The Annals of Applied Probability
2013, Vol. 23, No. 5, 2020–2052
DOI: 10.1214/12-AAP893

MAXIMIZING FUNCTIONALS OF THE MAXIMUM IN THE SKOROKHOD EMBEDDING PROBLEM AND AN APPLICATION TO VARIANCE SWAPS

BY DAVID HOBSON AND MARTIN KLIMMEK[1]

University of Warwick

The Azéma–Yor solution (resp., the Perkins solution) of the Skorokhod embedding problem has the property that it maximizes (resp., minimizes) the law of the maximum of the stopped process. We show that these constructions have a wider property in that they also maximize (and minimize) expected values for a more general class of bivariate functions $F(W_\tau, S_\tau)$ depending on the joint law of the stopped process and the maximum. Moreover, for monotonic functions g, they also maximize and minimize $\mathbb{E}[\int_0^\tau g(S_t)\,dt]$ amongst embeddings of μ, although, perhaps surprisingly, we show that for increasing g the Azéma–Yor embedding minimizes this quantity, and the Perkins embedding maximizes it.

For $g(s) = s^{-2}$ we show how these results are useful in calculating model independent bounds on the prices of variance swaps.

Along the way we also consider whether μ_n converges weakly to μ is a sufficient condition for the associated Azéma–Yor and Perkins stopping times to converge. In the case of the Azéma–Yor embedding, if the potentials at zero also converge, then the stopping times converge almost surely, but for the Perkins embedding this need not be the case. However, under a further condition on the convergence of atoms at zero, the Perkins stopping times converge in probability (and hence converge almost surely down a subsequence).

1. Introduction. Let $W = (W_t)_{t\geq 0}$ be Brownian motion, null at 0, and μ a centered probability measure. Then the Skorokhod embedding problem (SEP) (Skorokhod [21]) is to find a stopping time τ such that the stopped process satisfies $W_\tau \sim \mu$. There are many classical solutions to this problem (for a survey, see Obłój [16]), and further solutions continue to appear in the literature, including most recently Hirsch et al. [9]. Further impetus to the investigation of old and new solutions is derived from the connections between solutions of the SEP and model independent bounds for the prices of options; for a survey, see Hobson [10].

Given the multiplicity of solutions to the SEP, it is natural to search for embeddings with additional optimality properties. In particular, if Ψ is a functional of the stopped Brownian path $(W_t)_{0\leq t\leq\tau}$, then these constructions aim to maximize

Received February 2012; revised September 2012.
[1]Supported from EPSRC funding.
MSC2010 subject classifications. Primary 60G40; secondary 60G44, 60J65, 91G20, 93E20.
Key words and phrases. Skorokhod embedding problem, Azema–Yor solution, Perkins solution, variance swaps.

Ψ over (a suitable subclass of) embeddings of μ. For example, if F is an increasing function, and $S_t = \sup_{s\le t} W_s$, then the Azéma–Yor solution [2] maximizes $\mathbb{E}[F(S_\tau)]$ over uniformly integrable embeddings, and the Perkins embedding [17] minimizes the same quantity.

Our goal in this paper is to extend this result to functions $F = F(W_\tau, S_\tau)$. Then, subject to regularity conditions, our first result (Theorem 5.3) is that:

> Suppose $F_s(w,s)/(s-w)$ is monotonic decreasing in w. Then $\mathbb{E}[F(W_\tau, S_\tau)]$ is minimized (resp., maximized) over uniformly integrable embeddings τ of μ by the Azéma–Yor (resp., Perkins) embedding.

This result is a tool in the derivation of our second result, Theorem 7.1, which, again subject to regularity conditions is as follows:

> Suppose g is increasing. Then $\mathbb{E}[\int_0^\tau g(S_u)\,du]$ is minimized (resp., maximized) over uniformly integrable embeddings τ of μ by the Azéma–Yor (resp., Perkins) embedding.

One approach to finding extremal values of $\mathbb{E}[F(W_\tau, S_\tau)]$ is to utilize the work of Kertz and Rösler [13], Vallois [23] and Rogers [20] who characterize the possible joint laws of (W_τ, S_τ). These characterizations take the form of constraints on the possible laws of (W_τ, S_τ), but that still leaves our problem as a constrained optimization problem. In fact, there are parallels between equation 3.2 of Theorem 3.1 of Rogers [20], and some of the quantities that arise in our study (see Remark 5.7), but we shall not make direct use of this connection.

At first sight the second result above may appear counterintuitive. After all, for increasing g the Azéma–Yor embedding maximizes the law of $g(S_\tau)$ so one might also expect it to maximize the law of $\int_0^\tau g(S_u)\,du$. However, the exact opposite is true, and the Azéma–Yor embedding minimizes the expected value of this quantity. We return to this issue in Remark 7.2, where we explain this phenomenon.

One of our tools for solving the above problems is to solve the problem in the case where μ has bounded support, and to approach the case of a general measure by approximation. In order to carry out this program we need to analyze when and whether convergence of probability measures is sufficient to guarantee that the associated Azéma–Yor and Perkins embeddings converge. This proves to be a delicate question. Under the additional (and necessary) hypothesis that $\int_{\mathbb{R}} |x|\mu_n(dx) \to \int_{\mathbb{R}} |x|\mu(dx)$, then indeed the Azéma–Yor embedding of μ_n converges almost surely to the Azéma–Yor embedding of μ. However, this need not be the case for the Perkins embedding, and the sequence of Perkins embeddings of μ_n may fail to converge on an almost sure basis.

We note that although the focus in this paper is on functionals involving the running maximum, there is a parallel set of results for functionals involving the running minimum process. The corresponding results can be easily proved by following the proofs given for the maximum and making the appropriate changes. Alternatively, given a Brownian motion W and a centered target law μ, let $\tilde{\mu}$ be the measure μ reflected around zero. Then, with $I_t = \inf_{s\le t} W_s$, the problem of minimizing $\mathbb{E}[F(W_\tau, I_\tau)]$ over embeddings τ of μ is equivalent to minimizing

$\mathbb{E}[F(-W_{\tilde{\tau}}, -S_{\tilde{\tau}})]$, over embeddings $\tilde{\tau}$ of $\tilde{\mu}$. See the next section and Section 8.1 for calculations along these lines.

2. A variance swap on squared returns. The original motivation for our study came from financial mathematics and the pricing of variance swaps, and one of the contributions of this article is to establish a link between variance swap bounds and Skorokhod embedding theory. The implications of this connection are the subject of related work [11]. Informed by the results presented here, but necessarily using different methods, Hobson and Klimmek [11] show how to construct model-independent bounds and hedging strategies for a general family of variance swaps. In this section we outline the link between variance swaps and the second result from the Introduction.

Let $X = (X_t)_{0\le t\le T}$ represent the discounted price of a financial asset. Under the assumption of no-arbitrage, there exists a measure under which X is a (local)-martingale. We may suppose that there exists a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that B is a $\mathbb{F}$-Brownian and such that $X_t = B_{A_t}$ for a (possibly discontinuous) time-change $t \to A_t$, null at 0. (If X is continuous, then the existence of such a time-change is guaranteed by the Dambis–Dubins–Schwarz theorem, and in general the existence is guaranteed by Monroe [15], Theorem 2.) Since X is a nonnegative price process we suppose it has starting value $X_0 = B_0 = x_0 > 0$.

Now suppose that we know the prices of put and call options with maturity T. Knowledge of put and call option prices with expiry time T is equivalent to knowledge of the marginal law of process at time T; see Breeden and Litzenberger [3]. Suppose that $X_T \sim \mu$ and that μ is centered at x_0, and has support in $\mathbb{R}^+$. We will determine bounds for the fair value of a variance swap given the terminal law μ. Note that if $X_T \sim \mu$, then A_T is a solution of the Skorokhod embedding problem for μ in B.

Following Demeterfi et al. [7] we define the pay-out $V = V((X_s)_{0\le s\le T})$ of an idealized variance swap as

$$V_T = \int_0^T \frac{d[X,X]_t}{(X_{t-})^2} = \int_0^T \left(\frac{dX_t^c}{X_{t-}}\right)^2 + \sum_{0\le t\le T}\left(\frac{\Delta X_t}{X_{t-}}\right)^2, \tag{2.1}$$

where $\Delta X_t = X_t - X_{t-}$, and X^c is the continuous part of X.

Let A^c be the continuous part of A. Note that $dA_t^c = (dX_t^c)^2 = d[X,X]_t^c$. Let $S^X = (S_t^X)_{t\ge 0}$ (resp., S^B) be the process of the running maximum of X (resp., B), and let I^X (resp., I^B) denote the corresponding infimum. Then we have $X_t \le S_t^X \le S_{A_t}^B$, and it follows that path-by-path with $\Delta B_{A_t} = B_{A_t} - B_{A_{t-}}$ that

$$\begin{aligned} V_T &\ge \int_0^T \frac{d[X,X]_t^c}{(S_{t-}^X)^2} + \sum_{0\le t\le T}\left(\frac{\Delta X_t}{S_{t-}^X}\right)^2 \\ &\ge \int_0^T \frac{dA_t^c}{(S_{A_{t-}}^B)^2} + \sum_{0\le t\le T}\left(\frac{\Delta B_{A_t}}{S_{A_{t-}}^B}\right)^2. \end{aligned} \tag{2.2}$$

We suppose that X has a second moment. Then $(X_t)_{0\le t\le T}$ is a square-integrable martingale and we find that

$$\begin{aligned}\mathbb{E}\biggl[\int_0^T \frac{dA_t^c}{(S_{A_{t-}}^B)^2} + \sum_{0\le t\le T}\biggl(\frac{\Delta B_{A_t}}{S_{A_{t-}}^B}\biggr)^2\biggr] &= \mathbb{E}\biggl[\int_0^T \frac{dA_t^c + \Delta A_t}{(S_{A_{t-}}^B)^2}\biggr] \\ &= \mathbb{E}\biggl[\int_0^T \frac{dA_t}{(S_{A_{t-}}^B)^2}\biggr] \\ &\ge \mathbb{E}\biggl[\int_0^{A_T} \frac{du}{(S_u^B)^2}\biggr].\end{aligned} \tag{2.3}$$

We say that τ is an embedding of μ if τ is a stopping time for which B_τ has law μ [we write $B_\tau \sim \mu$ or $\mu = \mathcal{L}(B_\tau)$]. Let $\mathcal{S} \equiv \mathcal{S}(B,\mu)$ be the set of stopping times which embed μ, and let $S_{\mathrm{UI}} = \mathcal{S}_{\mathrm{UI}}(B,\mu)$ be the subset of $\mathcal{S}(B,\mu)$ for which $(B_{t\wedge\tau})_{t\ge 0}$ is uniformly integrable. The inequalities above imply that the fair value of V_T is bounded below by

$$\inf_{\tau\in\mathcal{S}_{\mathrm{UI}}(B,\mu)} \mathbb{E}\biggl[\int_0^\tau \frac{du}{(S_u^B)^2}\biggr]. \tag{2.4}$$

Similarly, using the inequality $I_{A_t}^B \le I_t^X \le X_t$ we find that the fair value of V_T is bounded above by

$$\sup_{\tau\in\mathcal{S}_{\mathrm{UI}}(B,\mu)} \mathbb{E}\biggl[\int_0^\tau \frac{du}{(I_u^B)^2}\biggr]. \tag{2.5}$$

This problem can be converted into a problem concerning the maximum S^B by a reflection argument; see Section 8.1.

Now let $G(b,s) = \frac{(s-b)^2}{s^2}$. Then by Itô's lemma,

$$G(B_\tau, S_\tau^B) = G(0,0) + \int_0^\tau \frac{du}{(S_u^B)^2} - \int_0^\tau \frac{2(S_u^B - B_u)}{(S_u^B)^2}\,dB_u.$$

It follows that if $\int_0^{\tau\wedge t} 2(S_u^B - B_u)(S_u^B)^{-2}\,dB_u$ is a uniformly integrable martingale, then

$$\mathbb{E}\biggl[\int_0^\tau \frac{du}{(S_u^B)^2}\biggr] = \mathbb{E}\biggl[\frac{(S_\tau^B - B_\tau)^2}{(S_\tau^B)^2}\biggr],$$

and the question of bounding the fair value of V_T is transformed into a question of maximizing or minimizing expressions of the form $\mathbb{E}[F(B_\tau, S_\tau)]$ over embeddings of μ. We return to the calculation of the variance swap bound in Section 8.1.

3. Preliminaries. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions and supporting a Brownian motion $W = (W_t)_{t\geq 0}$ with $W_0 = 0$, and sufficiently rich that $\mathcal{F}_0$ contains a further uniform random variable which is independent of W. Let μ be a centered probability measure. To exclude trivialities we assume that μ is not δ_0, the unit mass at 0. We say that τ is an embedding of μ if τ is a stopping time for which W_τ has law μ [we write $W_\tau \sim \mu$ or $\mu = \mathcal{L}(W_\tau)$] and we say that τ is uniformly integrable if the family $(W_{t\wedge\tau})_{t\geq 0}$ is uniformly integrable.

Let $\mathcal{S} \equiv \mathcal{S}(W, \mu)$ be the set of stopping times which embed μ, and let $\mathcal{S}_{\mathrm{UI}} \equiv \mathcal{S}_{\mathrm{UI}}(W, \mu)$ be the subset of $\mathcal{S}(W, \mu)$ consisting of uniformly integrable stopping times. For $\mathcal{S}_{\mathrm{UI}}(W, \mu)$ to be nonempty we must have that μ is centered [i.e., $\int_{\mathbb{R}} |x|\mu(dx) < \infty$ and $\int_{\mathbb{R}} x\mu(dx) = 0$]. In this context (Brownian motion and centered target laws) a result of Monroe [14] gives that a stopping time is uniformly integrable if and only if it is minimal (in the sense that if τ is minimal and $\sigma \leq \tau$ with $W_\sigma \sim W_\tau$, then $\sigma \equiv \tau$ almost surely). The class of minimal stopping times is a natural class of "good" (in the sense of small) stopping times.

For the Brownian motion W, started at 0, we write H_x for the first hitting time of x, and for a set A, $H_A = \inf\{u \geq 0 : W_u \in A\}$.

For a process $(Y_t)_{t\geq 0}$ and a stopping time σ, we write $Y^\sigma = (Y_t^\sigma)_{t\geq 0}$ for the stopped process $Y_t^\sigma = Y_{\sigma\wedge t}$.

Given a centered probability measure μ, let X_μ be a random variable with law μ, and define $C(x) \equiv C_\mu(x) = \mathbb{E}[(X_\mu - x)^+]$ and $P(x) \equiv P_\mu(x) = \mathbb{E}[(x - X_\mu)^+]$. Then C and P are monotonic convex functions with $C(0) = P(0)$. Then $U(x) = U_\mu(x) = \mathbb{E}[|X_\mu - x|] = C(x) + P(x)$ is (minus) the potential associated with μ. Conversely any convex function U with $\lim_{x\to\pm\infty}(U(x) - |x|) = 0$ is the potential of some centered probability measure μ (Chacon [4]).

If μ has an atom at zero, then we write μ^* for the measure obtained by omitting the atom at 0, and then rescaling to get a probability measure. Thus $\mu^*(A) = \mu(A \setminus \{0\})/(1 - \mu(\{0\}))$. Finally, we write $\hat{x} = \hat{x}_\mu$ for the upper limit on the support of μ [so $\hat{x}_\mu = \sup\{x : C_\mu(x) > 0\}$] and $\check{x} = \check{x}_\mu$ for the corresponding lower limit $\check{x}_\mu = \inf\{x : P_\mu(x) > 0\}$.

3.1. *The Azéma–Yor solution.* For $x \geq 0$, up to the upper limit on the support of μ, define $\beta = \beta_\mu$ by

$$\beta(x) = \arg\min_{y<x} \frac{C_\mu(y)}{x - y}. \tag{3.1}$$

Then β is an increasing function with $\beta(x) < x$, see Figure 1. Where the arg min is not uniquely defined it is not important which value we choose. However, we fix one by insisting that β is right-continuous, or equivalently by choosing the largest value for which the minimum is attained. Observe that at $x = 0$, β takes the value of the infimum of the support of μ. For x equal to, or to the right of, the upper limit on the support of μ we set $\beta(x) = x$.

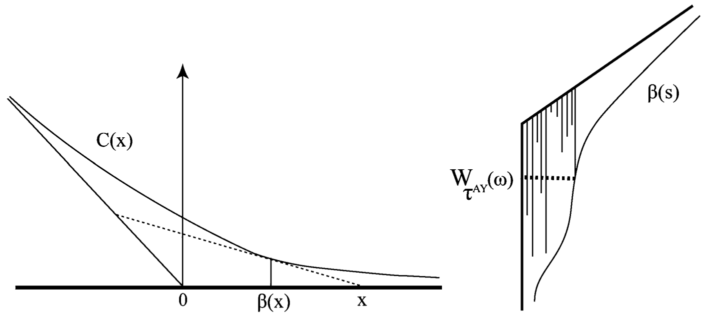

FIG. 1. *For each x, the value of $\beta(x)$ is determined by finding the tangent line to C_μ originating at x: $\beta(x)$ is the horizontal co-ordinate of the point of contact between the tangent line and C_μ.* [*If C_μ includes a straight line section, then this point of contact may not be uniquely defined, in which case we take $\beta(x)$ to be the largest value of the horizontal co-ordinate at which contact occurs.*] *The stopping time τ_β associated to this construction is given by the first time that an excursion from the maximum crosses below β.*

For an increasing function $\beta : \mathbb{R}^+ \mapsto \mathbb{R}$ with $\beta(x) \leq x$ let τ_β be given by

$$\tau_\beta = \inf\{u : W_u \leq \beta(S_u)\}. \tag{3.2}$$

Then $\tau^{\mathrm{AY}} \equiv \tau_\mu^{\mathrm{AY}}$, the Azéma–Yor stopping time for μ, is given by $\tau_\mu^{\mathrm{AY}} \equiv \tau_{\beta_\mu}$. Thus we have $\tau_{\beta_\mu} \in \mathcal{S}_{\mathrm{UI}}(W, \mu)$, and moreover, for F increasing, τ_{β_μ} maximizes $\mathbb{E}[F(S_\tau)]$ over $\tau \in \mathcal{S}_{\mathrm{UI}}(W, \mu)$ (Azéma–Yor [1, 2], Rogers [19]).

Note that τ_{β_μ} does not maximize this quantity over *all* embeddings, but it does give the maximum over uniformly integrable (i.e., minimal) embeddings.

Let $b \equiv b_\mu$ be the right-continuous inverse to β. Then b is the barycenter function and for $x < \hat{x}_\mu$, $b(x)$ is given by

$$b(x) = \mathbb{E}[X_\mu | X_\mu \geq x]. \tag{3.3}$$

The barycentre $b(x)$ is defined up to the upper limit of the support of μ and is a nonnegative, nondecreasing function with $b(x) \geq x$. We set $b(x) = x$ for $x \geq \hat{x}_\mu$. (The reverse barycentre $\check{b}(x) = \mathbb{E}[X | X \leq x]$ is defined analogously to the barycentre.)

It is more standard to define the barycenter function as in (3.3) and to set β to be the inverse barycenter function, but the two approaches are equivalent, and our approach via potentials allows for a unified treatment with the Perkins construction in the next section.

If μ has an interval with no mass, then b is constant over that interval, and β has a jump. If μ has an atom at x then b has a jump at x [unless the atom is at the upper limit $\hat{x}$ of the support of μ in which case $b(\hat{x}) = \hat{x}$] and β is constant over a range of s. From the definition of τ_β [see (3.2)] and excursion theory (see Rogers [20], equation 2.13), we have

$$\exp\left(-\int_0^s \frac{dr}{r - \beta(r)}\right) = \mathbb{P}(S_{\tau_\beta} \geq s) \tag{3.4}$$

and then also $\mathbb{P}(S_{\tau_\beta} \geq s) = \mathbb{P}(W_{\tau_\beta} \geq \beta(s)) = \mu(\beta(s), \infty)$. Note that it does not matter which convention we use for $\beta(s)$ here since μ places no mass on $(\beta(s-), \beta(s+))$.

EXAMPLE 3.1. If $\mu = U[-1, 1]$, then $C_\mu(x) = (x-1)^2/4$ and $P_\mu(x) = (x+1)^2/4$ (at least for $-1 = \check{x}_\mu \leq x \leq \hat{x} = 1$). Then the barycenter function is given by $b(x) = (x+1)/2$ for $-1 \leq x \leq 1$ and hence $\beta(s) = 2s - 1$ for $0 \leq s \leq 1$. It follows that $S_{\tau_\mu^{\mathrm{AY}}} \equiv b(W_{\tau_\mu^{\mathrm{AY}}})$ is uniformly distributed on $[0, 1]$.

LEMMA 3.2. *If μ places mass on (x, ∞), then $(r - \beta(r))^{-1}$ is integrable over $[0, x]$.*

PROOF. This follows immediately from (3.4) and $\mathbb{P}(S_{\tau_\beta} \geq x) \geq \mathbb{P}(W_{\tau_\beta} \geq x) > 0$. $\square$

3.2. *The Perkins solution.* For $x > 0$ define $\alpha_\mu^+ = \alpha^+ : \mathbb{R}_+ \to \mathbb{R}_-$ by

$$\alpha^+(x) = \arg\min_{y<0} \frac{C_\mu(x) - P_\mu(y)}{x - y} \tag{3.5}$$

and for $x < 0$ define $\alpha_\mu^- = \alpha^- : \mathbb{R}_- \to \mathbb{R}_+$ by

$$\alpha^-(x) = \arg\max_{y>0} \frac{P_\mu(x) - C_\mu(y)}{y - x}. \tag{3.6}$$

Then $\alpha^\pm$ are monotonic functions, see Figure 2. If the arg min (resp., the arg max) is not uniquely defined, we take the largest value (in modulus) for which the minimum (resp., the maximum) is attained; in this way $\alpha^+ : \mathbb{R}_+ \mapsto \mathbb{R}_-$ is right-continuous and α^- is left-continuous. Again, none of the subsequent analysis will depend on this convention. For convenience we will sometimes write α as shorthand for $\alpha^\pm$.

If P_μ (resp., C_μ) is differentiable at $\alpha^+(x)$ [resp., $\alpha^-(x)$], then $\alpha^+(x)$ [resp., $\alpha^-(x)$] satisfies

$$\frac{C_\mu(x) - P_\mu(\alpha^+(x))}{x - \alpha^+(x)} = P'_\mu(\alpha^+(x)) \tag{3.7}$$

[resp., $P_\mu(x) - C_\mu(\alpha^-(x)) = C'_\mu(\alpha^-(x))(x - \alpha^-(x))$].

Let $a^\pm$ be the inverse to $\alpha^\pm$ and let $\bar{a}(w) = w$ for $w > 0$ and $\bar{a}(w) = a^+(w)$ for $w < 0$. Recall the definition of I as the infimum process for W so that $I_t = \inf_{s \leq t} W_s$.

For a pair of monotonic functions $\alpha^+ : \mathbb{R}_+ \mapsto \mathbb{R}_-$ (nonincreasing) and $\alpha^- : \mathbb{R}_- \mapsto \mathbb{R}_+$ (nondecreasing) define the stopping time

$$\tau_\alpha = \inf\{u > 0 : W_u \leq \alpha^+(S_u)\} \wedge \inf\{u > 0 : W_u \geq \alpha^-(I_u)\}.$$

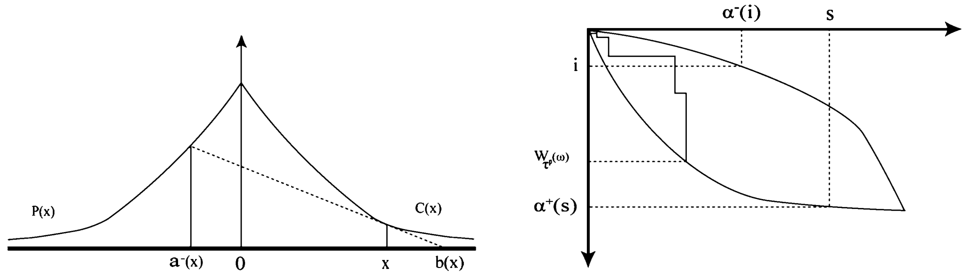

FIG. 2. *Suppose that μ has no atoms. Then for $x > 0$, $a^-(x)$ is the horizontal co-ordinate of the point where the tangent line to C at $(x, C(x))$ intersects with P. Alternatively, it is the horizontal co-ordinate of the point where the tangent line to C emanating from $(b(x), 0)$ intersects with P. [We may instead consider the inverse α^- of a^-: for $y < 0$, $\alpha^-(y) > 0$ is the horizontal co-ordinate of the point such that the tangent to C at $\alpha^-(y)$ crosses P at $(y, P(y))$.] These definitions extend naturally to the case where the convex function C has kinks or straight-line segments. Similarly, $a^+(z)$ is found by drawing tangents to P emanating from the reverse barycenter function evaluated at $z < 0$ and determining intersection points with C. The stopping rule associated with this construction is to stop the Brownian motion when its running maximum or minimum exit the region determined by α^+ and α^-.*

Suppose μ does not have an atom at zero. Then the Perkins [17] embedding $\tau^{\mathrm{P}} \equiv \tau^{\mathrm{P}}_\mu \equiv \tau^{\mathrm{P}}(\mu)$ is given by $\tau^{\mathrm{P}}_\mu = \tau_{\alpha_\mu}$.

If μ has an atom at zero, then we use independent randomization to set $\tau^{\mathrm{P}} = 0$ with probability $\mu(\{0\})$; and otherwise $\tau^{\mathrm{P}} = \tau_{\alpha_\mu}$. More precisely, in the case where μ has an atom at zero we set the Perkins embedding to be

$$\tau^{\mathrm{P}} = \begin{cases} 0, & \text{if } Z \le \mu(0), \\ \tau_{\alpha_\mu}, & \text{if } Z > \mu(0), \end{cases}$$

where Z is a uniform random variable which is measurable with respect to $\mathcal{F}_0$. Here $\alpha^\pm_\mu$ are the quantities defined in (3.5) and (3.6) for μ. Note that if μ^* is obtained from μ by removing any mass at zero, and rescaling to give a probability measure, then although C_{μ^*} and P_{μ^*} are scalar multiples of C_μ and P_μ, respectively, nonetheless we have $\alpha^\pm_{\mu^*} \equiv \alpha^\pm_\mu$.

Note that if μ has an atom at zero, then we need $\mathcal{F}_0$ to be nontrivial in order to be able to define the Perkins embedding. Note further that since there are potentially many uniform random variables Z which are measurable with respect to $\mathcal{F}_0$, if $\mu(\{0\}) > 0$, then the Perkins embedding is not unique. Sometimes it is convenient to think about the Perkins embedding associated with an identified $\mathcal{F}_0$ random variable Z, in which case we write $\tau^{\mathrm{P},Z}_\mu$ instead of just τ^{P}_μ.

The results of Perkins [17] show that $\tau^{\mathrm{P}}_\mu \in \mathcal{S}_{\mathrm{UI}}(W, \mu)$ and moreover, for F increasing, τ^{P} minimizes $\mathbb{E}[F(S_\tau)]$ over $\tau \in \mathcal{S}(W, \mu)$, and not just $\mathcal{S}_{\mathrm{UI}}(W, \mu)$ (Perkins [17], although the representation via (3.5) and (3.6) is due to Hobson and Pedersen [12]).

EXAMPLE 3.3. If $\mu = U[-1, 1]$, then $P = P_\mu$ and $C = C_\mu$ are as given in Example 3.1 and, from (3.7), $\alpha^+(s)$ is the unique root of the equation $P'(\alpha)(s - \alpha) = C(s) - P(\alpha)$. It is easily verified that this root is given by $\alpha^+(s) = s - 2\sqrt{s}$. Similarly, $\alpha^-(i) = i + 2\sqrt{|i|}$. It can be shown that $\mathbb{P}(S_{\tau_\alpha} \geq s) = \mathbb{P}(W_{\tau_\alpha} \geq s) = \mathbb{P}(W_{\tau_\alpha} \leq s - 2\sqrt{s}) = 1 - \sqrt{s}$.

EXAMPLE 3.4. Notwithstanding the above example, in general it is difficult to derive an explicit form for the stopping boundary associated with the Perkins stopping time. Here we give a second example where analytic expressions, albeit complicated ones, can be derived.

Suppose the target law is a centered Pareto distribution with support $[-1, \infty)$ and density function $f(x) = 2(x + 2)^{-3}$. Then for $k \geq -1$, $C(k) = (2 + k)^{-1}$ and $P(k) = k + (2 + k)^{-1}$, and for $k < -1$, $C(k) = -k$, $P(k) = 0$.

Then, for the Azéma–Yor embedding, β solves $C(\beta) = (s - \beta)|C'(\beta)|$ and $\beta(s) = (s/2) - 1$.

For the Perkins embedding, $\alpha^+(s)$ solves $P'(\alpha^+) = (C(s) - P(\alpha^+))/(s - \alpha^+)$, and we have (after some algebra)

$$\alpha^+(s) = \frac{-2s^2 - 5s + \sqrt{s^4 + 6s^3 + 12s^2 + 8s}}{2s - 1 + s^2}.$$

The expression for α^- is $\alpha^-(i) = \frac{-3i - 2i^2 + \sqrt{-(i^4 + 6i^3 + 12i^2 + 8i)}}{2i + 1 + i^2}$.

If μ has an interval in $\mathbb{R}_+$ (resp., $\mathbb{R}_-$) with no mass, then α^- (resp., α_+) has a jump (unless that interval is contiguous with zero, in which case $\alpha^\pm$ starts at a nonzero value). If μ has an atom in $(0, \infty)$ [resp., $(-\infty, 0)$], then α^- (resp., α_+) is constant over a range of values.

LEMMA 3.5. *Suppose $x > 0$. If μ places mass on $[x, \infty)$, then $(r - \alpha^+(r))^{-1}$ is integrable over $(0, x)$.*

PROOF. We have $(W_u \geq \alpha^+(S_u); \forall u \leq H_x) \supseteq (\tau_\alpha \geq H_x) \supseteq (W_{\tau_\alpha} \geq x)$, and then by excursion theory [recall (3.4)],

$$\exp\left(-\int_0^x \frac{dr}{r - \alpha^+(r)}\right) = \mathbb{P}\big(W_u \geq \alpha^+(S_u); \forall u \leq H_x\big) \geq \mu\big([x, \infty)\big) > 0. \qquad \square$$

4. Convergence of measures and convergence of embeddings. Let $(\mu_n)_{n\geq 1}$ be a sequence of measures, and write U_n, β_n and α_n as shorthand for U_{μ_n}, β_{μ_n} and α_{μ_n}, with a similar convention for other functionals.

Suppose that, for each n, μ_n is centered and that $(\mu_n)_{n\geq 1}$ converges weakly to μ, where μ is also centered. Then it does not follow that $U_n \to U_\mu$, nor that $\beta_n \to \beta_\mu$, nor that $\alpha_n \to \alpha_\mu$. However, with the correct additional hypotheses, then these types of convergence are equivalent.

Our first key result is the following.

PROPOSITION 4.1. *Let* (μ_n) *be a sequence of measures such that* $\mu_n \Rightarrow \mu$ *and* $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$. *Then* $b_n(x) \to b(x)$ *at continuity points* $x < \hat{x}$ *of* b.

PROOF. Chacon [4] shows that if $\mu_n \Rightarrow \mu$ and $U_n(0) \to U(0)$, then $U_n \to U$ pointwise. Since $C_n(x) = (U_n(x) + x)/2$ it follows trivially that $C_n \to C$ pointwise, where $C_n(x) = C_{\mu_n}(x)$ and $C(x) = C_\mu(x)$.

Recall that x is a discontinuity point of b if and only if there is an atom of μ at x. Suppose $x < \hat{x}$ is a continuity point of b. Then (3.3) gives $b(x) = x + \frac{C(x)}{\mu([x,\infty))}$ and

$$b_n(x) = x + \frac{C_n(x)}{\mu_n([x,\infty))} \to x + \frac{C(x)}{\mu([x,\infty))} = b(x). \qquad \square$$

COROLLARY 4.2. *Let* (μ_n) *be a sequence of measures such that* $\mu_n \Rightarrow \mu$ *and* $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$. *Then* $\beta_n(s) \to \beta(s)$ *at continuity points* $s < \hat{x}$ *of* β. *Moreover, if* $\hat{x} < \infty$, *then for each* $z > \hat{x}$, $\liminf \beta_n(z) \geq \hat{x}$.

PROOF. Since $b_n(\hat{x} - \varepsilon) < \hat{x} + \varepsilon$ for sufficiently large n we have for these same n that $\beta_n(\hat{x} + \varepsilon) \geq \hat{x} - \varepsilon$. $\square$

COROLLARY 4.3. *Under the assumptions of Proposition* 4.1, $\tau_{\beta_n} \to \tau_\beta$ *almost surely.*

PROOF. Let D be the set of discontinuity points of β. If $S_{\tau_\beta} \notin D$, then $W_{\tau_\beta} = \beta(S_{\tau_\beta})$, and it follows that

$$(\omega : \tau_{\beta_n} \not\to \tau_\beta) \subseteq (\omega : S_{\tau_\beta} \in D) \cup \bigl(\omega : S_{\tau_\beta} \notin D, W_{\tau_\beta} = \beta(S_{\tau_\beta}), \tau_{\beta_n} \not\to \tau_\beta\bigr).$$

For any stopping time σ write: let $H_x^\sigma = \inf\{u \geq \sigma : W_u = x\}$.

Case 1: $\hat{x} = \infty$. Note that since β is increasing, D is countable and $\mathbb{P}(S_{\tau_\beta} \in D) = 0$.

First we argue that on $(\omega : S_{\tau_\beta} = x)$ we have that for sufficiently large n, $S_{\tau_{\beta_n}} \geq x$: since there are only countably many values of $s < x$ on which the value of W_u gets below $S_u = s$, and on each of these excursions W stays above $\beta(S)$, for sufficiently large n, W must stay above $\beta_n(S)$ also.

Hence $\liminf_n S_{\tau_{\beta_n}} \geq S_{\tau_\beta}$ almost surely. Then on $\{\omega : S_{\tau_\beta} = x \notin D, W_{\tau_\beta} = \beta(x)\}$, we have $\tau_{\beta_n}(\omega) \to \tau_\beta(\omega)$ unless $\inf\{W_u : \tau_\beta \leq u \leq H_{S_{\tau_\beta}}^{\tau_\beta}\} = W_{\tau_\beta} = \beta(x)$ and $\beta_n(x) < \beta(x)$. But, almost surely, on any interval of positive length Brownian motion goes below its starting value. In particular, the set $(\omega : S_{\tau_\beta} \notin D, W_{\tau_\beta} = \beta(S_{\tau_\beta}), \tau_{\beta_n} \not\to \tau_\beta)$ has probability zero.

Case 2: $\hat{x} < \infty$ *and* $\mu(\{\hat{x}\}) = 0$. The only paths for which issues of convergence might be different to the previous case are those for which $S_{\tau_\beta} = \hat{x}$. But since μ has no atom at $\hat{x}$, $\mathbb{P}(S_{\tau_\beta} = \hat{x}) = \mathbb{P}(W_{\tau_\beta} = \hat{x}) = 0$ and $\tau_{\beta_n} \to \tau_\beta$ almost surely.

Case 3: $\hat{x} < \infty$ *and* $\mu(\{\hat{x}\}) > 0$. In this case $\beta(\hat{x}-) := \lim_{y\uparrow\hat{x}} \beta(y) < \beta(\hat{x}) = \hat{x}$. We show that on the set $(S_{\tau_\beta} = \hat{x})$ we have $\lim \tau_{\beta_n} = \tau_\beta$, almost surely. Off the set $(S_{\tau_\beta} = \hat{x})$ convergence follows exactly as in the previous cases.

First we argue that $\limsup_n S_{\tau_{\beta_n}} \le \hat{x}$ almost surely. Fix $z > \hat{x}$, then given $0 < \varepsilon < z - \hat{x}$, there exists N such that for $n \ge N$, $\beta_n(\hat{x} + \varepsilon) > \hat{x} - \varepsilon$. Hence, for sufficiently large n,

$$\bigl(\omega : S_{\tau_{\beta_n}}(\omega) \ge z\bigr) \subseteq \bigl(\omega : \inf\{W_u : H_{\hat{x}+\varepsilon} \le u \le H_z\} \ge \hat{x} - \varepsilon\bigr).$$

But

$$\mathbb{P}\bigl(\inf\{W_u : H_{\hat{x}+\varepsilon} \le u \le H_z\} \ge \hat{x} - \varepsilon\bigr) \le \exp\biggl(-\int_{\hat{x}+\varepsilon}^{z} \frac{dy}{y - (\hat{x} - \varepsilon)}\biggr) = \frac{2\varepsilon}{z - \hat{x} + \varepsilon}.$$

By choosing ε small compared with $(z - \hat{x})$ we deduce that $\limsup_n S_{\tau_{\beta_n}} \le z$ for any $z > \hat{x}$.

Now we argue that on $S_{\tau_\beta} = \hat{x}$ we have $\liminf W_{\tau_{\beta_n}} \ge \hat{x}$ almost surely. Coupled with the result from the previous paragraph we can then conclude that on $W_{\tau_\beta} = \hat{x}$ we have $\tau_{\beta_n} \to H_{\hat{x}} = \tau_\beta$.

Given δ and $\varepsilon < \hat{x} - \beta(\hat{x}-) - \delta$, there exists N such that for all $n > N$, $\beta_n(\hat{x} - \varepsilon) < \beta(\hat{x}-) + \varepsilon < \hat{x} - \delta$. Then

$$\begin{aligned}\bigl(\omega : W_{\tau_{\beta_n}}(\omega) < \hat{x} - \delta, S_{\tau_\beta}(\omega) = \hat{x}\bigr) &\subseteq \bigl(\omega : \inf\{W_u : H_{\hat{x}-\varepsilon} \le u \le H_{\hat{x}}\} \le \hat{x} - \delta\bigr) \\ &\quad \cup (\omega : S_{\tau_{\beta_n}} < \hat{x} - \varepsilon, S_{\tau_\beta} = \hat{x}).\end{aligned}$$

By similar arguments to those in case 1 we can prove that the final event has small probability. Moreover, using that the fact that the probability that an event occurs is smaller than the expected number of times that it occurs,

$$\mathbb{P}\bigl(\omega : \inf\{W_u : H_{\hat{x}-\varepsilon} \le u \le H_{\hat{x}}\} \le \hat{x} - \delta\bigr) \le \int_{\hat{x}-\varepsilon}^{\hat{x}} \frac{dy}{y - (\hat{x} - \delta)} = \ln\bigl(\delta/(\delta - \varepsilon)\bigr).$$

By choosing ε compared to δ this probability can be made arbitrarily small. $\square$

Note that if $\tau_{\beta_n} \to \tau_\beta$ almost surely, then by the continuity of Brownian motion $W_{\tau_{\beta_n}} \to W_{\tau_\beta}$ almost surely and $\mu_n \Rightarrow \mu$.

We can summarize the results as follows:

PROPOSITION 4.4. *Suppose that* $(\mu_n)_{n\ge1}$ *and* μ *are centered and that* $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$. *Then the following are equivalent*:

(i) $\mu_n \Rightarrow \mu$ *and* $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$;

(ii) $U_n(x) \to U_\mu(x)$ *for each* $x \in \mathbb{R}$;

(iii) $\beta_n \to \beta$ *at continuity points* s *of* β, *provided* s *is less than or equal to the upper limit on the support of* μ;

(iv) $\tau_{\beta_n} \xrightarrow{a.s.} \tau_\beta$;

(v) $W_{\tau_{\beta_n}} \xrightarrow{a.s.} W_{\tau_\beta}$.

Now we want to prove a similar result for the Perkins embedding.

LEMMA 4.5. *Let $(\mu_n)_{n\geq 1}$ be a sequence of centered probability measures such that $\mu_n \Rightarrow \mu$ and $\mathbb{E}[|X_n|] \to \mathbb{E}[|X_\mu|]$. Then $a_n^\pm(x) \to a^\pm(x)$ at continuity points $x \in (\check{x}, \hat{x}) \setminus \{0\}$ of a. Moreover $\alpha_n^\pm(x) \to \alpha^\pm(x)$ at nonzero continuity points $\check{x} < x < \hat{x}$ of μ.*

PROOF. We prove the result for (a_n^+, a^+), the other case being similar. The extension from $a^\pm$ to $\alpha^\pm$ follows as in Corollary 4.2.

Again we have that $x < 0$ is a discontinuity point of a^+ if and only if there is an atom of μ at x. Suppose that x is not an atom of μ. Then, recall (3.7), $a^+(x)$ is the unique solution in z of $P(x) + P'(x)(z - x) = C(z)$. Moreover, for any $\tilde{a}_n(x) \in (a_n^+(x+), a_n^+(x-))$,

$$P_n(x) + P_n'(x+)(\tilde{a}_n(x) - x) \geq C_n(\tilde{a}_n(x)),$$
$$P_n(x) + P_n'(x-)(\tilde{a}_n(x) - x) \leq C_n(\tilde{a}_n(x)).$$

Suppose $a_n^+(x) \to \gamma$ (down a subsequence if necessary). Then since $P_n'(x\pm) \to P'(x)$,

$$P(x) + P'(x)(\gamma - x) \geq C(\gamma) \geq P(x) + P'(x)(\gamma - x).$$

Hence $\gamma = a^+(x)$ and $a_n^+(x) \to a(x)$. $\square$

PROPOSITION 4.6. *Suppose that $(\mu_n)_{n\geq 1}$ and μ are centered and that $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$.*

(a) *Suppose there exists an open interval I containing 0 such that $\mu_n(I) = \mu(I) = 0$. Then the following are equivalent*:
   (i) *$\mu_n \Rightarrow \mu$ and $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$*;
   (ii) *$U_n(x) \to U_\mu(x)$ for each $x \in \mathbb{R}$*;
   (iii) *$\alpha_n^\pm \to \alpha^\pm$ at continuity points of $\alpha^\pm$ which lie within the range of the support of μ*;
   (iv) *$\tau_{\mu_n}^{\mathrm{P}} \xrightarrow{a.s.} \tau_\mu^{\mathrm{P}}$*;
   (v) *$W_{\tau_{\mu_n}^{\mathrm{P}}} \xrightarrow{a.s.} W_{\tau_\mu^{\mathrm{P}}}$.*

(b) *More generally, suppose $\mu_n \Rightarrow \mu$ and $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$. Then, $\alpha_n^\pm \to \alpha^\pm$ at continuity points of $\alpha^\pm$ which lie within the range of the support of μ.*

   Suppose further that $\mu_n(\{0\}) \to \mu(\{0\})$. Then there exists a sequence of Perkins embeddings of μ_n such that $\tau_{\mu_n}^{\mathrm{P}}$ converges in probability to a Perkins embedding τ_μ^{P} of μ. In particular, if Z_n converges in probability to Z, then the Perkins embeddings $(\tau_{\mu_n}^{\mathrm{P},Z_n})_{n\geq 1}$ converge in probability to the Perkins embedding $\tau_\mu^{\mathrm{P},Z}$ of μ.

Thus, if $\mu_n \Rightarrow \mu$, $\mathbb{E}[|X_{\mu_n}|] \to \mathbb{E}[|X_\mu|]$ and $\mu_n(\{0\}) \to \mu(\{0\})$, then if $(\tau_{\mu_n}^{\mathrm{P},Z_n})_{n\geq 1}$ is a sequence of Perkins embeddings of $(\mu_n)_{n\geq 1}$, then there exists a subsequence n_k along which $\lim \tau_{\mu_{n_k}}^{\mathrm{P},Z_{n_k}}$ exists almost surely, and is a Perkins embedding of μ.

PROOF. For part (a) the equivalence of (i) and (ii) follows as before. Lemma 4.5 gives that (ii) implies (iii). It follows from the pathwise construction of τ_{α_n} (and the existence of the interval I which is not charged by μ_n so that $\tau_{\mu_n}^{\mathrm{P}} \equiv \tau_{\alpha_n}$) that $\tau_{\mu_n}^{\mathrm{P}} \to \tau_\mu^{\mathrm{P}}$ almost surely and hence we have (iii) implies (iv). The continuity of Brownian motion allows us to deduce (v), from which (i) follows immediately.

For part (b) the statement about the convergence of $\alpha_n^{\pm}$ follows as before. For the other results, suppose first that $\mu(\{0\}) = 0$ and $\mu_n(\{0\}) = 0$ for all sufficiently large n. Recall that $\tau_\alpha = \inf\{u : W_u \leq \alpha^+(S_u) \text{ or } W_u \geq \alpha^-(I_u)\}$ and for $\eta > 0$ define the stopping time

$$\rho_{\alpha,\eta} = \tau_{\alpha_\eta},$$

where $\alpha_\eta^+(s) = \min\{\alpha^+(s), -\eta\}$, $\alpha_\eta^-(i) = \max\{\alpha^-(i), \eta\}$. Note that $\rho_{\alpha,\eta}$ is the Perkins embedding of a law which places no mass on $(-\eta, \eta)$.

We have that $\alpha_n \to \alpha$ at nonzero continuity points. Let $\alpha_{n,\eta}^{\pm} = \mp \max\{\mp\alpha_n^{\pm}(s), \eta\}$ and let $\rho_{\alpha_n,\eta}$ be the Perkins embedding for $B_{\tau_{\alpha_n,\eta}}$. Then $\alpha_{n,\eta}^{\pm} \to \alpha_\eta^{\pm}$ at continuity points and by the pathwise construction of $\rho_{\alpha_n,\eta}$, we have $\rho_{\alpha_n,\eta} \to \rho_{\alpha,\eta}$ almost surely. In particular, given $\delta, \varepsilon > 0$ there exists N_0 such that for all $n \geq N_0$

$$\mathbb{P}\big(|\rho_{\alpha_n,\eta} - \rho_{\alpha,\eta}| > \varepsilon\big) < \delta/2.$$

Note that on $|W_{\tau_\alpha}| > \eta$ we have $\rho_{\alpha,\eta} = \tau_\alpha$ with a similar statement for α_n. We can choose $\eta > 0$ so that $\mu([-2\eta, 2\eta]) < \delta/6$ and then N_1 so that for $n \geq N_1$, $\mu_n([-\eta, \eta]) < \delta/3$. Then

$$\begin{aligned}\big(|\tau_{\alpha_n} - \tau_\alpha| > \varepsilon\big) \subseteq{}& \big(|W_{\tau_\alpha}| \leq \eta\big) \cup \big(|W_{\tau_{\alpha_n}}| \leq \eta\big)\\ &\cup \big(|\tau_{\alpha_n} - \tau_\alpha| > \varepsilon, |W_{\tau_\alpha}| > \eta, |W_{\tau_{\alpha_n}}| > \eta\big)\end{aligned}$$

and the set $(|\tau_{\alpha_n} - \tau_\alpha| > \varepsilon)$ has probability at most δ.

It follows that $\tau_{\alpha_n} \to \tau_\alpha$ in probability, and hence that there is almost sure convergence down a subsequence. Furthermore, down the same subsequence $W_{\tau_{\alpha_n}} \to W_{\tau_\alpha}$ almost surely.

Now suppose that $\mu(\{0\}) = 0$ and that $\lim \mu_n(\{0\}) = 0$. Recall the definition of μ_n^* as the measure μ_n with probability mass at zero removed, and then rescaled to be a probability measure, and note that $\alpha_{\mu_n^*} \equiv \alpha_{\mu_n}$. Then also $\mu_n^* \Rightarrow \mu$ and $U_{\mu_n^*} \to U_\mu$ pointwise.

Then, $\tau_{\mu_n}^{\mathrm{P},Z_n} = 0$ for $Z_n \leq \mu_n(\{0\})$ and $\tau_{\mu_n}^{\mathrm{P},Z_n} = \tau_{\alpha_n}$ otherwise, so that $\tau_{\mu_n}^{\mathrm{P},Z_n} \to \tau_\alpha$ in probability. Moreover, down a subsequence, $\tau_{\mu_n}^{\mathrm{P},Z_n} \to \tau_\alpha$ almost surely.

It remains to consider the case where $\mu(\{0\}) > 0$. For $\varepsilon < 1$, writing $A_n = (Z_n \le \mu_n(\{0\}), Z > \mu(\{0\}))$ and $B_n = (Z_n > \mu_n(\{0\}), Z \le \mu(\{0\}))$,

$$\left(\left|\tau_{\mu_n}^{\mathrm{P},Z_n} - \tau_{\mu}^{\mathrm{P},Z}\right| > \varepsilon\right) \subseteq A_n \cup B_n \cup \left(Z_n > \mu_n(\{0\}), Z > \mu(\{0\}), |\tau_{\alpha_n} - \tau_\alpha| > \varepsilon\right)$$

and $\tau_{\mu_n}^{\mathrm{P},Z_n} \to \tau_{\mu}^{\mathrm{P},Z}$ in probability. As before, there is almost sure convergence down a subsequence. □

REMARK 4.7. One easy and natural way to guarantee that $Z_n \to Z$ is to take $Z_n = Z$ with probability one, or in other words to use the same independent randomization variable for each embedding.

REMARK 4.8. Suppose that μ is less than or equal to ν in convex order (we write $\mu \le_{\mathrm{cx}} \nu$). Then $U_\mu \le U_\nu$. However, it does not follow that $\beta_\mu \ge \beta_\nu$, and so it does not follow that $\tau_\mu^{\mathrm{AY}} \le \tau_\nu^{\mathrm{AY}}$. Similarly, we do not have that $|\alpha_\mu^{\pm}| \le |\alpha_\nu^{\pm}|$ nor $\tau_\mu^{\mathrm{P}} \le \tau_\nu^{\mathrm{P}}$.

Nonetheless, given μ it is possible to choose μ_n increasing to μ in convex order and such that the barycenters are decreasing, and hence the stopping times $\tau_{\mu_n}^{\mathrm{AY}}$ are monotonically increasing and converge to μ. This idea is used extensively in Azéma and Yor [2], see also Revuz and Yor [18], Section VI.5, and also below in the proof of Theorem 7.1.

Similar remarks apply for the Perkins embedding.

EXAMPLE 4.9. In Proposition 4.4 it does not hold that $\beta_n(s) \to \beta(s)$ for s beyond the upper limit on the support of μ.

Suppose $\mu = \frac{1}{2}(\delta_1 + \delta_{-1})$ and $\mu_n = (1 - n^{-2})\frac{1}{2}(\delta_1 + \delta_{-1}) + n^{-2}\frac{1}{2}(\delta_n + \delta_{-n})$. Then $U_\mu(0) = 1$ and $U_n(0) = 1 + n^{-1} - n^{-2} \to 1$.

We have b_n is piecewise constant, and $b_n(x) = 0$ for $x < -n$, $b_n(x) = n/(2n^2 - 1)$ for $-n \le x < -1$, $b_n(x) = 1 + n^{-1} - n^{-2}$ for $-1 \le x < 1$ and $b_n(x) = n$ for $1 \le x < n$. Then $\beta_n(s) \to \beta_\infty(s)$ where $\beta_\infty(s) = -1$ for $s \le 1$ and $\beta_\infty(s) = 1$ for $s > 1$. In contrast, $\beta(s) = -1$ for $s < 1$ and $\beta(s) = s$ for $s \ge 1$.

EXAMPLE 4.10. If $\alpha_n \to \alpha_\mu$, but $U_n(0) \not\to U_\mu(0)$, then in general $\mu_n \not\Rightarrow \mu$.

Suppose $\mu = p(\delta_1 + \delta_{-1}) + (1 - 2p)\delta_0$ and $\mu_n = q(\delta_1 + \delta_{-1}) + (1 - 2q)\delta_0$. Then $\alpha_n \equiv \alpha_\mu$ but $\mu_n \not\Rightarrow \mu$ unless $p = q$.

EXAMPLE 4.11. Suppose $\alpha_n \to \alpha_\mu$ at continuity points of α_μ and $U_n(0) \to U_\mu(0)$, but $\mu_n(\{0\})$ does not tend to $\mu(\{0\})$. Then it does not follow that τ_{α_n} converges in probability, although even then we may still have $\mu_n \Rightarrow \mu$.

Let $\mu = \frac{1}{4}(\delta_1 + \delta_{-1}) + \frac{1}{2}\delta_0$, and for $n > 1$ let μ_n consist of masses of size

$$\left\{\frac{n+1}{4n}; \frac{1}{2}; \frac{n-1}{4n}\right\}$$

at $\{-1, 1/n, 1\}$, respectively. Then $\alpha^{\pm}(x) = \mp 1$, $\alpha_n^+(x) = -1$ and $\alpha_n^-(x) = 1/n$ for $-1/n \le x < 0$ and $\alpha_n^-(x) = 1$ for $x < -1/n$. Further, $\tau_\alpha = H_{\pm 1}$ and

$$\tau_{\alpha_n} = \begin{cases} H_{1/n}, & \text{if } H_{1/n} < H_{-1/n}; \\ H_{-1}, & \text{if } H_{1/n} > H_{-1/n} \text{ and } H_{-1} < H_1; \\ H_1, & \text{if } H_{1/n} > H_{-1/n} \text{ and } H_1 < H_{-1}. \end{cases}$$

Then, if E_n is the event that $\tau_{\alpha_n} = H_{1/n}$, then $\mathbb{P}(E_n) = 1/2$ and for $n > m$,

$$\mathbb{P}(E_n | E_m) = \mathbb{P}(E_m | E_n) = \mathbb{P}_{1/n}(H_{1/m} < H_{-1/m}) = \frac{n+m}{2n}.$$

Hence $\mathbb{P}(E_n \cap E_m^c) = (n-m)/4n$ which does not tend to zero as $n \to \infty$ for fixed m. Hence

$$\mathbb{P}\big(|\tau_{\alpha_n} - \tau_{\alpha_m}| > \varepsilon\big) \ge \mathbb{P}\big(H_{\pm 1} - H_{\pm 1/2} > \varepsilon, E_n \cap E_m^c\big) \not\to 0$$

and the sequence $(\tau_{\alpha_n})_{n \ge 1}$ is not Cauchy in probability.

EXAMPLE 4.12. Suppose $\alpha_n \to \alpha_\mu$ at continuity points of α_μ and $U_n(0) \to U_\mu(0)$ and $\mu_n(\{0\}) = 0 = \mu(\{0\})$. If there is no interval I containing 0 on which $\mu_n(I) = 0 = \mu(I)$, then it need not follow that $\tau_{\alpha_n} \to \tau_\alpha$ almost surely, although there is convergence in probability by Proposition 4.6(b).

Let $\mu = U\{-1, +1\}$ and for $n > 2$ let μ_n consist of masses of size

$$\left\{ \frac{n(1+2^{-n})}{2(1+n)}; \frac{1}{1+n}; \frac{n(1-2^{-n})}{2(1+n)} \right\}$$

at $\{-1, n2^{-n}, 1\}$, respectively. Then $\alpha^{\pm}(x) = \mp 1$, $\alpha_n^+(x) = -1$ and $\alpha_n^-(x) = n2^{-n}$ for $-2^{-n} \le x < 0$ and $\alpha_n^-(x) = 1$ for $x < -2^{-n}$. Further, $\tau_\alpha = H_{\pm 1}$ and

$$\tau_{\alpha_n} = \begin{cases} H_{n2^{-n}}, & \text{if } H_{n2^{-n}} < H_{-2^{-n}}; \\ H_{-1}, & \text{if } H_{n2^{-n}} > H_{-2^{-n}} \text{ and } H_{-1} < H_1; \\ H_1, & \text{if } H_{n2^{-n}} > H_{-2^{-n}} \text{ and } H_1 < H_{-1}. \end{cases}$$

Then, if E_n is the event that $\tau_{\alpha_n} \ne \tau_\alpha$, then $\mathbb{P}(E_n) = 1/(n+1)$ and for $n > m$,

$$\begin{aligned} \mathbb{P}(E_m \cap E_n) = \mathbb{P}(E_n)\mathbb{P}(E_m | E_n) &= \frac{1}{(1+n)} \frac{n2^{-n} + 2^{-m}}{m2^{-m} + 2^{-m}} \\ &= \mathbb{P}(E_n)\mathbb{P}(E_m) + \frac{n2^{m-n}}{(1+n)(1+m)}. \end{aligned}$$

Then by the Kochen–Stone lemma (Durrett [8], Exercise 1.6.19), E_n happens infinitely often, almost surely. In particular, almost surely, $\tau_{\mu_n}^{\mathrm{P}}$ does not converge.

5. Objective functions as terminal values. Our goal is to prove that for a suitable class of bivariate functions $F(w,s)$, the Azéma–Yor and Perkins embeddings, which are well known to maximize and minimize $\mathbb{E}[F(W_\tau, S_\tau)]$ in the special case where F does not depend on w and F is increasing in s, continue to optimize this quantity even when there is nontrivial dependence on w.

ASSUMPTION 5.1. Throughout we assume that $F:\{(w,s)\in\mathbb{R}\times\mathbb{R}_+; w\le s\}\mapsto\mathbb{R}_+$ is a continuous function and hence is bounded on compact sets. We further assume that the partial derivative F_s exists and is continuous.

We are interested in functions F which are monotonic in the following sense (note that in our terminology increasing is a synonym for nondecreasing).

DEFINITION 5.2. F satisfies F-MON↑ or F-MON↓ if:

F-MON↑ $F_s(w,s)/(s-w)$ is monotonic increasing in w.
F-MON↓ $F_s(w,s)/(s-w)$ is monotonic decreasing in w.

For $r\le\hat{x}\le\infty$ and $\eta\in\{\beta,\alpha^+\}$ define

$$\lambda_\eta(r)=\frac{F_s(\eta(r),r)}{r-\eta(r)},$$

$\Lambda_\eta(s)=\int_0^s\lambda_\eta(r)\,dr$ and $\Lambda_\eta^{(1)}(s)=\int_0^s r\lambda_\eta(r)\,dr$. Set $\bar{\Lambda}_\eta=\sup_{s<\hat{x}}|\Lambda_\eta(s)|$. Define $\Phi_\eta(w,s)=\int_0^s\lambda_\eta(r)(r-w)\,dr$; whence $\Phi_\eta(w,s)=\Lambda_\eta^{(1)}(s)-w\Lambda_\eta(s)$. Finally, define $\xi_\beta(w)$ by

$$\xi_\beta(w)=F\bigl(w,b(w)\bigr)-\Phi_\beta\bigl(w,b(w)\bigr)$$

and $\xi_{\alpha^+}(w)$ by

$$\xi_{\alpha^+}(w)=F\bigl(w,\bar{a}(w)\bigr)-\Phi_{\alpha^+}\bigl(w,\bar{a}(w)\bigr),$$

where $\bar{a}(w)=w$ for $w\ge 0$ and $\bar{a}(w)=a^+(w)$ for $w<0$. Note that $\xi_\beta(w)$ [resp., $\xi_\alpha(w)$] does not depend on the convention chosen for $b(w)$ [resp., $a^+(w)$].

5.1. *Target laws with bounded support.* In this section we suppose μ has bounded support so that $\check{x}$ and $\hat{x}$ are finite. This assumption will be relaxed in the next section.

THEOREM 5.3. *Suppose that μ has bounded support and that F-MON↑ holds. Then*

$$\sup_{\tau\in\mathcal{S}_{\mathrm{UI}}(W,\mu)}\mathbb{E}\bigl[F(W_\tau,S_\tau)\bigr]=\mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{AY}}},S_{\tau_\mu^{\mathrm{AY}}})\bigr],\tag{5.1}$$

$$\inf_{\tau\in\mathcal{S}_{\mathrm{UI}}(W,\mu)}\mathbb{E}\bigl[F(W_\tau,S_\tau)\bigr]=\mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{P}}},S_{\tau_\mu^{\mathrm{P}}})\bigr].\tag{5.2}$$

REMARK 5.4. In the case where μ has no atoms (so that the arg min in (3.1) is strictly increasing and $\mathbb{E}[X|X \geq x] = \mathbb{E}[X|X > x]$), then we can write

$$\mathbb{E}\big[F(W_{\tau_\beta}, S_{\tau_\beta})\big] = \int_{\mathbb{R}} F\big(w, b_\mu(w)\big)\mu(dw). \tag{5.3}$$

This formula need not hold if μ has atoms.

In cases where μ has a strictly positive density ρ on $(\check{x}, \hat{x})$ and β is differentiable, the expression in (5.3) can be rewritten as

$$\begin{aligned}\mathbb{E}\big[F(W_{\tau_\beta}, S_{\tau_\beta})\big] &= \int_{\mathbb{R}} F\big(\beta(s), s\big)\mathbb{P}(S_{\tau_\beta} \in ds) \\ &= \int_{\mathbb{R}} F\big(\beta(s), s\big)\rho\big(\beta(s)\big)\beta'(s)\, ds,\end{aligned} \tag{5.4}$$

where we use the fact that in the atom-free case

$$\mu\big([\beta(s), \infty)\big) = \mathbb{P}\big(W_{\tau_\beta} \geq \beta(s)\big) = \mathbb{P}(S_{\tau_\beta} \geq s).$$

A similar remark applies to $\mathbb{E}[F(W_{\tau_\mu^{\mathrm{P}}}, S_{\tau_\mu^{\mathrm{P}}})] = \int_{\mathbb{R}} F(w, \bar{a}(w))\mu(dw)$.

REMARK 5.5. The requirement that the infimum in (5.2) is taken over $\tau \in \mathcal{S}_{\mathrm{UI}}(W, \mu)$ (and not over all embeddings) is necessary, as can be seen by considering $F(w, s) = -(s - w)^3$. However, if we restrict attention to functions F which are increasing in s, then we may replace the infimum in (5.2) with an infimum over all embeddings.

The key to the proof of the theorem is the following lemma.

LEMMA 5.6. *Suppose F satisfies F-MON↑. Then, for all $w \leq s$*

$$\xi_{\alpha^+}(w) + \Phi_{\alpha^+}(w, s) \leq F(w, s) \leq \xi_\beta(w) + \Phi_\beta(w, s)$$

with equality on the left at $w = s$ and $w = \alpha^+(w)$ and equality on the right at $w = \beta(s)$.

PROOF. For $\eta \in \{\beta, \alpha^+\}$ define

$$L_\eta(w, s) = \left[F(w, s) - \xi_\eta(w) - \int_0^s \lambda_\eta(r)(r - w)\, dr\right]. \tag{5.5}$$

We will show that $L_{\alpha^+}(w, s) \geq 0$ with equality at $w = s$ and $w = \alpha^+(s)$, and $L_\beta(w, s) \leq 0$ with equality at $w = \beta(s)$.

Consider the latter inequality first:

$$\begin{aligned}L_\beta(w, s) &= F(w, s) - \xi_\beta(w) - \int_0^s \lambda_\beta(r)(r - w)\, dz \\ &= F(w, s) - F\big(w, b(w)\big) + \int_0^{b(w)} dr\, F_s\big(\beta(r), r\big)\frac{r - w}{r - \beta(r)}\end{aligned}$$

$$\begin{aligned}&-\int_0^s dr\, F_s(\beta(r),r)\frac{r-w}{r-\beta(r)}\\&=\int_{b(w)}^s\left\{\frac{F_s(w,r)}{r-w}-\frac{F_s(\beta(r),r)}{r-\beta(r)}\right\}(r-w)\,dr.\end{aligned}$$

If $b(w) < r < s$, then since β is increasing, $w < \beta(r)$ and by F-MON↑ the integrand is negative. If $s < r < b(w)$, then $w > b(r)$ and the integrand is positive. Thus $L_\beta(w,s) \le 0$ as required. Clearly, there is equality at $s = b(w)$.

For L_{α^+} a similar calculation to the one above shows that

$$L_{\alpha^+}(w,s)=\int_{\bar{a}(w)}^s\left\{\frac{F_s(w,r)}{r-w}-\frac{F_s(\alpha^+(r),r)}{r-\alpha^+(r)}\right\}(r-w)\,dr.$$

To see that $L_{\alpha^+}(w,s) \ge 0$, consider $w \ge 0$ and $w < 0$ separately. For $w \ge 0$, $\bar{a}(w) = w$ and for $w < r < s$, $\alpha^+(r) \le \alpha^+(w) \le w$ so that the integrand is positive and $L_{\alpha^+}(w,s) \ge 0$. For $w < 0$, $\bar{a}(w) = a(w)$, and then if $a(w) < r < s$, we have $w > \alpha^+(r)$ and the integrand is positive. Otherwise if $s < r < a(w)$, $w < \alpha^+(r)$ and the integrand is negative. In either case, allowing for the limits on the integral, $L_{\alpha^+}(w,s) \ge 0$. Equality holds at $w = s$ and $w = \alpha^+(s)$. □

REMARK 5.7. Essentially, the idea behind Lemma 5.6 and the proof of Theorem 5.3 is to interpret the embedding property and Doob's (in)-equality for the martingale W as linear constraints on the possible joint laws of (W_τ, S_τ), with associated Lagrange multipliers. Thus, if the joint law is given by $\nu(dw,ds)$, then $\int_{s\ge r}(w-r)\nu(dw,ds) = 0$ (which is equivalent to (3.2) in Rogers [20]). There is an identity of this form for each r and when they are integrated against a family of Lagrange multipliers $\lambda_\eta(r)$ we obtain

$$0=\int_0^\infty \lambda_\eta(r)\int_{s\ge r}(w-r)\nu(dw,ds)=\int \nu(dw,ds)\int_{0\le r\le s}\lambda_\eta(r)(w-r)\,dr.$$

The integrand of this last expression appears as the last term in (5.5).

It remains to prove Theorem 5.3. The main idea for the proof of the theorem is that provided that $\bar{\Lambda}_\beta$ and $\bar{\Lambda}_{\alpha^+}$ are finite, then for $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$ both $(\Phi_{\alpha^+}(W_t^\tau, S_t^\tau))_{t\ge0}$ and $(\Phi_\beta(W_t^\tau, S_t^\tau))_{t\ge0}$ are uniformly integrable martingales. [By Itô's formula, $d\Phi_\eta(W_t,S_t) = -\Lambda_\eta(S_t)\,dW_t$ since the finite variation term involves the product $(S_t - W_t)\,dS_t$ and when S is increasing we must also have $S_t - W_t = 0$.] It follows that $\mathbb{E}[\Phi_\beta(W_\tau,S_\tau)] = 0$ and

$$\mathbb{E}[\xi_{\alpha^+}(W_\tau)] \le \mathbb{E}[F(W_\tau,S_\tau)] \le \mathbb{E}[\xi_\beta(W_\tau)],$$

which, given the forms of ξ_α and ξ_β leads to the first result given in the Introduction.

REMARK 5.8. The processes $(\Phi_{\alpha^+}(W_t^\tau, S_t^\tau))_{t\ge 0}$ and $(\Phi_\beta(W_t^\tau, S_t^\tau))_{t\ge 0}$ belong to the class of Azéma–Yor martingales. A martingale $M = (M_t)_{t\ge 0}$ is an Azéma–Yor martingale if $M_t = G(S_t^X) - (S_t^X - X_t)g(S_t)$ for X a martingale and $G' = g$; see [2].

REMARK 5.9. An alternative derivation of (the right inequality of) Lemma 5.6 is to look for pathwise inequalities $F(W_t, S_t) \le \xi(W_t) + M_t$ such that M_t is a Markovian function of W_t and S_t and such that there is equality at $S_t = b(W_t)$.

If $M_t = \Phi(W_t, S_t)$ and Φ is appropriately differentiable, then M must be an Azéma–Yor martingale $\Phi(W_t, S_t) = -H(S_t) + H'(S_t)(S_t - W_t)$ for some H. Further, if there is to be equality at $s = b(w)$, then we must have $\xi(w) = F(w, b(w)) - \Phi(w, b(w))$. Then we want conditions on F such that there is an inequality $F(w,s) \le \xi(w) + \Phi(w,s)$, or equivalently

$$\begin{aligned}\int_{b(w)}^s F_s(w,r)\,dr &= F(w,s) - F\big(w, b(w)\big)\\ &\le \Phi(w,s) - \Phi\big(w,b(w)\big) = \int_{b(w)}^s \Phi_s(w,r)\,dr\\ &= \int_{b(w)}^s H''(r)(r-w)\,dr.\end{aligned}$$

From this it follows that a sufficient condition is $F_s(w,r) \le H''(r)(r-w)$ for $r > b(w)$ and the reverse inequality for $r < b(w)$, which holds if F-MON↑ holds and $H''(s) = F_s(\beta(s), s)/(s - \beta(s))$.

PROOF OF THEOREM 5.3. Consider first the bound associated with the Azéma–Yor embedding. $\bar\Lambda_\beta$ depends on the combination of μ and F.

Suppose that μ has an atom at $\hat{x}$. By Lemma 3.2 $(r - \beta(r))^{-1}$ is integrable near zero so that if μ has an atom at $\hat{x}$, then $r - \beta(r)$ is bounded below for $r < \hat{x}$ and $\bar\Lambda_\beta < \infty$. Since $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$ implies $(W_t^\tau)_{t\ge 0}$ is bounded, and since $\Lambda_\beta(s)$ and $\Lambda^{(1)}(s)$ are bounded, we have that $\Phi_\beta(W_t^\tau, S_t^\tau)$ is a bounded local martingale and hence $\mathbb{E}[\Phi_\beta(W_t^\tau, S_t^\tau)] = 0$, which can be re-expressed as $\mathbb{E}[\Lambda_\beta^{(1)}(S_\tau)] = \mathbb{E}[W_\tau \Lambda_\beta(S_\tau)]$. In view of Lemma 5.6 we have

$$F(W_\tau, S_\tau) \le \xi_\beta(W_\tau) + \Phi_\beta(W_\tau, S_\tau). \tag{5.6}$$

Thus

$$\mathbb{E}\big[F(W_\tau, S_\tau)\big] \le \int \xi_\beta(w)\mu(dw).$$

Note that for $\tau = \tau_\beta$, we have equality in (5.6) and hence equality in this last expression.

Now suppose there is no atom at $\hat{x}$. Fix $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$ and let $\sigma_n = \tau \wedge H_{\check{x}-1/n}$ and $\mu_n = \mathcal{L}(W_{\sigma_n})$. Then $U_{\mu_n} \to U_\mu$ for each x and by bounded convergence we have both

$$\mathbb{E}\bigl[F(W_\tau, S_\tau)\bigr] = \mathbb{E}\bigl[\lim F(W_{\sigma_n}, S_{\sigma_n})\bigr] = \lim \mathbb{E}\bigl[F(W_{\sigma_n}, S_{\sigma_n})\bigr]$$

and

$$\mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{AY}}}, S_{\tau_\mu^{\mathrm{AY}}})\bigr] = \mathbb{E}\bigl[\lim F(W_{\tau_{\mu_n}^{\mathrm{AY}}}, S_{\tau_{\mu_n}^{\mathrm{AY}}})\bigr] = \lim \mathbb{E}\bigl[F(W_{\tau_{\mu_n}^{\mathrm{AY}}}, S_{\tau_{\mu_n}^{\mathrm{AY}}})\bigr].$$

The result follows from the previous case on comparing σ_n with $\tau_{\mu_n}^{\mathrm{AY}}$.

The proof of (5.2) is identical except that there is no need to separate the case where there is an atom at $\hat{x}$, since by Lemma 3.5 $(r - \alpha^+(r))^{-1}$ is integrable near zero and hence the fact that μ has bounded support is sufficient for $\bar{\Lambda}_{\alpha^+} < \infty$. $\square$

There are a parallel pair of results based on F-MON↓, the proofs of which are very similar.

LEMMA 5.10. *Suppose F satisfies F-MON↓. Then, for all $w \le s$*

$$\xi_\beta(w) + \Phi_\beta(w,s) \le F(w,s) \le \xi_{\alpha^+}(w) + \Phi_{\alpha^+}(w,s)$$

with equality on the right at $s = w$ and $s = a(w)$ and equality on the left at $s = b(w)$.

THEOREM 5.11. *Suppose F-MON↓ holds. Then*

$$\inf_{\tau \in \mathcal{S}(W,\mu)} \mathbb{E}\bigl[F(W_\tau, S_\tau)\bigr] = \mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{AY}}}, S_{\tau_\mu^{\mathrm{AY}}})\bigr],$$

$$\sup_{\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)} \mathbb{E}\bigl[F(W_\tau, S_\tau)\bigr] = \mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{P}}}, S_{\tau_\mu^{\mathrm{P}}})\bigr].$$

EXAMPLE 5.12. Suppose $\mu = U[-1,1]$ and $F(w,s) = (s-w)^c$ for $c > -1$ (with $c \neq 0$). Then for $c \ge 2$ F-MON↓ holds, for $0 < c \le 2$ F-MON↑ holds and for $-1 < c < 0$, F-MON↓ holds again.

Write B^{AY} and B^{P} for the bounds associated with the Azéma–Yor and Perkins embeddings.

Recall the expressions for β and α from Examples 3.1 and 3.3.

For the Azéma–Yor embedding, $\beta(s) = 2s - 1$ and the law of the S_{τ_β} is a uniform on [0, 1]. The associated bound (as a function of the parameter c) is given by

$$B^{\mathrm{AY}}(c) = \mathbb{E}\bigl[F(W_{\tau_\mu^{\mathrm{AY}}}, S_{\tau_m^{\mathrm{AY}} u})\bigr] = \int_{-1}^{1} \bigl(b(w) - w\bigr)^c \frac{dw}{2} = \int_0^1 \bigl(s - \beta(s)\bigr)^c\, ds$$
$$= \int_0^1 (1-s)^c\, ds = \frac{1}{c+1}.$$

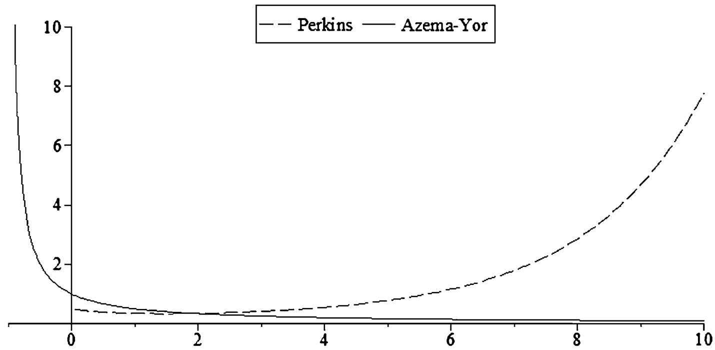

FIG. 3. *All uniformly integrable embeddings have the same expected value when $c = 2$. Note the reversal of the bounds at $c = 2$: for $0 < c < 2$ Theorem 5.3 applies while for $c > 2$ Theorem 5.11 applies. For $c < 0$, the Perkins bound is infinite and the Azéma–Yor bound is finite. The Perkins bound as a function of c is discontinuous at $c = 0$.*

For the Perkins bound, note that for $c < 0$, $F(s,s) = \infty$, and so $B^{\mathrm{P}}(c) = 0$. For $c > 0$, $F(s,s) = 0$ and using the substitution $w = \alpha^+(s) = s - 2\sqrt{s}$,

$$B^{\mathrm{P}}(c) = \mathbb{E}[F(W_{\tau_\mu^{\mathrm{P}}}, S_{\tau_\mu^{\mathrm{P}}})] = \int_{-1}^{0} (a^+(w) - w)^c \frac{dw}{2}$$
$$= \frac{2^c}{(c+1)(c+2)}.$$

Results for a range of c are plotted in Figure 3. Observe that for $c = 2$, $B^{\mathrm{AY}}(2) = B^{\mathrm{P}}(2) = 1/3$ and all uniformly integrable embeddings for the terminal law are consistent with the same expected payoff. The reason for this will become clear in Section 7 and will correspond to the choice $g \equiv 1$.

In fact Assumption 5.1 is not satisfied for $-1 < c < 1$. Nonetheless, for c in this range and $\varepsilon > 0$ we can let $F_\varepsilon(w,s) = h_\varepsilon(s-w)$ where $h_\varepsilon(x) = x^c$ for $x \geq \varepsilon$ and $h_\varepsilon(x) = \varepsilon^c + c\varepsilon^{c-1}(x-c)$ for $x < \varepsilon$. Then F_ε does satisfy Assumption 5.1, and F and F_ε satisfy F-MON↑ or F-MON↓ together. Then arguments of Theorem 5.3 provide the upper and lower bounds for F_ε, and letting $\varepsilon \downarrow 0$ we obtain the pictured bounds for F.

EXAMPLE 5.13. Suppose again that $\mu = U[-1,1]$. Let $F(w,s) = \frac{(s-w)^2}{s^c}$. Note that for each c either F-MON↑ or F-MON↓ (or both) holds, so that the Azéma–Yor and Perkins embeddings give extremal values for $\mathbb{E}[F(W_\tau, S_\tau)]$. Consider the Azéma–Yor bound as a function of the parameter c (defined for $c < 1$),

$$B^{\mathrm{AY}}(c) = \int_{-1}^{1} \frac{(b(w)-w)^2}{b(w)^c} \frac{dw}{2} = \int_0^1 \frac{(s-\beta(s))^2}{s^c} ds = \int_0^1 \frac{(s-1)^2}{s^c} ds$$
$$= \frac{2}{(1-c)(2-c)(3-c)}.$$

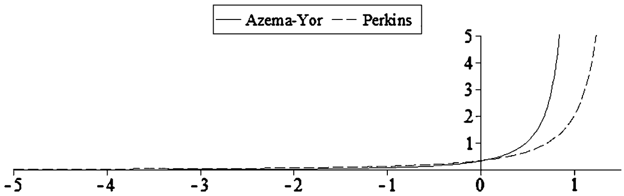

FIG. 4. *For $1 < c < 3/2$ the Azéma–Yor upper bound is infinite while the Perkins lower bound is finite.*

For the Perkins bound we have (for $c < 3/2$)

$$\begin{aligned} B^{\mathrm{P}}(c) &= \int_{-1}^{0} \frac{(a^+(w) - w)^2}{a^+(w)} \frac{dw}{2} \\ &= \int_0^1 \frac{2\sqrt{s}}{s^c}(1 - \sqrt{s})\, ds \\ &= \frac{1}{(3/2 - c)(2 - c)}. \end{aligned}$$

Observe that the expressions for $B^{\mathrm{AY}}(\cdot)$ and $B^{\mathrm{P}}(\cdot)$ co-incide at $c = 0$ where both F-MON↑ and F-MON↓ hold. See Figure 4.

6. General centered target measures.

THEOREM 6.1. *Fix $\tau \in \mathcal{S}_{\mathrm{UI}}(W, \mu)$. Suppose, in addition to Assumption 5.1, that $F \geq 0$, that*

$$\mathbb{E}\big[F(W_{H_{\pm n}}, S_{H_{\pm n}}); \tau \geq H_{\pm n}\big] \to 0 \tag{6.1}$$

and that if $(\mu_n)_{n\geq 1}$ is any sequence of measures which is increasing in convex order for which $\mu_n \Rightarrow \mu$, $U_{\mu_n}(0) \to U_\mu(0)$ and $\mu_n(\{0\}) \to \mu(\{0\})$, then both

$$\mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu_n}}, S_{\tau^{\mathrm{AY}}_{\mu_n}})\big] \to \mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu}}, S_{\tau^{\mathrm{AY}}_{\mu}})\big] \tag{6.2}$$

and

$$\mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu_n}}, S_{\tau^{\mathrm{P}}_{\mu_n}})\big] \to \mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu}}, S_{\tau^{\mathrm{P}}_{\mu}})\big]. \tag{6.3}$$

Then if F-MON↑ holds,

$$\mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu}}, S_{\tau^{\mathrm{P}}_{\mu}})\big] \leq \mathbb{E}\big[F(W_\tau, S_\tau)\big] \leq \mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu}}, S_{\tau^{\mathrm{AY}}_{\mu}})\big],$$

whereas, if F-MON↓ holds, then

$$\mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu}}, S_{\tau^{\mathrm{AY}}_{\mu}})\big] \leq \mathbb{E}\big[F(W_\tau, S_\tau)\big] \leq \mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu}}, S_{\tau^{\mathrm{P}}_{\mu}})\big].$$

PROOF. Suppose F-MON↑ holds (the proof for F-MON↓ is similar). Given $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$, let $\sigma_n = \tau \wedge H_{\pm n}$, $\mu_n = \mathcal{L}(W_{\sigma_n})$ and define $\tau^{\mathrm{AY}}_{\mu_n}$ and $\tau^{\mathrm{P}}_{\mu_n}$ to be the Azéma–Yor and Perkins stopping times associated with μ_n.

We have, using monotone convergence, (6.1), Theorem 5.3 and finally (6.2),

$$\begin{aligned}
\mathbb{E}\big[F(W_\tau, S_\tau)\big] &= \mathbb{E}\big[\lim F(W_{\sigma_n}, S_{\sigma_n}); \sigma_n = \tau \le H_{\pm n}\big] \\
&= \lim \mathbb{E}\big[F(W_{\sigma_n}, S_{\sigma_n}) I_{\{\tau \le H_{\pm n}\}}\big] \\
&= \lim \mathbb{E}\big[F(W_{\sigma_n}, S_{\sigma_n}) I_{\{\tau < H_{\pm n}\}} + F(W_{H_{\pm n}}, S_{H_{\pm n}}) I_{\{\tau \ge H_{\pm n}\}}\big] \\
&= \lim \mathbb{E}\big[F(W_{\sigma_n}, S_{\sigma_n})\big] \\
&\le \lim \mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu_n}}, S_{\tau^{\mathrm{AY}}_{\mu_n}})\big] = \mathbb{E}\big[F(W_{\tau^{\mathrm{AY}}_{\mu}}, S_{\tau^{\mathrm{AY}}_{\mu}})\big].
\end{aligned}$$

Similarly

$$\lim \mathbb{E}\big[F(W_{\sigma_n}, S_{\sigma_n})\big] \ge \lim \mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu_n}}, S_{\tau^{\mathrm{P}}_{\mu_n}})\big] = \mathbb{E}\big[F(W_{\tau^{\mathrm{P}}_{\mu}}, S_{\tau^{\mathrm{P}}_{\mu}})\big]. \qquad \square$$

COROLLARY 6.2. *Suppose that $F(w,s) \le A(1 + |w|^k + s^k)$ for $k \ge 1$ and that μ has finite $k + \varepsilon$ moment, for some positive ε. Then the hypotheses (6.1), (6.2) and (6.3) are all satisfied, and the conclusions of Theorem 6.1 hold.*

PROOF. By Doob's submartingale inequality for $(|W_{t\wedge\tau}|^{k+\varepsilon})_{t\ge 0}$, for any $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$,

$$m^{k+\varepsilon}\mathbb{P}(\tau > H_{\pm m}) < \mathbb{E}\big[|W_\tau|^{k+\varepsilon}\big] < \infty.$$

Then

$$\mathbb{E}\big[F(W_{H_{\pm n}}, S_{H_{\pm n}}); \tau \ge H_{\pm n}\big] \le A\big(1 + 2n^k\big)\mathbb{P}(\tau > H_{\pm n}) \to 0.$$

For (6.2) we have that $\tau_{\beta_n} \to \tau_\beta$ almost surely. Moreover, since $\mu_n \le_{\mathrm{cx}} \mu$ there exists a stopping time (ρ_n say) with $\rho_n \ge \tau_{\beta_n}$ and $\rho_n \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$. For such a ρ_n, $\mathbb{E}[|W_{\rho_n}|^{k+\varepsilon}] = \int_{\mathbb{R}} |x|^{k+\varepsilon}\mu(dx) < \infty$ by hypothesis, and then (defining $W^*_t = \sup_{s\le t}|W_s|$) by Doob's L^p inequality $\mathbb{E}[(W^*_{\rho_n})^{k+\varepsilon}] \le D < \infty$ for some constant D, independent of n.

Set $F_n = F(W_{\tau^{\mathrm{AY}}_{\mu_n}}, S_{\tau^{\mathrm{AY}}_{\mu_n}})$ and $F = F(W_{\tau^{\mathrm{AY}}_{\mu}}, S_{\tau^{\mathrm{AY}}_{\mu}})$, then $F_n \to F$ almost surely. The goal is to show that $\mathbb{E}[F_n] \to \mathbb{E}[F]$ which will follow if $\sup_n \mathbb{E}[(F_n)^p] < \infty$, for then $(F_n)_{n\ge 1}$ is uniformly integrable. We have that if $|w| \le x$ and $s \le x$, then with $p = 1 + k/\varepsilon$,

$$F(w,s)^p \le A^p\big(1 + 2x^k\big)^p \le A^p 3^p\big(1 + x^{kp}\big).$$

Hence

$$\mathbb{E}\big[F_n^p\big] \le A^p 3^p\big(1 + \mathbb{E}\big[\big(W^*_{\tau_n}\big)^{kp}\big]\big) \le A^p 3^p\big(1 + \mathbb{E}\big[\big(W^*_{\rho_n}\big)^{kp}\big]\big) \le A^p 3^p(1 + D) < \infty.$$

For (6.3), consider a subsequence $n(k)$. Then down a further subsequence $\tau^{\mathrm{P}}_{\mu_n} \to \tau^{\mathrm{P}}_{\mu}$ almost surely and down this subsequence (6.3) holds by identical arguments as in the case for the Azéma–Yor embedding. Hence (6.3) holds. $\square$

7. Objective functions as running costs. Our original aim in studying functions $F(w,s)$ was as an aid in the analysis of the expected values of integrals of the form $\int_0^\tau g(S_t)\,dt$. Motivated by the variance swap problem in mathematical finance we asked:

> Given g and μ, what is the range of possible values of $\mathbb{E}[\int_0^\tau g(S_u)\,du]$ over embeddings τ of μ in Brownian motion?

Our aim is to reduce this problem to the case previously considered, but to use the extra structure to prove more powerful results under weaker hypotheses.

The expected value of $\int_0^\tau g(S_u)\,du$ is intimately related to the value of $\mathbb{E}[G(W_\tau, S_\tau)]$ where $G(w,s) = (s-w)^2 g(s)$. Indeed, if g is continuously differentiable, then by Itô's lemma,

$$G(W_\tau, S_\tau) = G(0,0) + \int_0^\tau g(S_u)\,du - \int_0^\tau 2(S_u - W_u)g(S_u)\,dW_u, \tag{7.1}$$

so that if $g(0)$ is finite [and then $G(0,0)=0$], and if

$$\left(\int_0^{\tau\wedge t} 2(S_u - W_u)g(S_u)\,dW_u\right)_{t\geq 0}$$

is a uniformly integrable martingale, then $\mathbb{E}[\int_0^\tau g(S_u)\,du] = \mathbb{E}[G(W_\tau, S_\tau)]$.

If g is increasing (resp., decreasing), then G satisfies G-MON$\downarrow$ (resp., G-MON$\uparrow$), and we can apply the results of previous sections to deduce that the Azéma–Yor and Perkins solutions give bounds $\mathbb{E}[\int_0^\tau g(S_u)\,du]$ over embeddings τ of μ.

THEOREM 7.1. *Suppose $g:\mathbb{R}_+ \mapsto \mathbb{R}_+$ is a positive function and that μ is centered.*

(i) *Suppose g is increasing. Then*

$$\inf_{\tau\in\mathcal{S}(W,\mu)} \mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] = \mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{AY}}} g(S_u)\,du\right]$$

and

$$\sup_{\tau\in\mathcal{S}_{\mathrm{UI}}(W,\mu)} \mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] = \mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{P}}} g(S_u)\,du\right].$$

(ii) *Suppose g is decreasing. Then*

$$\inf_{\tau\in\mathcal{S}(W,\mu)} \mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] = \mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{P}}} g(S_u)\,du\right]$$

and

$$\sup_{\tau\in\mathcal{S}_{\mathrm{UI}}(W,\mu)} \mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] = \mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{AY}}} g(S_u)\,du\right].$$

REMARK 7.2. As we remarked in the Introduction, at first sight this result is counterintuitive. Given increasing g, the Azéma–Yor stopping time maximizes $\mathbb{E}[g(S_\tau)]$ over $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$, and it seems plausible that it might also maximize $\mathbb{E}[\int_0^\tau g(S_u)\,du]$. In fact the exact opposite is true. The explanation is that for the Azéma–Yor embedding there is co-monotonicity[2] between S_τ and W_τ, and conditional on $S_\tau \geq s$, the stopping time occurs quite soon [and certainly before W drops below $\beta(s)$], whereas for the Perkins embedding, conditional on $S_\tau \geq s$, there are paths which will only be stopped when W goes below $\alpha^+(s)$. Thus, for increasing g when we wish to maximize the time (before τ) for which S is large, this is best achieved by the Perkins embedding: although relatively few paths will have large S (most will have already been stopped) those with a large maximum will spend a long time after first hitting s before being stopped.

EXAMPLE 7.3. Recall Example 5.13. Suppose $\mu = U[-1,1]$ and $g(s) = s^{-c}$. Then, for $c<0$, $(1-c)^{-1}(2-c)^{-1}(3-c)^{-1} \leq \mathbb{E}[\int_0^\tau S_u^{-c}\,du] \leq (2-c)^{-1}(3/2-c)^{-1}$.

For $0<c<1$, $(2-c)^{-1}(3/2-c)^{-1} \leq \mathbb{E}[\int_0^\tau S_u^{-c}\,du] \leq (1-c)^{-1}(2-c)^{-1}(3-c)^{-1}$, for $1 \leq c < 3/2$, $(2-c)^{-1}(3/2-c)^{-1} \leq \mathbb{E}[\int_0^\tau S_u^{-c}\,du] \leq \infty$ and for $c \geq 3/2$, $\mathbb{E}[\int_0^\tau S_u^{-c}\,du] = \infty$ for all embeddings τ.

Note that for $c=0$, $\mathbb{E}[\tau]$ is independent of τ and equal to the variance of μ.

EXAMPLE 7.4. Recall the calculations from Example 3.4. Let the target law μ with support $[-1,\infty)$ satisfy $\mu(dx) = \frac{2}{(x+2)^3}\,dx$. Let $g(s) = \frac{1}{c+s}$ for $c>0$ which is decreasing in s.

The Azéma–Yor upper bound can be calculated explicitly to be

$$\begin{aligned} B^{\mathrm{AY}}(c) &= \int_{-1}^{\infty} \frac{(b(w)-w)^2}{b(w)+c} \frac{2}{(w+2)^3}\,dw \\ &= \frac{2(\log(c)-\log(2))}{c-2}. \end{aligned}$$

The expression for the Perkins lower bound is given by

$$B^{\mathrm{P}}(c) = \int_{-1}^{\infty} \frac{(a^+(w)-w)^2}{a^+(w)+c} \frac{2}{(w+2)^3}\,dw.$$

The expression for α^+ is too complicated for the expression above to have an analytic representation. However, the values can be computed numerically for different c.

[2]A pair of random variables X and Y is co-monotonic if $\mathbb{P}(X \leq x, Y \leq y) = \min\{\mathbb{P}(X \leq x), \mathbb{P}(Y \leq y)\}$ for all x and y.

The rest of this section is devoted to a proof of Theorem 7.1. We split the proof into four separate parts.

PROOF OF THEOREM 7.1(i): LOWER BOUND. Suppose first that g is monotonic increasing and that we are interested in minimizing the quantity $\mathbb{E}[\int_0^\tau g(S_u)\,du]$ over embeddings τ of μ in W. Note that it is sufficient to restrict attention to $\mathcal{S}_{\mathrm{UI}}(W,\mu)$: for nonminimal $\tau \in \mathcal{S}(W,\mu)$ there exists $\tilde{\tau} \le \tau$ with $\tilde{\tau} \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$, and then $\int_0^\tau g(S_u)\,du \ge \int_0^{\tilde{\tau}} g(S_u)\,du$ for each $\omega \in \Omega$.

Suppose temporarily that g is bounded and continuously differentiable. Later we will relax this assumption. Then $G(w,s) = (s-w)^2 g(s)$ satisfies G-MON↓.

For $\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)$ let $\sigma_n = \tau \wedge H_{\pm n}$, let $\mu_n = \mathcal{L}(W_{\sigma_n})$, β_n be the inverse barycenter of μ_n and finally let $\tau^{\mathrm{AY}}_{\mu_n}$ be the Azéma–Yor stopping rule associated with the law μ_n so that $\tau^{\mathrm{AY}}_{\mu_n} = \tau_{\beta_n} = \inf\{u : W_u \le \beta_n(S_u)\}$. Then, by Proposition 4.4, since $U_{\mu_n} \uparrow U_\mu$, $\tau_{\beta_n} \to \tau_\beta$ almost surely.

If a stopping time ρ is such that $\rho \le H_{\pm n}$, then $\mathbb{E}[\rho] < \infty$ and for $u \le \rho$, $(S_u - W_u)g(S_u)$ is bounded. Then if $M_t = \int_0^t (S_u - W_u)g(S_u)\,dW_u$, we have that $(M_t^\rho)_{t\ge 0}$ is an L^2 bounded martingale for which

$$\mathbb{E}[M_\infty^\rho] = \mathbb{E}\left[\int_0^\rho (S_u - W_u)g(S_u)\,du\right] = 0. \tag{7.2}$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[\int_0^{\sigma_n} g(S_u)\,du\right] &= \mathbb{E}\left[(S_{\sigma_n} - W_{\sigma_n})^2 g(S_{\sigma_n})\right] \\
&\ge \mathbb{E}\left[(S_{\tau_{\beta_n}} - W_{\tau_{\beta_n}})^2 g(S_{\tau_{\beta_n}})\right] \\
&= \mathbb{E}\left[\int_0^{\tau_{\beta_n}} g(S_u)\,du\right],
\end{aligned}$$

where we have used (7.1) and (7.2) twice and Theorem 5.11. Then it follows from the Fatou lemma that

$$\begin{aligned}
\lim \mathbb{E}\left[\int_0^{\sigma_n} g(S_u)\,du\right] &\ge \lim \mathbb{E}\left[\int_0^{\tau_{\beta_n}} g(S_u)\,du\right] \\
&\ge \mathbb{E}\left[\liminf \int_0^{\tau_{\beta_n}} g(S_u)\,du\right]
\end{aligned} \tag{7.3}$$

and by monotone convergence and the fact that $\tau_{\beta_n} \to \tau_\beta$ almost surely,

$$\mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] \ge \mathbb{E}\left[\int_0^{\tau_\beta} g(S_u)\,du\right]$$

as required.

Finally we remove the temporary assumptions on g. Given g is monotonic increasing we can find an increasing sequence of bounded, continuously differentiable (increasing) functions g_m which approximate g from below. Then, by monotone convergence

$$\mathbb{E}\biggl[\int_0^{\tau} g(S_u)\,du\biggr] = \lim_m \mathbb{E}\biggl[\int_0^{\tau} g_m(S_u)\,du\biggr] \geq \lim_m \mathbb{E}\biggl[\int_0^{\tau_\beta} g_m(S_u)\,du\biggr]$$
$$= \mathbb{E}\biggl[\int_0^{\tau_\beta} g(S_u)\,du\biggr].$$

Note that this same argument will apply in all four parts of Theorem 7.1, and henceforth without loss of generality we will assume that g is continuously differentiable and bounded by $\bar{g}$.

PROOF OF THEOREM 7.1(ii): LOWER BOUND. *Case* 1: There exists an open interval $I \subseteq [-1, 1]$ containing 0 with $\mu(I) = 0$.

Given $\tau \in \mathcal{S}(W, \mu)$, let $\sigma_m = \tau \wedge H_{\pm m}$. Let $\mu_m = \mathcal{L}(W_{\sigma_m})$. Write τ_m^{P} for the Perkins embedding of μ_m. Note that $\mu_m \Rightarrow \mu$, $U_{\mu_m}(0) \to U_\mu(0)$ and $\mu_m(I) = 0$. Then, $\tau_m^{\mathrm{P}} = \tau_{\alpha_m}$ and by Proposition 4.6(a), $\tau_{\alpha_m} \to \tau_\alpha$ almost surely. Then exactly as in (7.3), but now using Theorem 5.3 to give that the lower bound is attained by the Perkins embedding, we conclude that $\mathbb{E}[\int_0^{\tau} g(S_u)\,du] \geq \mathbb{E}[\int_0^{\tau_\mu^{\mathrm{P}}} g(S_u)\,du]$.

Case 2: General μ. Given any subsequence, by Proposition 4.6(b) we may take a further subsequence down which $\tau_m^{\mathrm{P}} \to \tau^{\mathrm{P}}$ almost surely. Then down this subsequence the result holds, as in case 1. Since the first subsequence was arbitrary we are done.

PROOF OF THEOREM 7.1(ii): UPPER BOUND. Now consider the upper bound in Theorem 7.1(ii). Rather than attempting to find a dominating random variable which will allow us to use the reverse Fatou lemma in place of the Fatou lemma above we will use a slightly different approach based on defining a sequence of intermediate stopping times.

Let τ be any element of $\mathcal{S}_{\mathrm{UI}}(W, \mu)$. Suppose g is bounded, continuously differentiable and monotonic decreasing, and that μ has support in a bounded interval $[\check{x}, \hat{x}]$. Then, as above, $\mathbb{E}[\int_0^{\tau} g(S_u)\,du] = \mathbb{E}[G(W_\tau, S_\tau)]$. Moreover, we can conclude from Theorem 5.3 that

$$\sup_{\tau \in \mathcal{S}_{\mathrm{UI}}(W,\mu)} \mathbb{E}\biggl[\int_0^{\tau} g(S_u)\,du\biggr] = \mathbb{E}\biggl[\int_0^{\tau_\beta} g(S_u)\,du\biggr].$$

It remains to remove the assumptions on μ.

Given ε, let $U_\varepsilon(x) = \max\{U_\mu(x) - \varepsilon, |x|\}$, and let $\check{x}_\varepsilon$ and $\hat{x}_\varepsilon$ be the left and right-hand endpoints of the interval $I_\varepsilon = \{x : U_\varepsilon(x) > |x|\}$.

Let $\sigma_\varepsilon = \tau \wedge \inf\{u : W_u \notin I_\varepsilon\}$. Let $\bar{\mu}_\varepsilon$ be the law of W_{σ_ε}, and let $\bar{U}_\varepsilon$ be the associated potential. Then $\bar{U}_\varepsilon = U_\varepsilon$ on I_ε^c and $U_\varepsilon \leq \bar{U}_\varepsilon \leq U_\mu$.

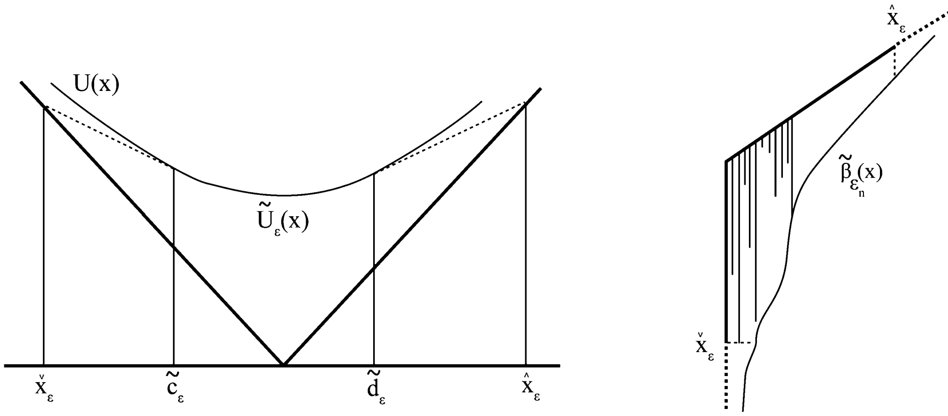

FIG. 5. *The potentials $\tilde{U}_\varepsilon$ increase monotonically as ε decreases. Moreover, over a range of x, depending on ε_n, we have $\tilde{\beta}_{\varepsilon_n}(x) \equiv \beta(x)$, and hence, the inverse barycentre functions converge monotonically.*

Now let $\tilde{U}_\varepsilon$ be the largest convex function such that $\tilde{U}_\varepsilon(x) = |x|$ on I_ε^c and $\tilde{U}_\varepsilon \le U_\mu$. It follows that $\tilde{U}_\varepsilon$ is actually equal to U on an interval $\tilde{I}_\varepsilon = [\tilde{c}_\varepsilon, \tilde{d}_\varepsilon]$. If ε is small enough, then $0 \in \tilde{I}_\varepsilon$. See Figure 5. Further, $U_\varepsilon \le \bar{U}_\varepsilon \le \tilde{U}_\varepsilon \le U$ and in terms of the associated measures $\mu_\varepsilon \le_{\mathrm{cx}} \bar{\mu}_\varepsilon \le_{\mathrm{cx}} \tilde{\mu}_\varepsilon \le_{\mathrm{cx}} \mu$, where $\tilde{\mu}_\varepsilon$ is such that $U_{\tilde{\mu}_\varepsilon} = \tilde{U}_\varepsilon$, and we recall that $\le_{\mathrm{cx}}$ denotes "less than or equal to in convex order." Then, by a theorem of Strassen [22] (or for a more explicit construction in our context, Chacon and Walsh [5]), given σ_ε there exists a stopping time $\tilde{\sigma}_\varepsilon$ such that $\sigma_\varepsilon \le \tilde{\sigma}_\varepsilon$ almost surely, and $\tilde{\mu}_\varepsilon = \mathcal{L}(W_{\tilde{\sigma}_\varepsilon})$.

Now consider a sequence ε_n decreasing to zero. Let $\tilde{\beta}_{\varepsilon_n}$ be the inverse barycentre associated with $\tilde{\mu}_{\varepsilon_n}$, and let $\tilde{\tau}_n$ be the Azéma–Yor stopping time associated with $\tilde{\beta}_{\varepsilon_n}$. The introduction of the stopping times $\tilde{\sigma}_{\varepsilon_n}$ gives extra structure which means that not only do the barycenters converge (as in Proposition 4.4), but also that they converge monotonically.

LEMMA 7.5. *$\tilde{\beta}_n \downarrow \beta$ and $\tilde{\tau}_n \uparrow \tau_\beta$ almost surely.*

PROOF. Write $\check{x}_n$ (resp., $\hat{x}_n, c_n, d_n$) for $\check{x}_{\varepsilon_n}$ (resp., $\hat{x}_{\varepsilon_n}, c_{\varepsilon_n}, d_{\varepsilon_n}$).

Then, for $s \le b(c_n)$, $\tilde{\beta}_n(s) = \check{x}_n \ge \beta(s)$, for $b(\tilde{c}_n) < s < \hat{x}_n$, $\tilde{\beta}_n(s) = \beta(s)$ and for $s \ge \hat{x}_n$, $\tilde{\beta}_n(s) = s \ge \beta(s)$.

Monotonicity in n of $\tilde{\tau}_n$ follows immediately. $\square$

It follows from the results for bounded target distributions that

$$\mathbb{E}\left[\int_0^{\sigma_n} g(S_u)\,du\right] \le \mathbb{E}\left[\int_0^{\tilde{\sigma}_n} g(S_u)\,du\right] = \mathbb{E}\left[G(W_{\tilde{\sigma}_n}, S_{\tilde{\sigma}_n})\right] \le \mathbb{E}\left[G(W_{\tilde{\tau}_n}, S_{\tilde{\tau}_n})\right]$$
$$= \mathbb{E}\left[\int_0^{\tilde{\tau}_n} g(S_u)\,du\right].$$

We have that the integral inside the first expectation converges monotonically to $\int_0^\tau g(S_u)\,du$, whereas the integral inside the final expression converges monotonically to $\int_0^{\tau_\beta} g(S_u)\,du$. Hence $\mathbb{E}[\int_0^\tau g(S_u)\,du] \le \mathbb{E}[\int_0^{\tau_\beta} g(S_u)\,du]$ as required.

PROOF OF THEOREM 7.1(i): UPPER BOUND. The final element of Theorem 7.1 is the upper bound in the case of monotonically increasing g. Recall that we suppose that g is continuously differentiable, and bounded by $\bar{g}$.

If μ has bounded support, then Theorem 5.11 applies directly, so we assume that the support of μ is unbounded.

If $\mu \notin L^2$, then for each $\tau \in \mathcal{S}(W,\mu)$ we have $\mathbb{E}[\tau] = \infty$ and using the fact that $\mathbb{E}[H_\varepsilon \wedge \tau_\mu^{\mathrm{P}}] \le \mathbb{E}[H_\varepsilon \wedge H_{\alpha^+(\varepsilon)}] < \infty$, we have that $\mathbb{E}[\int_0^{\tau_\mu^{\mathrm{P}}} g(S_u)\,du] \ge g(\varepsilon)\mathbb{E}[\int_{H_\varepsilon \wedge \tau_\mu^{\mathrm{P}}}^{\tau_\mu^{\mathrm{P}}} du] = \infty$, and there is nothing to prove.

So suppose $\mu \in L^2$. Then the area between the curves $U_\mu(x)$ and $|x|$ is finite.

Let $U_\varepsilon(x) = \max\{U_\mu(x) - \varepsilon, |x|\}$ and related quantities be defined as above.

This time, since $\tilde{U}_\varepsilon \equiv U_\mu$ on $\tilde{I}_\varepsilon$ we have that $\alpha_{\tilde{\mu}_\varepsilon} = \alpha_\mu$ on some sub-interval $\acute{I}_\varepsilon \subseteq \tilde{I}_\varepsilon$ of the form $\acute{I}_\varepsilon = [\acute{c}_\varepsilon, \acute{d}_\varepsilon]$, and as $\varepsilon \downarrow 0$, $\acute{I}_\varepsilon$ increases to the support of μ.

Now

$$\mathbb{E}\left[\int_0^\tau g(S_u)\,du\right] = \lim_{\varepsilon \downarrow 0} \mathbb{E}\left[\int_0^{\sigma_\varepsilon} g(S_u)\,du\right]$$

and

$$\mathbb{E}\left[\int_0^{\sigma_\varepsilon} g(S_u)\,du\right] \le \mathbb{E}\left[\int_0^{\tilde{\sigma}_\varepsilon} g(S_u)\,du\right] \le \mathbb{E}\left[\int_0^{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon)} g(S_u)\,du\right].$$

But

$$\begin{aligned}\mathbb{E}\left[\int_0^{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon)} g(S_u)\,du\right] &= \mathbb{E}\left[\int_0^{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) \wedge H_{\acute{c}_\varepsilon} \wedge H_{\acute{d}_\varepsilon}} g(S_u)\,du\right] \\ &\quad + \mathbb{E}\left[\int_{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) \wedge H_{\acute{c}_\varepsilon} \wedge H_{\acute{d}_\varepsilon}}^{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon)} g(S_u)\,du\right].\end{aligned}$$

Since $\alpha_{\tilde{\mu}_\varepsilon} = \alpha_\mu$ on $\acute{I}_\varepsilon$ and we have that $\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) \wedge H_{\acute{c}_\varepsilon} \wedge H_{\acute{d}_\varepsilon}$ is monotonically increasing as $\varepsilon \downarrow 0$ and hence the first term on the right-hand side converges to $\mathbb{E}[\int_0^{\tau^{\mathrm{P}}(\mu)} g(S_u)\,du]$. Meanwhile, the second term is bounded by $\bar{g}\mathbb{E}[\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) - \tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) \wedge H_{\acute{c}_\varepsilon} \wedge H_{\acute{d}_\varepsilon}]$. This last quantity is at most $\bar{g}$ multiplied by the area between the potentials U_μ and $U_{\acute{\mu}_\varepsilon}$ where $\acute{\mu}_\varepsilon = \mathcal{L}(W_{\tau^{\mathrm{P}}(\tilde{\mu}_\varepsilon) \wedge H_{\acute{c}} \wedge H_{\acute{d}}})$. However, as ε tends to zero this area tends to zero. Hence $\mathbb{E}[\int_0^\tau g(S_u)\,du] \le \mathbb{E}[\int_0^{\tau^{\mathrm{P}}(\mu)} g(S_u)\,du]$. □

8. An application and extensions.

8.1. *Variance swap on the sum of squared returns.* We now return to the question which originally motivated this paper which was to find model-independent bounds for variance swaps given the terminal law of the underlying asset price process or equivalently, call prices with expiry T for all strikes. Using the results developed in this article we will show how to bound the idealized variance swap based on squared returns, introduced in Section 2. The results in this article motivated further work on model-independent bounds and hedging strategies for variance swaps in a general setting; see Hobson and Klimmek [11].

As in Section 2, let $X = (X_t)_{0\le t\le T}$ be a square-integrable martingale started at $X_0 = x_0$ with $X_T \sim \mu$, where μ is centered at x_0 and supported on $\mathbb{R}_+$. Recall from (2.1) the definition for the payoff of an idealized variance swap $V_T = V((X_s)_{0\le s\le T}) = \int_0^T (X_{t-})^{-2}\, d[X, X]_t$. By (2.4) and (2.5) we have

$$\inf_{\tau\in S_{\mathrm{UI}}(B,\mu)} \mathbb{E}\left[\int_0^\tau \frac{du}{(S_u^B)^2}\right] \le \mathbb{E}[V_T] \le \sup_{\tau\in S_{\mathrm{UI}}(B,\mu)} \mathbb{E}\left[\int_0^\tau \frac{du}{(I_u^B)^2}\right].$$

Let $\tilde{\mu}$ be the measure μ reflected around 0, so that $\tilde{\mu}$ is a measure on $\mathbb{R}_-$, and observe that

$$\sup_{\tau\in S_{\mathrm{UI}}(B,\mu)} \mathbb{E}\left[\int_0^\tau \frac{du}{(I_u^B)^2}\right] = \sup_{\tau\in S_{\mathrm{UI}}(\tilde{B},\tilde{\mu})} \mathbb{E}\left[\int_0^\tau \frac{du}{(S_u^{\tilde{B}})^2}\right],$$

where $\tilde{B}$ is a Brownian motion started at $-x_0$, with maximum process $S^{\tilde{B}}$. Now we apply Theorem 7.1 to see that

$$\mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{P}}} \frac{du}{(S_u^B)^2}\right] \le \mathbb{E}[V_T] \le \mathbb{E}\left[\int_0^{\tau_{\tilde{\mu}}^{\mathrm{P}}} \frac{du}{(S_u^{\tilde{B}})^2}\right].$$

Note that the Perkins embedding for $\tau_{\tilde{\mu}}$ is determined by the monotonic functions $\alpha_{\tilde{\mu}}^{\pm}$ where $\alpha_{\tilde{\mu}}^{\pm}(x) = -\alpha_{\mu}^{\mp}(-x)$.

EXAMPLE 8.1. Suppose that $X_0 = 1$ and $\mu = U[0, 2]$. Shifting the quantities calculated in Example 3.1 to allow for the starting value $X_0 = 1$ it is clear that $\alpha_\mu^+ : [1, 2] \to [0, 1]$ is defined $\alpha_\mu^+(s) = s - 2\sqrt{s-1}$ and $\alpha_\mu^- : [0, 1] \to [1, 2]$ is defined $\alpha_\mu^-(i) = i + \sqrt{1-i}$. Hence the lower bound can be calculated,

$$\mathbb{E}\left[\int_0^{\tau_\mu^{\mathrm{P}}} \frac{du}{S_u^2}\right] = \mathbb{E}\left[\left(1 - \frac{B_{\tau_\mu^{\mathrm{P}}}}{S_{\tau_\mu^{\mathrm{P}}}}\right)^2\right] = \int_0^1 \left(1 - \frac{x}{a_\mu^+(x)}\right)^2 \frac{dx}{2} = \frac{\pi}{2} - 2\log 2.$$

For the upper bound, first considering $g_\varepsilon(s) = s^{-2} \wedge \varepsilon^{-2}$ and then letting $\varepsilon \downarrow 0$,

$$\mathbb{E}\left[\int_0^{\tau_{\tilde{\mu}}^{\mathrm{P}}} \frac{du}{\tilde{S}_u^2}\right] = \mathbb{E}\left[\left(1 - \frac{\tilde{B}_{\tau_{\tilde{\mu}}^{\mathrm{P}}}}{\tilde{S}_{\tau_{\tilde{\mu}}^{\mathrm{P}}}}\right)^2\right] = \int_0^1 \left(1 - \frac{x}{a_\mu^-(x)}\right)^2 \frac{dx}{2} = \infty.$$

8.2. *Extension to diffusions.* Suppose that $(X_t)_{t\geq 0}$ is a time-homogeneous diffusion on $I \subseteq \mathbb{R}$. More specifically, let $\sigma : I \to (0,\infty)$ and $b : I \to \mathbb{R}$ be Lipschitz functions, and define $(X_t)_{t\geq 0}$ to be the solution to

$$dX_t = \sigma(X_t)\,dB_t + b(X_t)\,dt, \qquad X_0 = x_0,$$

where $(B_t)_{t\geq 0}$ is a Brownian motion.

Let $s : I \to \mathbb{R}$ be the strictly increasing and C^2 scale function of X,

$$s(x_0) = 0, \qquad s'(x) = \exp\biggl(-\int_0^x 2\frac{b(u)}{\sigma(u)^2}\,du\biggr)$$

and let $h = s^{-1}$.

Consider the problem of maximizing (or minimizing) $\mathbb{E}[F(X_\tau, S_\tau^X)]$ over minimal embeddings τ of μ. Since $M_t = s(X_t)$ is a local martingale it follows that it can be represented as $M_t = W_{A(t)}$, for some (continuous) time-change $t \to A(t)$. Define the measure ν by $\nu(G) = \mu(s^{-1}(G))$ for Borel sets $G \subseteq s(I)$. Notice that σ is a minimal embedding of ν in W if and only if $\tau = A^{-1}(\sigma)$ is a minimal embedding of ν in M and hence a minimal embedding of μ in X.

Define the function $\hat{F}$ by $\hat{F}(w,s) = F(h(w), h(s))$. Then

$$F\bigl(X_\tau, S_\tau^X\bigr) = \mathbb{F}\bigl(h(W_{A_\tau}), h(S_{A_\tau})\bigr) = \hat{F}(W_{A_\tau}, S_{A_\tau}). \tag{8.1}$$

LEMMA 8.2. *Suppose F satisfies F-MON↑. Then $\hat{F}$ satisfies $\hat{F}$-MON↑ if $F_s < 0$ and h is concave or if $F_s > 0$ and h is convex.*

Similarly, suppose F satisfies F-MON↓. Then $\hat{F}$ satisfies $\hat{F}$-MON↓ if $F_s < 0$ and h is convex or if $F_s > 0$ and h is concave.

PROOF. The result follows from the expression

$$\frac{\hat{F}_s(x,s)}{s-x} = \frac{h'(s)F_s(h(x),h(s))}{h(s)-h(x)}\,\frac{h(s)-h(x)}{s-x}. \tag{8.2}$$ □

Note that h is convex (concave) when s is concave (convex), and since $2s''(x)/s'(x) = -\sigma(x)^2/b(x)$, the scale function is concave if $b(x) > 0$ for all x.

PROPOSITION 8.3. *Suppose $\nu = \mu \circ h$ is centered about zero, and suppose $b > 0$. Suppose F satisfies F-MON↑ and is increasing in s. Then*

$$\sup_{\tau \in S_{\mathrm{UI}}(X,\mu)} \mathbb{E}\bigl[F\bigl(X_\tau, S_\tau^X\bigr)\bigr] = \mathbb{E}\bigl[\hat{F}(W_{\tau_\nu^{\mathrm{AY}}}, S_{\tau_\nu^{\mathrm{AY}}})\bigr],$$

$$\inf_{\tau \in S_{\mathrm{UI}}(X,\mu)} \mathbb{E}\bigl[F\bigl(X_\tau, S_\tau^X\bigr)\bigr] = \mathbb{E}\bigl[\hat{F}(W_{\tau_\nu^{\mathrm{P}}}, S_{\tau_\nu^{\mathrm{P}}})\bigr].$$

REMARK 8.4. Whilst necessary to apply the results of the Brownian setting, the assumption that $\nu \equiv \mu \circ h$ is centered is not as innocuous as might first appear, and in the setting of a transient diffusion it is natural to wish to consider embeddings for target laws which, after transformation by the scale function, are not centered. For example, let X be a three-dimensional Bessel process, started at one. Then $s(x) = -1/x + 1$ and $h(m) = 1/(1-m)$. Now let μ be any probability measure on $\mathbb{R}^+$ with $\int_{\mathbb{R}^+} x^{-1}\mu(dx) \leq 1$. Then, there exists a minimal embedding of μ in X, but only if $\int_{\mathbb{R}^+} x^{-1}\mu(dx) = 1$ does this embedding correspond to a uniformly integrable embedding of $M \equiv 1 - X^{-1}$.

See Cox and Hobson [6] (and the references therein) for a further discussion of this issue, and of the construction of embeddings in Brownian motion of noncentered target laws.

Acknowledgments. Both authors thank a pair of referees for their detailed comments on an earlier version of this article.

REFERENCES

[1] AZÉMA, J. and YOR, M. (1979). Le problème de Skorokhod: Compléments à "Une solution simple au problème de Skorokhod." In *Séminaire de Probabilités, XIII* (*Univ. Strasbourg, Strasbourg*, 1977/78). *Lecture Notes in Math.* **721** 625–633. Springer, Berlin. MR0544832

[2] AZÉMA, J. and YOR, M. (1979). Une solution simple au problème de Skorokhod. In *Séminaire de Probabilités, XIII* (*Univ. Strasbourg, Strasbourg*, 1977/78). *Lecture Notes in Math.* **721** 90–115. Springer, Berlin. MR0544782

[3] BREEDEN, D. T. and LITZENBERGER, R. H. (1978). Prices of state-contingent claims implicit in option prices. *J. Bus.* **51** 621–651.

[4] CHACON, R. V. (1977). Potential processes. *Trans. Amer. Math. Soc.* **226** 39–58. MR0501374

[5] CHACON, R. V. and WALSH, J. B. (1976). One-dimensional potential embedding. In *Séminaire de Probabilités, X* (*Prèmiere Partie, Univ. Strasbourg, Strasbourg, Année Universitaire* 1974/1975). *Lecture Notes in Math.* **511** 19–23. Springer, Berlin. MR0445598

[6] COX, A. M. G. and HOBSON, D. G. (2006). Skorokhod embeddings, minimality and noncentred target distributions. *Probab. Theory Related Fields* **135** 395–414. MR2240692

[7] DEMETERFI, K., DERMAN, E., KAMAL, M. and ZOU, J. (1999). A guide to volatility and variance swaps. *The Journal of Derivatives* **6** 9–32.

[8] DURRETT, R. (1991). *Probability: Theory and Examples*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA. MR1068527

[9] HIRSCH, F., PROFETA, C., ROYNETTE, B. and YOR, M. (2011). Constructing self-similar martingales via two Skorokhod embeddings. In *Séminaire de Probabilités XLIII. Lecture Notes in Math.* **2006** 451–503. Springer, Berlin. MR2790387

[10] HOBSON, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance* 2010. *Lecture Notes in Math.* **2003** 267–318. Springer, Berlin. MR2762363

[11] HOBSON, D. G. and KLIMMEK, M. (2012). Model-independent hedging strategies for variance swaps. *Finance Stoch.* DOI:10.1007/s00780-012-0190-3.

[12] HOBSON, D. G. and PEDERSEN, J. L. (2002). The minimum maximum of a continuous martingale with given initial and terminal laws. *Ann. Probab.* **30** 978–999. MR1906424

[13] KERTZ, R. P. and RÖSLER, U. (1990). Martingales with given maxima and terminal distributions. *Israel J. Math.* **69** 173–192. MR1045372

[14] MONROE, I. (1972). On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.* **43** 1293–1311. MR0343354

[15] MONROE, I. (1978). Processes that can be embedded in Brownian motion. *Ann. Probab.* **6** 42–56. MR0455113

[16] OBŁÓJ, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1** 321–390. MR2068476

[17] 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). *Progr. Probab. Statist.* **12** 172–223. Birkhäuser, Boston, MA. MR0896743

[18] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357

[19] ROGERS, L. C. G. (1989). A guided tour through excursions. *Bull. Lond. Math. Soc.* **21** 305–341. MR0998631

[20] ROGERS, L. C. G. (1993). The joint law of the maximum and terminal value of a martingale. *Probab. Theory Related Fields* **95** 451–466. MR1217446

[21] SKOROKHOD, A. V. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, MA. MR0185620

[22] STRASSEN, V. (1967). Almost sure behavior of sums of independent random variables and martingales. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (*Berkeley*, *Calif.*, 1965/66) Vol. II: Contributions to Probability Theory, Part 1, pp. 315–343. Univ. California Press, Berkeley, CA. MR0214118

[23] VALLOIS, P. (1994). Sur la loi du maximum et du temps local d'une martingale continue uniformement intégrable. *Proc. Lond. Math. Soc.* (3) **69** 399–427. MR1281971

STATISTICS DEPARTMENT
UNIVERSITY OF WARWICK
CV47AL, COVENTRY
UNITED KINGDOM
E-MAIL: d.hobson@warwick.ac.uk
m.klimmek@warwick.ac.uk