

A unifying class of Skorokhod embeddings: connecting the Azéma–Yor and Vallois embeddings

A.M.G. COX¹ and D.G. HOBSON²

¹*Department of Mathematics, University of York, York YO10 5DD, UK.
E-mail: amgc500@york.ac.uk*

²*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK.
E-mail: dgh@maths.bath.ac.uk*

We consider the Skorokhod embedding problem in Brownian motion. In particular, we give a solution based on the local time at zero of a variably skewed Brownian motion related to the underlying Brownian motion. Special cases of the construction include the Azéma–Yor and Vallois embeddings. In turn, the construction has an interpretation in the Chacon–Walsh framework.

Keywords: Azéma–Yor embedding; Brownian motion; embedding; stopping time; variably skewed Brownian motion

1. Introduction

The Skorokhod embedding problem was first proposed, and then solved, by Skorokhod (1965), and may be described as follows. Given a Brownian motion $(B_t)_{t \geq 0}$ and a centred target law μ can we find a ‘small’ stopping time T such that B_T has distribution μ ? Skorokhod gave an explicit construction of the stopping time T in terms of independent random variables, and in doing so showed that any zero-mean probability measure may be embedded in Brownian motion using a uniformly integrable stopping time. Since the problem was posed many more solutions have been given; see, for example, Dubins (1968), Root (1969), Bass (1983) and the comprehensive survey article of Oblój (2004).

In this paper we are concerned with the connection between two of the well-known constructions due to Azéma and Yor (1979a) and Vallois (1983), and the relationship between these constructions and the construction of Chacon and Walsh (1976). The Azéma and Yor (1979a) construction defines a stopping time

$$T_{AY} = \inf\{t \geq 0 : B_t \leq \psi(S_t)\},$$

where S_t is the maximum of the Brownian motion B and ψ is an increasing function dependent on μ . This is an embedding of μ and has a nice optimality property – namely, that it maximizes the distribution of the maximum for embeddings under which $B_{t \wedge T}$ is uniformly integrable (see Azéma and Yor 1979b). The other well-known construction of interest to us

here is the embedding of Vallois (1983). This construction uses the local time L_t^B of the Brownian motion at 0. In particular, for certain functions α, β ,

$$T_V = \inf\{t \geq 0 : B_t \notin (\beta(L_t^B), \alpha(L_t^B))\}$$

is an embedding. In Vallois (1992) it is shown that for a convex function Ψ a specific choice of α and β minimizes $\mathbb{E}\Psi(L_T^B)$ among the class of all embeddings of μ . Moreover, the same choice of α and β is optimal for each Ψ . It is this specific choice of α and β , for which α is decreasing and β increasing, which interests us here.

It is not surprising, given Lévy’s result on the equivalence in law between $(S_t, S_t - B_t)$ and $(L_t^B, |B_t|)$, that there is a strong connection between the two constructions (not least when μ is symmetric); however, the aim of this paper is to demonstrate that both embeddings can be seen as examples of a much broader class of embeddings.

The embeddings we construct have an interpretation within the balayage construction of Chacon and Walsh (1976). As shown in Meilijson (1983), the Azéma–Yor construction is already known to fit in this framework. We shall show how our construction can be interpreted in a similar way. Given any centred probability measure μ , we shall define a family \mathcal{A}_μ of admissible curves γ in the Chacon–Walsh picture. In turn, for any $\gamma \in \mathcal{A}_\mu$ we define a function $G \equiv G_{\mu,\gamma} : \mathbb{R}_+ \rightarrow \mathbb{R}$ with Lipschitz constant 1. Our embeddings will be characterized by this function which will be used to define a variably skewed Brownian motion.

Skew Brownian motion (Harrison and Shepp 1981) can be thought of as a Brownian motion with excursions being positive with probability p and negative with probability $1 - p$. For a variably skewed Brownian motion (Barlow *et al.* 2000) the probability of positive excursions will vary with the local time at zero of the process. Given a Brownian motion B_t , a variably skewed Brownian motion X_t is the solution to

$$X_t = B_t - G(L_t^X) \tag{1}$$

where L_t^X is the local time of X_t at 0 and G is a function with Lipschitz constant at most 1. Our embedding will then be a stopping time of the form

$$T = \inf\{t \geq 0 : X_t \notin (b(L_t^X), a(L_t^X))\} = \inf\{t \geq 0 : B_t \notin (\beta(L_t^X), \alpha(L_t^X))\} \tag{2}$$

for suitable pairs of functions (a, b) and (α, β) . When $G \equiv 0$, X_t is just B_t (this remark can be restated in a more complicated fashion by saying that X is a skew Brownian motion with constant skewness parameter 1/2) and the stopping time is the same as Vallois’ stopping time. When $G(l) = l$ for all l , we find that X_t has only negative excursions and $L_t^X = \sup_{s \leq t} B_s$; in particular, the embedding is now of the form

$$T = \inf\{t \geq 0 : B_t \leq \psi(S_t)\}$$

for some ψ and we recover the Azéma–Yor stopping time. Further, for each μ , there are choices of curves $\gamma \in \mathcal{A}_\mu$ which lead to $G(l) = 0$ and $G(l) = l$.

The paper proceeds as follows. We shall begin by establishing the important quantities in the construction, and some relationships between them; then we show in Theorem 6 that there is a strong solution to (1), and hence that, given the Brownian motion, we may indeed construct X . The main result of this paper is Theorem 10, where we use excursion

arguments to confirm that the stopping time is an embedding, and that $B_{t \wedge T}$ is uniformly integrable. In Section 4 we discuss the connection with the Azéma–Yor and Vallois embeddings, and describe a new embedding in the general framework which maximizes $\mathbb{E}\Psi(L_T^x)$ for Ψ convex and L_T^x the local time of the Brownian motion at a non-zero level x .

2. Construction

Our aim is to show that we can construct embeddings of a target distribution μ in a Brownian motion $(B_t)_{t \geq 0}$ with $B_0 = 0$, where the embedding is determined by a curve $\gamma : [0, \zeta] \rightarrow \mathbb{R} \times \mathbb{R}_+$, $\gamma(s) = (F(s), h(s))$ and what is in some sense the local time of the Brownian motion on the function F . The parametrization of the curve is essentially arbitrary, but for simplicity we might choose s to be the arc length. The construction we provide will have as special cases the embeddings of Azéma and Yor (1979a) and Vallois (1983). In particular, we will see that the Azéma–Yor construction corresponds to the choice $F(s) = h(s)$ and the Vallois construction to the choice $F \equiv 0$.

In order to perform the construction, we need to make some assumptions on the behaviour of γ . The exact form of the conditions is dependent on the target distribution, and further notation is required to state them clearly; however, the assumptions have a straightforward graphical interpretation, which we now describe, and which is illustrated in Figure 1.

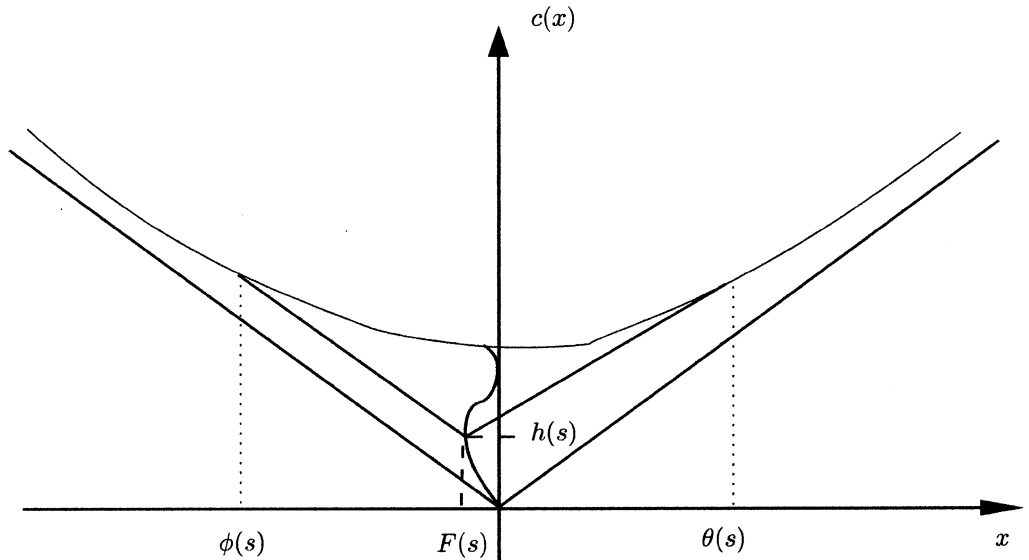


Figure 1. The ‘potential’ $c(x)$ for a distribution μ . The embedding is determined by the curve $\gamma = (F, h)$ and $\theta(s)$ and $\phi(s)$ are the x -coordinates of the points at which the tangents to c intersect γ at $\gamma(s)$.

As with many embeddings, there are strong connections with potential theory. We define the function

$$c(x) = \int_{\mathbb{R}} |y - x| \mu(dy). \tag{3}$$

This is related to a common definition of the potential of a Brownian motion (in one-dimension) via $c(x) = -u_{\mu}(x)$, where u_{μ} is the potential of μ (see Chacon 1977). We shall denote throughout the left and right derivatives of a function f by f'_-, f'_+ respectively.

Proposition 1. *Let μ be a centred target distribution with finite first moment. Then the function c as defined in (3) has the following useful properties:*

- (i) c is a positive, continuous, convex function such that $c(x) \geq |x|$.
- (ii) As $x \rightarrow \pm\infty$, $c(x) - |x| \rightarrow 0$.
- (iii) c has left and right derivatives given by

$$c'_-(x) = 2\mu((-\infty, x)) - 1, \quad c'_+(x) = 2\mu((-\infty, x]) - 1,$$

so in particular, c is differentiable Lebesgue-almost everywhere. The set $\{x : \mu(\{x\}) > 0\}$ is precisely the set where c is not differentiable.

We now discuss the set of admissible curves γ . We suppose that the functions F and h , with $F(0) = h(0) = 0$, are both absolutely continuous, and hence have derivatives Lebesgue-almost everywhere, which we suppose in addition satisfy $\text{Leb}(\{s : h'(s) = F'(s) = 0\}) = 0$. We define $\zeta = \sup\{s > 0 : h(s) \leq c(F(s))\}$, and require that the curve $\gamma(s) = (F(s), h(s))$ hits c , at least in the limit $s \rightarrow \zeta$ (although this could occur at $\pm\infty$).

The curve γ is key to defining the other primary variables. The functions θ, ϕ are defined in terms of c, F and h so that for $s \geq 0$,

$$\theta(s) = \operatorname{argmin}_{x > F(s)} \left\{ \frac{c(x) - h(s)}{x - F(s)} \right\}, \tag{4}$$

$$\phi(s) = \operatorname{argmax}_{x < F(s)} \left\{ \frac{h(s) - c(x)}{F(s) - x} \right\}. \tag{5}$$

The functions $\theta(s)$ and $\phi(s)$ have the interpretations that they are the x -coordinates of the points on c where the tangents pass through $(F(s), h(s))$. Since the argmin and argmax in (4) and (5) may not be uniquely defined (but note that from the convexity of c it follows that the candidate argmin and argmax are restricted to an interval in each case), we replace (4) and (5) with the definitions

$$\theta(s) = \sup \left\{ x \in \mathbb{R} : c(\lambda) \geq h(s) + (\lambda - F(s)) \frac{c(x) - h(s)}{x - F(s)} \quad \forall \lambda \leq x \right\}, \tag{6}$$

$$\phi(s) = \inf \left\{ x \in \mathbb{R} : c(\lambda) \geq h(s) + (\lambda - F(s)) \frac{h(s) - c(x)}{F(s) - x} \quad \forall \lambda \geq x \right\}. \tag{7}$$

This is equivalent to taking the argmin to be the supremum of the minimizing values in (4),

and the argmax to be the infimum of the maximizing values in (5). From the definitions (6) and (7) we have that θ and ϕ are left-continuous.

In order for the construction to make sense, we must restrict γ so that for any particular value of s the future of the curve will lie inside the (closed) area constructed via the tangents at $\theta(s)$ and $\phi(s)$ and the potential c . We note that under this restriction, the curve can have $h'(s) < 0$, which is why we cannot parametrize the curve by its height. Our first aim is to express this idea mathematically. Define

$$R(s) = \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)} = \inf_{x > F(s)} \frac{c(x) - h(s)}{x - F(s)}, \quad (8)$$

$$S(s) = \frac{h(s) - c(\phi(s))}{F(s) - \phi(s)} = \sup_{x < F(s)} \frac{h(s) - c(x)}{F(s) - x}, \quad (9)$$

$$\Gamma(s) = \frac{R(s) - S(s)}{2}. \quad (10)$$

The fact that the pair of definitions in (8) are equivalent follows from the fact that the lines joining $(F(s), h(s))$ to $(\theta(s), c(\theta(s)))$ and $(\phi(s), c(\phi(s)))$ are supporting tangents to c . In particular, $R(s)$ and $S(s)$ are the gradients of these tangents. As well as Γ representing half the difference in the slopes it will also have an interpretation as the probability that the embedding has not yet occurred. The conditions we require on γ are then

$$\frac{h'_+(s)}{F'_+(s)} \geq R(s), \quad F'_+(s) > 0,$$

$$\frac{h'_+(s)}{F'_+(s)} \leq S(s), \quad F'_+(s) < 0,$$

and $h'_+ > 0$ when $F'_+(s) = 0$. Since we have $R(s) > S(s)$, these conditions can be summarized as

$$h'_+(s) \geq \max(F'_+(s)S(s), F'_+(s)R(s)). \quad (11)$$

We will say that a curve $\gamma = (F, h)$ with this property is admissible, and, given μ , the family of admissible curves will be labelled \mathcal{A}_μ . By construction, for $\gamma \in \mathcal{A}_\mu$, the corresponding functions θ and ϕ will be monotonic. It follows from their representations as slopes of tangents that R and S are also monotonic, and $\Gamma(s)$ is strictly decreasing.

The construction of the stopping time can be considered via the transformations shown in Figure 2. We want to be able to make sense of the 'local time' of B_t on F , which we denote by Y_t . For the present we are not specific about the process Y except that it is an increasing process which only grows when ' B is on F ' in a sense which will become clear. The full definition of Y can be obtained from reversing the construction process we are about to describe.

The curves in Figure 2(a) (starting with the highest) represent θ , F and ϕ . In order to define the candidate stopping time we need to transform the horizontal and vertical scales of this graph. Firstly, we map the horizontal axis by H^{-1} (where H will be defined later), so that the curves in Figure 2(b) are plots of

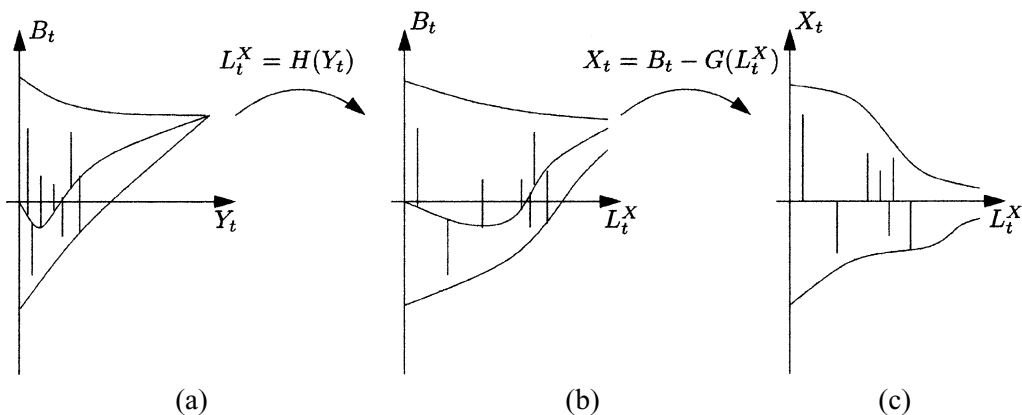


Figure 2. Three different interpretations of the stopping time. (a) shows the functions $\theta(s)$, $\phi(s)$ and $F(s)$; the horizontal axis represents the current ‘position along the curve’ or the ‘local time’ of B on F labelled by Y_t – our primary aim is to understand how to interpret this. In (b) we transform the horizontal scale using a function H to be determined, so that the horizontal axis is the local time L_t^X . In (c) we transform the vertical scale by subtracting $G = F \circ H^{-1}$, so that the vertical axis is $X_t = B_t - G(L_t^X)$. The vertical lines are the excursions of B from F , or excursions of X from zero. In each interpretation the stopping time can be thought of as the first time an excursion goes outside the region bounded by the upper and lower curves.

$$(F(H^{-1}(l)), \theta(H^{-1}(l)), \phi(H^{-1}(l))) \equiv (G(l), \alpha(l), \beta(l)).$$

Secondly, we transform the vertical scale by subtracting $G(l)$ from the vertical coordinate, so that Figure 2(c) shows $a(l) = \alpha(l) - G(l)$ and $b(l) = \beta(l) - G(l)$. Also shown in Figure 2 (by the vertical lines) are the excursions of the Brownian motion away from F . In Figures 2(a) and 2(b) these excursions are plotted in the coordinates (Y_t, B_t) and (L_t^X, B_t) respectively. In Figure 2(c) we have a process X_t making excursions from 0, and it now makes sense to define the horizontal axis to be the (symmetric) local time of X_t at 0. In particular, if the process in Figure 2(c) has local time L_t^X , then X must satisfy

$$X_t = B_t - G(L_t^X).$$

We will later show (Theorem 6) that this equation has a strong solution. The candidate embedding is then of the form

$$T = \inf\{t \geq 0 : X_t \notin (b(L_t^X), a(L_t^X))\},$$

where $a(l) = \theta(H^{-1}(l)) - F(H^{-1}(l))$ and $b(l) = \phi(H^{-1}(l)) - F(H^{-1}(l))$. This stopping rule can then be re-expressed in the form (2).

Having described the form and philosophy of the construction, it remains to determine the transformation function H , and then to show that the algorithm outlined above leads to a genuine embedding of the target law μ . The advantage of representing the construction via the diagrams in Figures 1 and 2 is to show how these representations are special cases of the Chacon and Walsh (1976) embedding.

The intuition is that the tails of the distribution are embedded first. The points on the curve, $(F(s), h(s))$, are used to define $\phi(s), \theta(s)$ and when the construction is in progress at a point represented by $(F(s), h(s))$ the mass remaining to be embedded is restricted to the interval $[\phi(s), \theta(s)]$. The potential of the remaining mass is then a rescaled version of the central part of Figure 1 in which $(F(s), h(s))$ becomes the origin and the lines joining $(F(s), h(s))$ to $(\phi(s), c(\phi(s)))$ and $(\theta(s), c(\theta(s)))$ become the lines $\pm x$ in the original Chacon–Walsh picture.

The conditions on the curve γ ensure that the function θ is non-increasing, and ϕ is non-decreasing. It also follows from the definitions that θ and ϕ are both monotonic and left-continuous. Note that a and β share these properties, but that a and b need not be monotonic.

The first results discuss the properties of the functions R and S introduced in (8) and (9). Note that (by Proposition 1(iii)) we have the inequalities

$$1 - 2\mu([\theta(s), \infty)) \leq R(s) \leq 1 - 2\mu((\theta(s), \infty)), \quad (12)$$

$$2\mu((-\infty, \phi(s))) - 1 \leq S(s) \leq 2\mu((-\infty, \phi(s))) - 1. \quad (13)$$

Lemma 2. *The functions R and S are absolutely continuous on closed subsets of $[0, \zeta)$ and*

$$R(s) = 1 - \int_0^s \frac{1}{\theta(z) - F(z)} [h'(s) - F'(z)R(z)] dz,$$

$$S(s) = -1 + \int_0^s \frac{1}{F(z) - \phi(z)} [h'(s) - F'(z)S(z)] dz.$$

Proof. The function $R(s)$ is the gradient of the tangent to c that hits γ at $(F(s), h(s))$. It is clear that R is non-increasing (by the assumptions on γ), and for every gradient there is a unique tangent which hits γ at some point; hence, R must be continuous.

Also, for δ small (we insist that $F(s - \delta) < \theta(s)$), we have from the absolute continuity of F and h that $F(s - \delta) \leq F(s) + A\delta$ and $h(s - \delta) \geq h(s) - B\delta$ for some $A, B > 0$, and by the last form of (8),

$$\begin{aligned} R(s - \delta) &\leq \frac{c(\theta(s)) - h(s) + B\delta}{\theta(s) - F(s) - A\delta} \\ &= \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)} + \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s) - A\delta} - \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)} + \frac{B\delta}{\theta(s) - F(s) - A\delta} \\ &= R(s) + \left[A \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)} + B \right] \frac{\delta}{\theta(s) - F(s) - B\delta}, \end{aligned}$$

so that R is locally Lipschitz on $[0, \zeta)$ and absolutely continuous on $[0, z]$ for $z < \zeta$. In particular, it is differentiable Lebesgue-almost everywhere, and if R' is a version of the derivative of R , then we must have

$$R(s) = 1 + \int_0^s R'(z) dz$$

for $s < \zeta$ (and hence, by continuity, for $s = \zeta$).

We can calculate the left derivative of $R(s)$, using the fact that θ is left-continuous and decreasing, to be:

$$\frac{d^-}{ds} R(s) = -\frac{h'_-(s)}{\theta(s) - F(s)} + \frac{F'_-(s)R(s)}{\theta(s) - F(s)} + \frac{\theta'_-(s)}{\theta(s) - F(s)} \left[c'_+(\theta(s)) - \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)} \right].$$

However, the second term will disappear: if there is no atom at $\theta(s)$, from (12) and Proposition 1(iii), we obtain

$$c'_+(\theta(s)) = \frac{c(\theta(s)) - h(s)}{\theta(s) - F(s)}, \tag{14}$$

while if there is an atom at $\theta(s)$, $\theta'_-(s) = 0$ (unless s is the largest s corresponding to the atom, however, then (14) still holds).

Consequently,

$$\frac{d^-}{ds} R(s) = -\frac{h'_-(s)}{\theta(s) - F(s)} + \frac{F'_-(s)R(s)}{\theta(s) - F(s)}$$

and the result holds. We can deduce the statement for $S(s)$ similarly. □

Proposition 3. *Define*

$$\begin{aligned} H(s) &= \int_0^s \frac{h'(z)}{\Gamma(z)} dz - \int_0^s F'(z) \frac{R(z) + S(z)}{R(z) - S(z)} dz \\ &= \int_0^s [2h'(z) - F'(z)(R(z) + S(z))] \frac{dz}{R(z) - S(z)}. \end{aligned} \tag{15}$$

Then H is strictly increasing and continuous on $[0, \zeta)$, with continuous inverse H^{-1} .

Proof. This follows from (11) and the fact that $R(z) \geq S(z)$, with equality only if $z = \zeta$. □

Proposition 4. *The function $G(l) = F \circ H^{-1}(l)$ is absolutely continuous with $|G'(l)| \leq 1$. Further, G is invariant under any increasing reparametrization of the curve γ .*

Proof. Both F and H^{-1} are differentiable almost everywhere, so that

$$G'(l) = \frac{F'(H^{-1}(l))}{H'(H^{-1}(l))}. \tag{16}$$

It is sufficient to show that $|G'(l)| \leq 1$, or equivalently that $-H' \leq F' \leq H'$. However,

$$H' - F' = \frac{2h'}{R - S} - \frac{2F'R}{R - S} = \frac{2}{R - S} (h' - F'R) \geq 0$$

and

$$H' + F' = \frac{2h'}{R-S} - \frac{2F'S}{R-S} = \frac{2}{R-S}(h' - F'S) \geq 0$$

by (11).

The statement that G is independent of the parametrization of the curve follows from the representation (15). \square

The following lemma follows immediately from Lemma 2 and Proposition 3.

Lemma 5. *We have the relationships*

$$\frac{H' - F'}{\theta - F} = -\frac{R'}{\Gamma}, \quad \frac{H' + F'}{F - \phi} = \frac{S'}{\Gamma}.$$

3. Verification

We have now completed the definitions of the auxiliary functions. It remains to show, firstly, that the process given by (1) is well defined, and secondly, that the stopping time in (2) embeds the law μ .

Theorem 6. *Consider the equation*

$$X_t = B_t - G(L_t^X), \tag{17}$$

where B_t is a Brownian motion, and L_t^X is the local time at 0 of the process X . There exists a strong solution to (17). Further, the solution is a variably skewed Brownian motion, with probability of positive excursions p given by

$$p(l) = \frac{1 - G'(l)}{2}$$

when the local time $L_t^X = l$.

The existence of a strong solution is shown in Barlow *et al.* (2000) under the slightly stronger condition that $|G'(l)| < 1$. They prove that the solutions of the equation are pathwise unique and that there exists a weak solution, which, by the Yamada–Watanabe theorem, implies the existence of a strong solution. (The form of the Yamada–Watanabe theorem given, for example, in Revuz and Yor 1999: IX.1.7 is stated for an ordinary stochastic differential equation but, as Barlow *et al.* note, can be extended to cover the case of interest.) The proof of Theorem 3.1 in Barlow *et al.* (2000) carries over without alteration to the case where $|G'(l)| \leq 1$, so it only remains to show the existence of a weak solution in the more general case. The statement about the probability of excursions follows on consideration of the case where $G(x) = \beta x$ for $|\beta| \leq 1$.

Consequently, Theorem 6 will follow from the following result:

Lemma 7. *There exists a weak solution to (17).*

The proof of this lemma will follow the proof of Proposition 2.2 of Barlow *et al.* (2000); the major difference is that it is necessary to use two functions n , m instead of the single function ϕ used by Barlow *et al.*

Proof. For $x \geq 0$, define $v(x)$ to be the solution of

$$\frac{dv}{dx} = \left(\frac{2}{1 + (G'(x))^2} \right)^{1/2}$$

with $v(0) = 0$, so that $1 \leq v'(x) \leq \sqrt{2}$. Define y to be the inverse of v . Then $1/\sqrt{2} \leq y'(x) \leq 1$. Define also

$$n(x) = \frac{y(x) - G(y(x))}{2}, \tag{18}$$

$$m(x) = \frac{y(x) + G(y(x))}{2}. \tag{19}$$

In particular, $0 \leq n' \leq y' \leq 1$ and $0 \leq m' \leq y' \leq 1$.

Now define

$$R(x, l) = \begin{cases} 2n'(l)x, & x \geq 0, \\ 2m'(l)x, & x < 0. \end{cases}$$

Let \tilde{B}_t, \tilde{L}_t be a Brownian motion and its local time at 0. By applying Revuz and Yor (1999: VI.4.3) separately to $R_+(\tilde{B}_t, \tilde{L}_t) = 2n'(\tilde{L}_t)(\tilde{B}_t \vee 0)$ and $R_-(\tilde{B}_t, \tilde{L}_t) = m'(\tilde{L}_t)(\tilde{B}_t \wedge 0)$ we deduce

$$\begin{aligned} R(\tilde{B}_t, \tilde{L}_t) &= 2 \int_0^t (n'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s \geq 0\}} + m'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s < 0\}}) d\tilde{B}_s + \int_0^t (n'(\tilde{L}_s) - m'(\tilde{L}_s)) d\tilde{L}_s \\ &= 2 \int_0^t (n'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s \geq 0\}} + m'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s < 0\}}) d\tilde{B}_s + n(\tilde{L}_t) - m(\tilde{L}_t), \\ |R(\tilde{B}_t, \tilde{L}_t)| &= 2 \int_0^t (n'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s \geq 0\}} - m'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s < 0\}}) d\tilde{B}_s + \int_0^t (n'(\tilde{L}_s) + m'(\tilde{L}_s)) d\tilde{L}_s \\ &= 2 \int_0^t (n'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s \geq 0\}} - m'(\tilde{L}_s) \mathbf{1}_{\{\tilde{B}_s < 0\}}) d\tilde{B}_s + n(\tilde{L}_t) + m(\tilde{L}_t). \end{aligned}$$

The stochastic integrals in both of the above equations have quadratic variation

$$A_t = 4 \int_0^t (n'(\tilde{L}_s)^2 \mathbf{1}_{\{\tilde{B}_s \geq 0\}} + m'(\tilde{L}_s)^2 \mathbf{1}_{\{\tilde{B}_s < 0\}}) ds.$$

Since n' and m' are bounded, $A_t < \infty$ almost surely, and because $n' \vee m' \geq y'/2 \geq 1/2\sqrt{2}$ we also have $A_t \uparrow \infty$ almost surely. Consequently, defining $T_t = \inf\{s : A_s > t\}$, we deduce that

$$W_t = 2 \int_0^{T_t} (n'(\tilde{L}_s)\mathbf{1}_{\{\tilde{B}_s \geq 0\}} + m'(\tilde{L}_s)\mathbf{1}_{\{\tilde{B}_s < 0\}})d\tilde{B}_s,$$

$$\tilde{W}_t = 2 \int_0^{T_t} (n'(\tilde{L}_s)\mathbf{1}_{\{\tilde{B}_s \geq 0\}} - m'(\tilde{L}_s)\mathbf{1}_{\{\tilde{B}_s < 0\}})d\tilde{B}_s$$

are Brownian motions.

Since $L_t = n(\tilde{L}_{T_t}) + m(\tilde{L}_{T_t})$ is increasing only on the set $\{t : R(\tilde{B}_{T_t}, \tilde{L}_{T_t}) = 0\}$ and $|R(\tilde{B}_{T_t}, \tilde{L}_{T_t})| = \tilde{W}_t + L_t$, L_t must be the symmetric local time at 0 of $R(\tilde{B}_{T_t}, \tilde{L}_{T_t})$.

However, by (18) and (19), we have

$$R(\tilde{B}_{T_t}, \tilde{L}_{T_t}) = W_t - G(L_t),$$

and hence a weak solution to (17). □

The important fact is that X_t is a strong solution, so that it is adapted to the filtration \mathcal{F}_t^B of the Brownian motion – in particular, given the Brownian motion, it is possible to construct X_t and its local time L_t^X , even if we do not have an explicit construction (see, however, Remark 3.5 of Barlow *et al.* 2000).

Given a curve γ , we can define θ and ϕ as in (6) and (7), and thence define R, S, H and $G = F \circ H^{-1}$. Define also

$$a(l) = \theta(H^{-1}(l)) - F(H^{-1}(l)), \tag{20}$$

$$b(l) = \phi(H^{-1}(l)) - F(H^{-1}(l)), \tag{21}$$

and let X be given by the solution to (17).

Definition 1. Define the stopping time

$$T = T_\gamma = \inf\{t \geq 0 : X_t \notin (b(L_t^X), a(L_t^X))\}. \tag{22}$$

T will be our embedding of μ .

The key idea will be that we can use excursion theory to calculate the distribution of B_T . We shall first need the following result:

Proposition 8. Let $Y_t = H^{-1}(L_t^X)$. Then the law of Y_T is given by

$$\mathbb{P}(Y_T \geq s) = \Gamma(s). \tag{23}$$

Proof. By an excursion theory argument (see, for example, Rogers 1989), the rate $r(l)$ at which positive excursions from zero hit $a(l)$ is the rate at which an excursion from 0 has maximum modulus $a(l)$ multiplied by the probability that the excursion is positive:

$$r(l) = \frac{1}{a(l)} \times \frac{1 - G'(l)}{2} = \frac{1 - G'(l)}{2(\theta(H^{-1}(l)) - F(H^{-1}(l)))}.$$

Similarly, the rate of negative excursions is

$$\frac{1}{|b(l)|} \times \frac{G'(l) + 1}{2} = \frac{G'(l) + 1}{2(F(H^{-1}(l)) - \phi(H^{-1}(l)))}.$$

Consequently,

$$\begin{aligned} \mathbb{P}(Y_T \geq s) &= \mathbb{P}(L_T^X \geq H(s)) \\ &= \exp \left\{ - \int_0^{H(s)} \left[\frac{1 - G'(l)}{2(\theta(H^{-1}(l)) - F(H^{-1}(l)))} + \frac{G'(l) + 1}{2(F(H^{-1}(l)) - \phi(H^{-1}(l)))} \right] dl \right\} \\ &= \exp \left\{ - \int_0^s \frac{1}{2} \left[\frac{H'(z) - F'(z)}{\theta(z) - F(z)} + \frac{F'(z) + H'(z)}{F(z) - \phi(z)} \right] dz \right\} \\ &= \exp \left\{ \int_0^s \frac{\Gamma'(z)}{\Gamma(z)} dz \right\} \\ &= \Gamma(s), \end{aligned}$$

where the penultimate equality is a consequence of Lemma 5. □

Theorem 9. *The stopping time T is a UI embedding; that is, $B_T \sim \mu$ and the process $B_{t \wedge T}$ is uniformly integrable.*

Remark 1. We note that T being a UI embedding is equivalent to T being minimal in the sense of Monroe (1972); that is, T is minimal if, whenever S is a stopping time with $S \leq T$ and $B_S \sim \mu$, we have $S = T$ almost surely. For further discussions on minimality, see Cox and Hobson (2003) and Cox (2005).

Proof. We first show that T is an embedding. Our proof of this fact will rely on excursion theory (Rogers 1989; Rogers and Williams 2000: VI.51), and as such, we note the following about the excursion point process of X_t : by Theorem 6, we know that there is an excursion point process $\tilde{\Xi}$ on $\mathbb{R}_+ \times U$, where U is the space of excursions, which is a Poisson random measure with intensity measure $\tilde{\Sigma}$ related to the Brownian excursion point process Ξ with intensity measure Σ by

$$\begin{aligned} \tilde{\Sigma}(dl, A_+) &= 2p(l)\Sigma(dl, A_+), \\ \tilde{\Sigma}(dl, A_-) &= 2(1 - p(l))\Sigma(dl, A_-), \end{aligned}$$

where the sets A_+ and A_- are subsets of the positive and negative excursions, respectively. In particular, if A_+ is the set of excursions with maximum greater than $a(l)$, then $\Sigma(dl, A_+) = 1/2a(l)$ and

$$\tilde{\Sigma}(dl, A_+) = \frac{p(l)}{a(l)}. \tag{24}$$

Let x be a point such that $x > F(\zeta)$. By definition,

$$\mathbb{P}(B_T \geq x) = \mathbb{P}(X_T + G(L_T^X) \geq x).$$

We will only stop at or above x if we have an excursion of X_t above $a(L_t^X)$ before the local time is too great – in particular, before $\theta(H^{-1}(l)) < x$. By (24), when the local time $L_t^X = l$, the rate of excursions of X which hit $a(l)$ is $p(l)a(l)^{-1}$. Hence, using (16), (20), (22) and (23),

$$\begin{aligned} \mathbb{P}(B_T \geq x) &= \int_{\{l:\theta(H^{-1}(l))\geq x\}} \mathbb{P}(L_T^X \geq l) \frac{p(l)}{a(l)} dl \\ &= \int_{\{z:\theta(z)\geq x\}} \mathbb{P}(Y_T \geq z) \frac{\frac{1}{2}(1 - F'(z)/H'(z))}{\theta(z) - F(z)} H'(z) dz \\ &= \int_{\{z:\theta(z)\geq x\}} \frac{H'(z) - F'(z)}{2(\theta(z) - F(z))} \Gamma(z) dz \\ &= \int_{\{z:\theta(z)\geq x\}} -\frac{R'(z)}{2} dz. \end{aligned}$$

Since R is continuous and decreasing, it will attain the lower bound in (12) on the set $\{z : \theta(z) \geq x\}$, so that we deduce (since $R(0) = 1$)

$$\mathbb{P}(B_T \geq x) = \frac{1 - (1 - 2\mu([x, \infty)))}{2} = \mu([x, \infty)).$$

A similar calculation shows that, for $x < F(\zeta)$,

$$\mathbb{P}(B_T \leq x) = \mu((-\infty, x])$$

and therefore that $B_T \sim \mu$.

To demonstrate uniform integrability, we show that $x\mathbb{P}(T > H_x) \rightarrow 0$ as $|x| \rightarrow \infty$, where $H_x = \inf\{t \geq 0 : B_t = x\}$ (see Azéma *et al.* 1980; Cox 2005). We consider the case $x > 0$. If either $F(s) = h(s)$ or $F(s) = -h(s)$ then we are reduced to the case of the Azéma and Yor (1979a) embedding (see Remark 3 below) and uniform integrability follows from standard results. Otherwise, the curve γ enters the interior of the region between $|x|$ and $c(x)$ in Figure 1 and then $\sup_{s \leq \zeta} |F(s)| \leq K$ for some K . Then, for $x > 2K$,

$$\mathbb{P}(T > H_x) = \int_{\{l:\theta(H^{-1}(l))\geq x\}} \mathbb{P}(L_T \geq l) \frac{p(l)}{x - F(H^{-1}(l))} dl.$$

But $\mathbb{P}(L_T \geq 1) \leq 1$, $p(l) \leq 1$ and $[x - F(H^{-1}(l))]^{-1} < (x - K)^{-1} < 2/x$ and so

$$x\mathbb{P}(T > H_x) \leq 2\text{Leb}(\{l : \theta(H^{-1}(l)) \geq x\}) = 2H(\theta^{-1}(x)),$$

and as $x \uparrow \infty$, θ^{-1} tends to zero, and, since H is continuous, $x\mathbb{P}(T > H_x) \rightarrow 0$. □

4. Examples

Remark 2. The choice $F \equiv 0$ and h increasing, so that the curve γ is simply the y -axis and $G(l) \equiv 0$, can easily be seen to reduce to the construction of Vallois (1992) with decreasing functions. (Vallois (1992) also has a construction with increasing functions; this is not a special case of the current construction.)

Remark 3. Taking $F(s) = h(s) = s$ (or indeed any increasing function of s), we see that $R(s) = 1$ and so $H(s) = F(s)$. Consequently, we have $G(l) = l$. Equation (17) reduces to

$$X_t = B_t - L_t^X$$

which is known to have solution $L_t^X = \sup_{s \leq t} B_s$, and X_t is a reflecting Brownian motion on $(-\infty, 0]$. In particular, for some functions ϕ_1, ϕ_2 , our stopping rule is now of the form

$$T = \inf\{t \geq 0 : X_t \leq \phi_1(L_t^X)\} = \inf\left\{t \geq 0 : B_t \leq \phi_2\left(\sup_{s \leq t} B_s\right)\right\},$$

the second form being notable as the same form as the Azéma–Yor stopping time; it is easy to check the functions are identical.

We finish by demonstrating how new embeddings can be constructed using the previous results, often with useful optimality properties. The embedding we construct will maximize the expected value of a convex function of the local time at a level $x \neq 0$. (The case where $x = 0$ is the Vallois construction of Remark 2.) We assume for convenience that $x > 0$. Then we consider the embedding constructed with $F(s) = s \wedge x$ and $h(s) = s$. The resulting embedding uses Azéma–Yor to embed the left tail of the distribution, until the moment that the maximum first reaches x , and then uses the Vallois construction to embed the rest of the distribution. The choice of the Azéma–Yor type embedding prior to the first hitting time of x will not affect the optimality of the construction, but fits in well with the ideas described in this paper.

Theorem 10. *Suppose $x \geq 0$, and let L^x be the local time of B at x . The stopping time T defined in (22) with $F(s) = s \wedge x$ and $h(s) = s$ maximizes $\mathbb{E}\Psi(L_T^x)$ for any convex function Ψ , over all embeddings τ of μ in B_t such that $B_{t \wedge \tau}$ is uniformly integrable.*

Proof. Without loss of generality, we may assume that $\Psi(0) = \Psi'(0) = 0$; the second assumption is possible since $\mathbb{E}L_T^x = 2\mathbb{E}(B_T - x)_+$, which is fixed for any UI embedding. Further, we shall also suppose that μ has a density on \mathbb{R} given by $f(y)$; however, it is clear that this assumption can be removed by considering a limiting argument.

We shall prove the theorem using a Lagrangian argument. Suppose that on the set $L_T^x > 0$ the optimal choice of embedding has a joint density function for the stopped distribution of (B_T, L_T^x) given by $\rho(b, l)$. Note that on this set we have to embed all the mass of μ to the right of x together with sufficient mass to the left of x to give a subprobability measure with mean x .

Then the problem of finding the optimal embedding can be restated as:

$$\begin{aligned} &\text{maximize} && \int_0^\infty \int_{-\infty}^\infty \Psi(l)\rho(b, l)db dl \\ &\text{subject to} && \int_0^\infty \rho(b, l)dl = f(b), \quad b > x, \end{aligned} \tag{25}$$

$$\int_0^\infty \rho(b, l) dl \leq f(b), \quad b < x, \tag{26}$$

$$\int_x^\infty (b - x)\rho(b, l)db = \frac{1}{2} \int_{-\infty}^\infty \int_l^\infty \rho(b, u)db du, \tag{27}$$

$$\int_{-\infty}^x (x - b)\rho(b, l)db = \frac{1}{2} \int_{-\infty}^\infty \int_l^\infty \rho(b, u)db du, \tag{28}$$

$$\rho(b, l) \geq 0.$$

The necessity and sufficiency of conditions (27) and (28) follow from Theorem 1.1 of Roynette *et al.* (2002). We shall construct the dual problem and show that, for a suitable choice of the Lagrangian multipliers and the conjectured optimal $\rho(b, l)$, the primal and dual have the same value. It will follow that we have indeed found the construction which maximizes $\mathbb{E}\Psi(L_T^x)$.

We introduce the multiplier $\lambda(b)$ corresponding to (25) for $b > x$ and (26) for $b < x$, and $\epsilon(l)$, $\eta(l)$ corresponding to (27) and (28), respectively. The resulting Lagrangian is:

$$\begin{aligned} &\int \int \rho(b, l) \left[\Psi(l) - \lambda(b) - \epsilon(l)(b - x)\mathbf{1}_{\{b > x\}} - \eta(l)(x - b)\mathbf{1}_{\{b < x\}} + \frac{1}{2} \int_0^l (\epsilon(u) + \eta(u)) du \right] db dl \\ &+ \int \lambda(b)f(b)db. \end{aligned}$$

The corresponding dual problem is therefore to minimize $\int \lambda(b)f(b)db$ subject to the term in square brackets being non-positive, and $\lambda(b) \geq 0$ for $b \leq x$ (as a consequence of the inequality in (26)).

Let $\delta < x$ satisfy $\int_\delta^\infty (b - x)f(b)db = 0$. Having embedded all the mass to the left of δ before the first hitting time H_x of x , on the set $T > H_x$ we then use the Vallois construction to embed μ restricted to $[\delta, \infty)$: let $\alpha(l) \geq x$ decreasing, and $\beta(l) \leq x$ increasing be the Vallois boundaries, then on $T > H_x$ we have $T = \inf\{t \geq H_x : B_t \notin (\beta(L_t), \alpha(L_t))\}$. On $T > H_x$, T embeds μ restricted to $[\delta, \infty)$.

Define the multipliers

$$\begin{aligned}
 A(m) &= \frac{1}{2} \int_0^m \left[\frac{1}{\alpha(k) - x} + \frac{1}{x - \beta(k)} \right] dk; \\
 \epsilon(l) &= \int_0^l \frac{e^{A(u)}}{\alpha(u) - x} \int_0^u \Psi''(m) e^{-A(m)} dm du; \\
 \eta(l) &= \int_0^l \frac{e^{A(u)}}{x - \beta(u)} \int_0^u \Psi''(m) e^{-A(m)} dm du; \\
 \lambda(b) &= \begin{cases} \int_0^{\alpha^{-1}(b)} \frac{e^{A(u)}(\alpha(u) - b)}{\alpha(u) - x} \int_0^u \Psi''(m) e^{-A(m)} dm du, & b > x, \\ \int_0^{\beta^{-1}(b)} \frac{e^{A(u)}(b - \beta(u))}{x - \beta(u)} \int_0^u \Psi''(m) e^{-A(m)} dm du, & b < x. \end{cases}
 \end{aligned}$$

Note, in particular, that $\lambda(b) \geq 0$ for $b < x$. With these definitions, and using the properties of the functions $\alpha(l)$ and $\beta(l)$, it is now possible to show that

$$\Psi(l) - \lambda(b) - \epsilon(l)(b - x)\mathbf{1}_{\{b > 0\}} - \eta(l)(x - b)\mathbf{1}_{\{b < 0\}} + \frac{1}{2} \int_0^l (\epsilon(u) + \eta(u)) du \leq 0,$$

so that we have a feasible solution to the dual problem, and also

$$\int \lambda(b) f(b) db = \mathbb{E} \Psi(L_T^x)$$

when T is the candidate embedding. Consequently the embedding is indeed optimal. □

Acknowledgements

The first author is grateful to the Nuffield Foundation for financial support. The second author is supported by an EPSRC Advanced Fellowship. Both authors thank the referees for their detailed reports which helped improve the clarity of the paper.

References

Azéma, J., Gundy, R.F. and Yor, M. (1980) Sur l'intégrabilité uniforme des martingales continues. In J. Azéma and M. Yor (eds), *Séminaire de Probabilités XIV*, Lecture Notes in Math. 784, pp. 53–61. Berlin: Springer-Verlag.

Azéma, J. and Yor, M. (1979a) Une solution simple au problème de Skorokhod. In C. Dellacherie, P.A. Meyer and M. Weil (eds), *Séminaire de Probabilités XIII*, Lecture Notes in Math. 721, pp. 90–115. Berlin: Springer-Verlag.

Azéma, J. and Yor, M. (1979b) Le problème de Skorokhod: compléments à 'Une solution simple au problème de Skorokhod'. In C. Dellacherie, P.A. Meyer and M. Weil (eds), *Séminaire de Probabilités XIII*, Lecture Notes in Math. 721, pp. 625–644. Berlin: Springer-Verlag.

- Barlow, M., Burdzy, K., Kaspi, H. and Mandelbaum, A. (2000) Variably skewed Brownian motion. *Electron. Comm. Probab.*, **5**, 57–66.
- Bass, R.F. (1983) Skorokhod imbedding via stochastic integrals. In J. Azéma and M. Yor (eds), *Séminaire de Probabilités XVII*, Lecture Notes in Math. 986, pp. 221–224. Berlin: Springer-Verlag.
- Chacon, R.V. (1977) Potential processes. *Trans. Amer. Math. Soc.*, **226**, 39–58.
- Chacon, R.V. and Walsh, J.B. (1976) One-dimensional potential embedding. In P.A. Meyer (ed.), *Séminaire de Probabilités X*, Lecture Notes in Math. 511, pp. 19–23. Berlin: Springer-Verlag.
- Cox, A.M.G. (2006) Extending Chacon-Walsh: Minimality and generalised starting distributions. In *Séminaire de Probabilités*, to appear.
- Cox, A.M.G. and Hobson, D.G. (2006) Skorokhod embeddings, minimality and non-centred target distributions. *Probab. Theory Related Fields*, **135**(3), 395–414.
- Dubins, L.E. (1968) On a theorem of Skorokhod. *Ann. Math. Statist.*, **39**, 2094–2097.
- Harrison, J.M. and Shepp, L.A. (1981) On skew Brownian motion. *Ann. Probab.*, **9**(2), 309–313.
- Meilijson, I. (1983) On the Azéma–Yor stopping time. In J. Azéma and M. Yor (eds), *Séminaire de Probabilités XVII*, Lecture Notes in Math. 986, pp. 225–226. Berlin: Springer-Verlag.
- Monroe, I. (1972) On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.*, **43**, 1293–1311.
- Oblój, J. (2004) The Skorokhod embedding problem and its offspring. *Probab. Surv.*, **1**, 321–390.
- Revuz, D. and Yor, M. (1999) *Continuous Martingales and Brownian Motion*, 3rd edn. Berlin: Springer-Verlag.
- Rogers, L.C.G. (1989) A guided tour through excursions. *Bull. London Math. Soc.*, **21**(4), 305–341.
- Rogers, L.C.G. and Williams, D. (2000) *Diffusions, Markov Processes, and Martingales. Vol. 2: Itô Calculus*. Cambridge: Cambridge University Press.
- Root, D.H. (1969) The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.*, **40**, 715–718.
- Roynette, B., Vallois, P. and Yor, M. (2002) A solution to Skorokhod’s embedding for linear Brownian motion and its local time. *Studia Sci. Math. Hungar.*, **39**(1–2), 97–127.
- Skorokhod, A.V. (1965) *Studies in the Theory of Random Processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass.
- Vallois, P. (1983) Le problème de Skorokhod sur \mathbf{R} : une approche avec le temps local. In J. Azéma and M. Yor (eds), *Séminaire de Probabilités XVII*, Lecture Notes in Math. 986, pp. 227–239. Berlin: Springer-Verlag.
- Vallois, P. (1992) Quelques inégalités avec le temps local en zero du mouvement brownien. *Stochastic Process. Appl.*, **41**(1), 117–155.

Received June 2005 and revised April 2006