1}{2}\right) \approx 0.5826\sqrt{\frac{\sigma^2 t}{n}}\,.$$

Thus the strong order of convergence of $\overline{B}_{t/n}$ to $\overline{B}$ is precisely 1/2.

Theorem 2 will be established by bare-hand calculations, without appealing to Theorem 1. We take advantage of Spitzer's identity in discrete random walk theory (see [2], Chapter 8):

$$\mathbb{E}\max_{k=0,\dots,n} B\left(\frac{k}{n}\right) = \sum_{k=1}^{n}\frac{1}{k}\mathbb{E}B^{+}\left(\frac{k}{n}\right). \tag{3.1}$$

Since $\max_{k=0,\dots,n} B(k/n) \uparrow \max_{0\le t\le 1} B(t)$, the monotone convergence theorem, in combination with a Riemann sum approximation of the r.h.s. of (3.1), yields

$$\mathbb{E}\max_{0\le t\le 1} B(t) = \int_0^1 \frac{1}{t}\mathbb{E}B^+(t)\,dt. \tag{3.2}$$

So for ε_n as in (2.2),

$$\mathbb{E}\varepsilon_n = \int_0^1 \frac{1}{t}\mathbb{E}B^+(t)\,dt - \sum_{k=1}^n \frac{1}{k}\mathbb{E}B^+\left(\frac{k}{n}\right). \tag{3.3}$$

To analyze the asymptotic behavior of (3.3), we will use some known variations of the Euler–MacLaurin summation formula; see, for example, Lyness and Ninham [25]. Part (a) of the following lemma is standard. For completeness we indicate a quick proof of (b) and (c) via (a).

LEMMA 5. (a) *If* $f \in C_2[0,1]$, *then*

$$\int_0^1 f(x)\,dx = \frac{1}{2n}f(0) + \frac{1}{n}\sum_{k=1}^n f\left(\frac{k}{n}\right) - \frac{1}{2n}f(1) + O(n^{-2}).$$

(b) *If* $f \in C_2[0,1]$ *and* $f'(x)$ *is continuously differentiable at* 0, *then*

$$\int_0^1 f(\sqrt{x})\,dx = \frac{1}{2n}f(0) + \frac{1}{n}\sum_{k=1}^n f\left(\sqrt{\frac{k}{n}}\right) - \frac{1}{2n}f(1) + O(n^{-3/2}).$$

(c) *If* $f \in C_2[0,1]$ *and* $f'(x)$ *is continuously differentiable at* 0, *then*

$$\int_0^1 \frac{1}{\sqrt{x}}f(x)\,dx = \frac{1}{n}\sum_{k=1}^n \frac{1}{\sqrt{k/n}}f\left(\frac{k}{n}\right) + \frac{-\zeta(1/2)f(0)}{\sqrt{n}} + o(n^{-1/2}).$$

PROOF. Part (a) is standard. For (b), let $g(x) = f(\sqrt{x}) - f'(0)\sqrt{x}$. Then $g(x)$ satisfies the assumptions of (a), and hence

$$\begin{aligned}\int_0^1 f(\sqrt{x})\,dx &= \int_0^1 g(x)\,dx + f'(0)\int_0^1 \sqrt{x}\,dx \\ &= \frac{1}{2n}f(0) + \frac{1}{n}\sum_{k=1}^n \left\{ f\left(\sqrt{\frac{k}{n}}\right) - f'(0)\sqrt{\frac{k}{n}} \right\} - \frac{1}{2n}\{f(1) - f'(0)\} \\ &\quad + O(n^{-2}) + f'(0)\int_0^1 \sqrt{x}\,dx.\end{aligned}$$

Now

$$\begin{aligned}\int_0^1 \sqrt{x}\,dx - \frac{1}{n}\sum_{k=1}^n \sqrt{\frac{k}{n}} &= \frac{1}{n^{3/2}}\left\{\int_0^n \sqrt{x}\,dx - \sum_{k=1}^n \sqrt{k}\right\}\\ &= \frac{1}{n^{3/2}}\sum_{k=1}^n \int_{k-1}^k (\sqrt{x} - \sqrt{k})\,dx\\ &= \frac{1}{n^{3/2}}\sum_{k=1}^n \int_{k-1}^k \left\{\frac{x-k}{2\sqrt{k}} + O(k^{-3/2})\right\}dx\\ &= -\frac{1}{n^{3/2}}\left\{\sum_{k=1}^n \frac{1}{4\sqrt{k}} + O(1)\right\}\\ &= -\frac{1}{n^{3/2}}\left\{\frac{\sqrt{n}}{2} + O(1)\right\}.\end{aligned}$$

Collecting terms, (b) follows.

For (c), write $f(x)/\sqrt{x} = f(0)/\sqrt{x} + g(\sqrt{x})$, where $g(x) = (f(x^2) - f(0))/x$. Then g satisfies the assumption of (b). Finally, use the standard formula

$$\lim_{n\to\infty} \sqrt{n}\left(\int_0^1 \frac{1}{\sqrt{x}}\,dx - \frac{1}{n}\sum_{k=1}^n \frac{1}{\sqrt{k/n}}\right) = -\zeta\left(\frac{1}{2}\right)$$

(see [1], formula 23.2.9). □

Proof of Theorem 2. Let $U \sim N(m, \sigma^2)$. Then by easy explicit calculus,

$$\mathbb{E}U^+ = m\Phi\left(\frac{m}{\sigma}\right) + \frac{\sigma}{\sqrt{2\pi}}e^{-m^2/2\sigma^2}.$$

Thus, letting $g(t) = \mu\Phi(\mu\sqrt{t}) + \sigma e^{-\mu^2 t/2}/\sqrt{2\pi t}$, Lemma 5 yields

$$\mathbb{E}\varepsilon_n = \int_0^1 g(t)\,dt - \frac{1}{n}\sum_{k=1}^n g\left(\frac{k}{n}\right).$$

By Lemma 5(b), the contribution from $\mu\Phi(\mu\sqrt{t})$ is $O(1/n)$, and Lemma 5(c) applied to the $e^{-\mu^2 t/2}/\sqrt{2\pi t}$ term yields the result. □

A result which is related to Theorem 2 but is somewhat less explicit was recently derived by Calvin [8].

Given Theorem 2, it is natural to conclude that $\mathbb{E}W$ must be $-\zeta(1/2)/\sqrt{2\pi}$. To make this rigorous, we apply Theorems 1 and 2 and the following result.

Lemma 6. *For any $\beta < \infty$, the family $\{\exp(\beta\sqrt{n}\,\varepsilon_n)\}$ is uniformly integrable. In particular, the $(\sqrt{n}\,\varepsilon_n)^p$ are uniformly integrable for $p < \infty$.*

Proof. We use the bound

$$\varepsilon_n \le \left(M - B\left(\frac{\lfloor n\tau\rfloor}{n}\right)\right)I\left(\tau \ge \frac{1}{2}\right) + \left(M - B\left(\frac{\lfloor n\tau\rfloor + 1}{n}\right)\right)I\left(\tau \le \frac{1}{2}\right).$$

Hence by Cauchy–Schwarz and a symmetry argument, it suffices to show that

$$\limsup_{n\to\infty} \mathbb{E}\left[\exp\left(2\beta\sqrt{n}\,(M - B(\lfloor n\tau\rfloor/n))\right);\ \tau \geq \frac{1}{2}\right] < \infty. \tag{3.4}$$

By Proposition 2,

$$\begin{aligned} &\mathbb{E}_{m,s,y}\exp(\alpha(M - B(\lfloor n\tau\rfloor/n))) \\ &\quad = \mathbb{E}[\exp(\alpha R(s - \lfloor ns\rfloor/n))\,|\,R(s) = m]. \end{aligned} \tag{3.5}$$

Now for $\alpha \geq 0$, $t \leq s$,

$$\begin{aligned} &\mathbb{E}[\exp(\alpha R(t))\,|\,R(s) = m] \\ &\quad = \mathbb{E}\left[\exp\left\{\alpha\sqrt{B_1^2(t) + B_2^2(t) + B_3^2(t)}\right\}\Big|\,B_1(s) = m,\, B_2(s) = B_3(s) = 0\right] \\ &\quad \leq \mathbb{E}\big[\exp\{\alpha(|B_1(t)| + |B_2(t)| + |B_3(t)|)\}\big|B_1(s) = m, \\ &\qquad\qquad B_2(s) = B_3(s) = 0\big] \\ &\quad = \mathbb{E}\big[\exp\{\alpha(|B_1(t)|)\}\big|B_1(s) = m\big]\big(\mathbb{E}\big[\exp\{\alpha(|B_1(t)|)\}\big|B_1(s) = 0\big]\big)^2. \end{aligned}$$

Further, if X is normal with mean $\mu \geq 0$ and variance σ^2, then

$$\mathbb{E}e^{\alpha|X|} \leq \mathbb{E}e^{\alpha X} + \mathbb{E}e^{-\alpha X} \leq 2e^{\alpha\mu + \alpha^2/\sigma^2},$$

so that, for $s \geq 1/2$,

$$\mathbb{E}\big[\exp\{\alpha|B_1(t)|\}\big|B_1(s) = m\big] \leq 2\exp\left\{2\alpha t m + \frac{\alpha^2 t}{2}\right\}.$$

Letting $\alpha = 2\beta\sqrt{n}$, $t = \tau - \lfloor n\tau\rfloor/n \leq 1/n$, we get

$$\begin{aligned} &\mathbb{E}\left[\exp\left(2\beta\sqrt{n}\left(M - B\frac{\lfloor n\tau\rfloor}{n}\right)\right);\ \tau \geq \frac{1}{2}\right] \\ &\qquad \leq 8\mathbb{E}\exp\left\{\frac{8\beta M}{\sqrt{n}} + 6\beta^2\right\} \to \exp(6\beta^2). \qquad \square \end{aligned}$$

We now turn to the order of weak convergence of $\overline{B}_h$ to $\overline{B}$ at time t. Suppose that $g \in \mathscr{C}_p^2(\mathbb{R}, \mathbb{R})$. Then

$$\sqrt{n}\left(g(\overline{B}(t)) - g(\overline{B}_{t/n}(t))\right) = g'(\theta_n(t))\cdot\sqrt{n}\,\varepsilon_n(t), \tag{3.6}$$

where $\theta_n(t)$ lies between $\overline{B}_{t/n}(t)$ and $\overline{B}(t)$. Since $\overline{B}_{t/n}(t) \to_{\mathscr{D}} \overline{B}(t)$, we conclude that

$$\sqrt{n}\left(g(\overline{B}(t)) - g(\overline{B}_{t/n}(t))\right) \to_{\mathscr{D}} g'(\overline{B}(t))\sqrt{\sigma^2 t}\,W$$

as $n \to \infty$, where $\overline{B}(t)$ and W are independent (see Remark 2). However,

$$\begin{aligned} \mathbb{E}\left|g'(\theta_n(t))\cdot\sqrt{n}\,\varepsilon_n(t)\right|^2 &\leq \sqrt{\mathbb{E}g'(\theta_n(t))^4}\cdot\sqrt{\mathbb{E}(\sqrt{n}\,\varepsilon_n(t))^4} \\ &\leq \sqrt{c + d\mathbb{E}\sup_{|s-t|\leq t/n}|\overline{B}(s)|^p}\cdot\sqrt{\mathbb{E}(\sqrt{n}\,\varepsilon_n(t))^4}, \end{aligned}$$

so that the r.h.s. of (3.6) is uniformly integrable. This establishes the proof of our next result.

PROPOSITION 3. *Suppose $g \in \mathscr{C}_p^2(\mathbb{R}, \mathbb{R})$. Then as $n \to \infty$,*

$$\mathbb{E}g\big(\overline{B}_{t/n}(t)\big) = \mathbb{E}g\big(\overline{B}(t)\big) + \frac{1}{\sqrt{n}}\mathbb{E}g'\big(\overline{B}(t)\big) \cdot \sqrt{\sigma^2 t}\,\mathbb{E}W + o\bigg(\frac{1}{\sqrt{n}}\bigg).$$

This result shows that $\overline{B}_{t/n}$ converges weakly to $\overline{B}$ with order $1/2$ at time t. In contrast to Theorem 2, which describes the strong order of convergence, Proposition 3 establishes a rate of convergence that is independent of how we define the joint distribution of $\overline{B}_h$ and $\overline{B}$.

A related question is how close the probability distribution $F_n(\cdot)$ of $\overline{B}_{t/n}$ is to $F(x) = \mathbb{P}(\overline{B}(t) \le x)$. Note that F is absolutely continuous and let $f(\cdot)$ be its Lebesgue density (explicit expressions for F and f are given in Asmussen [2], but need not concern us here). From Proposition 3, we obtain, by formal integration by parts, that

$$F_n(x) = F(x) + \sqrt{\frac{\sigma^2 t}{n}}\,\mathbb{E}W \cdot f(x) + o\bigg(\frac{1}{\sqrt{n}}\bigg), \tag{3.7}$$

and, by formal differentiation, that

$$f_n(x) = f(x) + \sqrt{\frac{\sigma^2 t}{n}}\,\mathbb{E}W \cdot f'(x) + o\bigg(\frac{1}{\sqrt{n}}\bigg), \tag{3.8}$$

$x \ge 0$. The rigorous proof requires, however, some uniform versions of earlier estimates, for example, that

$$\sup_{0 \le y \le n} \Big|\mathbb{P}\big(\sqrt{n}\,\varepsilon_n \ge z \big| B(1) = y\big) - \mathbb{P}(W \ge z)\Big| \to 0,$$

$$\sup_{0 \le y \le x + n^{\delta}} \Big|\mathbb{E}\Big[\big(\sqrt{n}\,\varepsilon_n\big)^2 \Big| B(1) = y\Big] - \mathbb{E}W^2\Big| \to 0.$$

4. Improved approximations.

4.1. *On bias and optimal efficiency.* Suppose that we want to estimate $\mathbb{E}g(\overline{B}(t))$ via simulation. The approach most commonly followed would involve simulating m i.i.d. replicates of the r.v. $g(\overline{B}_h(t))$, where m and h are chosen suitably. It seems reasonable to choose these two parameters so as to maximize the accuracy of the resulting estimator for a given computer budget c.

To study the trade-off between m and h from an asymptotic standpoint, we exploit the following framework. Suppose that we wish to estimate a parameter α via independent replications of an r.v. $V(h)$. Suppose that, as $h \downarrow 0$, the following hold:

(i) $V(h) \to_{\mathscr{D}} V$;
(ii) $\mathbb{E}V(h)^2 \to \mathbb{E}V^2 < \infty$;
(iii) $\alpha(h) = \alpha + \beta h^p + o(h^p)$, where $\alpha(h) = \mathbb{E}V(h)$, $\beta \ne 0$ and $p > 0$;
(iv) the computer time required to generate $V(h)$ is given by $\tau(h)$, where $\tau(h)$ is deterministic and satisfies $\tau(h) = \gamma h^{-q} + o(h^{-q})$ where $\gamma, q > 0$.

Consider the estimator

$$\alpha(c,h) = \frac{1}{n(c,h)} \sum_{i=1}^{n(c,h)} V_i(h),$$

where the $V_i(h)$ are i.i.d. copies of $V(h)$ and $n(c,h) = \lfloor c/\tau(h) \rfloor$. Then Duffie and Glynn [13] establish the following properties of $\alpha(c, h_c)$ under conditions (i)–(iv):

(i) If $h_c c^{1/(q+2p)} \to \infty$ or if $h_c c^{1/(q+2p)} \to 0$ as $c \to \infty$, then

$$c^{p/(q+2p)} | \alpha(c, h_c) - \alpha | \to_{\mathscr{D}} + \infty.$$

(ii) If $h_c c^{1/(q+2p)} \to x$, where $0 < x < \infty$, then

$$c^{p/(q+2p)} \big(\alpha(c, h_c) - \alpha \big) \to_{\mathscr{D}} \sigma \sqrt{\frac{\gamma}{x^q}}\, N(0,1) + \beta x^p,$$

where $\sigma^2 = \operatorname{Var} V$.

In our setting, $\alpha = \mathbb{E} g(\overline{B}(t))$ and $V(1/n) = g(\overline{B}_{t/n}(t))$. If $g \in \mathscr{C}_p^1(\mathbb{R}, \mathbb{R})$, then conditions (i)–(iv) are satisfied with $p = 1/2$ and $q = 1$ [of course, in reality $\tau(h)$ is random and only approximately deterministic]. The above result therefore asserts that the best possible convergence rate is of order $c^{-1/4}$ in the computer time budget c and is achieved when $h_c \approx x c^{-1/2}$. This is to be contrasted with the rate of $c^{-1/3}$ which is typical of Euler schemes for SDE's without boundary conditions.

In contrasting the Euler approximation $g(\overline{B}_{t/n}(t))$ to other approximations of $g(\overline{B}(t))$, the above results indicate that the key parameter is the order of bias as reflected in the parameter p. Proposition 3 suggests the approximation

$$g\big(\overline{B}_{t/n}(t)\big) - \frac{1}{\sqrt{n}} g'\big(\overline{B}_{t/n}(t)\big) \sqrt{\sigma^2 t}\, \mathbb{E} W. \tag{4.1}$$

If $g \in \mathscr{C}_p^1(\mathbb{R}, \mathbb{R})$, then Proposition 3 shows that the bias of this estimator for $\alpha = \mathbb{E} g(\overline{B}(t))$ is of order $1/n$, so that $p = 1$ in condition (iii). The estimator $\alpha(c, h_c)$ based on the approximation then has the best possible convergence rate $c^{-1/3}$ (achieved when $h_c \approx x c^{-1/3}$).

4.2. *Adaptive step size.* This means that (for a fixed budget c) the time step $h_c = h_c(x)$ is chosen according to the current value x of $\{\overline{B}(t)\}$. We shall only comment briefly upon the two-step case

$$h_c(x) = \begin{cases} h_c^{(1)}, & x < x_c, \\ h_c^{(2)}, & x \ge x_c. \end{cases}$$

We obviously want $h_c^{(1)} \ll h_c^{(2)}$ and, to obtain an improved bias reduction, $h_c^{(1)} = o(c^{-1/2})$. A large x_c cannot be optimal for sample size reasons, and a small one cannot because it makes the probability of a step from level $x \ge x_c$ to 0 nonvanishing and thereby makes the bias of order $\sqrt{h_c^{(2)}}$ rather than $\sqrt{h_c^{(1)}}$.

Rough calculation, which we omit, indicates that the best possible convergence rate for such a two-step scheme is close to $c^{-3/10}$ and is attained by taking x_c close to $c^{-1/5}$, $h_c^{(1)}$ close to $c^{-3/5}$ and $h_c^{(2)}$ close to $c^{-2/5}$.

4.3. *Extrapolation.* The low-bias approximation (4.1) takes advantage of the fact that Proposition 3 provides a precise description of the principal bias term for one-dimensional RBM. We now present an approach toward bias reduction that holds promise for higher-dimensional reflected diffusions and essentially amounts to applying the concept of Richardson or Romberg extrapolation (see page 285 of Kloeden and Platen [19]). Fix $t > 0$. Proposition 3 establishes the existence of a constant β such that, for $g \in \mathscr{C}_p^1(\mathbb{R}, \mathbb{R})$,

$$\mathbb{E}g\big(\overline{B}_{t/n}(t)\big) = \mathbb{E}g\big(\overline{B}(t)\big) + \frac{\beta}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right). \tag{4.2}$$

As mentioned in the Introduction, this bias expansion is consistent with the theory that has recently been developed for higher-dimensional reflected diffusions. It therefore seems likely that such bias expansions hold in great generality for diffusions that undergo reflection. Assuming that (4.2) is valid, it is an easy matter to verify that

$$\mathbb{E}\Big[2g\big(\overline{B}_{t/n}(t)\big) - g\big(\overline{B}_{4t/n}(t)\big)\Big] = \mathbb{E}g\big(\overline{B}(t)\big) + o\left(\frac{1}{\sqrt{n}}\right)$$

as $n \to \infty$. Thus, the estimator

$$2g\big(\overline{B}_{t/n}(t)\big) - g\big(\overline{B}_{4t/n}(t)\big) \tag{4.3}$$

enjoys significantly better bias properties than does $g(\overline{B}_{t/n}(t))$ itself. When n is a multiple of 4, we note that $\overline{B}_{4t/n}(t)$ can be computed along the subgrid $\{4t/n\colon i = 0, \dots, n/4\}$ of the grid $\{t/n\colon i = 0, \dots, n\}$. Furthermore, one can (and should) use common Brownian increments to compute both $\overline{B}_{4t/n}(t)$ and $\overline{B}_{t/n}(t)$, in order to reduce both the computational burden and the variance (this is an application of the method of common random numbers; see Bratley, Fox and Schrage [7] for details).

It seems reasonable to expect that the $o(1/\sqrt{n})$ term in (4.2) would have the form $\gamma/n + o(1/n)$ for some constant γ. Then the bias of the estimator (4.3) is of order $1/n$ so that $p = 1$ in condition (iii) and a convergence rate of $c^{-1/3}$ is available.

4.4. *Liu's method.* Another approach to bias reduction was proposed recently by Liu [24]. The method was developed for general reflecting diffusions with normal reflection at the boundary (note, however, that the RBM's of Reiman [29] typically give rise to nonnormal reflection). To illustrate the method, let $\overline{B}'$ be Liu's approximation to the one-dimensional RBM $\overline{B}$ and write

$$\overline{B}'_i = \overline{B}'_{t/n}\left(\frac{it}{n}\right), \qquad \Delta B_i = B\left(\frac{(i+1)t}{n}\right) - B\left(\frac{it}{n}\right).$$

The algorithm then takes the following recursive form.

ALGORITHM A. If $\overline{B}'_i + \mu t/n + \sigma \Delta B_{i+1} \geq 0$, let $\overline{B}'_{i+1} \leftarrow \overline{B}'_i + \mu t/n + \sigma \Delta B_{i+1} \geq 0$; else let

$$\overline{B}'_{i+1} \leftarrow \begin{cases} \overline{B}'_i + \dfrac{\mu t}{n} - \sigma \Delta B_{i+1}, & \overline{B}'_i + \dfrac{\mu t}{n} \geq 0, \\ |\sigma \Delta B_i|, & \overline{B}'_i + \dfrac{\mu t}{n} < 0. \end{cases}$$

To define $\overline{B}'_{t/n}$ between grid points, let $\overline{B}'_{t/n}(s) = \overline{B}'_{t/n}(t/n\lfloor ns/t \rfloor)$. Liu showed that, under some smoothness conditions,

$$\mathbb{E}g\left(\overline{B}'_{t/n}(t)\right) = \mathbb{E}g\left(\overline{B}(t)\right) + O\left(\frac{1}{n}\right). \tag{4.4}$$

Assuming that the $O(1/n)$ term has the form $\beta'/n + o(1/n)$ for some constant β', we have $p = 1$ in condition (iii) and the best possible rate of convergence for schemes based on this approximation is $c^{-1/3}$. Of course, this would also suggest use of an extrapolation approximation similar to (4.3), providing an improved convergence rate of order $c^{-2/5}$ if the $o(1/n)$ is in fact of order n^{-2}.

4.5. *Exact algorithms in one dimension.* The improved approximations described above all were assessed in terms of weak error. It should be clear that no improvement from the point of strong convergence error should necessarily be expected. Fortunately, it turns out that for one-dimensional RBM $\overline{B}$ enough is known about the process to give an exact algorithm for generating $\overline{B}$ at the grid points.

If $\mu = 0$, the problem is simple: we can just simulate $\overline{B}$ as $\overline{B}(t) = |B(t)|$, where the BM B can be simulated at a discrete grid as a random walk with normally distributed increments. The bias at the discrete grid is 0 for both B and $\overline{B}$.

If $\mu \neq 0$, we may proceed by noting that for fixed $T > 0$ the joint density of

$$\left(B(T),\ \max_{0 \leq t \leq T} B(t)\right) \tag{4.5}$$

is known. For simulation purposes, a convenient representation of this distribution is to note that marginally $B(T)$ is normal $(\mu T, T)$ and that, by Lemma 3,

$$F_y(x) = \mathbb{P}\left(\max_{0 \leq t \leq T} B(t) - y \leq x \,\middle|\, B(T) = y\right) = 1 - e^{-2x(y+x)/T}.$$

By easy calculus,

$$F_y^{-1}(z) = \frac{-2y + \sqrt{4y^2 - 8T\log(1-z)}}{4}.$$

Thus, we may first generate $B(T)$ as normal $(-\mu T, T)$ and next let

$$\max_{0 \le t \le T} B(t) = \frac{B(T)}{2} + \frac{\sqrt{B(T)^2 - 2T\log(U)}}{2},$$

where U is uniform on $(0,1)$.

Thus, an algorithm for unbiased simulation of BM B, the maximum M and thereby RBM $X = M - B$ at the epochs $t = 0, 1/n, 2/n \dots$ is obtained. (This algorithm was independently obtained by Lépingle [22].)

ALGORITHM B.

1. Let $t \leftarrow 0$, $B \leftarrow 0$, $X \leftarrow 0$, $M \leftarrow 0$.
2. Generate (T_1, T_2) from the density (4.5) with $T = 1/n$.
3. Let $t \leftarrow t + 1/n$, $M \leftarrow \max(M, B + T_2)$, $B \leftarrow B + T_1$, $X \leftarrow M - B$.
4. Return to step 2.

For Algorithm B, both the weak and strong error of the approximation vanishes at the grid point. In particular, the rate of convergence associated with estimating $\mathbb{E} f(\overline{B}(t))$ is of order $c^{-1/2}$. However, because of the computational complexity associated with the random variate generation used in Algorithm B, and because this method fails to generalize to higher dimensions, we believe that the other improved approximations developed in this section are also of value.

We conclude this section with the discussion of a related algorithm that is exact on a random grid associated with a Poisson process that is run independently of the RBM $\overline{B}$. It takes advantage of the following well-known lemma.

LEMMA 7. *Let T be an exponential r.v. with rate λ which is independent of $\{B(t)\}$. Then the r.v.'s $\max_{0 \le t \le T} B(t) - B(T)$ and $\max_{0 \le t \le T} B(t)$ are independent and exponentially distributed with rates η and ω, respectively, where*

$$\eta = -\mu + \sqrt{\mu^2 + 2\lambda}, \qquad \omega = \mu + \sqrt{\mu^2 + 2\lambda}.$$

Thus, an algorithm for unbiased simulation of BM B, the maximum M and thereby RBM $X = M - B$ at the epochs t of a Poisson(λ) grid is obtained as follows.

ALGORITHM C.

1. Let $t \leftarrow 0$, $B \leftarrow 0$, $X \leftarrow 0$, $M \leftarrow 0$.
2. Generate T, S_1, S_2 as exponential r.v.'s with rates $1, \eta, \omega$, respectively.
3. Let $t \leftarrow t + T$, $M \leftarrow \max(M, B + S_2)$, $B \leftarrow B + S_2 - S_1$, $X \leftarrow M - B$.
4. Return to step 2.

APPENDIX A

PROOF OF PROPOSITION 1. For (i), note that

$$\begin{aligned}\max_{0\le s\le h\lfloor t/h\rfloor}&(B(s)-B_h(s))\\&=\max_{1\le k\le\lfloor t/h\rfloor}\max_{0\le s\le h}\sigma(B^*(h(k-1)+s)-B^*(h(k-1)))+s\mu\\&=_{\mathscr{D}}\sqrt{h}\left(\max_{1\le k\le\lfloor t/h\rfloor}\max_{0\le s\le 1}\sigma B_k^*(s)+s\mu\sqrt{h}\right),\end{aligned}$$

where $=_{\mathscr{D}}$ denotes equality in distribution and $B_1^*, B_2^*, \ldots$ are i.i.d. copies of B^*. Recall that $\max_{0\le s\le 1} B_k^*(s) =_{\mathscr{D}} |B_k^*(1)|$ (e.g., Karlin and Taylor [17], page 346) and that

$$\left(\max_{1\le k\le n}|B_k^*(1)|-\sqrt{2\log n}\right)\to_{\mathscr{D}} 0$$

as $n\to\infty$ (e.g., Barlow and Proschan [5]). Since $s\mu\sqrt{h}\to 0$ uniformly in $s\in[0,1]$, (i) easily follows when the maximum is taken over $[0,h\lfloor t/h\rfloor]$. But the same argument works when the maximum is taken over $[0,h\lceil t/h\rceil]$, proving (i).

For (ii), we use the fact that

$$\begin{aligned}g(B(t))&=g(B_h(t))+g'(B_h(t))(B(t)-B_h(t))\\&\quad+\tfrac{1}{2}g''(\theta_h(t))(B(t)-B_h(t))^2,\end{aligned}$$

where $\theta_h(t)$ lies between $B(t)$ and $B_h(t)$. Since $B(t)-B_h$ is independent of $B_h(t)$, the second term on the r.h.s. vanishes in expectation. On the other hand,

$$\begin{aligned}g''(\theta_h(t))(B(t)-B_h(t))^2&=g''(B_h(t))(B(t)-B_h(t))^2\\&\quad+(g''(\theta_h(t))-g''(B_h(t)))(B(t)-B_h(t))^2.\end{aligned}$$

The first term on the r.h.s. has expectation

$$\mathbb{E}g''(B_h(t))\sigma^2(t-h\lfloor t/h\rfloor).$$

The second term, when multiplied by h^{-1}, is dominated by

$$|g''(\theta_h(t))-g''(B(t))|\cdot h^{-1}\cdot\max_{0\le s\le h}\sigma^2(B^*(h\lfloor t/h\rfloor+s)-B^*(h\lfloor t/h\rfloor))^2,$$

which converges to 0 in probability. Since $g\in\mathscr{C}_p^2(\mathbb{R},\mathbb{R})$, uniform integrability holds and (ii) follows. □

APPENDIX B

Generating W by simulation. Since the definition of W involves the minimum over an infinite time horizon, the simulation of W is nontrivial and we shall therefore give an outline of the ideas behind the algorithm. For simplicity, we redefine W as $\min_{n=0,1,2,\dots} R(U+n)$ since the generation of the two-sided minimum just involves a replication with U replaced by $1-U$. Throughout, w is the candidate for the minimum at the present stage of the simulation (the minimum at the times $U+n$ for the n considered so far). The key description of BES(3) starting from $R(0) > 0$ can be found in [39] and states that the minimum r is uniform on $(0, R(0))$, that $\{R(t)\}$ evolves as BM until hitting r and as $\{r + R^*(t)\}$ thereafter, where $\{R^*(t)\}$ is BES(3) starting from $R^*(0) = 0$. The implication for the simulation is that we can run BES(3) as BES(3) segments alternating with Brownian segments.

The r.v.'s occurring in the construction (all independent) are as follows:

(a) $U, U_1, U_2, \dots$ uniform $(0,1)$ r.v.'s;
(b) $\tau_1(l_1), \tau_2(l_2), \dots$ Brownian passage times, from level 0 to level $-l_i$ for $\tau_i(l_i)$ which can be generated as l_i^2/X_i^2, where X_i is standard normal;
(c) $B_1(t_1), M_1(t_1)), (B_2(t_2), M_2(t_2)), \dots,$ random vectors with

$$(B_i(t_i), M_i(t_i)) =_{\mathscr{D}} \Big(B(t_i), \max_{t \le t_i} B(t)\Big);$$

(d) $B_1^*(t_1, l_1), B_2^*(t_2, l_2), \dots,$ r.v.'s with

$$B_i^*(t_i, l_i) =_{\mathscr{D}} \Big(B(t_i) \Big| \max_{t \le t_i} B(t) \le l_i\Big);$$

the $B_i^*(t_i, l_i)$ can be generated by acceptance–rejection from the $(B_i(t_i), M_i(t_i))$.

Step 0 is the initialization and consists of generating U and $\mathbf{B}(U)$, where $\{\mathbf{B}(t)\}$ is a three-dimensional BM underlying the BES(3). We then let $R(U) = \|\mathbf{B}(U)\|$, $w = R(U)$.

Step 1 consists of a BES(3) segment, generating $\mathbf{B}(U+1), \mathbf{B}(U+2), \dots$ and updating $R(U+1), R(U+2), \dots, w$ in an obvious manner. The procedure is stopped when $R(U+k) > Cw$ for some $C > 1$ (we took $C = 5$). Then the post-$(U+k)$ minimum is generated as $U_1 R(U+k)$. If $U_1 R(U+k) > w$, we are done and let $W = w$; otherwise, we proceed to step 2.

Step 2 consists of a Brownian segment, starting from a value of the form $R(U+k)$ and terminating when the generated post-$(U+k)$ minimum $U_1 R(U+k)$ is hit, and is the most intricate; the difficulties are to avoid discretization of BM and the fact that the expected length of such a segment is infinite, whereas it is desirable that the algorithm has a finite mean number of steps. This is achieved as follows. Let $\tau_1 = \tau_1(l_1)$ be the length of the Brownian segment [here $l_1 = (1-U_1)R(U+k)$]. We do not generate τ_1 directly, but start instead by generating $\tau_2(l_2)$, where $l_2 = R(U+k) - w$. Next let $t_1 = \lfloor \tau_2 \rfloor + 1 - \tau_2$, $l_3 = w - U_1 R(U+k)$. If $\tau_3(l_3) < t_1$, the Brownian segment is finished since then $U_1 R(U+k)$ is hit at time $\tau_1 = \tau_2 + \tau_3 \in [\lfloor \tau_2 \rfloor, \lfloor \tau_2 \rfloor + 1]$ (w is unchanged). Otherwise, we generate $B^*(t_1, l_3)$, let $R(\lfloor \tau_2 \rfloor + 1) = w - B^*(t_1, l_3)$ and replace w by $w - B^*(t_1, l_3)$. Now continue in this way by incrementing time in units $1 = t_2 = t_3 = \cdots$ until the BM either attempts to go below $U_1 R(U+k)$ (in which case the Brownian segment is finished) or goes above w, in which case we reset it to w as in the initial step of the Brownian segment. A geometric trials argument shows that the procedure will terminate in a finite expected number of steps.

Step 3 is the repetition of steps 1 and 2 until eventually the simulated future minimum in step 1 exceeds w [the probability of this in each step is at least $\mathbb{P}(CU > 1) > 0$ so that the necessary number of repetitions is geometrically bounded].

Acknowledgments. We are grateful to Hermann Brunner and Erik Jørgensen for helpful hints and references.

REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. (1972). *Handbook of Mathematical Functions*, 10th ed. Dover, New York.

[2] ASMUSSEN, S. (1987). *Applied Probability and Queues*. Wiley, New York.

[3] ASMUSSEN, S. (1989). Validating the heavy-traffic performance of regenerative simulation. *Stochastic Models* **5** 617–628.

[4] ASMUSSEN, S. (1992). Queueing simulation in heavy traffic. *Math. Oper. Res.* **17** 84–111.

[5] BARLOW, R. E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Academic Press, New York.

[6] BIANE, PH. and YOR, M. (1988). Quelques précisions sur le méandre brownien. *Bull. Sci. Math.* **112** 101–109.

[7] BRATLEY, P., FOX, B. L. and SCHRAGE, L. E. (1987). *A Guide to Simulation*. Springer, New York.

[8] CALVIN, J. M. (1992). Average performance of passive algorithms for global optimization. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology.

[9] CHITASHVILI, R. J. and LAZRIEVA, N. L. (1981). Strong solutions of stochastic differential equations with boundary conditions. *Stochastics* **5** 225–309.

[10] CSÁKI, E., FÖLDES, A. and SALMINEN, P. (1987). On the joint distribution of the maximum and its location for a linear diffusion. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 179–194.

[11] DAI, J. G. and HARRISON, J. M. (1991). Steady-state analysis of RBM in a rectangle: numerical methods and a queueing application. *Ann. Appl. Probab.* **1** 1–15.

[12] DENISOV, I. V. (1984). A random walk and a Wiener process near a maximum. *Theory Probab. Appl.* **28** 821–824.

[13] DUFFIE, J. D. and GLYNN, P. W. (1995). Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.* **5** 897–905.

[14] FITZSIMMONS, P. (1985). Excursions above the minimum for diffusions. Unpublished manuscript.

[15] FIZSIMMONS, P., PITMAN, J. and YOR, M. (1992). Markovian bridges: construction, Palm interpretation and splicing. In *Seminar on Stochastic Processes* (E. Çinlar, K. L. Chung and M. J. Sharpe, eds.) 101–134. Birkhäuser, Boston.

[16] IMHOF, J. P. (1984). Density factorizations for Brownian motion, meander and the three-dimensional Bessel process, and applications. *J. Appl. Probab.* **21** 500–510.

[17] KARLIN, S. and TAYLOR, H. M. (1975). *A Second Course in Stochastic Processes*. Academic Press, New York.

[18] KINKLADZE, G. N. (1983). Thesis, Tbilisi.

[19] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer, New York.

[20] KNUTH, D. E. (1984). An algorithm for Brownian zeroes. *Computing* **33** 89–94.

[21] KURTZ, T. G. and PROTTER, P. (1991). Wong–Zakai corrections, random evolutions, and simulation schemes for sde's. Unpublished manuscript.

[22] LÉPINGLE, D. (1992). Un schéma d'Euler pour équations différentielles réfléchies. *C. R. Acad. Sci. Paris Sér. I Math.* **316** 601–605.

[23] LÉVY, P. (1939). Sur certains processus stochastiques homogènes. *Compositio Math.* **7** 283–339.

[24] LIU, Y. (1993). Numerical approaches to reflected diffusion processes. Unpublished manuscript.

[25] LYNESS, J. N. and NINHAM, R. W. (1967). Numerical quadrature and asymptotic expansions. *Math. Comp.* **21** 162–178.

[26] MIL'STEIN, G. M. (1974). Approximate integration of stochastic differential equations. *Theory Probab. Appl.* **19** 557–562.

[27] PARDOUX, E. and TALAY, D. (1985). Discretization and simulation of stochastic differential equations. *Acta Appl. Math.* **19** 557–562.

[28] PETTERSON, R. (1993). Approximations for stochastic differential equations with reflecting convex boundaries. Unpublished manuscript.

[29] REIMAN, M. (1984). Open queueing networks in heavy traffic. *Math. Oper. Res.* **9** 441–458.

[30] REVUZ, D. and YOR, M. (1991). *Continuous Martingales and Brownian Motion.* Springer, New York.

[31] ROOTZÉN, H. (1980). Limit distributions for the error in approximation of stochastic integrals. *Ann. Probab.* **8** 241–251.

[32] RUMELIN, W. (1982). Numerical treatment of stochastic differential equations. *SIAM J. Numer. Anal.* **19** 604–613.

[33] SALMINEN, P. (1984). Brownian excursions, revisited. In *Seminar on Stochastic Processes 1983* (E. Çinlar, K. L. Chung and R. K. Getoor, eds.). Birkhäuser, Boston.

[34] SHEPP, L. (1979). The joint density of the maximum and its location for a Wiener process with drift. *J. Appl. Probab.* **16** 423–427.

[35] SLOMINSKI, L. (1994). On approximation of solutions of multidimensional SDE's with reflecting boundary conditions. *Stochastic Process. Appl.* **50** 423–427.

[36] TALAY, D. and TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* **8** 483–509.

[37] WHITT, W. (1985). Simulating Brownian motion. Unpublished manuscript.

[38] WILLIAMS, D. (1974). Path decomposition and continuity of local time for one dimensional diffusions I. *Proc. London Math. Soc.* (3) **28** 738–768.

[39] WILLIAMS, D. (1979). *Diffusions, Markov Processes and Martingales.* Wiley, New York.

S. ASMUSSEN
INSTITUTE OF ELECTRONIC SYSTEMS
AALBORG UNIVERSITY
FR. BAJERSV. 7
DK-9220 AALBORG Ø
DENMARK

P. GLYNN
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022

J. PITMAN
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720