J. Appl. Prob. **34**, 632–642 (1997)
Printed in Israel

AVERAGE CASE BEHAVIOR OF RANDOM SEARCH FOR THE MAXIMUM

JAMES M. CALVIN,* *New Jersey Institute of Technology*
PETER W. GLYNN,** *Stanford University*

Abstract

This paper is a study of the error in approximating the global maximum of a Brownian motion on the unit interval by observing the value at randomly chosen points. One point of view is to look at the error from random sampling for a given fixed Brownian sample path; another is to look at the error with both the path and observations random. In the first case we show that for almost all Brownian paths the error, normalized by multiplying by the square root of the number of observations, does not converge in distribution, while in the second case the normalized error does converge in distribution. We derive the limiting distribution of the normalized error averaged over all paths.

BROWNIAN MOTION; GLOBAL OPTIMIZATION; AVERAGE COMPLEXITY

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J65
SECONDARY 68Q25

1. Introduction

Let f be a function defined on the unit interval. An obvious approach to searching for the global maximum of f is to randomly sample the unit interval, and to approximate the global maximum of f by the maximum function evaluation observed. Specifically, let $(t_i : i > 0)$ be a sequence of i.i.d. uniform random variables. We approximate $M = \max(f(x) : 0 \leqq x \leqq 1)$ by $M_n = \max(f(t_i) : 1 \leqq i \leqq n)$, where the error in our approximation is denoted by $\Delta_n = M - M_n$. We note that this type of algorithm is non-adaptive in the sense that the algorithm's current search pattern is not modified on the basis of previously observed values. Nevertheless, this algorithm bears some similarity to the random re-start algorithms that are frequently used in global optimization. In addition, such algorithms appear particularly natural when dealing with extremely non-smooth surfaces (such as might be the case in trying to determine the maximal deviation from prescribed tolerance of some finely machined surface).

Received 3 January 1995; revision received 20 March 1996.
* Postal address: Department of Computer and Information Science, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-1982, USA.
** Postal address: Department of Operations Research, Stanford University, Stanford, CA 94305-4022, USA.
The research of JMC is supported by the National Science Foundation under grant DDM-9010770.
The research of PWG is supported by the National Science Foundation under grant DDM-9101580 and the Army Research Office under Contract No. DAAL03-91-G-0319.

Several approaches have been used to study the error resulting from approximating the global maximum of a function with randomly selected observations. An approach based on extreme value theory is given by de Haan (1981); see also Zhigljavsky (1991). Under certain conditions on the function f there exists a deterministic positive sequence a_n such that $a_n \Delta_n$ converges in distribution to a random variable with a Weibull distribution.

In this paper we study the average error in approximating the global maximum by uniform sampling in the sense that we take an average with respect to a probability on objective functions. The probability we consider is the Wiener measure on the continuous functions on the unit interval. That is, let $(B(t) : 0 \leqq t \leqq 1)$ be a Brownian motion. We consider the problem of approximating $M = \max_{0 \leqq t \leqq 1} B(t)$ by the maximum of B at a randomly selected set of points. Our goal is to analyze the asymptotic behavior of the normalized error random variable $\sqrt{n}\,\Delta_n = \sqrt{n}(M - \max_{1 \leqq i \leqq n} B(t_i))$ as $n \to \infty$. Note that there are two sources of randomness in the error random variable: the Brownian path and the random observation points. When the path is fixed, we are in the setting of the previous paragraph, but the extreme value theory does not apply in this case. More precisely, we will show that, for almost all Brownian paths, the normalized error under random sampling does not converge in distribution. The situation is different when we treat the path, as well as the observations, as random. This might be thought of as looking at the average error over many independent searches of different objective functions. In this case $\sqrt{n}\,\Delta_n$ converges in distribution, and we derive the limiting distribution.

This topic is a particular instance of the problem of analyzing the average error for methods that non-adaptively approximate the global maximum. This problem has been previously studied for the Brownian motion case. Ritter (1990) showed that, for the best non-adaptive method, the average error decreases at rate $n^{-1/2}$ in the number of observations n. Calvin (1995) compared the average error for deterministic uniformly spaced observations with the expected error with random uniform sampling. Al-Mharmah and Calvin (1996) show that the optimal sampling density for minimizing the error for Brownian motion is the Beta distribution with parameters (2/3, 2/3). Asmussen *et al.* (1994) describe the limiting distribution of $\sqrt{n}\,\Delta_n$ for the deterministic grid with $t_i = i/n$.

In Section 2 we establish basic results and derive the limiting distribution of the normalized error random variable (averaged over all paths). In Section 3 we show that for almost all Brownian paths, the normalized error under random sampling does not converge in distribution.

2. Limiting error distribution

In this section we establish a basic result (Theorem 1) that will be used throughout the paper, and use the result to determine the limiting distribution of the normalized error random variable.

First we establish some notation. Let $(B(t) : 0 \leqq t \leqq 1)$ be a standard Brownian motion defined on a probability space $(\Omega_1, \mathscr{F}_1, P_1)$, and let $\{t_1, t_2, \cdots\}$ be a sequence of independent, uniform (0, 1) random variables defined on a probability space $(\Omega_2, \mathscr{F}_2, P_2)$. Set $(\Omega, \mathscr{F}, P)=(\Omega_1 \times \Omega_2, \mathscr{F}_1 \times \mathscr{F}_2, P_1 \times P_2)$, and define

$$\Delta_n = \Delta_n(\omega_1, \omega_2) = M(\omega_1) - \max_{1 \leqq i \leqq n} B(\omega_1, t_i(\omega_2)).$$

We will at different times consider Δ_n as a random variable defined on Ω or on Ω_2 with ω_1 fixed.

Conditioning on the location of the global maximizer and the value of the Brownian motion at time 1, the segments of the process before and after the global maximizer are diffusion bridges that can be described in terms of Bessel bridges. Analyzing the error random variable under uniform random sampling then reduces to a study of the occupation measures of Bessel bridges. Theorem 1 below is sufficient for the analysis of the limiting distribution of the normalized error random variable.

For $d \geqq 1$, the d-dimensional Bessel process is the diffusion that is identical in law to the modulus of a d-dimensional Brownian motion. Let $\tilde{\Omega}$ be the set of continuous functions $\omega : \mathscr{R}^+ \to \mathscr{R}$ and $\tilde{\Omega}_t$ the continuous functions $\omega : [0, t] \to \mathscr{R}$. Define the coordinate mappings $X_t : \tilde{\Omega} \to \mathscr{R}$ by $X_t(\omega) = \omega(t)$. For each $t \geqq 0$ let $\tilde{\mathscr{F}}_t = \sigma(X(s); 0 \leqq s \leqq t)$, and $\tilde{\mathscr{F}} = \vee_{t \geqq 0} \tilde{\mathscr{F}}_t$. Let R be the law on $(\tilde{\Omega}, \tilde{\mathscr{F}})$ under which the coordinate process is the three-dimensional Bessel process, and let $R^T_{0,v}$ be the law on $(\tilde{\Omega}_T, \tilde{\mathscr{F}}_T)$ under which the coordinate process is a Bessel bridge from 0 to $v > 0$ in time T. Let P be the law for a Brownian motion, and let P^{abs}_x be the law for a Brownian motion starting at x and absorbed at 0.

Theorem 1. *Let $T > 0$ and $v > 0$ be fixed. For any choice of $k \geqq 1$ and*

$$0 < y_1 < y_2 < \cdots < y_k, \quad 0 < u_1 < u_2 < \cdots < u_k,$$

(1)
$$\begin{aligned} &\lim_{n \to \infty} R^T_{0,v}\left(n \int_{t=0}^{T} I_{\{X(t) \leqq y_1/\sqrt{n}\}} dt \leqq u_1, \cdots, n \int_{t=0}^{T} I_{\{X(t) \leqq y_k/\sqrt{n}\}} dt \leqq u_k \right) \\ &= R\left(\int_{t=0}^{\infty} I_{\{X(t) \leqq y_1\}} dt \leqq u_1, \cdots, \int_{t=0}^{\infty} I_{\{X(t) \leqq y_k\}} dt \leqq u_k \right). \end{aligned}$$

Proof. The Bessel process has the Brownian scaling property: if X is a Bessel process starting from x then, for any $c > 0$, the process $\{cX(t/c^2) : t \geqq 0\}$ is a Bessel process starting from cx (Proposition 1.10, p. 416 of Revuz and Yor (1991)). Similarly, if X is a Bessel bridge from a at time 0 to b at time T, then $\{cX(t/c^2) : 0 \leqq t \leqq c^2 T\}$ is a Bessel bridge starting from ca at time 0 and arriving at cb at time $c^2 T$. Therefore,

(2)
$$\begin{aligned} &R^T_{0,v}\left(n \int_{t=0}^{T} I_{\{X(t) \leqq y_1/\sqrt{n}\}} dt \leqq u_1, \cdots, n \int_{t=0}^{T} I_{\{X(t) \leqq y_k/\sqrt{n}\}} dt \leqq u_k \right) \\ &= R^{nT}_{0,v\sqrt{n}}\left(\int_{t=0}^{nT} I_{\{X(t) \leqq y_1\}} dt \leqq u_1, \cdots, \int_{t=0}^{nT} I_{\{X(t) \leqq y_k\}} dt \leqq u_k \right). \end{aligned}$$

Let T_x be the first hitting time to x, and L_x the last time state x is visited:

$$T_x = \inf\{t : X(t) = x\}, \quad L_x = \sup\{t : X(t) = x\}.$$

A Bessel bridge from level 0 to b in time t is the time reversal of a Brownian motion starting from b and absorbed at 0, conditional on absorption at time t. Therefore, for n sufficiently large that $y_k < v\sqrt{n}$, the last probability in (2) is equal to

$$P^{\text{abs}}_{v\sqrt{n}}\left(\int_{t=0}^{nT} I_{\{X(t)\leqq y_1\}}dt \leqq u_1, \cdots, \int_{t=0}^{nT} I_{\{X(t)\leqq y_k\}}dt \leqq u_k \mid T_0 = nT\right)$$

(3)
$$= \int_{s=0}^{nT} P^{\text{abs}}_{v\sqrt{n}}\left(\int_{t=0}^{nT} I_{\{X(t)\leqq y_i\}}dt \leqq u_i;\ 1 \leqq i \leqq k \mid T_0 = nT,\ T_{y_k} \in ds\right) P^{\text{abs}}_{\sqrt{n}v}(T_{y_k} \in ds \mid T_0 = nT)$$

$$= \int_{\sigma=0}^{nT} P^{\text{abs}}_{y_k}\left(\int_{t=0}^{\sigma} I_{\{X(t)\leqq y_i\}}dt \leqq u_i;\ 1 \leqq i \leqq k \mid T_0 = \sigma\right) P^{\text{abs}}_{\sqrt{n}v}(T_{y_k} \in nT - d\sigma \mid T_0 = nT)$$

$$= \int_{\sigma=0}^{\infty} P^{\text{abs}}_{y_k}\left(\int_{t=0}^{\sigma} I_{\{X(t)\leqq y_i\}}dt \leqq u_i;\ 1 \leqq i \leqq k \mid T_0 = \sigma\right) P^{\text{abs}}_{\sqrt{n}v}(T_{y_k} \in nT - d\sigma \mid T_0 = nT).$$

Now

(4)
$$R\left(\int_{t=0}^{\infty} I_{\{X(t)\leqq y_1\}}dt \leqq u_1, \cdots, \int_{t=0}^{\infty} I_{\{X(t)\leqq y_k\}}dt \leqq u_k\right)$$

$$= \int_{\sigma=0}^{\infty} R\left(\int_{t=0}^{\sigma} I_{\{X(t)\leqq y_1\}}dt \leqq u_1, \cdots, \int_{t=0}^{\sigma} I_{\{X(t)\leqq y_k\}}dt \leqq u_k \mid L_{y_k} = \sigma\right) P(L_{y_k} \in d\sigma),$$

because of the transience of the Bessel process. Finally, use the fact that the integrands in (3) and (4) are the same (since $P^{\text{abs}}_y(\cdot \mid T_0 = t) = R(\rho \circ \cdot \mid L_y = t)$ on $\tilde{\Omega}_t$, where ρ is the time reversal map sending $\omega \in \tilde{\Omega}_t$ to $(\omega(t-s) : 0 \leqq s \leqq t)$; see Salminen (1983), Remark 8), and the integrator of the first integral converges weakly to the integrator of the second one as $n \to \infty$. To establish the last fact, note that

$$P^{\text{abs}}_{\sqrt{n}v}(T_y \in nT - d\sigma \mid T_0 = nT)$$

$$= \frac{P_{\sqrt{n}v}(T_y \in nT - d\sigma,\ T_0 = nT)}{P_{\sqrt{n}v}(T_0 = nT)}$$

$$= \frac{P_{\sqrt{n}v}(T_y \in nT - d\sigma) P_y(T_0 = \sigma)}{P_{\sqrt{n}v}(T_0 = nT)}$$

$$= \frac{\dfrac{\sqrt{n}v - y}{nT - \sigma} \dfrac{1}{\sqrt{2\pi(nT-\sigma)}} \exp\left(-\dfrac{(\sqrt{n}v - y)^2}{2(nT-\sigma)}\right) \dfrac{y}{\sigma} \dfrac{1}{\sqrt{2\pi\sigma}} \exp\left(-\dfrac{y^2}{2\sigma}\right)}{\dfrac{\sqrt{n}v}{nT} \dfrac{1}{\sqrt{2\pi nT}} \exp\left(-\dfrac{(\sqrt{n}v)^2}{2nT}\right)}$$

$$= \frac{y(v-y/\sqrt{n})T^{3/2}}{(T-\sigma/\sqrt{n})\sigma v\sqrt{2\pi}\sqrt{\sigma}\sqrt{T-\sigma/n}} \exp\left(-\frac{v^2-2vy/\sqrt{n}+y^2/n}{2(T-\sigma/n)} - \frac{y^2}{2\sigma} + \frac{v^2}{2T}\right)$$

$$\to \frac{y}{\sqrt{2\pi\sigma^3}} \exp\left(-\frac{y^2}{2\sigma}\right)$$

as $n \to \infty$, which is the density of L_y under R (see Revuz and Yor (1991), Corollary 4.6, p. 294).

Specializing to $k=1$ we obtain the Laplace transform of the limiting distribution of the normalized occupation time.

Proposition 2. *Fix v, $T>0$, and for $z \geqq 0$ let*

$$A(z) = \int_{t=0}^{T} I_{\{X(t) \leqq z\}} dt$$

be the occupation time of $[0, z]$ *up to time T. Then, under $R^T_{0,v}$, $nA(n^{-1/2}y)$ converges in distribution to $\Psi(y)$, where*

$$E \exp(-\lambda\Psi(y)) = \frac{1}{\cosh(y\sqrt{2\lambda})}. \tag{5}$$

Notice that the limit distribution does not depend on T or v.

Proof. The result follows from Theorem 1, along with the fact that the local time process of the three-dimensional Bessel process is the square of a two-dimensional Bessel process. This fact is discussed, for example, in Williams (1974). (We take local time to be twice that in Williams, so that it is an occupation density.) Therefore, under R,

$$\int_{t=0}^{\infty} I_{\{X(t) \leqq y\}} dy \tag{6}$$

has the same distribution as

$$\int_{z=0}^{y} V(z) dz, \tag{7}$$

where $\{V(t): t \geqq 0\}$ is the square of a two-dimensional Bessel process. From Corollary 1.8, p. 414 of Revuz and Yor (1991),

$$E \exp\left(-\lambda \int_{z=0}^{y} V(z) dz\right) = \frac{1}{\cosh(y\sqrt{2\lambda})}, \tag{8}$$

which establishes (5).

Turning our attention back to the Brownian motion B on $[0, 1]$, let $T = \inf\{t : B(s) \leqq B(t)\ \forall s \in [0, 1]\}$ denote the (first) location where the global maximum is attained. For $0 < T < 1$, define the pre-T and post-T deviation segments

$$\begin{aligned} X_1(s) &= B(T) - B(T-s), \qquad 0 \leqq s \leqq T, \\ X_2(s) &= B(T) - B(T+s), \qquad 0 \leqq s \leqq 1-T. \end{aligned} \tag{9}$$

Given T, $B(1)$, and $B(T)$, X_1 and X_2 are independent Bessel bridges over $[0, T]$ and $[0, 1-T]$, respectively (see Asmussen *et al.* 1994).

The next theorem gives the limiting distribution of the normalized error. We emphasize that in this theorem the error is averaged over all Brownian sample paths, i.e. we view Δ_n as a random variable on Ω, and not Ω_2 with ω_1 fixed.

Theorem 3. *Conditional on $T = t$, $B(T) = m$ and $B(1) = u$, for $y > 0$,*

$$P(\sqrt{n}\Delta_n \leqq y) \to \tanh^2(\sqrt{2}y) \tag{10}$$

as $n \to \infty$.

Proof. Let

$$A_1(z) = \int_{t=0}^{T} I_{\{X_1(t) \leqq z\}} dt, \quad A_2(z) = \int_{t=0}^{1-T} I_{\{X_2(t) \leqq z\}} dt. \tag{11}$$

The time spent by the Brownian motion within z of its global maximum is then the sum of $A_1(z)$ and $A_2(z)$, which are (conditionally) independent. Therefore, for $y > 0$,

$$P(\sqrt{n}\Delta_n > y) = P(\Delta_n > y/\sqrt{n}) = E\left(1 - \frac{1}{n} n[A_1(y/\sqrt{n}) + A_2(y/\sqrt{n})]\right)^n. \tag{12}$$

By Proposition 2,

$$n[A_1(y/\sqrt{n}) + A_2(y/\sqrt{n})] \xrightarrow{\mathscr{D}} \Psi_1 + \Psi_2 \tag{13}$$

say, where the Ψ_i are independent with the Laplace transform given by (5). By the Skorokhod representation theorem, there exist random variables Z_n and Z defined on some probability space such that

$$Z_n \stackrel{\mathscr{D}}{=} n[A_1(y/\sqrt{n}) + A_2(y/\sqrt{n})], \qquad Z \stackrel{\mathscr{D}}{=} \Psi_1 + \Psi_2, \tag{14}$$

and $Z_n \to Z$ almost surely. Therefore,

$$\begin{aligned} P(\sqrt{n}\Delta_n > y) &= E\left(1 - \frac{1}{n} n[A_1(y/\sqrt{n}) + A_2(y/\sqrt{n})]\right)^n \\ &= E(1 - Z_n/n)^n \\ &\to E\mathrm{e}^{-Z} \end{aligned}$$

by the dominated convergence theorem, since $0 \leqq Z_n \leqq n$ and so $0 \leqq (1 - Z_n/n)^n \leqq 1$. By Proposition 2,

(15) $$Ee^{-Z} = Ee^{-\Psi_1 - \Psi_2} = \frac{1}{\cosh^2(\sqrt{2}y)}.$$

Therefore,

$$P(\sqrt{n}\Delta_n \leqq y) \to \tanh^2(\sqrt{2}y).$$

Note that the limiting distribution of the normalized error is independent of the location of the maximum T, the maximum value m, and the value of B at 1.

Let Δ be a random variable with the limiting normalized error distribution,

$$P(\Delta \leqq y) = \tanh^2(\sqrt{2}y).$$

The mean of the limiting error distribution is

(16) $$E(\Delta) = \int_{y=0}^{\infty} (1 - F(y))dy = \int_{y=0}^{\infty} \frac{dy}{\cosh^2(\sqrt{2}y)} = \frac{1}{\sqrt{2}}.$$

3. Asymptotic behavior of error for fixed sample path

Next we consider the question of convergence in distribution of $\sqrt{n}\Delta_n(\omega_1, \cdot\,)$, i.e. we fix a Brownian path and the only randomness is in the observations. With the path fixed, Δ_n is the minimum of i.i.d. non-negative random variables $M(\omega_1) - B(t_i)$. To put things in the usual notational setting of extreme value theory, let us denote the cumulative distribution function of these i.i.d. random variables by $F(\omega_1, \cdot\,)$. In terms of our previous notation established in (11),

$$F(\omega_1, z) = P_2(M(\omega_1) - B(t_i(\omega_2)) \leqq z) = A_1(\omega_1, z) + A_2(\omega_1, z),$$

where we have added the argument ω_1 to the A_i to emphasize the dependence of the time spent within z of the maximum on the path ω_1. By the Fisher–Tippet theorem of extreme value theory, if $\sqrt{n}\Delta_n(\omega_1, \cdot\,)$ converges in distribution, the limit must be one of the three extreme value distributions. Two of these have support that includes the negative axis, and so the only possible limit law is the Weibull. A necessary and sufficient condition for convergence of $\sqrt{n}\Delta_n(\omega_1)$ to a Weibull distribution is that

(17) $$\lim_{t \downarrow 0} \frac{F(\omega_1, ty)}{F(\omega_1, t)} = y^{\alpha(\omega_1)}$$

for all $y > 0$, for some $\alpha(\omega_1) > 0$ (see Leadbetter *et al.* 1983). Of course, $\alpha(\omega_1)$ could be different for different paths ω_1. The next theorem establishes that this limit does not exist for P_1-a.e. ω_1.

Theorem 4. *For almost all Brownian paths the error under uniform sampling does not converge in distribution, i.e.*

(18) $$P_1(\omega_1 : \sqrt{n}\Delta_n(\omega_1, \cdot\,) \text{ converges in distribution}) = 0.$$

Proof. We use the notation of (9) and (11). For the convergence in distribution to occur at ω_1, it is necessary that

$$(19) \qquad \lim_{t \downarrow 0} \frac{\int_{s=0}^{T} I_{\{X_1(s) \leqq ty\}} ds + \int_{s=0}^{1-T} I_{\{X_2(s) \leqq ty\}} ds}{\int_{s=0}^{T} I_{\{X_1(s) \leqq t\}} ds + \int_{s=0}^{1-T} I_{\{X_2(s) \leqq t\}} ds} = y^{\alpha(\omega_1)}$$

for all $y > 0$ and some $\alpha(\omega_1) > 0$. We will show that the limit in (19) does not exist for P_1-a.e. ω_1. Since, as previously noted, conditional on T, $B(T)$, and $B(1)$, X_1 and X_2 are independent Bessel bridges, it suffices to show that the limit in (19) does not exist almost surely under the probability under which X_1 and X_2 are independent Bessel bridges.

To be precise, let U denote the space of continuous $\mathscr{R}^2_+$-valued functions $u(u_1, u_2)$. Endow U with the σ-fields

$$\mathscr{G}^0_t = \sigma\{(u_1(s), u_2(s)) : 0 \leqq s \leqq t\}, \qquad 0 \leqq t \leqq 1.$$

Define a mapping $Y = (Y_1, Y_2) : (\Omega_1, \mathscr{F}_1) \to (U, \mathscr{G}^0_1)$ by

$$Y_1(\omega_1, t) = T(\omega_1)^{-1/2} X_1(\omega_1, tT(\omega_1)), \qquad 0 \leqq t \leqq 1,$$

$$Y_2(\omega_1, t) = (1 - T(\omega_1))^{-1/2} X_2(\omega_1, t(1 - T(\omega_1))), \qquad 0 \leqq t \leqq 1.$$

Let $Q^0_{z_1,z_2}$ be the probability on $\mathscr{G}^0_1$ such that u_1 and u_2 are independent Bessel bridges with $u_i(0) = 0$ and $u_i(1) = z_i$ for $i = 1, 2$. Let $\{\mathscr{G}_t, 0 \leqq t \leqq 1\}$ denote the augmentation of the filtration $\{\mathscr{G}^0_t, 0 \leqq t \leqq 1\}$, i.e. $\mathscr{G}_t = \sigma\{\mathscr{G}^0_t \cup \mathscr{N}\}$, where $\mathscr{N} = \{N \subset U : N \subset M \in \mathscr{G}^0_1, Q^0_{z_1,z_2}(M) = 0\}$. Denote by Q_{z_1,z_2} the extension of $Q^0_{z_1,z_2}$ to $\mathscr{G}_1$.

We next show that the augmented filtration is right-continuous, using a modification of the proof that the augmented filtration for a Feller process is right continuous; see Revuz and Yor (1991), p. 87, Prop. 2.10. Since $\mathscr{G}_t$ and $\mathscr{G}_{t+}$ are Q_{z_1,z_2} complete, it is enough to establish that, for $t \in [0, 1)$ and positive $Z \in \mathscr{G}^0_1$,

$$(20) \qquad E(Z \mid \mathscr{G}_t) = E(Z \mid \mathscr{G}_{t+}) \qquad Q_{z_1,z_2} \text{ a.s.}$$

By the monotone class theorem, it is enough to prove (20) for

$$Z = \prod_{i=1}^{n} f_i(u(t_i)), \qquad f_i \in C_0, \quad 0 \leqq t_1 < t_2 < \cdots < t_n < 1,$$

where C_0 denotes the space of continuous functions vanishing at infinity (we can take $t_n < 1$ since $u(1) = z$ is fixed).

Let $t \in [0, 1)$ and k be such that $t_{k-1} \leqq t < t_k$. For h sufficiently small (say $h \leqq (t_k - t)/2$), since

$$E(Z \mid \mathscr{G}_t) = E(Z \mid \mathscr{G}^0_t) \qquad Q_{z_1,z_2} \text{ a.s.}$$

for every t,

$$E(Z \mid \mathscr{G}_{t+h}) = E(Z \mid \mathscr{G}^0_{t+h}) = \prod_{i=1}^{k-1} f_i(u(t_i)) g(h, u(t+h)), \qquad Q_{z_1,z_2} \text{ a.s.},$$

where

$$g(h,u)=\int_{\mathbb{R}_+^2} dx_k p_{t_k-t-h}(u,x_k)f_k(x_k)\int_{\mathbb{R}_+^2} dx_{k+1}p_{t_{k+1}-t_k}(x_k,x_{k+1})f_{k+1}(x_{k+1})$$

$$\cdots\int_{\mathbb{R}_+^2} dx_n p_{t_n-t_{n-1}}(x_{n-1},x_n)\frac{p_{1-t_n}(x_n,z)}{p_{1-t-h}(u,z)}f_n(x_n)$$

$$=\int_{\mathbb{R}_+^2}\cdots\int_{\mathbb{R}_+^2}\Psi_h(u;x_k,x_{k+1},\cdots,x_n)f_k(x_k)\cdots f_n(x_n)dx_k\cdots dx_n$$

and

$$\Psi_h(u,x_k,\cdots,x_n)=\frac{p_{t_k-t-h}(u,x_k)p_{t_{k+1}-t_k}(x_k,x_{k+1})\cdots p_{t_n-t_{n-1}}(x_{n-1},x_n)p_{1-t_n}(x_n,z)}{p_{1-t-h}(u,z)}$$

is the density of $(X(t_k),\cdots,X(t_n))$ when X is a Bessel bridge from u at time $t+h$ to z at time 1. Here p is the transition density of the bivariate three-dimensional Bessel process

$$p_t(x,y)=p_t((x_1,x_2),(y_1,y_2))=\hat{p}_t(x_1,y_1)\hat{p}_t(x_2,y_2),$$

where $\hat{p}$ is the transition density of the three-dimensional Bessel process

$$\hat{p}_t(v,w)=\begin{cases}\dfrac{w}{v}\dfrac{1}{\sqrt{2\pi t}}\left\{\exp\left(-\dfrac{(v-w)^2}{2t}\right)-\exp\left(-\dfrac{(v+w)^2}{2t}\right)\right\} & \text{if } v>0,\\[2ex] \left(\dfrac{2}{\pi t^3}\right)^{1/2}w^2\exp\left(-\dfrac{w^2}{2t}\right) & \text{if } v=0,\end{cases}$$

for $w>0$. Observe that the density $\Psi_h(u;x_k,\cdots,x_n)$ is jointly continuous in $0\leqq h\leqq (t_k-t)/2$, $u,x_k,\cdots,x_n\in\mathbb{R}_+^2$, and so by Scheffé's theorem (since $f_k(\cdot)f_{k+1}(\cdot)\cdots f_n(\cdot)$ is in C_0 and therefore bounded), $g(h,u)$ is jointly continuous in $0\leqq h\leqq (t_k-t)/2$, $u\in\mathbb{R}_+^2$.

By path continuity, $u_{t+h}\to u_t$ as $h\downarrow 0$, so using Theorem 2.3, Ch. 2, of Revuz and Yor (1991) (since Z is bounded), we obtain

$$\begin{aligned}E(Z\mid\mathcal{G}_{t+})&=\lim_{h\downarrow 0}E(Z\mid\mathcal{G}_{t+h})\\&=\lim_{h\downarrow 0}\prod_{i=1}^{k-1}f_i(u(t_i))g(h,u(t+h))\\&=\prod_{i=1}^{k-1}f_i(u(t_i))g(0,u(t))\\&=E(Z\mid\mathcal{G}_t)\qquad Q_{z_1,z_2}\text{ a.s.,}\end{aligned}$$

which completes the proof of right continuity of the augmented filtration.

Fix $\varepsilon > 0$, and let

$$B_n(\varepsilon) = \left\{ u \in U : \min_{\varepsilon \leqq s \leqq 1} u_1(s) > 1/n, \min_{\varepsilon \leqq s \leqq 1} u_2(s) > 1/n \right\},$$

$$B_\varepsilon = \bigcup_n B_n(\varepsilon) = \left\{ u \in U : \min_{\varepsilon \leqq s \leqq 1} u_1(s) > 0, \min_{\varepsilon \leqq s \leqq 1} u_2(s) > 0 \right\}.$$

Since $\mathscr{G}_0$ is complete with respect to Q_{z_1,z_2} and $Q_{z_1,z_2}(B_\varepsilon) = 1$ (since the Bessel bridge almost surely does not return to 0), $B_\varepsilon \in \mathscr{G}_0$.

Now Q_{z_1,z_2} is a regular conditional probability for Y given $T = \tau \in (0, 1)$, $B(T) = z_1\sqrt{\tau}$, $B(1) = B(T) - z_2\sqrt{1-\tau}$ (Fitzsimmons 1985, Fitzsimmons *et al.* 1992). Therefore, to establish that

$$P_1\left(\lim_{t \downarrow 0} \frac{\int_{s=0}^{T} I_{\{X_1(s) \leqq ty\}} ds + \int_{s=0}^{1-T} I_{\{X_2(s) \leqq ty\}} ds}{\int_{s=0}^{T} I_{\{X_1(s) \leqq t\}} ds + \int_{s=0}^{1-T} I_{\{X_2(s) \leqq t\}} ds} \text{ exists} \right) = 0, \tag{21}$$

it suffices to show that

$$Q_{z_1,z_2}\left(\lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right) = 0. \tag{22}$$

Set

$$A(\tau) = \left\{ \lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right\}.$$

Since $Q_{z_1,z_2}(B_\varepsilon) = 1$, $Q_{z_1,z_2}(A_\tau) = Q_{z_1,z_2}(A_\tau \cap B_\varepsilon)$. Using (in the third equality below) the fact that, on $B_n(\varepsilon)$, $u_i(s) > 1/n$ for $\varepsilon \leqq s \leqq 1$ we obtain

$$\begin{aligned}
A_\tau \cap B_\varepsilon &= \bigcup_n (A_\tau \cap B_n(\varepsilon)) \\
&= \bigcup_n \left(\left\{ \lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{1} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{1} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right\} \cap B_n(\varepsilon) \right) \\
&= \bigcup_n \left(\left\{ \lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right\} \cap B_n(\varepsilon) \right) \\
&= \left\{ \lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right\} \cap \left(\bigcup_n B_n(\varepsilon) \right) \\
&= \left\{ \lim_{t \downarrow 0} \frac{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq ty/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq ty/\sqrt{1-\tau}\}} ds}{\tau \int_{s=0}^{\varepsilon} I_{\{u_1(s) \leqq t/\sqrt{\tau}\}} ds + (1-\tau) \int_{s=0}^{\varepsilon} I_{\{u_2(s) \leqq t/\sqrt{1-\tau}\}} ds} \text{ exists} \right\} \cap B_\varepsilon \\
&\in \mathscr{G}_\varepsilon.
\end{aligned}$$

Since this is true for any $\varepsilon > 0$, setting

$$B=\bigcap_{m\geqq 1} B_{1/m}=\{u\in U: u(t)>0,\ 0<t\leqq 1\}$$

we have that $A_\tau\cap B\in\mathscr{G}_{0+}=\mathscr{G}_0$ (since the augmented filtration is right continuous). Therefore, by Blumenthal's 0–1 law $Q_{z_1,z_2}(A_\tau\cap B)$ is 0 or 1, and so the same is true of $Q_{z_1,z_2}(A_\tau)$ since $Q_{z_1,z_2}(B)=1$. If $Q_{z_1,z_2}(A_\tau)=0$, then (22) is established. If $Q_{z_1,z_2}(A_\tau)=1$, then

$$\lim_{t\downarrow 0}\frac{\tau\int_{s=0}^{1} I_{\{u_1(s)\leqq ty/\sqrt{\tau}\}}ds+(1-\tau)\int_{s=0}^{1} I_{\{u_2(s)\leqq ty/\sqrt{1-\tau}\}}ds}{\tau\int_{s=0}^{1} I_{\{u_1(s)\leqq t/\tau\}}ds+(1-\tau)\int_{s=0}^{1} I_{\{u_2(s)\leqq t/\sqrt{1-\tau}\}}ds}=a(y,\tau)$$

Q_{z_1,z_2} a.s., where $a(y,\tau)$ is independent of u, and so the only way that the convergence in distribution can take place is if $a(y,\tau)=y^{\alpha(\tau)}$. But if this were the case, then we would have that, for almost every ω_1,

$$\lim_{t\downarrow 0}\frac{A_1(\omega_1,ty)+A_2(\omega_1,ty)}{A_1(\omega_1,t)+A_2(\omega_1,t)}=y^{\alpha(T)}.$$

This would imply that, conditional on T, $B(T)$, and $B(1)$, $\sqrt{n}\Delta_n$ (defined on Ω) must converge in distribution to a Weibull random variable, which we know not to be the case by Theorem 3. We have therefore established (22), and the proof is complete.

Acknowledgement

We are grateful to an anonymous referee for numerous corrections and suggested improvements.

References

Al-Mharmah, H. and Calvin, J. (1996) Optimal random non-adaptive algorithm for global optimization of Brownian motion. *J. Global Opt.* **8**, 81–90.

Asmussen, S., Glynn, P. and Pitman, J. (1994) Discretization error in simulation of one-dimensional reflecting Brownian motion. *Technical Report 94–1.* Stanford University.

Calvin, J. (1995) Average performance of non-adaptive algorithms for global optimization of Brownian motion. *J. Math. Anal. Appl.* **191**, 608–617.

de Haan, L. (1981) Estimation of the minimum of a function using order statistics. *J. Amer. Statist. Assoc.* **76**, 467–469.

Fitzsimmons, P. J. (1985) Excursions above the minimum for diffusions. *Unpublished manuscript.*

Fitzsimmons, P. J., Pitman, J. W. and Yor, M. (1992) Markovian bridges: construction, Palm interpretation, and splicing. In *Sem. Stochastic Processes 1992.* Birkhäuser, Boston.

Leadbetter, M., Lindgren, G. and Rootzén, H. (1983) *Extremes.* Springer, New York.

Revuz, D. and Yor, M. (1991) *Continuous Martingales and Brownian Motion.* Springer, Berlin.

Ritter, K. (1990) Approximation and optimization on the Wiener space. *J. Complexity* **6**, 337–364.

Salminen, P. (1983) Brownian excursions revisited. In *Sem. Stochastic Processes 1983.* ed. E. Çinlar, K. L. Chung and R. K. Getoor. Birkhäuser, Boston.

Williams, D. (1974) Path decomposition and continuity of local time for one-dimensional diffusions. I. *Proc. London Math. Soc.* **28**, 738–768.

Zhigljavsky, A. (1991) *Theory of Global Random Search.* Kluwer, Dordrecht.