Skorokhod embeddings, minimality and non-centred target distributions

A. M. G. Cox*and D. G. Hobson†
Department of Mathematical Sciences,
University of Bath,
Bath BA2 7AY, U. K.

October 17, 2003

Abstract

In this paper[1] we consider the Skorokhod embedding problem for target distributions with non-zero mean. In the zero-mean case, uniform integrability provides a natural restriction on the class of embeddings, but this is no longer suitable when the target distribution is not centred. Instead we restrict our class of stopping times to those which are minimal, and we find an condition involving stopping times which is equivalent to minimality.

We then apply these results, firstly to the problem of embedding non-centred target distributions in Brownian motion, and secondly to embedding general target laws in a diffusion.

We find an embedding (which reduces to the Azema-Yor embedding in the zero-target mean case) which maximises the law of $\sup_{s\leq T} B_s$ among the class of minimal embeddings of a general target distribution μ in Brownian motion. We construct a minimal embedding of μ in a diffusion X which maximises the law of $\sup_{s\leq T} h(X_s)$ for a general function h.

1 Introduction

The Skorokhod embedding problem was first proposed, and then solved, by Skorokhod (1965), and may be described thus:

> Given a Brownian motion $(B_t)_{t\geq 0}$ and a centred target law μ can we find a stopping time T such that B_T has distribution μ?

---

*e-mail: mapamgc@maths.bath.ac.uk; web: www.bath.ac.uk/~mapamgc/

†e-mail: dgh@maths.bath.ac.uk; web: www.bath.ac.uk/~masdgh/)

[1]**MSC 2000 subject classifications.** Primary: 60G40, 60J60; Secondary: 60G44, 60J65.
Keywords: Brownian Motion, Diffusion, Embedding, Azema-Yor Embedding, Stopping Time, Minimal Stopping Time.

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. Since the problem was posed many more solutions have been given, see for example Dubins (1968), Root (1969) and Chacon and Walsh (1976). With different solutions comes the question of optimal properties of the embeddings, and various optimal embeddings have been found — for example the embedding minimising the variance of T (Rost, 1976), the embedding stochastically minimising the law of the local time at zero (Vallois, 1992), or the embedding stochastically minimising the law of the maximum (Perkins, 1986).

An obvious extension of the problem is to consider more general classes of processes. Here the question of the existence of an embedding becomes more interesting. For a Markov process and an arbitrary target measure necessary and sufficient conditions are given by Rost (1971) and a construction in this general case is given by Bertoin and Le Jan (1992). In the case of diffusions on $\mathbb{R}$ simpler necessary and sufficient conditions are given in Grandits and Falkner (2000), Pedersen and Peskir (2001) and Cox and Hobson (2002), along with some constructions.

The work we present here was motivated by the following question:

> Given a diffusion $(X_t)_{t\geq 0}$ and a target distribution μ_X for which an embedding exists, which embedding maximises the law of $\sup_{s\leq T} X_s$ (respectively $\sup_{s\leq T} |X_s|$)?

For Brownian motion, the question has been solved by Azéma and Yor (1979) (respectively Jacka (1988)) under the condition that $B_{t\wedge T}$ is a UI-martingale.

There are several considerations that need to be made when moving from the Brownian case to the diffusion case. Firstly, the mean-zero assumption that is made by Azéma and Yor (1979) and Jacka (1988) is no longer natural since we are no longer necessarily dealing with a martingale. The second aspect that needs to be considered is with what restriction should we replace the UI condition? That such a condition is desirable may be seen by considering a recurrent diffusion. Here the maximisation problem can easily seen to be degenerate by considering first running the diffusion until it hits a level x, allowing it to return to the origin and then using the reader's favourite embedding. Clearly this dominates the unmodified version of the reader's favourite embedding.

In Pedersen and Peskir (2001) an integrability condition on the maximum (specifically that $\mathbb{E}(\sup_{s\leq T} s(X_s)) < \infty$ where s is the scale function of X) was suggested to replace the UI condition in the Brownian case. In this work we propose using the following class of stopping times introduced by Monroe (1972) to provide us with a natural restriction on the set of admissible embeddings.

Definition 1. A stopping time T for the process X is *minimal* if whenever $S \leq T$ is a stopping time such that X_S and X_T have the same distribution then $S = T$ a.s..

The class of minimal stopping times provides us with a natural link to the uniformly integrable Brownian case as a consequence of the following result:

Theorem 2. *(Monroe, 1972, Theorem 3) Let S be a stopping time such that $\mathbb{E}(B_S) = 0$. Then S is minimal if and only if the process $B_{t\wedge S}$ is uniformly integrable.*

It will turn out that the minimality idea fits well with the problem of embedding in diffusions. Our approach to embedding in diffusions will be to map the diffusion into natural scale (so that, up to a time change, it resembles Brownian motion) and use techniques developed for embedding Brownian motion. Using this method on a transient diffusion one finds that the state space and target distribution for the Brownian motion is restricted to a half-line (or sometimes a finite interval). We will show minimality to be equivalent to stopping the Brownian motion before it leaves this interval, so that a minimal stopping time is necessarily before the first explosion time of X.

When we map from the problem of embedding μ_X in X to the Brownian motion the target law μ we obtain for B is the image of μ_X under the scale function. The key point is that there is no reason why this target law should have mean zero. Thus, unlike most of the other studies of Skorokhod embeddings in Brownian motion we are interested in non-centred target distributions, and non-UI stopping times. One of our main results is to recharacterise the minimality condition on T in terms of a condition on $\mathbb{E}(B_T|\mathcal{F}_S)$ for stopping times $S \leq T$. In fact most of the paper will concentrate on embedding non-centred target distributions in B, and we will only return to the diffusion case in a short final section.

The paper will proceed as follows. In Section 2 we construct the classical Azema-Yor embedding (see Azéma and Yor (1979)) to introduce the reader to the construction we will use later. Then in Section 3 we prove some results concerning minimality of stopping times for non-centred target distributions, giving an equivalent condition to minimality in terms of the process. In particular, given a non-minimal embedding T, we show how to construct a new (minimal) stopping time $T' \leq T$ which embeds μ. Next, in Section 5 we construct an extension of the Azema-Yor embedding for non-centred target distributions and show both that it is minimal, and that it retains the optimality properties of the original Azema-Yor embedding. In Section 6 we use these stopping times to construct an embedding maximising the distribution of $\sup_{s\leq T} h(B_s)$ for a general function h. Finally in Section 7 we apply these results to the problem of embedding optimally in diffusions.

2 The Azema-Yor embedding

We begin by introducing the Azema-Yor embedding (see Azéma and Yor (1979)) and a notation which we will use in later sections.

Let μ be our target distribution on $\mathbb{R}$, with mean $m = 0$, and let B_t be a Brownian motion with $B_0 = 0$. Our goal is to embed μ in B — i.e. find a stopping time T such that the stopped process B_T has law μ.

Write

$$M_t = \sup_{s\leq t} B_s \tag{1}$$

$$J_t = \inf_{s\leq t} B_s \tag{2}$$

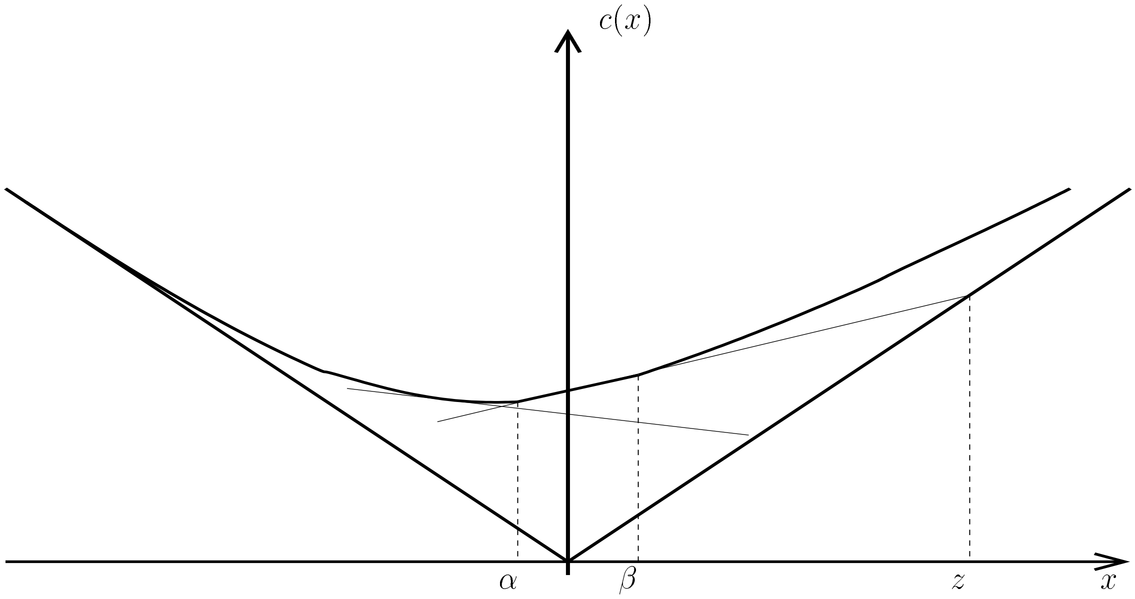

Figure 1: A graph of $c(x)$. The support of the corresponding distribution μ is bounded below. Further $\mu((\alpha,\beta)) = 0$ and there is an atom at β. Consequently there are multiple tangents to c at β. We have that $\alpha = b(z)$. Note that $u(u^{-1}(\beta)) = \alpha$, but that $c(\alpha) - \alpha u^{-1}(\alpha) = c(\beta) - \beta u^{-1}(\beta)$.

respectively for the maximum and minimum processes of B, and define

$$c(x) := \mathbb{E}^{\mu}|X - x|. \tag{3}$$

The definition of c ensures that $c(x)$ is a convex function which is asymptotic to, and greater than or equal to, the function $|x|$. For $\theta \in [-1, 1]$,

$$u(\theta) := \inf\{y \in \mathbb{R} : c(y) + \theta(x - y) \le c(x), \forall x \in \mathbb{R}\}. \tag{4}$$

We will later want to want to use the inverse function, which we will define to be

$$u^{-1}(y) = \inf\{\theta : u(\theta) \ge y\}.$$

For an interpretation of these and subsequent quantities we refer the reader to Figure 1. Our interpretation of θ in (4) is that it is the gradient of a tangent to c, and then $u(\theta)$ is the smallest x at which there exists a tangent to c at x with gradient θ.

Let

$$z_+(\theta) := \frac{c(u(\theta)) - \theta u(\theta)}{1 - \theta}, \tag{5}$$

and define also

$$b(w) := u(z_+^{-1}(w))$$

for $0 \le w \le \sup\{\mathrm{supp}(\mu)\}$. The function b is well defined and left-continuous since $z_+(\theta)$ is a continuous bijection $z_+ : [-1, 1] \to [0, \sup\{\mathrm{supp}(\mu)\}]$ (if $\sup\{\mathrm{supp}(\mu)\} =$

∞, $u(1) = \infty$ and we take $z_+(1) = \infty$). We interpret $z_+(\theta)$ as the x-co-ordinate of the intersection of the line $y = x$ and the tangent with gradient θ. It follows that $b(w)$ is the x-value of the left-most point on $(x, c(x))$ with the property that the tangent through this point hits the line $y = x$ at w, see Figure 1 for a pictorial representation of this idea.

Lemma 3 (The Azema-Yor Embedding). *For μ, B as above, define the stopping time*

$$T_{AY} := \inf\{t > 0 : B_t \leq b(M_t)\}. \tag{6}$$

Then T_{AY} is a minimal embedding of μ in B. Further T_{AY} has the property that if T is another minimal stopping time which embeds μ in B, then

$$\mathbb{P}(M_T \geq y) \leq \mathbb{P}(M_{T_{AY}} \geq y). \tag{7}$$

Remark 4. The above theorem is the standard statement of the Azema-Yor embedding (e.g. Perkins (1986, Theorem 2.5)) except for the minimality condition on T. In other statements of the result T is required to be a stopping time for which the process $B_{T\wedge t}$ is a UI-martingale. However we are able to replace this condition with a minimality condition due to Theorem 2.

Classically (see e.g. Pedersen and Peskir (2001)), the above result is stated for the stopping time

$$T'_{AY} := \inf\{t > 0 : M_t \geq \Psi(B_t)\},$$

where Ψ is the *barycentre* function:

$$\Psi(x) = \begin{cases} \frac{1}{\mu([x,\infty))} \int_{\{u \geq x\}} u \, \mu(du) & \mu([x,\infty)) > 0 \\ x & \mu([x,\infty)) = 0. \end{cases}$$

Here, we do not prove the result but show simply that $T_{AY} = T'_{AY}$.

Since

$$\begin{aligned} T_{AY} &= \inf\{t > 0 : B_t \leq b(M_t)\} \\ &= \inf\{t > 0 : b^{-1}(B_t) \leq M_t\} \end{aligned}$$

where $b^{-1}(y) = \inf\{z : b(z) \geq y\}$, it is sufficient to show that $\Psi = b^{-1}$. Note that $b^{-1}(y) = z_+(u^{-1}(y))$. Further, from the definition of c, we have that

$$\begin{aligned} u^{-1}(y) &= c'_-(y) \\ &= 1 - 2\mu([y,\infty)), \end{aligned} \tag{8}$$

where c_- denotes the left-derivative of c (which exists by the convexity of c).

Now $u(u^{-1}(y)) = y$ unless there exits some $z < y$ for which $\mu((z, y)) = 0$, in which case however it is still true that $c(u(u^{-1}(y))) - u(u^{-1}(y))u^{-1}(y) = c(y) - yu^{-1}(y)$, since $c'(w) = c'_-(y)$ for all $w \in (z, y)$. Collecting all these observations together we have

$$\begin{aligned}
b^{-1}(y) &= z_+(u^{-1}(y)) \\
&= \frac{c(u(u^{-1}(y))) - u(u^{-1}(y))u^{-1}(y)}{1 - u^{-1}(y)} \\
&= \frac{c(y) - yu^{-1}(y)}{1 - u^{-1}(y)} \\
&= \frac{\int |w - y| \, \mu(dw) - y(1 - 2\mu([y, \infty)))}{2\mu([y, \infty))} \\
&= \frac{1}{\mu([y, \infty))} \int_{\{w \geq y\}} w \, \mu(dw) \\
&= \Psi(y), \qquad (9)
\end{aligned}$$

and we are done.

Remark 5. Thus we have shown that the embedding we define in (6) is a representation of the Azema-Yor embedding and hence, by the well known properties of this embedding established in Blackwell and Dubins (1963), is optimal in the sense that it maximises the distribution of the maximum among all stopping times embedding μ in B which are minimal.

3 Minimal embeddings for non-centred distributions

The main result of this section is the following:

Theorem 6. *Let T be a stopping time of Brownian motion which embeds a distribution μ where $m = \int_{\mathbb{R}} x \, \mu(dx) < 0$. Then the following conditions are equivalent:*

(i) T is minimal for μ;

(ii) for all stopping times $S \leq T$,

$$\mathbb{E}(B_T | \mathcal{F}_S) \leq B_S \quad a.s.. \qquad (10)$$

In the case where supp$(\mu) \subseteq [\alpha, \infty)$ for some $\alpha < 0$ then the above conditions are also equivalent to the condition:
(iii)

$$\mathbb{P}(T \leq H_\alpha) = 1, \qquad (11)$$

where $H_\alpha = \inf\{t > 0 : B_t = \alpha\}$ is the hitting time of the level α.

Remark 7. Of course the Theorem may be restated in the case where $m > 0$ by considering the process $-B_t$. We will use this observation extensively in Section 5.

Remark 8. The second condition we give here can be thought of as analogous to the condition

$$\forall S \leq T, \qquad\qquad \mathbb{E}(B_T|\mathcal{F}_S) = B_S \quad \text{a.s.}$$

in the case where $m = 0$. The proof of the corresponding result in Monroe (1972) shows that this is the key idea in showing that uniform integrability is equivalent to minimality. In the case we are interested in there is no equivalent notion to correspond with uniform integrability, so we use (10) instead.

We will prove this theorem in several stages. We begin with a lemma whose corollary shows that the intermediate stopping time condition implies minimality. The lemma has the form given because we use this form in a later proof. For our current purposes it is the subsequent and immediate corollary which is most important.

Throughout this section it is to be understood that μ is a distribution with negative mean and T a stopping time embedding μ. Given a stopping time S let θ_S be the shift operator — the map for which $B_t(\theta_S(\omega)) = B_{S+t}(\omega)$.

Lemma 9. *Suppose that for all embeddings S of μ with $S \leq T$ we have*

$$\mathbb{E}(B_T|\mathcal{F}_{S_\epsilon}) \leq B_{S_\epsilon} \quad a.s. \quad \forall \epsilon \geq 0 \tag{12}$$

where $S_\epsilon = (S + H_{B_S - \epsilon} \circ \theta_S) \wedge T$, and $S_0 = S$. Then T is minimal.

Proof. Let $S \leq T$ be a stopping time such that $\mathbb{E}(B_T|\mathcal{F}_S) \leq B_S$ almost surely and such that S embeds μ. For $a \in \mathbb{R}$,

$$\begin{aligned} \sup_{A \in \mathcal{F}_T} \mathbb{E}(a - B_T; A) &= \mathbb{E}(a - B_T; B_T \leq a) \\ &= \mathbb{E}(a - B_S; B_S \leq a) \\ &\leq \mathbb{E}(a - B_T; B_S \leq a) \qquad\qquad (13) \\ &\leq \sup_{A \in \mathcal{F}_T} \mathbb{E}(a - B_T; A) \end{aligned}$$

where we use (12) to deduce (13). However since we have equality in the first and last expressions, we must also have equality throughout and so

$$\{B_T < a\} \subseteq \{B_S \leq a\} \subseteq \{B_T \leq a\}.$$

Since this holds for all $a \in \mathbb{R}$ we must have $B_T = B_S$ a.s..

Now suppose that $S \neq T$ with positive probability, and consider the stopping time S_ϵ. For small enough $\varepsilon > 0$, $S_\varepsilon < T$ with positive probability, and on the $\mathcal{F}_{S_\varepsilon}$-measurable set $\{B_{S_\varepsilon} = B_S - \varepsilon\}$ we have $B_{S_\varepsilon} = \mathbb{E}(B_T|\mathcal{F}_{S_\varepsilon}) - \varepsilon$ which contradicts (12). Consequently if (12) holds, $S \leq T$ and $S \sim \mu$ implies $S = T$ a.s.. □

Corollary 10. *Suppose that for all stopping times $S \leq T$,*

$$\mathbb{E}(B_T|\mathcal{F}_S) \leq B_S \quad a.s.. \tag{14}$$

Then T is minimal.

For the converse we need to show that if T is minimal then for any stopping time $S \leq T$ and $A \in \mathcal{F}_S$

$$\mathbb{E}(B_T; A) \leq \mathbb{E}(B_S; A).$$

We first prove a lemma which will allow us to consider only sets $A \subseteq \{S \leq H_{-k}\}$.

Lemma 11. *Suppose $S \leq T$ are stopping times and, for $A \in \mathcal{F}_S$, write*

$$A_k = \{S \leq H_{-k}\} \cap A.$$

If

$$\mathbb{E}(B_T; A_k) \leq \mathbb{E}(B_S; A_k) \tag{15}$$

for all k and all $A \in \mathcal{F}_S$, then

$$\mathbb{E}(B_T; A) \leq \mathbb{E}(B_S; A) \tag{16}$$

for all $A \in \mathcal{F}_S$.

Proof. Suppose (16) does not hold for some $A \in \mathcal{F}_S$, so that

$$\mathbb{E}(B_T; A) > \mathbb{E}(B_S; A). \tag{17}$$

We consider the cases $\mathbb{E}(|B_S|; A) < \infty$ and $\mathbb{E}(|B_S|; A) = \infty$ separately. In the first case, by dominated convergence,

$$\begin{aligned} \mathbb{E}(B_T; A_k) &\rightarrow \mathbb{E}(B_T; A) \\ \mathbb{E}(B_S; A_k) &\rightarrow \mathbb{E}(B_S; A) \end{aligned}$$

which would violate (15) for sufficiently large k.

In the second case, from $\mathbb{E}(B_T) = m$ and (17), we must have $\mathbb{E}(B_S^+; A) < \infty$, so $\mathbb{E}(B_S^-; A) = \infty$, where $B_t^+ = B_t \vee 0$ and $B_t^- = B_t \wedge 0$ — the positive and negative parts of B_t respectively. So, for $K > \mathbb{E}(|B_T|; A)$ we can find values β, β' such that

$$\mathbb{E}(B_S; \beta < B_S < \beta', A) = -K$$

and by the bounded convergence theorem

$$\mathbb{E}(B_S; \beta < B_S < \beta', A_k) \rightarrow -K.$$

Finally, by (15),

$$-K < -\mathbb{E}(|B_T|; A) \leq \mathbb{E}(B_T; \beta < B_S < \beta', A_k) \leq \mathbb{E}(B_S; \beta < B_S < \beta', A_k),$$

and we deduce a contradiction. □

Another result that we need to use is the following.

Lemma 12. *If T is minimal then, for all $\gamma < 0$ and $A \in \mathcal{F}_{T \wedge H_\gamma}$,*

$$\mathbb{E}(B_T - B_{T \wedge H_\gamma}; A) \leq 0.$$

Proof. Let $f(\gamma) = \mathbb{E}(B_T - B_{T\wedge H_\gamma}) = \mathbb{E}(B_T - \gamma; T > H_\gamma)$. We begin by showing that $f(\gamma) \leq 0$ for $\gamma \leq 0$.

Note that $f(0) = m < 0$. Recalling (2), since

$$\{T \in (H_{\gamma-\varepsilon}, H_{\gamma+\varepsilon}) \setminus \{H_\gamma\}\} = \{J_T \in (\gamma - \varepsilon, \gamma + \varepsilon), T \neq H_\gamma\}$$

and $\mathbb{P}(J_T = \gamma, T \neq H_\gamma) = 0$, we have that

$$\mathbb{P}(T \in (H_{\gamma-\varepsilon}, H_{\gamma+\varepsilon}) \setminus \{H_\gamma\}) \leq \mathbb{P}(J_T \in (\gamma - \varepsilon, \gamma) \cup (\gamma, \gamma + \varepsilon)).$$

Further $\mathbb{P}(J_T \in A)$ is a probability measure on $\mathbb{R}$, so it follows from the bounded convergence theorem that

$$\mathbb{P}(T \in (H_{\gamma-\varepsilon}, H_{\gamma+\varepsilon}) \setminus \{H_\gamma\}) \to 0$$

as $\varepsilon \to 0$. The continuity of $f(\gamma)$ follows from the dominated convergence theorem and the fact that $\mathbb{E}|B_T| < \infty$. As a corollary if $f(\gamma_0) > 0$ for some $\gamma_0 < 0$, then there exists $\gamma_1 \in (\gamma_0, 0)$ such that $f(\gamma_1) = 0$.

Given this γ_1, and conditional on $T > H_{\gamma_1}$, let $T'' = T - H_{\gamma_1}$, $W_t = B_{H_{\gamma_1}+t} - \gamma_1$, and $\mu'' = \mathcal{L}(W_{T''})$. Suppose that T'' is not minimal, so there exists $S'' \leq T''$ with law μ''. If we define

$$S = \begin{cases} T & \text{on } T \leq H_{\gamma_1} \\ H_{\gamma_1} + S'' & \text{on } T > H_{\gamma_1} \end{cases}$$

then S embeds μ and $S \leq T$ but $S \neq T$, contradicting the minimality of T. Hence T'' is minimal. But then by Theorem 2, $W_{t\wedge T''}$ is uniformly integrable and so, for $\gamma < \gamma_1$

$$\mathbb{E}(W_{T''} - (\gamma - \gamma_1); T'' > H^W_{\gamma-\gamma_1}) = 0$$

or equivalently

$$f(\gamma) = \mathbb{E}(B_T - \gamma; T > H_\gamma) = 0.$$

Hence $f(\gamma) \leq 0$ for all $\gamma \in (-\infty, 0]$.

Now fix $\gamma < 0$ and take $A \in \mathcal{F}_{T\wedge H_\gamma}$. We have $\mathbb{E}(B_T - B_{T\wedge H_\gamma}) \leq 0$ so if

$$\mathbb{E}(B_T - B_{T\wedge H_\gamma}; A) = \mathbb{E}(B_T - \gamma; (T > H_\gamma) \cap A) > 0 \tag{18}$$

then there exists $A' \supset A$ such that $\mathbb{E}(B_T - B_{T\wedge H_\gamma}; A') = 0$. Conditioning on A', we can now repeat the above argument to deduce that $\{(B_{H_\gamma+t} - \gamma)\mathbf{1}_{A'}\mathbf{1}_{T>H_\gamma}\}_{0\leq t\leq(T-H_\gamma)}$ is uniformly integrable, which is a contradiction to (18). ∎

Proposition 13. *Let T be a stopping time of Brownian motion which embeds a distribution μ with support on $\mathbb{R}$ and mean $m < 0$. If T is minimal for μ then for all stopping times $S \leq T$*

$$\mathbb{E}(B_T|\mathcal{F}_S) \leq B_S \quad a.s.. \tag{19}$$

Proof. By Lemma 11, we need only show (15) holds for all $A \in \mathcal{F}_S$, and all k. Fix k. Then $S \leq H_{-k}$ on A_k, and the result will follow if we can prove the following:

$$\begin{aligned} \mathbb{E}(B_T; A_k) &\leq \mathbb{E}(B_{T \wedge H_{-k}}; A_k) \\ &\leq \mathbb{E}(B_S; A_k). \end{aligned} \tag{20}$$

The first inequality is a direct application of Lemma 12. The second follows from the combination of the observation that $S \leq H_{-k}$ on A_k, (and hence $S = T \wedge S \leq T \wedge H_{-k}$) with the fact that $B_{T \wedge H_{-k} \wedge t}$ is a supermartingale. □

This result gives us equivalence between (i) and (ii) in Theorem 6. We are left with showing that if μ has support bounded below then (iii) is also equivalent. So assume that the target distribution μ has support contained in $[\alpha, \infty)$ and that T is an embedding of μ. In that case it is easy to show that (11) is equivalent to (10). To deduce the forward implication, note that $B_{t \wedge H_\alpha}$ is a continuous supermartingale, bounded below and therefore if $S \leq T \leq H_\alpha$,

$$\mathbb{E}(B_T | \mathcal{F}_S) \leq B_S.$$

The reverse implication follows from considering the stopping time $H_{\alpha-\epsilon} = \inf\{t \geq 0 : B_t \leq \alpha - \varepsilon\}$, for then if $A = \{\omega : H_{\alpha-\epsilon} < T\}$ and $S = H_{\alpha-\varepsilon} \wedge T$,

$$(\alpha - \epsilon)\mathbb{P}(A) = \mathbb{E}(B_{H_{\alpha-\epsilon}}; A) = \mathbb{E}(B_S; A) \geq \mathbb{E}(B_T; A) \geq \alpha \mathbb{P}(A)$$

which is only possible if $\mathbb{P}(A) = 0$.

This completes the proof of Theorem 6. We finish this section with a remark on an application of the theorem.

Remark 14. Suppose we are embedding a target distribution with negative mean m. Let T be such an embedding which consists of running the Brownian motion until it hits m, and thereafter using a (shifted) Azema-Yor embedding to embed the (zero-mean) shifted target distribution applied to $-B$, so that we maximise $\mathbb{P}(J \leq x)$ for $x < 0$. This is an embedding we will look at more closely in Section 5 and has been studied by Pedersen and Peskir (2001). We show that T is minimal.

It is clear that T has the property that $\mathbb{E}(B_T; A) = m\mathbb{P}(A)$ for any set $A \in \mathcal{F}_{H_m}$, since $(B_{T \wedge (H_m + t)})_{t \geq 0}$ is a UI process.

To prove the minimality property we show that $\mathbb{E}(B_T | \mathcal{F}_S) \leq B_S$ for all stopping times $S \leq T$. Let S be such a stopping time, and suppose $A \in \mathcal{F}_S$. Then A may be written as the disjoint union, $A = (A \cap \{S \leq H_m\}) \sqcup (A \cap \{S > H_m\})$ and

$$\mathbb{E}(B_T; A) = \mathbb{E}(B_T; A \cap \{S \leq H_m\}) + \mathbb{E}(B_T; A \cap \{S > H_m\}).$$

For the first term we can deduce that

$$\mathbb{E}(B_T; A \cap \{S \leq H_m\}) = m\mathbb{P}(A, S \leq H_m) \leq \mathbb{E}(B_S; A \cap \{S \leq H_m\}).$$

For the second term, and again as a consequence of the fact that on $\{t \geq H_m\}$ the process $B_{T \wedge t}$ is UI,

$$\mathbb{E}(B_T; A \cap \{S > H_m\}) = \mathbb{E}(B_S; A \cap \{S > H_m\}).$$

Combining the results for the two terms gives

$$\mathbb{E}(B_T; A) \le \mathbb{E}(B_S; A)$$

as required.

4 Constructing minimal stopping times

This section is concerned with the following question:

> Suppose condition (ii) of Theorem 6 fails, so T is not minimal. Can we find a stopping time $T^* \le T$ which embeds μ and is itself minimal?

The answer to this question is in the positive, and we will demonstrate this by providing just such a construction for a given non-minimal stopping time.

The construction will be carried out in three parts. First we construct T^* in the case where μ has support on $[\alpha, \infty)$ and an atom at α. We then use limiting arguments to show that we can drop the assumption of an atom at α, and finally that we can embed μ with support on $\mathbb{R}$.

So we first assume that our support is restricted to $[\alpha, \infty)$. We will show that if T is a stopping time embedding μ, and T has the property that

$$\mathbb{P}(T > H_\alpha) > 0 \tag{21}$$

then we can construct a new stopping time $T^* \le T$ which is less than H_α almost surely, and which embeds μ.

In the construction we make, we find it necessary to introduce independent randomisation. While this is undesirable, it is central to the method we use here; the randomisation will be used to 'kill' the process at rate ν. The set $\mathcal{M}^T$ we introduce below is a set of suitable 'killing measures'.

We begin by considering the set of positive measures $\mathcal{M}^+$ on $\mathcal{B}(\mathbb{R})$ and define an ordering on $\mathcal{M}^+$ by:

$$\rho \preceq \nu \text{ iff } \rho(A) \le \nu(A) \text{ for all } A \in \mathcal{B}(\mathbb{R}).$$

Then $\mathcal{M}^+$ is a lattice where

$$\begin{aligned}
(\rho \vee \nu)(A) &= \sup\{\rho(B) + \nu(A \setminus B) : B \subseteq A\};\\
(\rho \wedge \nu)(A) &= \inf\{\rho(B) + \nu(A \setminus B) : B \subseteq A\}.
\end{aligned}$$

Also (see Doob (1984, A.IV.4)) if $\{\rho_i : i \in I\}$ is an arbitrary subset of $\mathcal{M}^+$, then there exists an order supremum of the set. To see this, we begin by assuming that the set contains every supremum of finitely many of its elements since adding these does not change the overall supremum. Define

$$\rho_*(A) = \sup_{i \in I} \rho_i(A);$$

it is clear that $\rho_* = \bigvee_{i\in I} \rho_i$ provided that ρ_* is a measure. If $A = \bigcup_{j=0}^\infty A_j$ is a countable union of disjoint measurable sets, then

$$\rho_i(A) = \sum_{j=0}^\infty \rho_i(A_j) \le \sum_{j=0}^\infty \rho_*(A_j)$$

for all i, and hence ρ_* is countably subadditive. Finite additivity of ρ_* is trivial (given the fact that $\{\rho_i : i \in I\}$ contains the suprema of each finite set of elements) and implies

$$\rho_*(A) = \sum_{j=0}^n \rho_*(A_j) + \rho_*\left(\bigcup_{j=n+1}^\infty A_j\right) \ge \sum_{j=0}^n \rho_*(A_j).$$

On letting $n \to \infty$, we conclude that ρ_* is countably superadditive as well as subadditive. Hence ρ_* is a measure.

Recall that T is a stopping time that embeds μ. Given a measure $\nu \in \mathcal{M}^+$ we define the stopping time T^ν as follows.

- Let X_- and X_+ be independent random variables, independent also of B and T, and both distributed uniformly on $[0,1]$.
- Define the levels

$$\begin{aligned} G_+^\nu &= \inf\{x \ge 0 : \nu([0,x]) \ge X_+\}, \\ G_-^\nu &= \sup\{x \le 0 : \nu([x,0)) \ge X_-\}, \end{aligned}$$

  and the stopping times $S_+^\nu = H_{G_+^\nu}$, $S_-^\nu = H_{G_-^\nu}$ and $S^\nu = S_+^\nu \wedge S_-^\nu$.
- Finally set

$$T^\nu = S^\nu \wedge T \wedge H_\alpha.$$

We now define the set $\mathcal{M}^T \subseteq \mathcal{M}^+$ to be

$$\mathcal{M}^T = \{\nu \in \mathcal{M}^+ : \nu((-\infty,\alpha]) = 0, \mathbb{P}\left(B_{T^\nu} \in A\right) \le \mu(A) \;\; \forall A \in \mathcal{B}((\alpha,\infty))\}.$$

Our aim is to show that the supremum of this set, $\nu_* = \bigvee_{\nu\in\mathcal{M}^T} \nu$, is a non-zero element of $\mathcal{M}^T$, and that the stopping time $T^* \equiv T^{\nu_*}$ associated with this measure embeds μ. Since $T^* \le H_\alpha$ it is minimal by Theorem 6.

Our analysis of $\mathcal{M}^T$ begins with a statement of some basic properties. Recall that we are assuming that μ has support in $[\alpha,\infty)$. Suppose also that μ has an atom at α.

Lemma 15. *If $\mathbb{P}(H_\alpha < T) > 0$ and $\mu(\{\alpha\}) > 0$ then the set $\mathcal{M}^T$ has the following properties:*

(i) there exists a measure $\bar\nu \in \mathcal{M}^+$ satisfying $\bar\nu(\mathbb{R}) < \infty$ and such that $\nu \in \mathcal{M}^T$ implies $\nu \preceq \bar\nu$;

(ii) if $\nu_1 \preceq \nu_2$ then $T^{\nu_2} \le T^{\nu_1}$;

(iii) if $\nu_1, \nu_2 \in \mathcal{M}^T$ then their supremum $\nu_1 \vee \nu_2 \in \mathcal{M}^T$;

(iv) if $\nu_n \in \mathcal{M}^T$ are a sequence of measures and $\nu_n \uparrow \nu \in \mathcal{M}^+$ in the sense that $\nu_n(A)$ is increasing for all $A \in \mathcal{B}(\mathbb{R})$ and

$$\lim_n \sup_A \{|\nu_n(A) - \nu(A)| : A \in \mathcal{B}(\mathbb{R})\} = 0,$$

then $\nu \in \mathcal{M}^T$.

(v) if $\nu \in \mathcal{M}^T$ and we define a measure ρ with support in (α, ∞) such that for $A \in \mathcal{B}(\alpha, \infty)$, $\rho(A) = \mu(A) - \mathbb{P}(B_{T^\nu} \in A)$ then $\nu' = \nu + \rho$ is also an element of $\mathcal{M}^T$.

Proof. (i) Suppose $\nu \in \mathcal{M}^T$ and $x > 0$. Then $\mathbb{E}\left(B_{H_x \wedge T^\nu}\right) = 0$, so

$$\begin{aligned} |m| &= \mathbb{E}(B_{H_x \wedge T^\nu}) - \mathbb{E}(B_{T^\nu}) \\ &= \mathbb{E}(B_{H_x} - B_{T^\nu}; H_x \leq T^\nu) \\ &\leq (x - \alpha)\mathbb{P}(H_x \leq T^\nu). \end{aligned}$$

It follows that

$$\mathbb{P}\left(H_x \leq T^\nu\right) \geq \frac{|m|}{x - \alpha}. \tag{22}$$

In particular, for all $x > 0$, $\mathbb{P}(T^\nu \geq H_x) > 0$ and hence it must be the case that $\nu([0, x)) < 1$ for all $x > 0$.

Define a measure $\bar{\nu}$ by

$$\bar{\nu}(dx) = \begin{cases} \frac{(x-\alpha)\mu(dx)}{|m|} & x \geq 0 \\ \frac{\mu(dx)}{\mu([\alpha,x])} & \alpha < x < 0 \\ 0 & x \leq \alpha \end{cases}$$

We interpret statements about measures such as the above as shorthand for statements about the integrals over general functions. So we will write $\nu_1(dx) \leq \nu_2(dx)$ to mean that $\int f d\nu_1 \leq \int f d\nu_2$ for all positive, measurable functions f. We aim to show that $\bar{\nu}$ is an upper bound for elements of $\mathcal{M}^T$. First note that $\bar{\nu}(\mathbb{R}) < \infty$, since μ has a well defined first moment, and by assumption μ has an atom at α.

Fix $\nu \in \mathcal{M}^T$, and for $x \geq 0$ consider the probability that we stop in dx under T^ν. By definition this probability is bounded above by $\mu(dx)$. Conversely, one way to stop at x, is to be stopped by T^ν on first reaching x. Hence

$$\mu(dx) \geq \mathbb{P}\left(H_x \leq T^\nu\right)\nu(dx) \geq \frac{|m|}{(x - \alpha)}\nu(dx),$$

and, for $x > 0$, $\nu(dx) \leq \bar{\nu}(dx)$.

If $x < 0$, since $\mathbb{P}\left(B_{T^\nu} \in A\right) \leq \mu(A)$ for all $A \in \mathcal{B}(\mathbb{R} \setminus \{\alpha\})$, we must have

$$\mathbb{P}\left(H_x \leq T^\nu\right) \geq \mu([\alpha, x]).$$

By similar arguments to above we find that $\nu \in \mathcal{M}^T$ must satisfy

$$\nu(dx) \leq \frac{\mu(dx)}{\mu([\alpha, x])} = \bar{\nu}(dx).$$

(ii) This is immediate since $\nu_1 \leq \nu_2$ implies $S^{\nu_1} \geq S^{\nu_2}$.

(iii) We wish to show that for any two measures $\nu_1, \nu_2 \in \mathcal{M}^T$, their supremum $\nu_\vee = \nu_1 \vee \nu_2$ is also in $\mathcal{M}^T$.

For $x > \alpha$ and a measure $\nu \in \mathcal{M}^T$

$$\{B_{T^\nu} \in dx\} = \{B_T \in dx, T \leq S^\nu \wedge H_\alpha\} \cup \{B_{S^\nu} \in dx, S^\nu < T \wedge H_\alpha\}. \tag{23}$$

If $\nu' \in \mathcal{M}^T$ and $\nu \preceq \nu'$, then $S^{\nu'} \leq S^\nu$ and

$$\{B_T \in dx, T \leq S^{\nu'} \wedge H_\alpha\} \subseteq \{B_T \in dx, T \leq S^\nu \wedge H_\alpha\}. \tag{24}$$

Now we consider the term $\{B_{S^\nu} \in dx, S^\nu < T \wedge H_\alpha\}$. Suppose $x > 0$ (the case $x < 0$ is similar). Fix a Brownian path $\omega = (B_t)_{0 \leq t \leq T}$, and let $J = J_{H_x \wedge H_\alpha} = \inf\{B_t; t \leq H_x \wedge H_\alpha\}$. The interesting case is when $T(\omega) > H_x(\omega)$. Conditional on such a path ω

$$\begin{aligned} \mathbb{P}(B_{S^\nu} \in dx, S^\nu < T \wedge H_\alpha | \omega) &= \mathbb{P}(G_+^\nu \in dx, G_-^\nu < J)\mathbf{1}_{\{J > \alpha\}} \\ &= \nu(dx)(1 - \nu([J, 0)))\mathbf{1}_{\{J > \alpha\}}. \end{aligned}$$

Without loss of generality suppose $\nu_1(dx) \geq \nu_2(dx)$. Then

$$\nu_\vee(dx)(1 - \nu_\vee([J, 0))) \leq \nu_1(dx)(1 - \nu_1([J, 0)))$$

and

$$\mathbb{P}(B_{S^{\nu_\vee}} \in dx, S^{\nu_\vee} < T \wedge H_\alpha | \omega) \leq \mathbb{P}(B_{S^{\nu_1}} \in dx, S^{\nu_1} < T \wedge H_\alpha | \omega).$$

Combining this result with the decomposition (23) and the set inequality (24) (with $\nu' = \nu_\vee$ and $\nu = \nu_1$) we deduce

$$\mathbb{P}(B_{T^{\nu_\vee}} \in dx) \leq \mathbb{P}(B_{T^{\nu_1}} \in dx) \leq \mu(dx).$$

Hence $\nu_\vee = \nu_1 \vee \nu_2 \in \mathcal{M}^T$.

(iv) Suppose now we have a sequence of measures $\nu_n \in \mathcal{M}^T$ such that $\nu_n \uparrow \nu$. Recall the definitions of G_+^ν and G_-^ν and the fact that $\overline{\nu}(\mathbb{R}) < \infty$. Define

$$\begin{aligned} G_+^n &= \inf\{x \geq 0 : \nu_n([0, x]) \geq X_+\} \\ G_-^n &= \sup\{x \leq 0 : \nu_n([x, 0)) \geq X_-\} \end{aligned}$$

where the random variables X_+ and X_- are the same random variables as those used in the definition of G_+^ν and G_-^ν. Here G_+^n is a shorthand for $G_+^{\nu_n}$. Then, for example, $\mathbb{P}(G_+^n \in A) = \nu_n(A)$ for $A \in \mathcal{B}([0, \infty))$. Consequently,

$$\mathbb{P}(G_+^n \in A) \uparrow \mathbb{P}(G_+^\nu \in A) \tag{25}$$

for $A \in \mathcal{B}([0,\infty))$, and similarly for G^ν_-, G^n_-. Now consider a fixed path and stopping time (so T could be determined by some independent randomisation as well as the path) of the Brownian motion, $\omega = (B_t)_{t \le T}$, and a set $A \in \mathcal{B}(\mathbb{R})$. Then there exists a set $F = F(\omega, A) \in \mathcal{B}([0,\infty)) \times \mathcal{B}((-\infty,0))$ such that

$$B_{T^\nu} \in A \iff (G^\nu_+, G^\nu_-) \in F.$$

Note that for a fixed Brownian path, the event $B_{T^\nu} \in A$ depends on the measure ν only via the random variables G^ν_+ and G^ν_-. In particular for a different measure such as ν_n we have $B_{T^{\nu_n}} \in A$ if and only if $(G^n_+, G^n_-) \in F$ for the same set $F = F(\omega, A)$.

Now

$$|\mathbb{P}(B_{T^\nu} \in A) - \mathbb{P}(B_{T^{\nu_n}} \in A)| \le \mathbb{E}|\mathbb{P}(B_{T^\nu} \in A|\omega) - \mathbb{P}(B_{T^{\nu_n}} \in A|\omega)|.$$

Further, since $\mathbb{P}(B_{T^\nu} \in A|\omega) = \mathbb{P}((G^\nu_+, G^\nu_-) \in F(\omega))$ we have that

$$\begin{aligned}&\mathbb{P}(B_{T^\nu} \in A|\omega) - \mathbb{P}(B_{T^{\nu_n}} \in A|\omega) \\ &\qquad = \mathbb{P}((G^\nu_+, G^\nu_-) \in F(\omega)) - \mathbb{P}((G^n_+, G^n_-) \in F(\omega))\end{aligned}$$

which tends to zero using (25) and its analogue for S_-. We conclude that $|\mathbb{P}(B_{T^\nu} \in A) - \mathbb{P}(B_{T^{\nu_n}} \in A)| \to 0$ and hence that $\nu \in \mathcal{M}^T$.

(v) Suppose that $\nu \in \mathcal{M}^T$ and define the measures ρ and ν' with support (α, ∞) via $\rho(A) = \mu(A) - \mathbb{P}(B_{T^\nu} \in A)$ and $\nu' = \nu + \rho$. We wish to show that ν' is also in $\mathcal{M}^T$.

As before, for $x > \alpha$ we have

$$\{B_{T^{\nu'}} \in dx\} = \{B_T \in dx, T \le S^{\nu'} \wedge H_\alpha\} \cup \{B_{S^{\nu'}} \in dx, S^{\nu'} < T \wedge H_\alpha\},$$

and

$$\{B_T \in dx, T \le S^{\nu'} \wedge H_\alpha\} \subseteq \{B_T \in dx, T \le S^\nu \wedge H_\alpha\}. \tag{26}$$

Now suppose $x > 0$; the case $x < 0$ is similar. Conditional on a path $\omega = (B_t)_{0 \le t \le T}$ with the property that $H_x < T$, and with $J = J_{H_x \wedge H_\alpha} = \inf\{B_t; t \le H_x \wedge H_\alpha\}$ as before,

$$\mathbb{P}(B_{S^{\nu'}} \in dx; S^{\nu'} < T \wedge H_\alpha|\omega) = \nu'(dx)(1 - \nu'([J,0)))\mathbf{1}_{\{(J>\alpha)\}}.$$

It follows that

$$\begin{aligned}\mathbb{P}(B_{S^{\nu'}} \in dx; S^{\nu'} < T \wedge H_\alpha|\omega) &= (\nu(dx) + \rho(dx))(1 - \nu'([J,0)))\mathbf{1}_{\{(J>\alpha)\}} \\ &\le \nu(dx)(1 - \nu([J,0)))\mathbf{1}_{\{(J>\alpha)\}} + \rho(dx) \\ &= \mathbb{P}(B_{S^\nu} \in dx; S^\nu < T \wedge H_\alpha|\omega) + \rho(dx).\end{aligned}$$

Averaging over the Brownian paths, and combining this result with (26) we find

$$\mathbb{P}\left(B_{T^{\nu'}} \in dx\right) \le \mathbb{P}\left(B_{T^\nu} \in dx\right) + \rho(dx) = \mu(dx)$$

and hence $\nu' \in \mathcal{M}^T$. $\square$

We now show that T^*, our candidate for the minimal reduction of T, is an embedding. Suppose $\mathbb{P}(T > H_\alpha) > 0$, and also for the moment suppose that $\mu(\{\alpha\}) > 0$, so we may apply Lemma 15. The zero measure is an element of $\mathcal{M}^T$ and therefore by Lemma 15(v), $\mathcal{M}^T$ contains a non-zero element. Now take an increasing sequence ν_i of measures such that $\nu_i \uparrow \nu_* = \bigvee \mathcal{M}^T$. We know that $\nu_* \in \mathcal{M}^T$, (Lemma 15(iv)) so the law of B_{T^*} is dominated by μ. Conversely the law of T^* also dominates μ by (v). Hence T^* must embed μ.

We now note that we may drop the assumption that μ has an atom at α. Suppose μ has no atom at α. By (21) the law $\tilde{\mu}$ of $B_{T\wedge H_\alpha}$ does. Define

$$\mu^{(n)} = \frac{1}{n}\tilde{\mu} + \frac{n-1}{n}\mu,$$

and let $\mathcal{M}_n^T$ be the associated set of measures. Then for $k \geq n$, $\mathcal{M}_n^T \subseteq \mathcal{M}_k^T$ and so $T^{\nu_*^{(k)}} \leq T^{\nu_*^{(n)}}$ (where we write $\nu_*^{(k)}$ for the maximal element of $\mathcal{M}_k^T$). So as $n \to \infty$, $T^{\nu_*^{(n)}} \downarrow T^*$, which must therefore embed μ. Since $T^* \leq H_\alpha$, it must also be minimal.

We now consider the case of measures μ where the support is not bounded below. Define the measure μ_n by

$$\begin{aligned} \mu_n((x,y)) &= \mu((x,y)) \qquad \forall x,y \geq -n; \\ \mu_n(\{-n\}) &= \mu((-\infty,-n]); \\ \mu_n((-\infty,-n)) &= 0. \end{aligned}$$

Then for sufficiently large n, $\int x\, \mu_n(dx) < 0$. Also, $\mathcal{L}(B_{T\wedge H_{-n}})$ is dominated by both μ and μ_n on $(-n,\infty)$.

First consider the problem of embedding μ_n in $B_{T\wedge H_{-n}\wedge t}$ — that is finding a stopping time $T_n \leq T \wedge H_{-n}$ such that $\mathcal{L}(B_{T_n}) = \mu_n$. The construction above tells us that there exists a measure ν_n for which $T_n = T \wedge H_{-n} \wedge H_{G_+^n} \wedge H_{G_-^n}$ embeds μ_n, and is minimal, where

$$\begin{aligned} G_+^n &= \inf\{x \geq 0 : \nu_n([0,x]) \geq X_+\}, \\ G_-^n &= \sup\{x \leq 0 : \nu_n([x,0)) \geq X_-\}, \end{aligned}$$

for independent random variables X_+, X_- uniformly distributed on $[0,1]$.

Let n_0 be such that $\int x \mu_{n_0}(dx) < 0$. For $n_2 \geq n_1 \geq n_0$ we have that $\nu_{n_2}|_{(-n_1,\infty)} \in \mathcal{M}_{n_1}^T$ since for $A \in \mathcal{B}((-n_1,\infty))$,

$$\mathbb{P}(B_{T_{n_2}\wedge H_{-n_1}} \in A) \leq \mu_{n_2}(A) = \mu_{n_1}(A).$$

Hence for $A \in \mathcal{B}((-n_0,\infty))$, $\nu_n(A)$ decreases as n increases and we may define a measure ν_∞ by

$$\nu_\infty = \liminf_{n\to\infty} \nu_n = \sup_{k\geq 0} \inf_{n\geq k} \nu_n,$$

where the final representation ensures that ν_∞ is a measure. Our goal is to show that $T^* = S^{\nu_\infty} \wedge T$ embeds μ, and to use the fact that T_n is minimal for μ_n to deduce that T^* is minimal for μ. To this end we want to construct a coupling of the stopping times $(T_n)_{n\geq n_0}$ and T^*.

Let $\tilde{\nu}_n$ with support $(-n,\infty)$ be given by

$$\begin{array}{rcll} \tilde{\nu}_n([0,x]) & = & \frac{\nu_n([0,x])-\nu_\infty([0,x])}{1-\nu_\infty([0,x])} & x \geq 0; \\ \tilde{\nu}_n([x,0)) & = & \frac{\nu_n([x,0))-\nu_\infty([x,0))}{1-\nu_\infty([x,0))} & x \in (-n,0). \end{array}$$

Let $\tilde{X}_+, X_+^\infty, \tilde{X}_-, X_-^\infty$ be independent random variables, independent also of B and uniformly distributed on $[0,1]$, and define the levels

$$\begin{array}{rcl} \tilde{G}_+^n & = & \inf\{x \geq 0 : \tilde{\nu}_n([0,x]) \geq \tilde{X}_+\}, \\ G_+^\infty & = & \inf\{x \geq 0 : \nu_\infty([0,x]) \geq X_+^\infty\}, \\ \tilde{G}_-^n & = & \sup\{x \leq 0 : \tilde{\nu}_n([x,0)) \geq \tilde{X}_-\}, \\ G_-^\infty & = & \sup\{x \leq 0 : \nu_\infty([x,0)) \geq X_-^\infty\}, \\ \bar{G}_+^n & = & \tilde{G}_+^n \wedge G_+^\infty, \\ \bar{G}_-^n & = & \tilde{G}_-^n \vee G_-^\infty. \end{array}$$

Note that $\tilde{\nu}_n([0,x]) \downarrow 0$ as $n \uparrow \infty$ and hence $\tilde{G}_+^n \uparrow \infty$ almost surely. For $x > 0$ we have

$$\mathbb{P}(\bar{G}_+^n > x) = \mathbb{P}(\tilde{G}_+^n > x)\mathbb{P}(G_+^\infty > x) = 1 - \nu_n([x,0))$$

so that $\bar{G}_+^n$ has the same law as G_+^n defined above.

Similar calculations can be made for $x < 0$, and, using the fact that the pair $(\bar{G}_+^n, \bar{G}_-^n)$ has the same law as (G_+^n, G_-^n), we can deduce that $T_n^* = T \wedge H_{-n} \wedge H_{\bar{G}_+^n} \wedge H_{\bar{G}_-^n}$ embeds μ_n and is minimal. Furthermore, from the fact that $\tilde{G}_+^n$ and $\tilde{G}_-^n$ increase to infinity almost surely we have

$$T_n^* \uparrow T^* = T \wedge H_{G_+^\infty} \wedge H_{G_-^\infty}. \tag{27}$$

Since T_n^* embeds μ_n we find that T^* embeds μ.

Finally we show that T^* is minimal.

Proposition 16. *Suppose that T_n embeds μ_n, μ_n converges weakly to μ_∞ and $T_n \uparrow T_\infty < \infty$, almost surely. Then T_∞ embeds μ_∞.*

If also $l_n \to l_\infty < \infty$ where $l_n = \int |x|\mu_n(dx)$ and $l_\infty = \int |x|\mu_\infty(dx)$, and T_n is minimal for μ_n, then T_∞ is minimal for μ_∞.

Proof. The first part is clear, so we restrict ourselves to proving the minimlaity of T_∞ under the stated assumptions.

Suppose that $S \leq T_\infty$ and $A' \in \mathcal{F}_S$. We want to show $\mathbb{E}(B_{T_\infty}; A') \leq \mathbb{E}(B_S; A')$ for certain stopping times. Let $A = A' \cap \{S < T_\infty\}$ and $A_n = A' \cap \{S < T_n\}$. Then

$$\mathbb{E}(B_{T_\infty}; A') \leq \mathbb{E}(B_S; A') \iff \mathbb{E}(B_{T_\infty}; A) \leq \mathbb{E}(B_S; A) \tag{28}$$

so we restrict our attention to sets A. Note that $A_n \uparrow A$.

Since T_n is minimal and $A_n \in \mathcal{F}_{S \wedge T_n}$

$$\mathbb{E}(B_{T_n}; A_n) \leq \mathbb{E}(B_{S \wedge T_n}; A_n) = \mathbb{E}(B_S; A_n)$$

so we deduce that both sides of (28) hold provided:

$$\lim_n \mathbb{E}(B_{T_n}; A_n) = \mathbb{E}(B_{T_\infty}; A), \tag{29}$$
$$\lim_n \mathbb{E}(B_S; A_n) = \mathbb{E}(B_S; A). \tag{30}$$

For (29) we consider $|\mathbb{E}(B_{T_\infty}; A) - \mathbb{E}(B_{T_n}; A_n)|$. Then

$$|\mathbb{E}(B_{T_\infty}; A) - \mathbb{E}(B_{T_n}; A_n)| \leq \mathbb{E}(|B_{T_\infty}|; A \setminus A_n) + \mathbb{E}(|B_{T_\infty} - B_{T_n}|; A_n)$$

and the first term on the right tends to zero by dominated convergence. For the second term we show $\mathbb{E}(|B_{T_\infty} - B_{T_n}|) \to 0$. Fix $\varepsilon > 0$. We have

$$|B_{T_\infty} - B_{T_n}| \leq |B_{T_n}| - |B_{T_\infty}| + 2|B_{T_\infty}|\mathbf{1}_{\{T_n \leq T_\infty - \varepsilon\}} + 2|B_{T_\infty} - B_{T_n}|\mathbf{1}_{\{T_n > T_\infty - \varepsilon\}}.$$

We take expectations and let $n \to \infty$. By the definition of μ_n the first two terms cancel each other out, while the third tends to zero by dominated convergence. For the last term

$$\mathbb{E}(|B_{T_\infty} - B_{T_n}|; T_n > T_\infty - \varepsilon) \leq \mathbb{E}(|B_{T_n+\varepsilon} - B_{T_n}|) = \mathbb{E}(|B_\varepsilon|) = \sqrt{\frac{\varepsilon}{2\pi}}.$$

Consequently, in the limit, $\mathbb{E}(|B_{T_\infty} - B_{T_n}|; T_n > T_\infty - \varepsilon) \to 0$ and (29) holds.

By Lemma 9, in order to prove that T_∞ is minimal it is sufficient to show that (28) holds for each $\tilde{S} \leq T_\infty$ that embeds μ and the associated $\tilde{S}_\varepsilon$. By the above arguments it is sufficient to show that (30) holds for $(\tilde{S}_\varepsilon)_{\varepsilon \geq 0}$. But $|B_{\tilde{S}_\varepsilon}| \leq |B_S| + |B_{T_\infty}| + \varepsilon$ so (30) follows easily in each case by dominated convergence. □

5 A maximal embedding for a non-centred target distribution subject to a condition on the maximum

In this section we are interested in finding an embedding to solve the following problem:

> Given a Brownian motion $(B_t)_{t\geq 0}$ and an integrable (but possibly not centred) target distribution μ with mean m, find a minimal stopping time T such that T embeds μ and
>
> $$\mathbb{P}(M_T \geq x)$$
>
> is maximised for all x and over all minimal stopping times T embedding μ.

We call an embedding with this property the max-max embedding, and denote it by T_{max}.

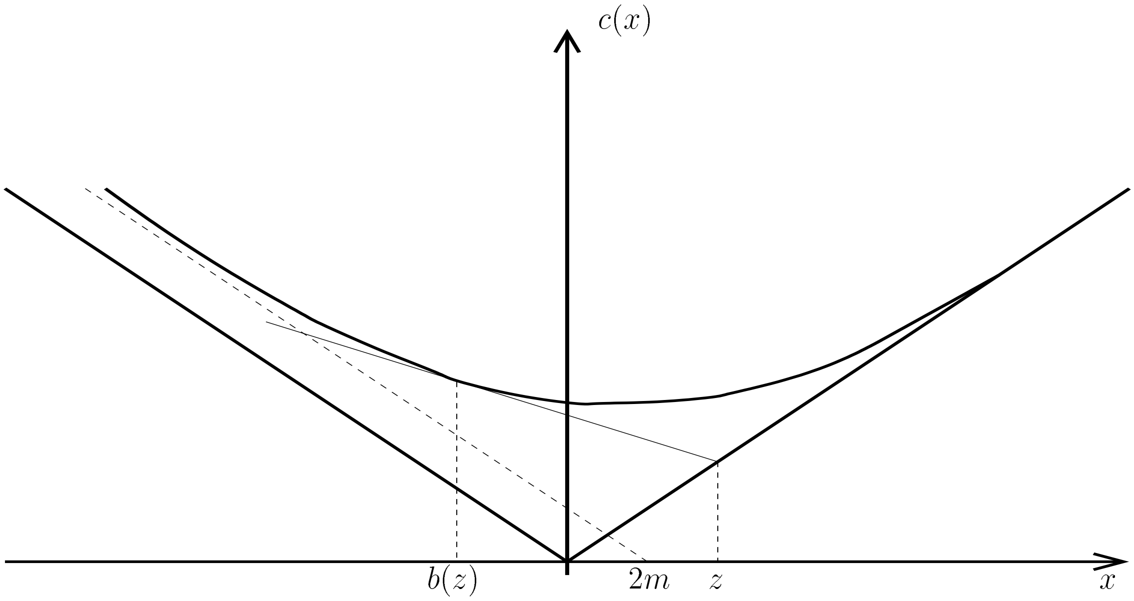

Figure 2: $c(x)$ for a μ with support bounded above, and positive non-zero mean m. Also shown is an intuitive idea of $b(z)$.

Without some condition on the class of admissible stopping times the problem is clearly degenerate — any stopping time may be improved upon by waiting for the first return of the process to 0 after hitting level x and then using the original embedding. For this improved embedding $\mathbb{P}(M_T \geq x) = 1$. Further, since no almost surely finite stopping time can satisfy

$$\mathbb{P}(M_T \geq x) = 1$$

for all $x > 0$, there can be no solution to the problem above in the class of all embeddings. As a consequence some restriction on the class of admissible stopping times is necessary for us to have a well defined problem.

Various conditions have been proposed in the literature to restrict the class of stopping times. In the case where $m = 0$, the condition on T that $B_{t \wedge T}$ is a UI martingale has been suggested Dubins and Gilat (1978), and in this case the maximal embedding is the Azema-Yor embedding. When $m = 0$ Monroe (1972) tells us that minimality and uniform integrability are equivalent conditions, so the Azema-Yor stopping time is the max-max embedding. For the case where $m > 0$, Pedersen and Peskir (2001) showed that $\mathbb{E}M_T < \infty$ is another suitable condition, with the optimal embedding being based on that of Azema and Yor. We argue that the class of minimal embeddings is the appropriate class for the problem under consideration since minimality is a natural and meaningful condition, which makes sense for all m (and which, for $m > 0$, includes as a subclass those embeddings with $\mathbb{E}(M_T) < \infty$.)

We now describe the construction of the candidate max-max stopping time. There is some difference in the proofs of embedding and maximality between the cases where $m > 0$ and $m < 0$, however the basic idea remains the same, and much

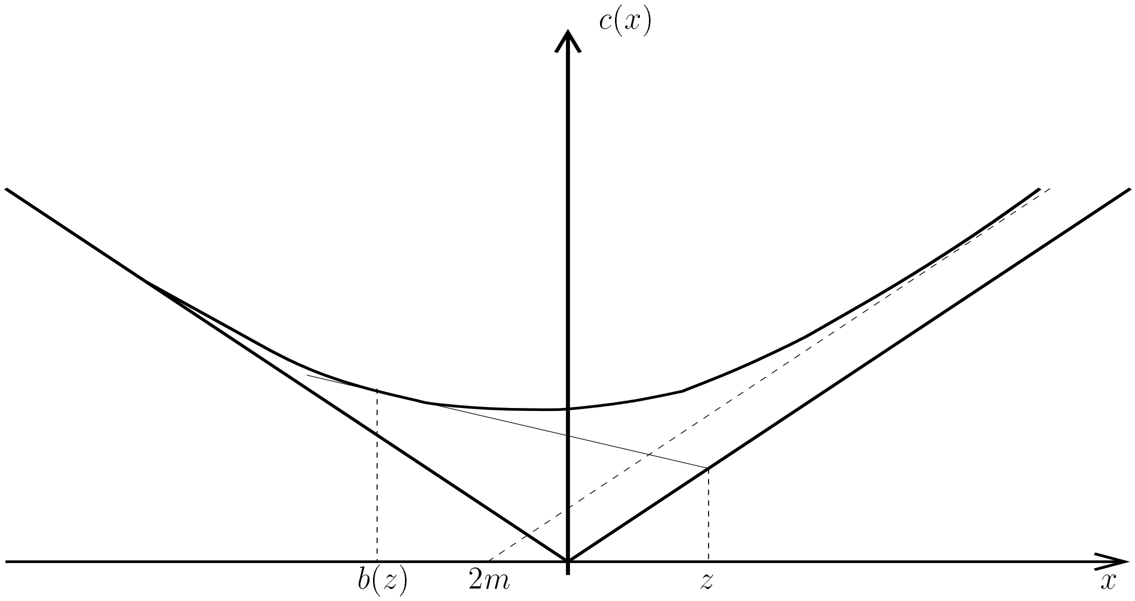

Figure 3: $c(x)$ for a distribution μ with strictly negative mean m. Now c is tangential to $x + |m|$ as $x \to \infty$.

of the following construction will apply for both cases. Figures 2 and 3 show how the constructions are related.

As a refinement of (3), define:

$$c(x) := \mathbb{E}^{\mu}|X - x| + |m|. \tag{31}$$

We note that as $x \to \pm\infty$, $c(x) - |x| \to |m| \mp m$. The refined function c has the same properties as before — it is convex and Lebesgue-almost everywhere differentiable. We maintain the same definitions for u, z_+ and b, so for $\theta \in [-1, 1]$, let

$$u(\theta) := \inf\{y \in \mathbb{R} : c(y) + \theta(x - y) \leq c(x), \forall x \in \mathbb{R}\},$$

$$z_+(\theta) := \frac{c(u(\theta)) - \theta u(\theta)}{1 - \theta},$$

and for $x \geq 0$

$$b(x) := u(z_+^{-1}(x)),$$

where z_+^{-1} is well defined. Finally we define the stopping time

$$T_{max} := \inf\{t > 0 : B_t \geq b(M_t)\}. \tag{32}$$

As mentioned above, for $m = 0$ this is exactly the Azema-Yor stopping time, while if $m > 0$, $b(x) = -\infty$ for $x < m$, and consequently $T_{max} \geq H_m$. So when $m > 0$ this embedding may be thought of as 'wait until the process hits m then use the Azema-Yor embedding.' This is also the embedding proposed by Pedersen and Peskir (2001) and discussed in Remark 14. Consequently, apart from the fact that we are considering a slightly more general class of stopping times, the original part of the subsequent theorem is the case in which $m < 0$ — the rest is included for completeness.

Theorem 17. *Let T be a stopping time of $(B_t)_{t\geq 0}$ which embeds μ and is minimal. Then for $x \geq 0$*

$$\mathbb{P}(M_T \geq x) \leq \left(\frac{1}{2} \inf_{\lambda < x} \frac{c(\lambda) - \lambda}{x - \lambda} \right). \tag{33}$$

Further T_{max} embeds μ, is minimal and attains equality in (33) for all $x \geq 0$.

Remark 18. Note that

$$\frac{c(\lambda) - \lambda}{x - \lambda} = 1 - \frac{x - c(\lambda)}{x - \lambda}. \tag{34}$$

We can relate the right-hand-side of (34) to the slope of a line joining (x, x) with $(\lambda, c(\lambda))$. In taking the infimum over λ we get a tangent to c and a value for the slope in $[-1, 1]$. Thus the bound on the right-hand-side of (33) lies in $[0, 1]$.

Remark 19. T_{max} has the property that it maximises the law of M_T over minimal stopping times which embed μ. If we want to minimise the law of the minimum, or equivalently we wish to maximise the law of $-J_T$, then we can deduce the form of the optimal stopping time by reflecting the problem about 0, or in other words by considering $-B$. Let T_{min} be the embedding which arises in this way, so that amongst the class of minimal stopping times which embed μ, the stopping time T_{min} maximises

$$\mathbb{P}(-J_T \geq x)$$

simultaneously for all $x \geq 0$.

The following lemma will be needed in the proof of the theorem.

Lemma 20. *Suppose $m < 0$ and T is minimal. Then for $x \geq 0$*

$$\mathbb{E}(B_{T \wedge H_x}) = 0.$$

Proof. For $x \geq 0$,

$$|B_{t \wedge T \wedge H_x}| \leq 2x - B_{t \wedge T \wedge H_x}$$

and thus

$$\mathbb{E}|B_{t \wedge T \wedge H_x}| \leq 2x - \mathbb{E}(B_{t \wedge T \wedge H_x}).$$

T is minimal so for the stopping time $S = t \wedge T \wedge H_x \leq T$, on taking expectations in (10), we get

$$\mathbb{E}(B_{t \wedge T \wedge H_x}) \leq \mathbb{E}(B_T) = m.$$

Thus $\mathbb{E}|B_{t \wedge T \wedge H_x}| \leq 2x + |m|$, and by dominated convergence

$$\mathbb{E}(B_{T \wedge H_x}) = \lim_{t \to \infty} \mathbb{E}(B_{t \wedge T \wedge H_x}) = 0.$$

□

We now turn to the proof of Theorem 17.

Proof. The following inequality for $x > 0$, $\lambda < x$ may be verified on a case by case basis:

$$\mathbf{1}_{\{M_T \geq x\}} \leq \frac{1}{x-\lambda}\left[B_{T\wedge H_x} + \frac{|B_T - \lambda| - (B_T + \lambda)}{2}\right]. \tag{35}$$

In particular, on $\{M_T < x\}$, (35) reduces to

$$0 \leq \begin{cases} 0 & \lambda \geq B_T \\ \frac{B_T - \lambda}{x-\lambda} & \lambda < B_T, \end{cases} \tag{36}$$

and on $\{M_T \geq x\}$ we get

$$1 \leq \begin{cases} \frac{x - B_T}{x-\lambda} & \lambda > B_T \\ 1 & \lambda \leq B_T. \end{cases} \tag{37}$$

Then taking expectations,

$$\mathbb{P}(M_T \geq x) \leq \frac{1}{x-\lambda}\left[\mathbb{E}(B_{T\wedge H_x}) + \frac{c(\lambda) - |m| - (m+\lambda)}{2}\right]. \tag{38}$$

If $m \leq 0$ then by Lemma 20 and the minimality of T we have $\mathbb{E}(B_{T\wedge H_x}) = 0$ and so

$$\mathbb{P}(M_T \geq x) \leq \frac{1}{2}\frac{c(\lambda) - \lambda}{x - \lambda}.$$

Conversely if $m > 0$, by Theorem 6 applied to $-B$,

$$m = \mathbb{E}(B_T) \geq \mathbb{E}(B_{T\wedge H_x}) \tag{39}$$

and so

$$\mathbb{P}(M_T \geq x) \leq \frac{1}{x-\lambda}\left[m + \frac{c(\lambda) - 2m - \lambda}{2}\right] = \frac{1}{2}\frac{c(\lambda) - \lambda}{x-\lambda}.$$

Since λ was arbitrary in either case, (33) must hold. It remains to show that T_{max} attains equality in (33), embeds μ and is minimal.

We begin by showing that it does attain equality in (33). Since

$$\frac{c(\lambda) - \lambda}{x - \lambda} = 1 + \frac{c(\lambda) - x}{x - \lambda}$$

the infimum in (33) is attained by a value λ^* with the property that a tangent of c at λ^* intersects the line $y = x$ at (x, x). By the definition of b we can choose $\lambda^* = b(x)$. In particular, since $\{M_{T_{max}} < x\} \subseteq \{B_{T_{max}} \leq b(x)\}$ and $\{M_{T_{max}} \geq x\} \subseteq \{B_{T_{max}} \geq b(x)\}$, the stopping time T_{max} attains equality almost surely in (36) and (37). Assuming that T_{max} is minimal, we are then done for $m \leq 0$. If $m > 0$ we do not always have equality in (39). If $x < m$ then $\mathbb{E}(B_{T_{max}\wedge H_x}) = x$, but then $\lambda^* = -\infty$ and so the extra term $(\mathbb{E}(B_{T_{max}\wedge H_x}) - m)/(x - \lambda^*) = 0$. As a result equality is again attained in (33). Otherwise, if $x \geq m$ then $T_{max} \geq H_m$ and the properties of the classical Azema-Yor embedding ensure that $\mathbb{E}(B_{T_{max}\wedge H_x}) = m$ and there is equality both in (39) and (33).

Fix a value of y which is less than the supremum of the support of μ, and recall that b^{-1} is defined to be left continuous. Then given (8) and the subsequent discussion (which remains valid even if $m \neq 0$), and since we now have equality in (33), we deduce:

$$
\begin{aligned}
\mathbb{P}(B_{T_{max}} \geq y) &= \mathbb{P}(M_{T_{max}} \geq b^{-1}(y)) \\
&= \frac{1}{2}\left[1 + \frac{c(b(b^{-1}(y))) - b^{-1}(y)}{b^{-1}(y) - b(b^{-1}(y))}\right] \\
&= \frac{1}{2}(1 - c'_{-}(y)) \\
&= \mu([y, \infty)).
\end{aligned}
$$

Hence T_{max} embeds μ.

In the case where $m > 0$, minimality of the stopping time is discussed in Remark 14, while the case where $m = 0$ is discussed in Remark 4. In the case where $m < 0$ we prove minimality by contradiction. Suppose there exists $S \leq T_{max}$ such that S embeds μ. By the construction in Section 4 we may assume that S is minimal.

Then the following must hold:

- $B_{T_{max}} = b(M_{T_{max}})$ by the definition of T_{max};
- $b(M_{T_{max}}) \geq b(M_S)$ since $T_{max} \geq S$ and $M_{T_{max}} \geq M_S$;
- $b(M_S) \leq B_S$ since $S \leq T_{max} = \inf\{u : B_u \leq b(M_u)\}$.

Hence

$$B_{T_{max}} = b(M_{T_{max}}) \geq b(M_S) \leq B_S.$$

If we can show $b(M_{T_{max}}) = b(M_S)$ a.s. then $B_{T_{max}} \leq B_S$ and since S and T_{max} embed the same law, $B_S = B_{T_{max}} = b(M_S)$. Thus $T_{max} \leq S$ and T_{max} is minimal.

For $y \in \mathbb{R}$, consider $B_{T_{max} \wedge H_{b^{-1}(y)}}$. This random variable is distributed according to μ on $(-\infty, y)$ with a mass of size $\mu([y, \infty))$ at $b^{-1}(y)$. By the definition of the barycentre $B_{T_{max} \wedge H_{b^{-1}(y)}}$ has mean zero.

Now consider $B_{S \wedge H_{b^{-1}(y)}}$. On the set $\{H_{b^{-1}(y)} < S\}$ we have $M_S \geq b^{-1}(y)$ and then $B_S \geq b(M_S) \geq b(b^{-1}(y))$. Since μ assigns no mass to $(b(b^{-1}(y)), y)$ we have $B_S \geq y$. Further since $B_S \sim \mu$, then $B_{S \wedge H_{b^{-1}(y)}}$ is distributed according to μ on $(-\infty, y)$ with perhaps some mass in $[y, b^{-1}(y))$ and the rest at $b^{-1}(y)$. However S is minimal, so $\mathbb{E}(B_{S \wedge H_{b^{-1}(y)}}) = 0$ by Lemma 20 and in fact $B_{S \wedge H_{b^{-1}(y)}}$ assigns no mass to $[y, b^{-1}(y))$. Thus

$$\mathbb{P}(M_S \geq b^{-1}(y)) = \mathbb{P}(M_{T_{max}} \geq b^{-1}(y))$$

and since $M_{T_{max}} \geq M_S$ we conclude that the sets $\{M_S \geq b^{-1}(y)\}$ and $\{M_{T_{max}} \geq b^{-1}(y)\}$ are almost surely equal.

Suppose now that $\mathbb{P}(b(M_{T_{max}}) > b(M_S)) > 0$. Then there exists a y such that $\mathbb{P}(b(M_{T_{max}}) > y > b(M_S)) > 0$. But

$$\{b(M_{T_{max}}) > y > b(M_S)\} \subseteq \{M_{T_{max}} > b^{-1}(y) \geq M_S\}$$

since b^{-1} is increasing and left-continuous, and the event on the right-hand side has zero probability. $\square$

6 An embedding to maximise the modulus

In Jacka (1988), Jacka shows how to embed a centred probability distribution in a Brownian motion so as to maximise $\mathbb{P}(\sup_{t\leq T}|B_t|\geq y)$. Our goal in this section is to extend this result to allow for non-centered target distributions with mean $m\neq 0$. In fact we solve a slightly more general problem. Let h be a measurable function; we will construct a stopping time T_{mod} which will maximise $\mathbb{P}(\sup_{t\leq T}|h(B_t)|\geq y)$ simultaneously for all y where the maximum is taken over the class of all minimal stopping times which embed μ. The reason for our generalisation will become apparent in the application in the next section.

Without loss of generality we may assume that h is a non-negative function with $h(0)=0$ and such that for $x>0$ both $h(x)$ and $h(-x)$ are increasing. To see this, observe that for arbitrary h we can define the function

$$\tilde{h}(x)=\begin{cases}\max_{0\leq y\leq x}|h(y)|-|h(0)| & x\geq 0;\\ \max_{x\leq y\leq 0}|h(y)|-|h(0)| & x<0.\end{cases}$$

Then $\tilde{h}$ has the desired properties and since

$$\sup_{s\leq T}|h(B_s)|=\sup_{s\leq T}|\tilde{h}(B_s)|+|h(0)|$$

the optimal embedding for $\tilde{h}$ will be an optimal embedding for h.

So suppose that h has the properties listed above. We want to find an embedding of μ in B which is minimal and which maximises the law of $\sup_{t\leq T}h(B_t)$. (Since h is non-negative we can drop the modulus signs.) Suppose also for definiteness that μ has a finite, positive mean $m=\int_{\mathbb{R}}x\mu(dx)>0$. In fact our construction will also be optimal when $m=0$ (the case covered by Jacka (1988)), but in order to avoid having to give special proofs for this case we will omit it.

We begin by making the definitions

$$z_+(\theta) \;:=\; \frac{c(u(\theta))-\theta u(\theta)}{1-\theta}, \tag{40}$$

$$z_-(\theta) \;:=\; \frac{c(u(\theta))-\theta u(\theta)}{1+\theta}, \tag{41}$$

and

$$\theta_0:=\inf\{\theta\in[-1,1]:h(z_+(\theta))\geq h(-z_-(\theta))\},$$

as pictured in Figure 4. Our optimal stopping time will take the following form. Run the process until it hits either $z_+(\theta_0)$ or $-z_-(\theta_0)$, and then embed the restriction of μ to $[u(\theta_0),\infty)$ or $(-\infty,u(\theta_0)]$ respectively (defining the target measures more carefully when there is an atom at $u(\theta_0)$). For the embeddings in the second part, we will use the constructions described in Section 5.

To be more precise about the measures we embed in the second step, define

$$p:=\mathbb{P}(H_{z_+(\theta_0)}<H_{-z_-(\theta_0)})=\frac{z_-(\theta_0)}{z_+(\theta_0)+z_-(\theta_0)},$$

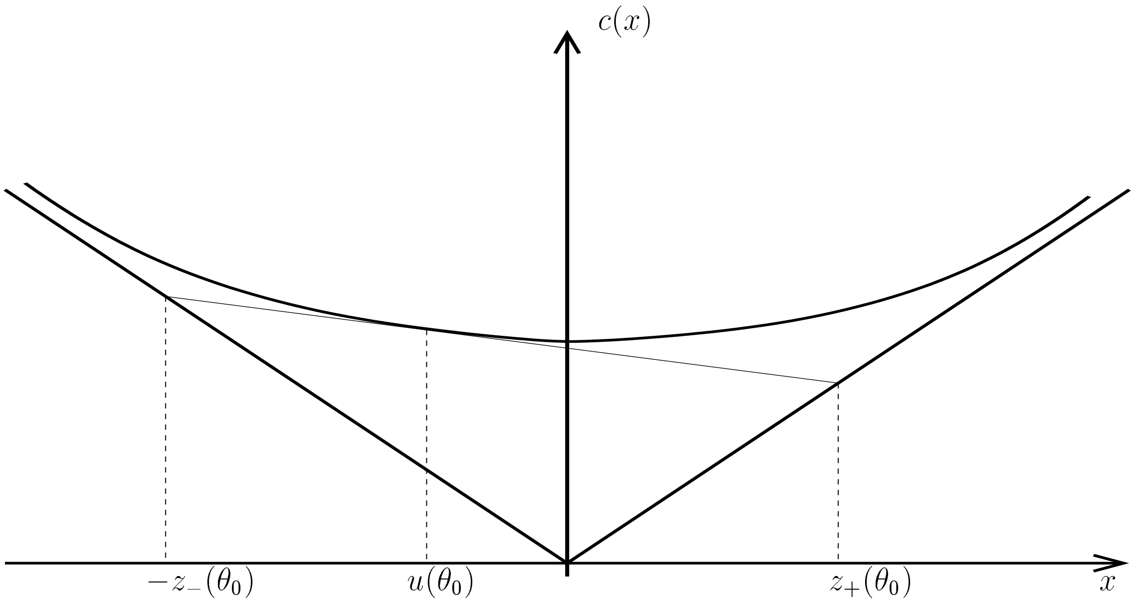

Figure 4: $c(x)$ for a distribution μ showing the construction of $z_+(\theta_0)$ and $z_-(\theta_0)$. The slope of the tangent is θ_0 where θ_0 has been chosen such that (assuming h is continuous) $h(z_+(\theta_0)) = h(-z_-(\theta_0))$.

and note

$$\theta_0 = \frac{z_+(\theta_0) - z_-(\theta_0)}{z_+(\theta_0) + z_-(\theta_0)} = 1 - 2p.$$

Then let μ_+ be the measure defined by

- $\mu_+(A) = p^{-1}\mu(A)$, $A \subseteq (u(\theta_0), \infty)$, A Borel;
- $\mu_+([u(\theta_0), \infty)) = 1$;
- $\mu_+((-\infty, u(\theta_0))) = 0$,

and similarly let μ_- be given by

- $\mu_-(A) = (1-p)^{-1}\mu(A)$, $A \subseteq (-\infty, u(\theta_0))$, A Borel;
- $\mu_-((-\infty, u(\theta_0)]) = 1$;
- $\mu_-((u(\theta_0), \infty)) = 0$.

The measure μ_+ (respectively μ_-) is obtained by conditioning a random variable with law μ to lie in the upper p^{th} (respectively lower $(1-p)^{\text{th}}$) quantile of its distribution.

Recall that

$$\begin{aligned} c(y) &= \int |w-y|\mu(dw) + |m| \\ &= 2\int_{\{w>y\}} (w-y)\mu(dw) - m + y + |m|. \end{aligned}$$

Then from the definition in (40) we have that

$$
\begin{aligned}
z_+(\theta_0) &= \frac{1}{2p}\Bigg(2\int_{\{w>u(\theta_0)\}}(w-u(\theta_0))\,\mu(dx) - m + u(\theta_0) \\
&\qquad + |m| - (1-2p)u(\theta_0)\Bigg) \\
&= \frac{1}{p}\int_{\{w>u(\theta_0)\}} w\,\mu(dw) + u(\theta_0)\left(1-\frac{1}{p}\int_{\{w>u(\theta_0)\}}\mu(dw)\right) \\
&\qquad + \frac{|m|-m}{2p} \\
&= \frac{1}{p}\int_{\{w>u(\theta_0)\}} w\,\mu(dw) + u(\theta_0)\left(1-\frac{1}{p}\mu((u(\theta_0),\infty))\right)
\end{aligned}
$$

where we have used $m = |m|$. In particular $z_+(\theta_0)$ is the mean of μ_+, since $\mu_+(\{u(\theta_0)\}) = 1 - \frac{1}{p}\mu((u(\theta_0),\infty))$. When we repeat the calculation for $z_-(\theta_0)$ we find that

$$
\begin{aligned}
-z_-(\theta_0) =& \frac{1}{1-p}\int_{\{w<u(\theta_0)\}} w\,\mu(dw) + u(\theta_0)\left(1-\frac{1}{1-p}\mu((-\infty,u(\theta_0))\right) \\
&- \frac{|m|+m}{2(1-p)}.
\end{aligned}
$$

Since $m > 0$ the final term does not disappear and $-z_-(\theta_0)$ is strictly smaller than the mean of μ_-.

We now describe the candidate stopping time $T_{mod} \equiv T^h_{mod}$. Note that this stopping time will depend implicitly on the function h via $z_\pm(\theta_0)$. Let

$$T_0 := \inf\{t > 0 : B_t \notin (-z_-(\theta_0), z_+(\theta_0))\},$$

and define

$$T_{mod} := \begin{cases} T^{\mu_+}_{max} \circ \theta_{T_0} & B_{T_0} = z_+(\theta_0) \\ T^{\mu_-}_{min} \circ \theta_{T_0} & B_{T_0} = -z_-(\theta_0). \end{cases}$$

Here we use θ_{T_0} to denote the shift operator, and $T^{\mu_+}_{max}$ is the stopping time constructed in Section 5 for a zero-mean target distribution, so that $T^{\mu_+}_{max}$ is a standard Azema-Yor embedding of the centred target law μ_+. (Recall that $z_+(\theta_0)$ is the mean of the corresponding part of the target distribution.) Similarly $T^{\mu_-}_{min}$ is the stopping time applied to $-B$ started at $-z_-(\theta_0)$ which maximises the law of the maximum of $-B$. In this case the mean of the target law μ_- is larger than $-z_-(\theta_0)$ so that in order to define $T^{\mu_-}_{min}$ we need to use the full content of Section 5 for embeddings of non-centred distributions.

The following theorem asserts that this embedding is indeed an embedding of μ, that it is minimal, and that it has the claimed optimality property.

Theorem 21. *Let μ be a target distribution such that $m > 0$. Then within the class of minimal embeddings of μ in B, the embedding T_{mod} as defined above has the property that it maximises*

$$\mathbb{P}\left(\sup_{t\leq T} h(B_t) \geq x\right)$$

simultaneously for all x.

Proof. By construction T_{mod} embeds μ. We need only show that it is optimal and minimal.

For $x \leq h(-z_-(\theta_0)) \wedge h(z_+(\theta_0))$ we know that the probability of the event $\{\sup_{t\leq T_{mod}} h(B_t) \geq x\}$ is one and so, for such x, T_{mod} is clearly optimal. Indeed if h is discontinuous at $-z_-(\theta_0)$ or $z_+(\theta_0)$ slightly more can be said. Note first that if $z_+(\theta_0)$ coincides with the supremum of the support of μ, then by Theorem 6(ii) and the minimality of T_{mod} (see below), the stopped Brownian motion can never go above $z_+(\theta_0)$. With this in mind let

$$L = \left(\lim_{y\uparrow -z_-(\theta_0)} h(y)\right) \wedge \left(\lim_{y\downarrow z_+(\theta_0)} h(y)\right)$$

if $z_+(\theta_0)$ is less than the supremum of the support of μ and

$$L = \left(\lim_{y\uparrow -z_-(\theta_0)} h(y)\right) \wedge h(z_+(\theta_0))$$

otherwise. Now take $x \leq L$. Then either $B_{T_0} = z_+(\theta_0)$ or $B_{T_0} = -z_-(\theta_0)$. If $B_{T_0} = z_+(\theta_0)$ then either $M_{T_{mod}} > z_+(\theta_0)$ almost surely and

$$\max_{0\leq t\leq T_{mod}} h(B_t) \geq \lim_{y\downarrow z_+(\theta_0)} h(y) \geq L$$

or $z_+(\theta_0)$ is the supremum of the support of μ and

$$\max_{0\leq t\leq T_{mod}} h(B_t) = h(B_{T_0}) \geq L.$$

Similar considerations apply for $B_{T_0} = -z_-(\theta_0)$ except that then $-J_{T_{mod}} > z_-(\theta_0)$ in all cases. We deduce that for $x \leq L$

$$\mathbb{P}\left(\sup_{t\leq T_{mod}} h(B_t) \geq x\right) = 1$$

and hence T_{mod} is optimal.

So suppose that $x > L$. For any stopping time T embedding μ, the following holds:

$$\mathbb{P}\left(\sup_{s\leq T} h(B_s) \geq x\right) \leq \mathbb{P}\left(h(M_T) \geq x\right) + \mathbb{P}\left(h(J_T) \geq x\right). \tag{42}$$

We will show that the embedding T_{mod} attains the maximal values of both terms on the right hand side, and further that for T_{mod} the two events on the right hand side are disjoint. Hence T_{mod} is optimal.

By the definition of θ_0, $x > (h(z_+(\theta_0))) \vee (h(-z_-(\theta_0)))$. It follows that

$$\mathbb{P}(h(M_{T_{mod}}) \geq x) = p\mathbb{P}(h(M_{T_{mod}}) \geq x | B_{T_0} = z_+(\theta_0))$$

and by the definition of T_{mod} and the properties of T_{max}, we deduce

$$\begin{aligned}\mathbb{P}(h(M_{T_{mod}}) \geq x) &= p\mathbb{P}(h(M_{T^\mu_{max}}) \geq x | M_{T^\mu_{max}} \geq z_+(\theta_0)) \\ &= \mathbb{P}(h(M_{T^\mu_{max}}) \geq x)\end{aligned}$$

where here T^μ_{max} is the embedding of Section 5 applied to μ. A similar calculation can be done for the minimum. In particular T_{mod} inherits its optimality property from the optimality of its constituent parts $T^{\mu+}_{max}$ and $T^{\mu-}_{min}$

Finally we note that T_{mod} is indeed minimal. Let $S \leq T_{mod}$ be a stopping time. We show that for $A \in \mathcal{F}_S$

$$\mathbb{E}(B_{T_{mod}}; A) \geq \mathbb{E}(B_S; A), \tag{43}$$

then minimality follows from Theorem 6. Observe that

$$B_T - B_S = B_T - B_{S\vee T_0} + B_{S \vee T_0} - B_S.$$

Then,

$$\mathbb{E}(B_T - B_{S\vee T_0}; A \cap \{T_0 = H_{z_+(\theta_0)}\}) = 0$$

by the properties of $T^{\mu+}_{max}$;

$$\mathbb{E}(B_T - B_{S\vee T_0}; A \cap \{T_0 = H_{-z_-(\theta_0)}\}) \geq 0$$

by the properties of $T^{\mu-}_{min}$ and

$$\mathbb{E}(B_{S\vee T_0} - B_S; A) = 0$$

since $B_{t \wedge T_0}$ is UI. If we add these various results then (43) follows. □

Remark 22. If the restrictions of h to $\mathbb{R}_+$ and $\mathbb{R}_-$ are strictly increasing then T_{mod} will be essentially the unique embedding which attains optimality in Theorem 21. If however h has intervals of constancy then other embeddings may also maximise the law of $\sup_{t \leq T} |h(B_t)|$.

7 Embeddings in diffusions

Our primary motivation in considering the embeddings of the previous sections was their use in the investigation of the following question:

> Given a regular (time-homogeneous) diffusion $(X_t)_{t\geq 0}$ and a target distribution μ_X, find (if possible) a minimal stopping time which embeds μ_X and which maximises the law of $\sup_{t\leq T} X_t$ (alternatively $\sup_{t\leq T} |X_t|$) among all such stopping times.

Note that in the martingale (or Brownian) case it is natural to consider centred target laws, at least in the first instance. However in the non-martingale case this restriction is no longer natural, and as we shall see below is completely unrelated to whether it is possible to embed the target law in the diffusion X. It was this observation which led us to consider the problem of embedding non-centred distributions in B.

The key idea is that we can relate the problem of embedding in a diffusion to the case where we are dealing with Brownian motion via the scale function. There exists a continuous, increasing function s such that $s(X_t)$ is a local martingale, and hence a time-change of Brownian motion. Then the requirement $X_{T^X} \sim \mu_X$ translates to finding an embedding of a related law in a Brownian motion B, and the criterion of maximising $\sup_{t \leq T} X_t$ also has an equivalent statement in terms of B.

We first recall the properties of the scale function (see e.g. Rogers and Williams (2000, V.45)). If $(X_t)_{t \geq 0}$ is a regular (time-homogeneous) diffusion on an interval $I \subseteq \mathbb{R}$ with absorbing or inaccessible endpoints and vanishing at zero, then there exists a continuous, strictly increasing scale function $s : I \to \mathbb{R}$ such that $Y_t = s(X_t)$ is a diffusion in natural scale on $s(I)$. We may also choose s such that $s(0) = 0$. In particular Y_t is (up to exit from the interior of $s(I)$) a time change of a Brownian motion with strictly positive speed measure. For definiteness we write $Y_t = B_{\tau_t}$.

We suppose also that our target distribution μ_X is concentrated on the interior of I. Then we may define a measure $\mu = \mu_Y$ on $s(I)^\circ$ by:

$$\mu(A) = \mu_Y(A) = \mu_X(s^{-1}(A)), \ A \in s(I)^\circ, \text{ Borel}$$

The original problem of embedding μ_X in X is equivalent to the problem of embedding μ_Y in Y before Y exits the interval $s(I)^\circ$. Since Y is a time change of a Brownian motion, we need only consider the problem of embedding μ_Y in a Brownian motion before exit from $s(I)^\circ$. If T is an embedding of μ in B then $T^X \equiv \tau^{-1}(T)$ is simultaneously an embedding of μ_Y in Y and μ_X in X.

The first question is when does any embedding exist? If we define the quantity $m = \int_I s(x)\mu_X(dx)$, then the following lemma (see Cox and Hobson (2002)) gives us necessary and sufficient conditions for an embedding to exist.

Lemma 23. *There are three different cases:*

(i) $s(I)^\circ = \mathbb{R}$, *in which case X is recurrent and we can embed any distribution μ_X on I° in X.*

(ii) $s(I)^\circ = (-\infty, \alpha)$ *(respectively $(-\alpha, \infty)$) for some $\alpha > 0$. Then we may embed μ_X in X if and only if m exists and $m \geq 0$ (resp. $m \leq 0$).*

(iii) $s(I)^\circ = (\alpha, \beta)$, *$\alpha < 0 < \beta$. Then we may embed μ_X in X if and only if $m = 0$.*

In each case it is clear that T^X is minimal for X if and only if T^X is minimal for Y if and only if $T = \tau(T^X)$ is minimal for B. Further, since $\mu = \mu_Y$ is concentrated on $s(I)^\circ$, if the stopping time T is minimal then T will occur before the Brownian motion leaves $s(I)^\circ$ (this is a consequence of Theorem 6 in case (ii) and Theorem 2 in case (iii)), and then T^X will be less than the first explosion time of X.

It is now possible to apply the results of previous sections to deduce a series of corollaries about embeddings of μ_X in X. Suppose that μ_X can be embedded in X or equivalently that μ_Y can be embedded in B, before the Brownian motion leaves $s(I)^\circ$. Let T_{max} and T^h_{mod} be the optimal embeddings of μ in B as defined in Sections 5 and 6. (Observe that from now on we make the dependence of T^h_{mod} on h explicit in the notation.) Then we can define T^X_{max} and $T^{X,h}_{mod}$ by

$$T^X_{max} = \tau^{-1} \circ T_{max} \qquad T^{X,h}_{mod} = \tau^{-1} \circ T^h_{mod}.$$

Corollary 24. *T^X_{max} is optimal in the class of minimal embeddings of μ_X in X in the sense that it maximises*

$$\mathbb{P}\left(\max_{t\leq T} X_t \geq y\right)$$

uniformly in $y \geq 0$.

Corollary 25. *$T^{X,h}_{mod}$ is optimal in the class of minimal embeddings of μ_X in X in the sense that it maximises*

$$\mathbb{P}\left(\max_{t\leq T}(h \circ s)(X_t) \geq y\right)$$

uniformly in $y \geq 0$.

Corollary 26. *$T^{X,|s^{-1}|}_{mod}$ is optimal in the class of minimal embeddings of μ_X in X in the sense that it maximises*

$$\mathbb{P}\left(\max_{t\leq T}|X_t| \geq y\right)$$

uniformly in $y \geq 0$.

References

Azéma, J. and M. Yor (1979), Une solution simple au problème de Skorokhod, *in* 'Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)', Springer, Berlin, pp. 90–115.

Bertoin, J. and Y. Le Jan (1992), 'Representation of measures by balayage from a regular recurrent point', *Ann. Probab.* **20**(1), 538–548.

Blackwell, David and Lester E. Dubins (1963), 'A converse to the dominated convergence theorem', *Illinois J. Math.* **7**, 508–514.

Chacon, R. V. and J. B. Walsh (1976), One-dimensional potential embedding, *in* 'Séminaire de Probabilités, X (Prèmiere partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)', Springer, Berlin, pp. 19–23. Lecture Notes in Math., Vol. 511.

Cox, A. M. G. and D. G. Hobson (2002), An optimal embedding for diffusions. University of Bath Mathematics Preprints 02/16, `http://www.maths.bath.ac.uk/MATHEMATICS/preprints.html`.

Doob, J. L. (1984), *Classical potential theory and its probabilistic counterpart*, Vol. 262 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, New York.

Dubins, L. E. (1968), 'On a theorem of Skorohod', *Ann. Math. Statist.* **39**, 2094–2097.

Dubins, Lester E. and David Gilat (1978), 'On the distribution of maxima of martingales', *Proc. Amer. Math. Soc.* **68**(3), 337–338.

Grandits, P. and N. Falkner (2000), 'Embedding in Brownian motion with drift and the Azéma-Yor construction', *Stochastic Process. Appl.* **85**(2), 249–254.

Jacka, S. D. (1988), 'Doob's inequalities revisited: a maximal H^1-embedding', *Stochastic Process. Appl.* **29**(2), 281–290.

Monroe, Itrel (1972), 'On embedding right continuous martingales in Brownian motion', *Ann. Math. Statist.* **43**, 1293–1311.

Pedersen, J. L. and G. Peskir (2001), 'The Azéma-Yor embedding in non-singular diffusions', *Stochastic Process. Appl.* **96**(2), 305–312.

Perkins, E. (1986), The Cereteli-Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion, *in* 'Seminar on stochastic processes, 1985 (Gainesville, Fla., 1985)', Birkhäuser Boston, Boston, MA, pp. 172–223.

Rogers, L. C. G. and D. Williams (2000), *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition.

Root, D. H. (1969), 'The existence of certain stopping times on Brownian motion', *Ann. Math. Statist.* **40**, 715–718.

Rost, H. (1971), 'The stopping distributions of a Markov Process', *Invent. Math.* **14**, 1–16.

Rost, H. (1976), Skorokhod stopping times of minimal variance, *in* 'Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)', Springer, Berlin, pp. 194–208. Lecture Notes in Math., Vol. 511.

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. (1992), 'Quelques inégalités avec le temps local en zero du mouvement brownien', *Stochastic Process. Appl.* **41**(1), 117–155.