Root's Barrier: Construction, Optimality and Applications to Variance Options

Alexander M. G. Cox* and Jiajie Wang†
Department of Mathematical Sciences,
University of Bath, Bath, U. K.

April 18, 2011

Abstract

Recent work of Dupire (2005) and Carr and Lee (2010) has highlighted the importance of understanding the Skorokhod embedding originally proposed by Root (1969) for the model-independent hedging of variance options. Root's work shows that there exists a *barrier* from which one may define a stopping time which solves the Skorokhod embedding problem. This construction has the remarkable property, proved by Rost (1976), that it minimises the variance of the stopping time among all solutions.

In this work, we prove a characterisation of Root's barrier in terms of the solution to a variational inequality, and we give an alternative proof of the optimality property which has an important consequence for the construction of subhedging strategies in the financial context.

1 Introduction

In this paper, we analyse the solution to the Skorokhod embedding problem originally given by Root (1969), and generalised by Rost (1976). Our motivation for this is recent work connecting the solution to this problem to questions arising in mathematical finance — specifically model-independent bounds for variance options — which has been observed by Dupire (2005), Carr and Lee (2010) and Hobson (2010). The financial motivation can be described as follows: consider a (discounted) asset which has dynamics under the risk-neutral measure

$$\frac{dS_t}{S_t} = \sigma_t \,\mathrm{d}W_t,$$

where the process σ_t is not necessarily known. We are interested in variance options, which are contracts where the payoff depends on the realised quadratic variation of the log-price process: specifically, we have

$$\mathrm{d}(\ln S_t) = \sigma_t \,\mathrm{d}W_t - \frac{1}{2}\sigma_t^2 \,\mathrm{d}t$$

*e-mail: `a.m.g.cox@bath.ac.uk`, web: `http://www.maths.bath.ac.uk/∼mapamgc/`
†e-mail: `j.wang2@bath.ac.uk`

and therefore

$$\langle \ln S \rangle_T = \int_0^T \sigma_t^2 \, \mathrm{d}t.$$

An option on variance is then an option with payoff $F(\langle \ln S \rangle_T)$. Important examples include variance swaps, which pay the holder $\langle \ln S \rangle_T - K$, and variance calls which pay the holder $(\langle \ln S \rangle_T - K)_+$. We shall be particularly interested in the case of a variance call, but our results will extend to a wider class of payoffs. Let $\mathrm{d}X_t = X_t \, \mathrm{d}\widetilde{W}_t$ for a suitable Brownian motion $\widetilde{W}_t$ and we can find a (continuous) time change τ_t such that $S_t = \widetilde{X}_{\tau_t}$, and so:

$$\mathrm{d}\tau_t = \frac{\sigma_t^2 S_t^2}{S_t^2} \, \mathrm{d}t.$$

Hence

$$\left(\widetilde{X}_{\tau_T}, \tau_T\right) = \left(S_T, \int_0^T \sigma_u^2 \, \mathrm{d}u\right) = \left(S_T, \langle \ln S \rangle_T\right).$$

Now suppose that we know the prices of call options on S_T with maturity T, and at all strikes (recall that σ_t is not assumed known). Then we can derive the law of S_T under the risk-neutral measure from the Breeden-Litzenberger formula. Call this law μ. This suggests that the problem of finding a lower bound on the price of a variance call (for an unknown σ_t) is equivalent to:

$$\text{find a stopping time } \tau \text{ to minimise } \mathbb{E}(\tau - K)_+, \text{ subject to } \mathcal{L}(\widetilde{X}_\tau) = \mu. \tag{1.1}$$

This is essentially the problem for which Rost has shown that the solution is given by Root's barrier. (In fact, the result trivially extends to payoffs of the form $F(\langle \ln S \rangle_T)$ where $F(\cdot)$ is a convex, increasing function.)

In this work, our aim is twofold: firstly, to provide a proof that Root's barrier can be found as the solution to a particular variational inequality, which can be thought of as the generalisation of an obstacle problem; secondly, we show that the lower bound which is implied by Rost's result can be enforced through a suitable hedging strategy, which will give an arbitrage whenever the price of a variance call trades below the given lower bound. To accomplish this second part of the paper, we will give a novel proof of the optimality of Root's construction, and from this construction we will be able to derive a suitable hedging strategy.

The use of Skorokhod embedding techniques to solve model-independent (or robust) hedging problems in finance can be traced back to the papers of Brown et al. (2001a). More recent results in this direction include Cox et al. (2008), Cox and Obłój (2011a) and Cox and Obłój (2011b). For a comprehensive survey of the literature on the Skorokhod embedding problem, we refer the reader to Obłój (2004). In addition, Hobson (2010) surveys the literature on the Skorokhod embedding problem with a specific emphasis on the applications in mathematical finance.

Variance options have been a topic of much interest in recent years, both from the industrial point of view, where innovations such as the VIX index have contributed to a large growth in products which are directly dependent on quantities derived from the quadratic variation, and also on the academic side, with a number of interesting contributions in the literature. The academic results go back to work of Dupire (1993) and Neuberger (1994), who noted that a variance swap — that is a contract which pays $\langle \ln S \rangle_T$, can be replicated model-independently using a contract paying the logarithm of the asset at maturity through the identity (from Itô's Lemma):

$$\ln(S_T) - \ln(S_0) = \int_0^T \frac{1}{S_t}\,\mathrm{d}S_t - \frac{1}{2}\langle \ln S \rangle_T . \tag{1.2}$$

More recently, work on options and swaps on volatility and variance, (in a model-based setting) includes Howison et al. (2004), Broadie and Jain (2008) and Kallsen et al. (2010). Other work (Keller-Ressel and Muhle-Karbe, 2010, Keller-Ressel, 2011) has considered the differences between the theoretical payoff ($\langle \ln S \rangle_T$) and the discrete approximation which is usually specified in the contract ($\sum_k \ln(S_{(k+1)\delta}/S_{k\delta})^2$). Finally, several papers have considered variants on the model-independent problems (Carr and Lee, 2010, Carr et al., 2011, Davis et al., 2010) or problems where the modelling assumptions are fairly weak. This latter framework is of particular interest for options on variance, since the markets for such products are still fairly young, and so making strong modelling assumptions might not be as strongly justified as it could be in a well-established market.

The rest of this paper is structured as follows: in Section 2 we review some known results and properties concerning Root's barrier. In Section 3, we establish a connection between Root's solution and an obstacle problem, and then in Section 4 we show that by considering an obstacle problem in a more general analytic sense (as a variational inequality), we are able to prove the equivalence between Root's problem and the solution to a variational inequality. In Section 5, we give a new proof of the optimality of Root's solution, and in Section 6 we show how this proof allows us to construct model-independent subhedges to give bounds on the price of variance options.

2 Features of Root's solution

Our interest is in Root's solution to the Skorokhod embedding problem. Simply stated, for a process $(X_t)_{t\geq 0}$, the Skorokhod embedding problem is to find a stopping time τ such that $X_\tau \sim \mu$. In this paper, we will consider firstly the case where $X_0 = 0$, and X_t is a continuous martingale and a time-homogeneous diffusion, and later the case where $X_0 \sim \nu$, is a centred, square integrable measure. In such circumstances, it is natural to restrict to the set of stopping times for which $(X_{t\wedge\tau})_{t\geq 0}$ is a uniformly integrable (UI) process. We will occasionally call stopping times for which this is true *UI stopping times*. In the case where μ is centered and has a second moment, this can be shown to be equivalent to the fact that $\mathbb{E}\tau < \infty$. For the case of a general starting measure, there is a natural restriction on the measures involved, which is that we require:

$$\mathrm{U}_\nu(x) := -\int_{\mathbb{R}} |y-x|\,\nu(\mathrm{d}y) \geq -\int_{\mathbb{R}} |y-x|\,\mu(\mathrm{d}y) =: \mathrm{U}_\mu(x), \tag{2.1}$$

for all $x \in \mathbb{R}$. By Jensen's inequality, such a constraint is clearly necessary for the existence of a suitable pair ν and μ; further, by Rost (1971), it is the only additional constraint on the measures we will need to impose. We shall write

$$S(\mu) = \{\tau : \tau \text{ is a stopping time}, X_\tau \sim \mu, (X_{t\wedge\tau})_{t\geq 0} \text{ is UI}\}. \tag{2.2}$$

There are a number of important papers concerning the construction of Root's barrier. The first work to consider the problem is Root (1969), and this paper proved the existence of a certain Skorokhod embedding when X_t is a Brownian motion. Specifically, Root showed that if X_t is a Brownian motion with $X_0 = 0$, and μ is the law of a centered random variable with finite variance, then there exists a stopping time τ, which is the first hitting time of a *barrier*, which is defined as follows:

Definition 2.1 (Root's Barrier)**.** A closed subset B of $[-\infty, +\infty] \times [0, +\infty]$ is a *barrier* if

1. $(x, +\infty) \in B$ for all $x \in [-\infty, +\infty]$;
2. $(\pm\infty, t) \in B$ for all $t \in [0, \infty]$;
3. if $(x, t) \in B$ then $(x, s) \in B$ whenever $s > t$.

In a subsequent paper Loynes (1970) proved a number of results relating to barriers. From our perspective, the most important are, firstly, that the barrier B can be written as: $B = \{(x, t) : t \geq R(x)\}$, where $R : \mathbb{R} \to [0, \infty]$ is a lower semi-continuous function (with the obvious extensions to the definition to cover $R(x) = \infty$); we will make frequent use of this representation. In addition, Loynes (1970, Theorem 1) says that Root's solution is essentially unique: if there are two barriers which embed the same distribution with a UI stopping time, then their corresponding stopping times are equal with probability one. The case where two different barriers can occur are then only the cases where, say $R(x_0) = 0$ for $x_0 > 0$, and then $R(x)$ is undetermined for all $x > x_0$.

The other important reference for our purposes is Rost (1976). This work vastly extends the generality of the results of Root and Loynes, and uses mostly potential-theoretic techniques. Rost works in the generality of a Markov process X_t on a compact metric space E, which satisfies the strong Markov property and is right-continuous. Then Rost recalls (from an original definition of Dinges (1974) in the discrete setting) the notion of *minimal residual expectation*:

Definition 2.2. A stopping time $\tau^* \in S(\mu)$ is of *minimal residual expectation* if, for each $t \in \mathbb{R}_+$, it minimises the quantity:

$$\mathbb{E}(\tau - t)_+ = \mathbb{E} \int_{\tau \wedge t}^{\tau} \mathrm{d}s = \int_t^{\infty} \mathbb{P}(\tau > s) \, \mathrm{d}s,$$

over all $\tau \in S(\mu)$.

Then Rost proves that (under (2.1)) there exists a stopping time of minimal residual expectation (Rost, 1976, Theorem 1), and that the hitting time of any barrier is of minimal residual expectation (Rost, 1976, Theorem 2). Finally, Rost also shows that the barrier stopping times are, to a degree, unique (Rost, 1976, Corollary to Theorem 2). The relevant result for our purposes (where there is a stronger form of uniqueness) is the corollary to Theorem 3 therein, which says that if X_t is a process for which the one-point sets are regular, then any stopping time of minimal residual expectation is Root's stopping time. The class of processes for which the one-point sets are regular include the class of time-homogenous diffusions we consider.

Note that a stopping time is of minimal residual expectation if and only if, for every convex, increasing function $F(t)$ (where, without loss of generality, we take $F(0) = F'_+(0) = 0$), it minimises the quantity:

$$\mathbb{E}F(\tau) = \mathbb{E}\int_0^\infty (\tau - t)_+ F''(dt),$$

this fact being a consequence of the above representation.

There are a number of important properties that the Root barrier possesses. Firstly, we note that, as a consequence of the fact that B is closed and the third property of Definition 2.1, the barrier is regular (that is, if we start at a point in the barrier, we will almost surely return to the barrier instantly) for the class of processes we will consider (time-homogeneous diffusions). This will have important analytical benefits. Secondly, for a point $(x, t) \notin B$, we know that if the stopped process at time t is at x, then we have not yet reached the stopping time for the embedding. This will help in our characterisation of the law of the stopped process (Lemma 3.2).

In the rest of this paper, we will then say that a barrier is either a lower semi-continuous function $R : \mathbb{R} \to \overline{\mathbb{R}}_+$, with $R(0) \neq 0$, or the complement of the corresponding connected open set $D = \{(x, t) : 0 < t < R(x)\} = \mathbb{R} \times (0, \infty) \setminus B$. As noted above, by Loynes (1970) this is equivalent to the barrier as defined in Definition 2.1. We will define the hitting time of the barrier as: $\tau_D = \inf\{t > 0 : (X_t, t) \notin D\}$. Note that the barrier B is closed and regular, so that $(X_{\tau_D}, \tau_D) \in B$ and $\mathbb{P}^{(x,t)}(\tau_D = 0) = 1$ whenever $(x, t) \in B$, where $\mathbb{P}^{(x,t)}$ is the law of our diffusion started at x at time t.

Finally, we give some examples where the barrier function can be explicitly calculated. We note that explicit examples appear to be the exception, and in general are hard to compute. Firstly, if μ is a Normal distribution, we easily see that $R(x)$ is a constant. Secondly, if μ consists of two atoms (weighted appropriately) at $a < 0 < b$ say, the corresponding barrier is

$$R(x) = \begin{cases} 0 & x \notin (a, b) \\ \infty & x \in (a, b) \end{cases}.$$

In this example, observe that the function $R(x)$ is not unique: we can choose any behaviour outside $[a, b]$, and achieve the same stopping time. Secondly, we note the that there are even more general solution to the Skorokhod embedding problem (without the uniform integrability condition) since there are also barriers of the form

$$R(x) = \begin{cases} t_a & x = a \\ t_b & x = b \\ \infty & x \notin \{a, b\} \end{cases},$$

which will embed the same law (provided $t_a, t_b > 0$ are chosen suitably), but which do not satisfy the uniform integrability condition. In general, a barrier can exhibit some fairly nasty features: consider for example the canonical measure on a middle third Cantor set C (scaled so that it is on $[-1, 1]$). Root's result tells us that there exists a barrier which embeds this distribution, and clearly the resulting barrier function must be finite only on the Cantor set, however the target distribution has no atoms, so that the 'spikes' in the barrier function can not be isolated (i.e. we must have $\liminf_{y\uparrow x} R(y) = \liminf_{y\downarrow x} R(y) = R(x)$ for all $x \in (-1, 1) \cap C$).

3 Connecting Root's Problem and an Obstacle Problem

We now consider alternative methods for describing Root's barrier. We will in general be interested in this question when our underlying process X_t is a solution to

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}W_t, \quad X_0 \sim \nu, \tag{3.1}$$

for a Brownian motion $(W_t)_{t\geq 0}$, and we will introduce our concepts in this general context. Initially, we assume that $\sigma : \mathbb{R} \to \mathbb{R}$ satisfies, for some positive constant K,

$$|\sigma(x) - \sigma(y)| \leq K|x - y|; \tag{3.2}$$

$$0 < \sigma^2(x) < K(1 + x^2); \tag{3.3}$$

$$\sigma \text{ is smooth.} \tag{3.4}$$

Recall that for the financial application we are interested in, we want the specific case $\sigma(x) = x$ to be included. Clearly, this case is currently excluded, however we will show in Section 4.3 that the results can be extended to include this case.

From standard results on SDEs, (3.2) and (3.3) imply that the unique strong solution X^a of (3.1) with $\nu = \delta_a$ is a strong Markov process with generator $\frac{1}{2}\sigma^2\partial_{xx}$ for any initial value $a \in \mathbb{R}$. Moreover, (3.4) implies that the operator $L := \frac{1}{2}\sigma^2\partial_{xx} - \partial_t$ is hypoelliptic (see Stroock (2008, Theorem 3.4.1)).

We will write Root's Skorokhod embedding problem as:

SEP(σ, ν, μ)**:** Find a lower-semicontinuous function $R(x)$ such that the domain $D = \{(x,t) : 0 < t < R(x)\}$ has $X_{\tau_D} \sim \mu$, and $(X_{t\wedge\tau_D})_{t\geq 0}$ is a UI process, where ν is the initial law of X_t, and σ the diffusion coefficient.

Our aim is to show that the problem of finding R is essentially equivalent to solving an obstacle problem. Assuming that the relevant derivatives exist, we shall show that the problem can be stated as:

OBS(σ, ν, μ)**:** Find a function $u(x,t) \in C^{1,1}(\mathbb{R} \times \mathbb{R}_+)$ such that

$$\mathrm{U}_\nu(x) = u(x,0) \tag{3.5a}$$

$$0 \geq \mathrm{U}_\mu(x) - u(x,t) \tag{3.5b}$$

$$0 \geq \frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\sigma(x)^2\frac{\partial^2 u}{\partial x^2}(x,t) \tag{3.5c}$$

$$0 > (\mathrm{U}_\mu(x) - u(x,t)) \implies 0 = \left(\frac{\partial u}{\partial t}(x,t) - \frac{1}{2}\sigma(x)^2\frac{\partial^2 u}{\partial x^2}(x,t)\right) \tag{3.5d}$$

where (3.5c) is interpreted in a distributional sense — that is, we require:

$$\int_{\mathbb{R}} \left(\phi(x)\frac{\partial u}{\partial t}(x,t) + \frac{1}{2}\sigma(x)^2\frac{\partial u}{\partial x}(x,t)\phi'(x)\right)\,\mathrm{d}x \leq 0$$

whenever $\phi \in C_K^\infty$ is a non-negative function. For (3.5d), we note that $(\mathrm{U}_\mu(x) - u(x,t)) = 0$ on a closed subset of $\mathbb{R} \times \mathbb{R}_+$, and hence $\frac{\partial^2 u}{\partial x^2}(x,t)$ would be continuous even if we were only to require (3.5d) to hold in a distributional sense.

In general, we do not expect u to be sufficiently nice that we can easily interpret all these statements, and one of the goals of this paper is to give a generalisation of $\mathbf{OBS}(\sigma, \nu, \mu)$ that will make sense more widely. Cases in which u may not be expected to be $C^{1,1}$ include the case where μ contains atoms (and therefore U_μ is not continuously differentiable). In addition, we specify this problem in $C^{1,1}$ since in general we would certainly not expect the second derivative to be continuous on the boundary between the two types of behaviour in (3.5d).

Theorem 3.1. *Suppose D is a solution to $\mathbf{SEP}(\sigma, \nu, \mu)$ and is such that:*

$$u(x,t) = -\mathbb{E}|X_{t\wedge\tau_D} - x| \in C^{1,1}(\mathbb{R} \times \mathbb{R}_+).$$

Then u solves $\mathbf{OBS}(\sigma, \nu, \mu)$.

This gives an initial connection between $\mathbf{OBS}(\sigma, \nu, \mu)$ and $\mathbf{SEP}(\sigma, \nu, \mu)$. We roughly expect solutions to Root's problem to be the unique solutions to the obstacle problem (of course, we do not currently know that such solutions exist or, when they do, are unique). This suggests that we can attempt to solve the obstacle problem to find the solution D to Root's problem. In particular, given a solution to $\mathbf{OBS}(\sigma, \nu, \mu)$, we can now identify D as $D = \{(x,t) : \mathrm{U}_\mu(x) < u(x,t), t > 0\}$.

Lemma 3.2. *For any $(x,t) \in D$, $\mathbb{P}\left(X_{t\wedge\tau_D} \in \mathrm{d}x\right) = \mathbb{P}\left(X_t \in \mathrm{d}x, t < \tau_D\right)$*

Proof. By the lower semi-continuity of R, since $(x,t) \in D$, there exists $h > 0$ such that

$$(x-h, x+h) \times [0, t+h) \subset D,$$

and hence, for any $y \in (x-h, x+h)$, $R(y) > t$. On the other hand, if $\tau_D \leq t$, we have

$$R(X_{\tau_D}) \leq \tau_D \leq t,$$

and hence, $X_{\tau_D} \notin (x-h, x+h)$. Therefore,

$$\begin{aligned}\mathbb{P}\left(X_{t\wedge\tau_D} \in \mathrm{d}x\right) &= \mathbb{P}\left(X_t \in \mathrm{d}x, t < \tau_D\right) + \mathbb{P}\left(X_{\tau_D} \in \mathrm{d}x, t \geq \tau_D\right)\\ &= \mathbb{P}\left(X_t \in \mathrm{d}x, t < \tau_D\right).\end{aligned}$$

□

Lemma 3.3. *The measure corresponding to $\mathcal{L}(X_t; t < \tau_D)$ has density $p^D(x,t)$ with respect to Lebesgue on D, and the density is smooth and satisfies:*

$$\frac{\partial}{\partial t} p^D(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma(x)^2 p^D(x,t)\right].$$

This result appears to be standard, but we are unable to find concise references. We give a short proof based on (Rogers and Williams, 2000, V.38.5).

Proof. First note that, as a measure, $\mathcal{L}(X_t; t < \tau_D)$ is dominated by the usual transition measure, so the density $p^D(x,t)$ exists.

Let (x_0, t_0) be a point in D, and we can therefore find an $\varepsilon > 0$ such that $A = (x_0 - \varepsilon, x_0 + \varepsilon) \times (t_0 - \varepsilon, t_0 + \varepsilon)$ satisfies $\bar{A} \subseteq D$. Then let f be a smooth function, supported on A, and by Itô's Lemma:

$$\begin{aligned} f(X_{t\wedge\tau_D}, t) =& f(X_0, 0) + \int_0^t \frac{\partial f}{\partial x}(X_{s\wedge\tau_D}, s)\,\mathrm{d}X_s \\ &+ \int_0^t \left(\frac{1}{2}\sigma(X_{s\wedge\tau_D})^2 \frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial t}\right) f(X_{s\wedge\tau_D}, s)\,\mathrm{d}s. \end{aligned}$$

Since f is compactly supported, taking $t > t_0 + \varepsilon$, the two terms on the left disappear, and the first integral term is a martingale. Hence, taking expectations, and interchanging the order of differentiation, we get:

$$\int_0^t \int p^D(y, s) \left(\frac{1}{2}\sigma(y)^2 \frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial t}\right) f(y, s)\,\mathrm{d}y\,\mathrm{d}s = 0.$$

Interpreting $p^D(y, s)$ as a distribution, we have

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma(x)^2 p^D(x, t)\right] - \frac{\partial}{\partial t} p^D(x, t) = 0,$$

for $(x, t) \in A$, and since the heat operator is hypoelliptic, we conclude that $p^D(x, t)$ is smooth in A (e.g. Stroock (2008, Theorem 3.4.1)). $\square$

We are now able to prove that any solution to Root's embedding problem is a solution to the obstacle problem.

Proof of Theorem 3.1. We first observe that $u(x, 0) = -\mathbb{E}|X_0 - x|$, and $X_0 \sim \nu$, so that $u(x, 0) = -\int |y - x|\,\nu(\mathrm{d}y)$ and (3.5a) holds. Secondly, since $(X_{t\wedge\tau_D})_{t\geq 0}$ is a UI process, by (conditional) Jensen's Inequality:

$$u(x, t) = -\mathbb{E}|x - X_{t\wedge\tau_D}| \geq -\mathbb{E}\left[\mathbb{E}\left[|x - X_{\tau_D}|\,\middle|\mathcal{F}_{t\wedge\tau_D}\right]\right] = \mathrm{U}_\mu(x),$$

and (3.5b) holds.

We now consider (3.5c). Suppose $(x, t) \in D$, and note that:

$$\frac{\partial u}{\partial x} = 1 - 2\mathbb{P}(X_{t\wedge\tau_D} < x) \tag{3.6}$$

and therefore (in D) by Lemma 3.3 the function u has a smooth second derivative in x. Further, we get:

$$\begin{aligned} \frac{1}{2}\int_0^t \sigma(x)^2 \frac{\partial^2 u}{\partial x^2}(x, s)\,\mathrm{d}s &= -\int_0^t \sigma(x)^2 p_D(x, s)\,\mathrm{d}s \\ &= \lim_{\varepsilon\downarrow 0} \mathbb{E}\left[\frac{1}{2\varepsilon}\int_0^{t\wedge\tau_D} \sigma(x)^2 \mathbf{1}_{[x-\varepsilon<X_s<x+\varepsilon]}\,\mathrm{d}s\right] \\ &= -\mathbb{E}L^x_{t\wedge\tau_D} \\ &= -\mathbb{E}|x - X_{t\wedge\tau_D}| + |x|, \end{aligned} \tag{3.7}$$

where L^x_t is the local time of the diffusion at x. It follows that u satisfies (3.5c) on D, and in fact attains equality there. On the other hand, if $(x, t) \notin D$, it follows from the definition of the barrier that if $\tau_D > t$, the diffusion cannot cross the line $\{(x,s) : s \geq t\}$ in the time interval $[t, \tau_D)$, and hence

$$X_{t\wedge\tau_D} > x \text{ and } t < \tau_D \implies X_{\tau_D} \geq x.$$

In addition, if $t \geq \tau_D$, then clearly $X_{t\wedge\tau_D} \geq x$ implies $X_{\tau_D} \geq x$. Finally, $X_{t\wedge\tau_D} = x$ implies $t \geq \tau_D$. Putting all this together, we observe that $X_{t\wedge\tau_D} \geq x \implies X_{\tau_D} \geq x$, and similarly $X_{t\wedge\tau_D} \leq x \implies X_{\tau_D} \leq x$.

In particular, using this fact and uniform integrability, we compute:

$$\begin{aligned}
&\mathbb{E}\left[|x - X_{t\wedge\tau_D}|\,\mathbf{1}_{\tau_D>t}\right] \\
&\quad= \mathbb{E}\left[(x - X_{t\wedge\tau_D})\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\leq x}\right] + \mathbb{E}\left[(X_{t\wedge\tau_D} - x)\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\geq x}\right] \\
&\quad= \mathbb{E}\left[(x - X_{\tau_D})\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\leq x}\right] + \mathbb{E}\left[(X_{\tau_D} - x)\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\geq x}\right] \\
&\quad= \mathbb{E}\left[|x - X_{\tau_D}|\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\leq x}\right] + \mathbb{E}\left[|X_{\tau_D} - x|\,\mathbf{1}_{\tau_D>t, X_{t\wedge\tau_D}\geq x}\right] \\
&\quad= \mathbb{E}\left[|x - X_{\tau_D}|\mathbf{1}_{\tau_D>t}\right].
\end{aligned}$$

and hence,

$$\begin{aligned}
-\mathbb{E}\,|x - X_{t\wedge\tau_D}| &= -\mathbb{E}\left[|x - X_{t\wedge\tau_D}|\,\mathbf{1}_{\tau_D>t}\right] - \mathbb{E}\left[|x - X_{t\wedge\tau_D}|\,\mathbf{1}_{\tau_D\leq t}\right] \\
&= -\mathbb{E}\left[|x - X_{\tau_D}|\,\mathbf{1}_{\tau_D>t}\right] - \mathbb{E}\left[|x - X_{\tau_D}|\,\mathbf{1}_{\tau_D\leq t}\right] \\
&= -\mathbb{E}\,|x - X_{\tau_D}| = \mathrm{U}_\mu(x),
\end{aligned}$$

so (3.5b) holds with equality when $(x,t) \notin D$. In particular, we can deduce that either (if $(x,t) \in D$) we have equality in (3.5c), or we have equality in (3.5b), in which case (3.5d) must hold. It remains to show that (3.5c) holds when $(x,t) \notin D$. However, to see this, consider $(x,t) \notin D$, and note first that $u(x,s) = u(x,t) = \mathrm{U}_\mu(x)$ whenever $s > t$, since $(x,s) \notin D$. Hence $\frac{\partial u}{\partial t}(x,t) = 0$. It is straightforward to check that $u(x,t)$ is concave in x, and therefore that $\frac{\partial^2 u}{\partial x^2}(x,t) \leq 0$, and (3.5c) also holds. □

This result connects Root's problem and the obstacle problem under a smoothness assumption on the function u. However, ideally we want a one-to-one correspondence. We know from the results of Rost (1976) that there always exists a solution to $\mathbf{SEP}(\sigma,\nu,\mu)$, and from Loynes (1970) that the solution is unique. Our aim is to show that a similar combination of existence and uniqueness hold for the corresponding analytic formulation. As already noted, we cannot make a strong smoothness assumption on the function $u(x,t)$ as required by $\mathbf{OBS}(\sigma,\nu,\mu)$ and so we need a weaker formulation of this problem. Generalisations of the obstacle problem are well understood, and commonly called variational inequalities. In the next section, we will reformulate the obstacle problem as a variational inequality, and we are able to state a problem for which existence and uniqueness are known due to existing results.

4 Root's Barrier and Variational Inequalities

We now study the relation between Root's Skorokhod embedding problem and a variational inequality. Our notation and definitions, and some of the key results which we will use, come from Bensoussan and Lions (1982).

4.1 Variational Inequalities

We begin with some necessary notation and results concerning evolutionary variational inequalities. Given a constant $\lambda > 0$ and a finite time $T > 0$, we define the Banach spaces $H^{m,\lambda} \subseteq L^2(\mathbb{R})$ and $L^2(0,T;H^{m,\lambda})$ with the norms:

$$\|g\|^2_{H^{m,\lambda}} = \sum_{k=0}^{m} \int_{\mathbb{R}} e^{-2\lambda|x|} \left| \frac{\partial^k g}{\partial x^k}(x) \right|^2 \mathrm{d}x;$$

$$\|w\|^2_{L^2(0,T;H^{m,\lambda})} = \int_0^T \|w(\cdot,t)\|^2_{H^{m,\lambda}} \,\mathrm{d}t,$$

where the derivatives $\frac{\partial^k g}{\partial x^k}(x)$ are to be interpreted as weak derivatives — that is, $\frac{\partial^k g}{\partial x^k}(x)$ is defined by the requirement that

$$\int_{\mathbb{R}} \phi(x) \frac{\partial^k g}{\partial x^k}(x) \,\mathrm{d}x = (-1)^k \int_{\mathbb{R}} g(x) \frac{\partial^k \phi}{\partial x^k}(x) \,\mathrm{d}x,$$

for all $\phi \in C^\infty_K(\mathbb{R})$, and C^∞_K is the set of compactly supported, smooth functions on $\mathbb{R}$. In particular, the spaces $H^{m,\lambda}$ and $L^2(0,T;H^{m,\lambda})$ are Hilbert spaces with respect to the obvious inner products. In addition, elements of the set $H^{1,\lambda}$ can always be taken to be continuous, and C^∞_K is dense in $H^{m,\lambda}$ (see e.g. Friedman (1963, Theorem 5.5.20)).

For functions $a(x,t),\ b(x,t),\ c(x,t) \in L^\infty(\mathbb{R} \times (0,T))$, we define an operator:

$$a_\lambda(t;v,w) = \int_{\mathbb{R}} e^{-2\lambda|x|} \left[a(x,t) \frac{\partial v}{\partial x} \frac{\partial w}{\partial x} + b(x,t) \frac{\partial v}{\partial x} w + c(x,t) v w \right] \mathrm{d}x,$$

for $v, w \in L^2(0,T;H^{1,\lambda})$. Moreover if $\partial a/\partial x$ exits, we define, for $v \in H^{2,\lambda}$,

$$A(t)v = -\frac{\partial}{\partial x}\left(a(x,t)\frac{\partial v}{\partial x} \right) + \left(b(x,t) + 2\lambda a(x,t)\,\mathrm{sgn}(x)\right) \frac{\partial v}{\partial x} + c(x,t)v.$$

And finally, for $v, w \in H^{0,\lambda}$,

$$(v,w)_\lambda = \int_{\mathbb{R}} e^{-2\lambda|x|} v w \,\mathrm{d}x,$$

so that, for suitably differentiable test functions $\phi(x)$ and $v \in H^{2,\lambda}$:

$$(\phi, A(t)v)_\lambda = a_\lambda(t;v,\phi).$$

Then we have the following restatement of Bensoussan and Lions (1982, Theorem 2.2, and Section 2.15, Chapter 3):

Theorem 4.1. *For any given $\lambda > 0$ and $T > 0$, suppose:*

1. *$a, b, c, \dfrac{\partial a}{\partial t}$ are bounded on $\mathbb{R} \times (0,T)$ with $a(x,t) \geq \alpha$ a.e. in $\mathbb{R} \times (0,T)$ for some $\alpha > 0$;*

2. *$f \in L^2(0,T;H^{0,\lambda}); \psi, \dfrac{\partial \psi}{\partial t} \in L^2(0,T;H^{1,\lambda})$;*

3. *$\bar{v} \in H^{1,\lambda}, \bar{v} \geq \psi(0)$;*

4. The set

$$\mathcal{X} := \Big\{ w \in L^2(0,T;H^{1,\lambda}) : \frac{\partial w}{\partial t} \in L^2\big(0,T;(H^{1,\lambda})^*\big), \\ w(t) \geq \psi(t) \ \text{a.e. } t \text{ in } [0,T] \Big\}$$

is non-empty, where $(H^{1,\lambda})^*$ denotes the dual space of $H^{1,\lambda}$.

Then there exists a unique function v such that:

$$v \in L^\infty(0,T;H^{1,\lambda}), \frac{\partial v}{\partial t} \in L^2(0,T;H^{0,\lambda}); \tag{4.1}$$

$$\left(\frac{\partial v}{\partial t}, w - v\right)_\lambda + a_\lambda(t;v,w-v) \geq (f, w-v)_\lambda, \quad \text{a.e. } t, \\ \forall w \in H^{1,\lambda} \text{ such that } w \geq \psi(t) \text{ a.e. } t \in (0,T); \tag{4.2}$$

$$v(\cdot,t) \geq \psi(t), \ \text{a.e. } t \in (0,T); \tag{4.3}$$

$$v(\cdot,0) = \bar{v}. \tag{4.4}$$

Moreover, if $v \in L^2(0,T;H^{2,\lambda})$, then v is a solution to the obstacle problem: find, $v \in L^2(0,T;H^{2,\lambda})$ such that v satisfies (4.3), (4.4), *and*

$$\frac{\partial v}{\partial t} + A(t)v - f \geq 0; \tag{4.5}$$

$$\left(\frac{\partial v}{\partial t} + A(t)v - f\right)(v - \psi) = 0, \tag{4.6}$$

almost everywhere in $\mathbb{R} \times (0,T)$.

Proof. For the most part, the Theorem is a restatement of Bensoussan and Lions (1982, Theorem 2.2, and Section 2.15, Chapter 3), where we have mapped $t \mapsto T - t$, and $v \mapsto -v$.

We therefore only need to explain the last part of the result. If we suppose $v \in L^2(0,T;H^{2,\lambda})$ and $\phi \in H^{1,\lambda}$, we have

$$\begin{aligned} a_\lambda(t;v,\phi) &= \int_{\mathbb{R}} e^{-2\lambda|x|} a(x,t) \frac{\partial v}{\partial x} \, \mathrm{d}\phi + \int_{\mathbb{R}} e^{-2\lambda|x|} \phi \left[b(x,t) \frac{\partial v}{\partial x} + c(x,t) v \right] \mathrm{d}x \\ &= \left[e^{-2\lambda|x|} a(x,t) \frac{\partial v}{\partial x} \phi \right]_{-\infty}^{\infty} + \int_{\mathbb{R}} e^{-2\lambda|x|} \phi \cdot A(t) v \, \mathrm{d}x, \end{aligned}$$

where the first term on the right-hand side vanishes since $v \in L^2(0,t;H^{1,\lambda})$ and $\phi \in H^{1,\lambda}$. Therefore, by (4.2), for any $w \in H^{1,\lambda}$ such that $w \geq \psi$ a.e. in $\mathbb{R}$,

$$\left(\frac{\partial v}{\partial t} + A(t)v - f, w - v\right)_\lambda \geq 0, \quad \text{a.e. } t.$$

Taking for example $w = v + \phi$, for a positive test function ϕ, we conclude that (4.5) holds. Moreover, let $w = \psi$ in the inequality above, we have

$$\int_{\mathbb{R}} e^{-2\lambda|x|} \left(\frac{\partial v}{\partial t} + A(t)v - f\right)(\psi - v) \, \mathrm{d}x \geq 0.$$

Then (4.6) follows from (4.3) and (4.5). □

4.2 Connection with Skorokhod's Embedding Problem

To connect our embedding problem $\mathbf{SEP}(\sigma,\nu,\mu)$ with the variational inequality, we need some assumptions on σ, μ and the starting distribution ν. Firstly, on $\sigma:\mathbb{R}\to\mathbb{R}_+$, we still assume (3.2) and (3.4) hold. In addition, we assume that:

$$\exists K>0, \text{ such that } \frac{1}{K}<\sigma<K \text{ on } \mathbb{R}. \tag{4.7}$$

On μ and ν, we still assume that $\mathrm{U}_\mu(x)\le \mathrm{U}_\nu(x)$ to ensure the existence of a solution to $\mathbf{SEP}(\sigma,\nu,\mu)$.

Under these assumptions, we can specify the coefficients in the evolutionary variational inequality, (4.4) and (4.5)–(4.6), to be:

$$\begin{aligned} a(x,t) &= \frac{\sigma^2(x)}{2}; \quad b(x,t)=\sigma(x)\sigma'(x)-\lambda\sigma^2(x)\mathrm{sgn}(x);\\ c(x,t) &= f(x,t)=0; \quad \psi(x,t)=\mathrm{U}_\mu(x); \quad \bar{v}=\mathrm{U}_\nu(x), \end{aligned} \tag{4.8}$$

then the corresponding operators are given by $A(t)=-\frac{\sigma^2(x)}{2}\frac{\partial^2}{\partial x^2}$ and

$$a_\lambda(t;v,w)=\int_{\mathbb{R}} e^{-2\lambda|x|}\left[\frac{\sigma^2(x)}{2}\frac{\partial v}{\partial x}\frac{\partial w}{\partial x}+\Big(\sigma(x)\sigma'(x)-\lambda\sigma^2(x)\mathrm{sgn}(x)\Big)\frac{\partial v}{\partial x}w\right]\mathrm{d}x.$$

We write the evolutionary variational inequality as:

$\mathbf{VI}(\sigma,\nu,\mu)$: Find a function $v:\mathbb{R}\times[0,T]\to\mathbb{R}$ satisfying (4.1)–(4.4), where all the coefficients are given in (4.8).

We also have a stronger formulation, that is:

$\mathbf{SVI}(\sigma,\nu,\mu)$: For given $T>0$, we seek a function v, in a suitable space, such that (4.3)–(4.6) hold, where all the coefficients are given in (4.8).

Our main result is then to show that finding the solution to $\mathbf{SEP}(\sigma,\nu,\mu)$ is equivalent to finding a (and hence the unique) solution to $\mathbf{VI}(\sigma,\nu,\mu)$:

Theorem 4.2. *Suppose* (3.2), (3.4) *and* (4.7) *hold. Also, let D and v be the solutions to $\mathbf{SEP}(\sigma,\nu,\mu)$ and $\mathbf{VI}(\sigma,\nu,\mu)$ respectively. Define $u(x,t):=-\mathbb{E}^\nu|x-X_{t\wedge\tau_D}|$ and D^T by:*

$$D^T:=\big\{(x,t)\in\mathbb{R}\times[0,T]; v(x,t)>\psi(x,t)\big\}. \tag{4.9}$$

Then we have $D^T=D\cap\mathbb{R}\times[0,T]$, and for all $(x,t)\in\mathbb{R}\times[0,T]$,

$$u(x,t)=v(x,t).$$

Moreover, if $u\in L^2(0,T;H^{2,\lambda})$ then u is also the solution to $\mathbf{SVI}(\sigma,\nu,\mu)$.

Proof. Let $\lambda>0$ be fixed, and suppose D is a solution to $\mathbf{SEP}(\sigma,\nu,\mu)$. We need to show u is a solution to $\mathbf{VI}(\sigma,\nu,\mu)$. First note that $\mathrm{U}_\mu(x)+|x|$ is continuous on $\mathbb{R}$, and converges to 0 as $x\to\pm\infty$, and hence is bounded. So $x\mapsto \mathrm{U}_\mu(x)+|x|\in L^\infty(0,T;H^{0,\lambda})$, and then $\mathrm{U}_\mu(x)\in L^\infty(0,T;H^{0,\lambda})$. Similarly, $\mathrm{U}_\nu(x)\in L^\infty(0,T;H^{0,\lambda})$. Since $0\ge \mathrm{U}_\nu(x)\ge u(x,t)\ge \mathrm{U}_\mu(x)$ for all $t\in[0,T]$, we have $u\in L^\infty(0,T;H^{0,\lambda})$. By (3.6), we also have $\left|\frac{\partial u}{\partial x}\right|\le 1$ since u is the potential of some probability distribution. Therefore we have $u\in L^\infty(0,T;H^{1,\lambda})$. By

Lemma 3.3 and the fact that u is constant (in time) outside D, $\left|\frac{\partial u}{\partial t}\right| \leq \sigma^2 p^\nu(x,t)$ a.e. on $\mathbb{R}\times[0,T]$ where $p^\nu(x,t)$ is the transition density of the diffusion process X starting from ν. Then by standard Gaussian estimates (e.g. Stroock (2008, Theorem 3.3.11)), we know there exists some constant $A>0$, depending only on K, such that

$$\begin{aligned}
&\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;H^{0,\lambda})} \\
&\quad \leq \int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} \frac{A}{1\wedge t}\exp\left\{-2\left(At-\frac{(x-y)^2}{At}\right)^- - 2\lambda|x|\right\}\mathrm{d}x\,\mathrm{d}t\,\nu(\mathrm{d}y) \\
&\quad = \int_{\mathbb{R}}\int_0^T \frac{A}{1\wedge t}\int_{y-At}^{y+At} e^{-2\lambda|x|}\,\mathrm{d}x\,\mathrm{d}t\,\nu(\mathrm{d}y) \\
&\qquad + \int_{\mathbb{R}}\int_0^T \frac{Ae^{2At}}{1\wedge t}\int_{\mathbb{R}\setminus(y-At,y+At)}\exp\left\{-\frac{2(x-y)^2}{At}-2\lambda|x|\right\}\mathrm{d}x\,\mathrm{d}t\,\nu(\mathrm{d}y) \\
&\quad \leq \int_{\mathbb{R}}\int_0^T \frac{A}{1\wedge t}\int_{-At}^{At} e^{-2\lambda|x|}\,\mathrm{d}x\,\mathrm{d}t\,\nu(\mathrm{d}y) \\
&\qquad + \int_{\mathbb{R}}\int_0^T \frac{Ae^{2At}}{1\wedge t}\int_{\mathbb{R}\setminus(y-At,y+At)}\exp\left\{-\frac{2(x-y)^2}{At}\right\}\mathrm{d}x\,\mathrm{d}t\,\nu(\mathrm{d}y) \\
&\quad = \frac{A}{\lambda}\int_0^T \frac{1}{1\wedge t}(1-e^{-2\lambda At})\,\mathrm{d}t + 2A\int_0^T \frac{e^{2At}}{1\wedge t}\int_{At}^\infty \exp\left\{-\frac{2z^2}{At}\right\}\mathrm{d}z\,\mathrm{d}t \\
&\quad \leq \frac{A}{\lambda}\int_0^T \frac{2A\lambda t}{1\wedge t}\,\mathrm{d}t + \frac{A^{3/2}\pi^{1/2}}{\sqrt{2}}\int_0^T \frac{e^{2At}\sqrt{t}}{1\wedge t}\,\mathrm{d}t < \infty,
\end{aligned}$$

where we have applied Hölder's inequality in the first line to get:

$$\left|\frac{\partial u}{\partial t}\right|^2 = \left|\int_{\mathbb{R}} p(t,y,x)\nu(\mathrm{d}y)\right|^2 \leq \int_{\mathbb{R}} p(t,y,x)^2\nu(\mathrm{d}y).$$

So $\frac{\partial u}{\partial t} \in L^2(0,T;H^{0,\lambda})$, and we have shown (4.1) holds.

By the same arguments used in the proof of Theorem 3.1, (4.3) and (4.4) hold. Now we consider (4.2). We begin by observing that, for any $\phi \in C_K^\infty$, if we write $\mu_t(\mathrm{d}x)$ for the law of $X_{t\wedge\tau_D}$, we have:

$$\begin{aligned}
\int_{\mathbb{R}} \frac{\partial\phi}{\partial x}\frac{\partial u}{\partial x}\,\mathrm{d}x &= \int_{\mathbb{R}} \frac{\partial \phi}{\partial x}\left(1-2\mathbb{P}(X_{t\wedge\tau_D}\leq x)\right)\mathrm{d}x \\
&= -2\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\partial\phi}{\partial x}\mathbf{1}_{\{y\leq x\}}\mu_t(\mathrm{d}y)\,\mathrm{d}x \\
&= 2\int_{\mathbb{R}} \phi(y)\mu_t(\mathrm{d}y) \\
&= 2\mathbb{E}\left[\phi(X_{t\wedge\tau_D})\right].
\end{aligned} \tag{4.10}$$

In addition, for any $w\in H^{1,\lambda}$, we can find a sequence $\{\phi_n\}\subset C_K^\infty$ such that

$$\lim_{n\to\infty}\|\phi_n-(w-u(\cdot,t))\|_{H^{1,\lambda}} = 0. \tag{4.11}$$

Moreover, $e^{-\lambda|x|}u(x,t)$ is bounded, and if $e^{-\lambda|x|}w$ is also bounded then we can in addition find a sequence $\{\phi_n\}\subset C_K^\infty$ such that $e^{-2\lambda|x|}\phi_n(x)\geq -K'$ for some constant K' independent of n. For any n, we therefore have

$$\int_{\mathbb{R}} e^{-2\lambda|x|} \frac{\sigma^2}{2} \frac{\partial u}{\partial x} \frac{\partial \phi_n}{\partial x} \, \mathrm{d}x = -\int_{\mathbb{R}} e^{-2\lambda|x|} \left(\sigma\sigma' - \lambda\sigma^2 \mathrm{sgn}(x)\right) \frac{\partial u}{\partial x} \phi_n \, \mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n \sigma^2 \, \mu_t(\mathrm{d}x). \tag{4.12}$$

On the other hand, since $\partial u/\partial t$ vanishes outside D, and, using the same arguments as (3.7) (which still hold on account of Lemma 3.3), is equal to $-\sigma(x)^2 p^D(x,t)$, we have, for almost every $t \in [0,T]$

$$\int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n \frac{\partial u}{\partial t} \, \mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n \sigma^2 \, \mu_t(\mathrm{d}x) = \int_{\mathbb{R}\setminus D_t} e^{-2\lambda|x|} \phi_n \sigma^2 \, \mu_t(\mathrm{d}x), \tag{4.13}$$

where $D_s := \{x \in \mathbb{R} : (x,s) \in D\}$. By (4.12) and (4.13),

$$\begin{aligned}
\left(\frac{\partial u}{\partial t}, \phi_n\right)_\lambda &+ a_\lambda(t; u, \phi_n) \\
&= \int_{\mathbb{R}} e^{-2\lambda|x|} \left[\frac{\partial u}{\partial t}\phi_n + \frac{\sigma^2}{2} \frac{\partial u}{\partial x} \frac{\partial \phi_n}{\partial x} + \left(\sigma\sigma' - \lambda\sigma^2 \mathrm{sgn}(x)\right) \frac{\partial u}{\partial x} \phi_n\right] \mathrm{d}x \\
&= \int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n \frac{\partial u}{\partial t} \, \mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n \sigma^2 \, \mu_t(\mathrm{d}x) \\
&= \int_{\mathbb{R}\setminus D_t} e^{-2\lambda|x|} \phi_n \sigma^2 \, \mu_t(\mathrm{d}x),
\end{aligned}$$

for almost every $t \in [0,T]$. Now suppose initially we have $e^{-\lambda|x|}w$ bounded, and choose a sequence ϕ_n as above. Then we can let $n \to \infty$ and apply Fatou's Lemma and the fact that $u = \psi$ on $\mathbb{R} \setminus D_t$ and $w \geq \psi$ to get:

$$\begin{aligned}
-\left(\frac{\partial u}{\partial t}, w - u\right)_\lambda &+ a_\lambda(t; u, w - u) \\
&= \int_{\mathbb{R}\setminus D_t} e^{-2\lambda|x|} (w - \psi) \sigma^2 \, \mu_t(\mathrm{d}x) \geq 0,
\end{aligned}$$

for almost every $t \in [0,T]$. So (4.2) holds when $e^{-\lambda|x|}w$ is bounded. The general case follows from noting that $\max\{w, -N\}$ converges to w in $H^{1,\lambda}$. We can conclude that u is a solution to $\mathbf{VI}(\sigma, \nu, \mu)$. In addition, the final statement of the theorem now follows from Theorem 4.1.

Conversely, suppose that we have already found the solution to $\mathbf{VI}(\sigma, \nu, \mu)$, denoted by $v(x,t)$. By Theorem 4.1 and the preceding argument, we have

$$-\mathbb{E}^\nu |x - X_{t \wedge \tau_D}| = v(x,t),$$

when $(x,t) \in \mathbb{R} \times [0,T]$. Finally, we need only note (from (3.7), and the line above) that whenever $(x,t) \in D$, we have $u(x,t) > \psi(x,t)$, and hence $D^T = D \cap \mathbb{R} \times [0,T]$. □

Remark 4.3. The constant λ which appears in the variational inequality can now be seen to be unimportant: if we consider two positive numbers $\lambda < \lambda^*$, then by Theorem 4.1, there exist v and v^* satisfying (4.1)–(4.4) with the parameters λ and λ^* respectively. According to Theorem 4.2,

$$u(x,t) = v(x,t) = v^*(x,t),$$

so $v = v^*$. Therefore, the description of Root's barrier by the strong variational inequality is not affected by the choice of the parameter $\lambda > 0$. We do however need $\lambda > 0$, since this assumption is used in e.g.(4.12) to ensure we can integrate by parts.

Remark 4.4. As noted in Bensoussan and Lions (1982), and which is well known, one can connect the solution to the variational inequality $\mathbf{VI}(\sigma, \nu, \mu)$ to the solution of a particular optimal stopping problem. In our context, the function v which arises in the solution to $\mathbf{VI}(\sigma, \nu, \mu)$ is also the function which arises from solving the problem:

$$v(x,t) = \sup_{\tau \le t} \mathbb{E}^x \left[\mathrm{U}_\mu(X_\tau) \mathbf{1}_{\{\tau < t\}} + \mathrm{U}_\nu(X_\tau) \mathbf{1}_{\{\tau = t\}} \right]. \tag{4.14}$$

This seems a rather interesting observation, and at one level extends a number of connections known to exist between solutions to the Skorokhod embedding problem, and solutions to optimal stopping problems: e.g. Peskir (1998), Obłój (2007) and Cox et al. (2008).

What is rather interesting, and appears to differ from these other situations, is that the above examples are all cases where the same stopping time is both a Skorokhod embedding, and a solution to the relevant optimal stopping problem. In the context here, we see that the optimal stopping problem is *not* solved by Root's stopping time. Rather, the problem given in (4.14) runs 'backwards' in time: if we keep t fixed, then the solution to (4.14) is:

$$\tau_D = \inf \{ s \ge 0 : (X_s, t-s) \notin D \} \wedge t.$$

In addition, our connection between these two problems is only through the analytic statement of the problem: it would be interesting to have a probabilistic explanation for the correspondence.

Remark 4.5. The above ideas also allow us to construct alternative embeddings which fail to be uniformly integrable. Consider using the variational inequality to construct the domain D in the manner described above, but with the function ψ chosen to be $\mathrm{U}_\mu(x) - c$, for some $c > 0$. Then one would expect $B = D^{\complement}$ to be a barrier, which is non-empty, so that $\tau_D < \infty$ a.s., and the functions $u(x,t)$ and $v(x,t)$ as defined in Theorem 4.2 to agree (for example by taking bounded approximations to D). In particular, $\lim_{t \to \infty} u(x,t) = \mathrm{U}_\mu(x) - c$. Since $X_{t \wedge \tau_D}$ is no longer uniformly integrable, we cannot simply infer that this holds in the limit, but we can consider for example

$$u(x,t) - u(z,t) = -\mathbb{E}\left[|X_{t \wedge \tau_D} - x| - |X_{t \wedge \tau_D} - z| \right]$$

which is a bounded function. Taking the limit as $t \to \infty$, we can deduce that

$$-\mathbb{E}\left[|X_{\tau_D} - x| - |X_{\tau_D} - z| \right] = \mathrm{U}_\mu(x) - \mathrm{U}_\mu(z).$$

From this expression, we can divide through by $(x - z)$ and take the limit as $x \downarrow z$ to get $2\mathbb{P}(X_{\tau_D} > z) - 1$. The law of X_{τ_D} now follows.

Note also that there is no reason that the distribution above needed to have the same mean as ν, and this can lead to constructions where the means differ. In general, these constructions will not give rise to a uniformly integrable embedding, but if we take two general (integrable) distributions, there is a natural choice, which is to find the smallest $c \in \mathbb{R}$ such that $\mathrm{U}_\nu(x) \geq \mathrm{U}_\mu(x) - c$. In such a case, we would expect the resulting construction to be *minimal* in the sense that there is no other construction of a stopping time which embeds the same distribution, and is almost surely smaller. See Monroe (1972) and Cox (2008) for further details regarding minimality.

4.3 Geometric Brownian motion

An important motivating example for our study is the financial application of Root's solution described in the introduction. In both Dupire (2005) and Carr and Lee (2010), the case $\sigma(x) = x$ plays a key role in both the pricing and the construction of a hedging portfolio. However, in the previous section, we only discussed the relation between Root's construction and variational inequalities under the assumptions (3.2), (3.4) and (4.7), where the last assumption is not satisfied by σ in this special case.

In this section, we study this special case: $\sigma(x) = x$, so that X_t is a geometric Brownian motion. In addition, we will assume that the process is strictly positive, so that ν and μ are supported on $(0, \infty)$. We therefore consider the Skorokhod embedding problem $\mathbf{SEP}(\sigma, \nu, \mu)$ with starting distribution ν, where ν and μ are integrable probability distributions satisfying

$$\mathrm{supp}(\mu) \subset (0, \infty), \quad \mathrm{supp}(\nu) \subset (0, \infty), \quad \mathrm{U}_\mu(x) \leq \mathrm{U}_\nu(x), \text{ and } \int x^2 \, \mathrm{d}\nu < \infty. \tag{4.15}$$

Note that the assumption that μ and ν are integrable, and $\mathrm{U}_\mu(x) \leq \mathrm{U}_\nu(x)$ together imply that $m := \int x \, \nu(\mathrm{d}x) = \int x \, \mu(\mathrm{d}x)$.

The solution to the stochastic differential equation

$$\mathrm{d}X_t = X_t \, \mathrm{d}W_t, \; X_0 = x_0$$

is the geometric Brownian motion $x_0 \exp\{W_t - t/2\}$, and, for $y > 0$, the transition density of the process is:

$$p_t(y, x) := \frac{1}{x} \frac{1}{\sqrt{2\pi t}} \mathbf{1}_{\{x>0\}} \exp\left\{ -\frac{(\ln x - \ln y + t/2)^2}{2t} \right\}. \tag{4.16}$$

By analogy with Theorem 3.1, if D is the solution to $\mathbf{SEP}(\sigma, \nu, \mu)$, then we would expect

$$\frac{\partial u}{\partial t} = \frac{x^2}{2} \frac{\partial^2 u}{\partial x^2}, \text{ on } D; \quad u(x, t) = \mathrm{U}_\mu(x) \text{ on } \mathbb{R} \times (0, \infty) \setminus D;$$

where u is defined as before by $u(x, t) = -\mathbb{E}|x - X_{t \wedge \tau_D}|$. However, if we follow the arguments in Section 4.2, we find that we need to set $a(x, t) = x^2/2$ in $\mathbf{VI}(\sigma, \nu, \mu)$, which would not satisfy the first condition of Theorem 4.1. To avoid this we will perform a simple transformation of the problem. We set

$$v(x, t) = u(e^x, t), \quad (x, t) \in \mathbb{R} \times [0, T].$$

Define the operator $A(t) := -\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial}{\partial x}$, then we have, when $(e^x, t) \in D$,

$$\frac{\partial v}{\partial t} + A(t)v = 0. \tag{4.17}$$

We state our main result of this section as follows:

Theorem 4.6. *Suppose $\sigma(x) = x$ on $(0, \infty)$ and μ and ν satisfy* (4.15)*. Moreover, assume D solves $\mathbf{SEP}(\sigma, \nu, \mu)$, and $u(x, t) := -\mathbb{E}|x - X_{t\wedge\tau_D}|$. Then $v(x, t) := u(e^x, t)$ is the unique solution to* (4.1)–(4.4) *where we set*

$$\begin{aligned} &a(x,t) = \frac{1}{2}; \quad b(x,t) = \frac{1}{2} - \lambda \cdot \operatorname{sgn}(x); \quad c(x,t) = f(x,t) = 0; \\ &\psi(x,t) = \mathrm{U}_\mu(e^x); \quad \bar{v} = \mathrm{U}_\nu(e^x); \quad \lambda > \frac{1}{2}. \end{aligned} \tag{4.18}$$

Proof. Much of the proof will follow the proof of Theorem 4.2. As before, (4.3) and (4.4) are clear. In addition, we note that $\psi - e^x$ is continuous and converges to 0 as $x \to \infty$ and converges to $\mathrm{U}_\mu(0) < \infty$ as $x \to -\infty$, so $x \mapsto \psi - e^x \in L^\infty(0, T; H^{0,\lambda})$. Hence $\psi \in L^\infty(0, T; H^{0,\lambda})$ since we have $\lambda > \dfrac{1}{2}$. Thus, $v \in L^\infty(0, T; H^{0,\lambda})$. Moreover, we can easily see $|\partial v/\partial x|$ is bounded by e^x. Therefore, $v \in L^\infty(0, T; H^{1,\lambda})$ when $\lambda > \dfrac{1}{2}$. On the other hand, since $|\partial v/\partial t|$ is bounded by $e^{2x} \int p_t(y, e^x)\, \nu(\mathrm{d}y)$, we have, by Hölder's inequality,

$$\left|\frac{\partial v}{\partial t}\right|^2 \le \int_{\mathbb{R}_+} \frac{1}{2\pi t} \exp\left\{-\frac{(x - \ln y + t/2)^2}{t} + 2x\right\} \nu(\,\mathrm{d}y),$$

and hence,

$$\begin{aligned} \left\|\frac{\partial v}{\partial t}\right\|_{L^2(0,T;H^{0,\lambda})} &\le \int_{\mathbb{R}_+} \int_0^T \int_{\mathbb{R}} \frac{e^{-2\lambda|x|}}{2\pi t} \exp\left\{-\frac{(x - \ln y + t/2)^2}{t} + 2x\right\} \mathrm{d}x\, \mathrm{d}t\, \nu(\mathrm{d}y) \\ &\le \int_{\mathbb{R}_+} \int_0^T \int_{\mathbb{R}} \frac{1}{2\pi t} \exp\left\{-\frac{(x - \ln y - t/2)^2}{t} + 2\ln y\right\} \mathrm{d}x\, \mathrm{d}t\, \nu(\mathrm{d}y) \\ &\le \int_{\mathbb{R}_+} y^2\, \nu(\mathrm{d}y) \int_0^T \frac{1}{2\sqrt{\pi t}}\, \mathrm{d}t < \infty. \end{aligned}$$

Hence (4.1) is verified.

Using (4.10), for $\phi \in C_K^\infty$ we get

$$\begin{aligned} \int_{\mathbb{R}} \left(\frac{\partial \phi}{\partial x}(x) + \phi(x)\right) \frac{\partial v}{\partial x}\, \mathrm{d}x &= \int_0^\infty \frac{\partial}{\partial y}\left[\phi(\ln(y))y\right] \frac{\partial u}{\partial x}(y,t)\, \mathrm{d}y \\ &= 2\mathbb{E}\left[\phi(\ln(X_{t\wedge\tau_D}))X_{t\wedge\tau_D}\right], \end{aligned} \tag{4.19}$$

and so we define the measure ν_t by

$$\int \phi(x)\, \nu_t(\mathrm{d}x) = \mathbb{E}\left[\phi(\ln(X_{t\wedge\tau_D}))X_{t\wedge\tau_D}\right].$$

Now take any $w \in H^{1,\lambda}$, and take $\{\phi_n\} \subset C_K^\infty$ satisfying (4.11). By (4.17) and (4.19), similar arguments to those used in the proof of Theorem 4.2 give

$$\int_{\mathbb{R}} e^{-2\lambda|x|} \frac{\partial v}{\partial x} \left(\frac{1}{2}\frac{\partial \phi_n}{\partial x} + \frac{1}{2}\phi_n - \lambda \cdot \operatorname{sgn}(x)\right) \mathrm{d}x = \int_{\mathbb{R}} e^{-2\lambda|x|} \phi_n\, \nu_t(\mathrm{d}x),$$

and

$$\int_{\mathbb{R}} e^{-2\lambda|x|}\frac{\partial v}{\partial t}\phi_n \,\mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n \,\nu_t(\mathrm{d}x) = \int_{\mathbb{R}\setminus\widetilde{D}_t} e^{-2\lambda|x|}\phi_n \,\nu_t(\mathrm{d}x),$$

for almost all $t \in [0,T]$, where $\widetilde{D}_t := \{x \in \mathbb{R} : (e^x, t) \in D\}$. Thus, for almost every $t \in [0,T]$,

$$\begin{aligned}\left(\frac{\partial v}{\partial t}, \phi_n\right)_\lambda &+ a_\lambda(t; v, \phi_n)\\ &= \int_{\mathbb{R}} \left(\frac{\partial v}{\partial t}\phi_n + \frac{1}{2}\frac{\partial \phi_n}{\partial x}\frac{\partial v}{\partial x} + \left(\frac{1}{2} - \lambda\cdot\mathrm{sgn}(x)\right)\phi_n\frac{\partial v}{\partial x}\right)\mathrm{d}x\\ &= \int_{\mathbb{R}\setminus\widetilde{D}_t} e^{-2\lambda|x|}\phi_n \,\nu_t(\mathrm{d}x).\end{aligned}$$

Finally, following the same arguments as in the proof of Theorem 4.2, we conclude (4.2) holds. Therefore v is a solution to (4.1)–(4.4) with coefficients determined by (4.18). The uniqueness is clear since it is easy to check the coefficients defined in (4.18) satisfy the conditions in Theorem 4.1. □

5 Optimality of Root's Solution

For a given distribution μ, Rost (1976) proves that Root's construction is optimal in the sense of 'minimal residual expectation'. It is easy to check that this is equivalent to the slightly more general problem:

$$\begin{aligned}\text{minimise} \quad & \mathbb{E}\big[F(\tau)\big]\\ \text{subject to:} \quad & \mathcal{L}(X_\tau) = \mu;\\ & \tau \text{ is a UI stopping time.}\end{aligned}$$

Here we assume μ is a given integrable and centred distribution, X is the diffusion process defined by (3.1), where the diffusion coefficient σ satisfies (3.2)–(3.4), with initial distribution $\mathcal{L}(X_0) = \nu$, and F is a given convex, increasing function with right derivative f and $F(0) = 0$.

Our aim in this section is twofold. Firstly, since Rost's original proof relies heavily on notions from potential theory, to give a proof of this result using probabilistic techniques. Secondly, we shall be able to give a 'pathwise inequality' which encodes the optimality in the sense that we can find a submartingale G_t, and a function $H(x)$ such that

$$F(t) \geq G_t + H(X_t) \tag{5.1}$$

and such that, for τ_D, equality holds in (5.1) and $G_{t\wedge\tau_D}$ is a UI martingale. It then follows that τ_D does indeed minimise $\mathbb{E}F(\tau)$ among all solutions to the Skorokhod embedding problem. The importance of (5.1) is that we can characterise the submartingale G_t, which will correspond in the financial setting to a dynamic trading strategy for constructing a sub-replicating hedging strategy for call-type payoffs on variance options.

We first define the key functions $G(x,t)$ and $H(x)$, where the submartingale in (5.1) is $G_t = G(X_t, t)$, and give key results concerning these functions.

We suppose that we have solved Root's problem for the given distributions, and hence have our barrier $B = D^{\complement}$. Define the function

$$M(x,t) = \mathbb{E}^{(x,t)} f(\tau_D), \tag{5.2}$$

where τ_D is the corresponding Root stopping time. In the following, we shall assume:

$$M(x,t) < \infty,\ \forall (x,t) \in \mathbb{R} \times \mathbb{R}_+. \tag{5.3}$$

We suppose also (at least initially) that (3.2)–(3.4) and (4.7) still hold. Note that $M(x,t)$ now has the following important properties. First, since f is right-continuous (it is the right derivative of F), $M(x,t) = f(t)$ whenever $(x,t) \notin D$ and $t > 0$. In addition, since f is increasing, for all x and t we have $M(x,t) \geq f(t)$.

Now define a function $Z(x)$ by:

$$Z(x) = 2 \int_0^x \int_0^y \frac{M(z,0)}{\sigma^2(z)} \, \mathrm{d}z \, \mathrm{d}y. \tag{5.4}$$

So in particular, we have $Z''(x) = 2\frac{M(x,0)}{\sigma^2(x)}$ and $Z(x)$ is a convex function. Define also:

$$G(x,t) = \int_0^t M(x,s) \, \mathrm{d}s - Z(x), \tag{5.5}$$

and

$$H(x) = \int_0^{R(x)} (f(s) - M(x,s)) \, \mathrm{d}s + Z(x). \tag{5.6}$$

Two key results concerning these functions are then:

Proposition 5.1. *We have, for all $(x,t) \in \mathbb{R} \times \mathbb{R}_+$:*

$$G(x,t) + H(x) \leq F(t). \tag{5.7}$$

And also:

Lemma 5.2. *Suppose that f is bounded and, for any $T > 0$:*

$$\mathbb{E}\left[\int_0^T Z'(X_s)^2 \sigma(X_s)^2 \, \mathrm{d}s\right] < \infty, \quad \mathbb{E}Z(X_0) < \infty. \tag{5.8}$$

Then the process

$$G(X_{t\wedge\tau_D}, t \wedge \tau_D) \text{ is a martingale,} \tag{5.9}$$

and

$$G(X_t, t) \text{ is a submartingale.} \tag{5.10}$$

Using these results, we are able to prove the following theorem, which gives us Rost's result regarding the optimality of Root's construction.

Theorem 5.3. *Suppose D solves* $\mathbf{SEP}(\sigma, \mu, \nu)$*, and equations* (5.3) *and* (5.8) *hold. Then*

$$\mathbb{E}F(\tau_D) \leq \mathbb{E}F(\tau) \tag{5.11}$$

whenever τ is a stopping time such that $X_\tau \sim \mu$.

Proof. We begin by considering the case where $\mathbb{E}\tau_D < \infty, \mathbb{E}\tau < \infty$ and f is bounded. Since $Z(x)$ is convex, by the Meyer-Itô formula (e.g. Protter (2005, Theorem IV.71)):

$$Z(X_t) = Z(X_0) + \int_0^t Z'(X_r)\,\mathrm{d}X_r + \frac{1}{2}\int_0^t Z''(X_r)\sigma^2(X_r)\,\mathrm{d}r.$$

By (5.8) and the fact that f is bounded (and hence also $M(X_s, 0)$ is bounded) we get

$$\mathbb{E}Z(X_{t\wedge\tau}) = \mathbb{E}Z(X_0) + \mathbb{E}\int_0^{t\wedge\tau} M(X_s, 0)\,\mathrm{d}s \le f(\infty)\mathbb{E}\tau + \mathbb{E}Z(X_0).$$

Applying Fatou's Lemma, we deduce that for any stopping time τ with finite expectation, $Z(X_\tau)$ is integrable. Moreover for such a stopping time, by convexity, $Z(X_{t\wedge\tau}) \le \mathbb{E}[Z(X_\tau)|\mathcal{F}_t]$, and so $G(X_{t\wedge\tau}, t\wedge\tau)$ is a submartingale which is bounded below by a UI martingale, and bounded above by $f(\infty)\tau$. It follows that $\mathbb{E}G(X_{t\wedge\tau}, t\wedge\tau) \to \mathbb{E}G(X_\tau, \tau)$ as $t \to \infty$. The same arguments hold when we replace τ by τ_D.

Since $R(X_{\tau_D}) \le \tau_D$ and if $t \in [R(x), \infty)$ then $\tau_D = t, \mathbb{P}^{(x,t)}$-a.s., so that $M(X_{\tau_D}, s) = f(s)$ for $s \ge \tau_D$, we have

$$\begin{aligned} G(X_{\tau_D}, \tau_D) + &\int_0^{R(X_{\tau_D})} (f(s) - M(X_{\tau_D}, s))\,\mathrm{d}s + Z(X_{\tau_D}) \\ &= \int_0^{\tau_D} M(X_{\tau_D}, s)\,\mathrm{d}s + \int_0^{R(X_{\tau_D})} (f(s) - M(X_{\tau_D}, s))\,\mathrm{d}s \\ &= \int_0^{\tau_D} M(X_{\tau_D}, s)\,\mathrm{d}s + \int_0^{\tau_D} (f(s) - M(X_{\tau_D}, s))\,\mathrm{d}s \\ &= \int_0^{\tau_D} f(s)\,\mathrm{d}s = F(\tau_D). \end{aligned} \tag{5.12}$$

On the other hand, since $X_{\tau_D} \sim X_\tau$, and observing that $G(X_{\tau_D}, \tau_D)$ and $F(\tau_D)$ are integrable, so too is $H(X_{\tau_D})$, and

$$\mathbb{E}H(X_{\tau_D}) = \mathbb{E}H(X_\tau).$$

In addition, by Lemma 5.2 and the limiting behaviour deduced above, we have:

$$\mathbb{E}G(X_{\tau_D}, \tau_D) = \lim_{t\to\infty} \mathbb{E}G(X_{t\wedge\tau_D}, t\wedge\tau_D) \le \lim_{t\to\infty} \mathbb{E}G(X_{t\wedge\tau}, t\wedge\tau) = \mathbb{E}G(X_\tau, \tau).$$

Putting these together, we get

$$\begin{aligned} \mathbb{E}F(\tau_D) &= \mathbb{E}\left[G(X_{\tau_D}, \tau_D) + H(X_{\tau_D})\right] \\ &\le \mathbb{E}\left[G(X_\tau, \tau) + H(X_\tau)\right] \\ &\le \mathbb{E}F(\tau). \end{aligned}$$

We now consider the case where at least one of τ or τ_D has infinite expectation. Note that if $F(\cdot) \not\equiv 0$ then there is some $\alpha, \beta \in \mathbb{R}$ with $\beta > 0$ such that $F(t) \ge \alpha + \beta t$, and hence we cannot have $\mathbb{E}\tau = \infty$ or $\mathbb{E}\tau_D = \infty$ without the corresponding term in (5.11) also being infinite. The only case which need concern us is the case where $\mathbb{E}\tau < \infty$, but $\mathbb{E}\tau_D = \infty$. Note however that τ_D remains UI, so $\mathbb{E}[X_{t\wedge\tau_D}|\mathcal{F}_t] = X_t$. In addition, from the arguments applied above, we know $Z(X_\tau)$ is integrable and since $X_\tau \sim X_{\tau_D}$, so too is $Z(X_{\tau_D})$. Then $H(X_\tau)$ and $H(X_{\tau_D})$ are both bounded above by an integrable random variable, so their expectations are well defined (although possibly not finite), and equal. Then, as above, $-\mathbb{E}[Z(X_{\tau_D})|\mathcal{F}_t] \leq -Z(X_{t\wedge\tau_D}) \leq G(X_{t\wedge\tau_D}, t\wedge\tau_D)$. We can deduce that $\mathbb{E}G(X_{\tau_D}, \tau_D) \leq \lim_{n\to\infty} \mathbb{E}G(X_{t\wedge\tau_D}, t\wedge\tau_D) = G(X_0, 0) \leq \mathbb{E}G(X_\tau, \tau)$. The remaining steps follow as previously, and it must follow that in fact $\mathbb{E}F(\tau_D) \leq \mathbb{E}F(\tau)$, which contradicts the assumption that $\mathbb{E}\tau < \infty$ and $\mathbb{E}\tau_D = \infty$.

To observe that the result still holds when f is unbounded, observe that we can apply the above argument to $f(t) \wedge N$, and $F_N(t) = \int_0^s f(s) \wedge N \,\mathrm{d}s$ to get $\mathbb{E}F_N(\tau_D) \leq \mathbb{E}F_N(\tau)$, and the conclusion follows on letting $N \to \infty$. □

We now turn to the proofs of our key results:

Proof of Proposition 5.1. If $t \leq R(x)$, then the left-hand side of (5.7) is:

$$\int_0^t f(s)\,\mathrm{d}s + \int_t^{R(x)} (f(s) - M(x,s))\,\mathrm{d}s = F(t) - \int_t^{R(x)} (M(x,s) - f(s))\,\mathrm{d}s$$

and we know $M(x,s) \geq f(s) \geq 0$, so that the inequality holds.

Now consider the case where $R(x) \leq t$. Then the left-hand side of (5.7) becomes:

$$\int_{R(x)}^t M(x,s)\,\mathrm{d}s + \int_0^{R(x)} f(s)\,\mathrm{d}s = \int_{R(x)}^t f(s)\,\mathrm{d}s + \int_0^{R(x)} f(s)\,\mathrm{d}s = F(t).$$

□

Proof of Lemma 5.2. We begin by noting that $Z(x)$ is convex, and therefore the Meyer-Itô formula (e.g. Protter (2005, Theorem IV.71)) gives:

$$Z(X_t) - Z(X_s) = \int_s^t Z'(X_r)\,\mathrm{d}X_r + \frac{1}{2}\int_s^t Z''(X_r)\sigma^2(X_r)\,\mathrm{d}r.$$

It follows from (5.8) that the first integral is a martingale. So we get:

$$\mathbb{E}\left[Z(X_t) - Z(X_s)|\mathcal{F}_s\right] = \int_s^t \mathbb{E}\left[M(X_r, 0)|\mathcal{F}_s\right]\,\mathrm{d}r, \quad s \leq t.$$

In addition, since $M(x,t) \geq f(t)$ and $f(t)$ is increasing, for $r, u \geq 0$ by the strong Markov property, writing $\widetilde{X}$ for an independent stochastic process with the same law as X and $\widetilde{\tau}_D$ for the corresponding hitting time of the barrier, we have:

$$\begin{aligned}
\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] &= \mathbf{1}_{\tau_D > r+u}\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] + \mathbf{1}_{\tau_D \leq r+u}\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] \\
&\leq \mathbf{1}_{\tau_D > r+u}\mathbb{E}^{(X_u^x, r+u)}[f(\widetilde{\tau}_D)] + \mathbf{1}_{\tau_D \leq r+u} f(r+u) \\
&\leq M(X_u^x, r+u).
\end{aligned}$$

When $r = 0$, we have $\mathbb{E}^{(x,0)}[f(\tau_D)|\mathcal{F}_u] \leq M(X_u^x, u)$. For $s, u \in [0,t]$,

$$\begin{aligned}
\mathbb{E}[M(X_t, u)|\mathcal{F}_s] &= \mathbb{E}^{X_s} M(\widetilde{X}_{t-s}, u) \\
&\geq \mathbb{E}^{(X_s, u-(t-s))}[f(\widetilde{\tau}_D)] \\
&\geq M(X_s, u-(t-s)),
\end{aligned} \tag{5.13}$$

when $u \geq t - s$. On the other hand, if $u < t - s$:

$$\begin{aligned}\mathbb{E}[M(X_t, u)|\mathcal{F}_s] &= \mathbb{E}[\mathbb{E}^{(X_{t-u},0)}\left[M(\widetilde{X}_u, u)\right]|\mathcal{F}_s] \\ &\geq \mathbb{E}[\mathbb{E}^{(X_{t-u},0)}\left[f(\widetilde{\tau}_D)\right]|\mathcal{F}_s] \\ &\geq \mathbb{E}[M(X_{t-u}, 0)|\mathcal{F}_s]. \end{aligned} \tag{5.14}$$

Then we can write:

$$\begin{aligned}\mathbb{E}\left[G(X_t, t)|\mathcal{F}_s\right] &= \int_0^t \mathbb{E}\left[M(X_t, u)|\mathcal{F}_s\right] \mathrm{d}u - \mathbb{E}\left[Z(X_t)|\mathcal{F}_s\right] \\ &= G(X_s, s) + \int_0^t \mathbb{E}\left[M(X_t, u)|\mathcal{F}_s\right] \mathrm{d}u - \int_0^s M(X_s, u)\,\mathrm{d}u \\ &\quad - \mathbb{E}\left[Z(X_t) - Z(X_s)|\mathcal{F}_s\right] \\ &\geq G(X_s, s) + \int_0^{t-s} \mathbb{E}\left[M(X_{t-u}, 0)|\mathcal{F}_s\right] \mathrm{d}u - \int_0^s M(X_s, u)\,\mathrm{d}u \\ &\quad - \int_s^t \mathbb{E}\left[M(X_u, 0)|\mathcal{F}_s\right] \mathrm{d}u + \int_{t-s}^t M(X_s, s - t + u)\,\mathrm{d}u \\ &\geq G(X_s, s) + \int_s^t \mathbb{E}\left[M(X_u, 0)|\mathcal{F}_s\right] du - \int_s^t \mathbb{E}\left[M(X_u, 0)|\mathcal{F}_s\right] \mathrm{d}u \\ &\quad + \int_0^s M(X_s, u)\,\mathrm{d}u - \int_0^s M(X_s, u)\,\mathrm{d}u \\ &\geq G(X_s, s).\end{aligned}$$

Where we have used (5.13) and (5.14) in the third line.

On the other hand, on $\{\tau_D \geq s\}$, from the definition of $M(x, t)$, and the Markov Property, we get:

$$\mathbb{E}\left[M(X_{t\wedge\tau_D}, t \wedge \tau_D - u)|\mathcal{F}_s\right] = M(X_s, s - u) \tag{5.15}$$

when $u \leq s$, and

$$\mathbb{E}\left[M(X_{t\wedge\tau_D}, t \wedge \tau_D - u)|\mathcal{F}_u\right] = M(X_u, 0) \tag{5.16}$$

when $u \in [s, t \wedge \tau_D]$. Then a similar calculation to above gives, for $s \leq \tau_D$:

$$\begin{aligned}&\mathbb{E}\left[G(X_{t\wedge\tau_D}, t \wedge \tau_D)|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[\int_0^{t\wedge\tau_D} M(X_{t\wedge\tau_D}, t \wedge \tau_D - u)\,\mathrm{d}u|\mathcal{F}_s\right] - \mathbb{E}\left[Z(X_{t\wedge\tau_D})|\mathcal{F}_s\right] \\ &= \int_0^s M(X_s, s - u)\,\mathrm{d}u + \mathbb{E}\left[\int_s^{t\wedge\tau_D} M(X_{t\wedge\tau_D}, t \wedge \tau_D - u)\,\mathrm{d}u|\mathcal{F}_s\right] \\ &\quad - Z(X_s) - \mathbb{E}\left[\int_s^{t\wedge\tau_D} M(X_u, 0)\,\mathrm{d}u|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[\int_s^t \mathbb{E}\left[M(X_{t\wedge\tau_D}, t \wedge \tau_D - u) - M(X_u, 0)|\mathcal{F}_u\right] \mathbf{1}_{\{u\leq\tau_D\}}\,\mathrm{d}u|\mathcal{F}_s\right] \\ &\quad + G(X_s, s) \\ &= G(X_s, s),\end{aligned}$$

where we have used (5.15) and (5.16). ☐

Remark 5.4. Note that the fact that our choice of D given in the solution is the domain D which arises in solving Root's embedding problem is only used in Theorem 5.3 to enforce the lower bound. In fact, we could choose any barrier B, and $D = B^{\complement}$ as our domain, and this would result in an lower bound, with corresponding functions G and H. The choice of Root's barrier gives the optimal lower bound, in that we can attain equality for some stopping time. In this context, it is worth recalling the lower bounds given by Carr and Lee (2010, Proposition 3.1) — here a lower bound is given which essentially corresponds to choosing the domain with $R(x) = Q$, for a constant Q. The arguments given above show that similar constructions are available for any choice of R, and the optimal choice corresponds to Root's construction.

Remark 5.5. Although the preceding section is written for a diffusion on $\mathbb{R}$, it is not hard to check that the case where $\sigma(x) = x$ can also be included without many changes. In this setting, we need to restrict the space variable to the space $(0, \infty)$ (so we assume that $\tau_D < \infty$ a.s.), and consider a starting distribution which is also supported on $(0, \infty)$, and with a corresponding change to (5.3).

6 Financial Applications

We now turn to our motivating financial problem: consider an asset price S_t defined on a complete probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, with:

$$\frac{dS_t}{S_t} = r_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t \tag{6.1}$$

under some probability measure $\mathbb{Q} \sim \mathbb{P}$, where $\mathbb{P}$ is the objective probability measure, and W_t a $\mathbb{Q}$-Brownian motion. In addition, we suppose r_t is the risk-free rate which we require to be known, but which need not be constant. In particular, let r_t, σ_t be locally bounded, predictable processes so that the integral in (6.1) is well defined, and so S_t is an Itô process. We suppose that the process σ_t is not known (or more specifically, we aim to produce conclusions which hold for all σ_t in the class described). Specifically, we shall suppose:

Assumption 6.1. *The asset price process, under some probability measure $\mathbb{Q} \sim \mathbb{P}$, is the solution to the SDE* (6.1)*, where r_t and σ_t are locally bounded, predictable processes.*

In addition, we need to make the following assumptions regarding the set of call options, which are initially traded:

Assumption 6.2. *We suppose that call options with maturity T, and at all strikes $\{K : K \geq 0\}$ are traded at time* 0*, and the prices, $C(K)$, are assumed to be known. In addition, we suppose call-put parity holds, so that the price of a put option with strike K is $P(K) = e^{-\int_0^T r_s \,\mathrm{d}s} K - S_0 + C(K)$. We make the additional assumptions that $C(K)$ is a continuous, decreasing and convex function, with $C(0) = S_0$, $C'_+(0) = -e^{-\int_0^T r_s \,\mathrm{d}s}$ and $C(K) \to 0$ as $K \to \infty$.*

Many of these notions can be motivated by arbitrage concerns (see e.g. Cox and Obłój (2011b)). That there are plausible situations in which these assumptions do not hold can be seen by considering models with bubbles (e.g.

Cox and Hobson (2005)), in which call-put parity fails, and $C(K) \not\to 0$ as $K \to \infty$. Let us define $B_t = e^{\int_0^t r_s \, \mathrm{d}s}$, and make the assumptions above. Following the perspective that the prices correspond to expectations under $\mathbb{Q}$, the implied law of $B_T^{-1} S_T$ (which we will denote μ) can be recovered by the Breeden-Litzenberger formula (Breeden and Litzenberger, 1978):

$$\mu((K,\infty)) = \mathbb{Q}^*(B_T^{-1} S_T \in (K,\infty)) = -2 B_T C'_+(B_T K). \tag{6.2}$$

Here we have used $\mathbb{Q}^*$ to emphasise the fact that this is only an *implied* probability, and not necessarily the distribution under the actual measure $\mathbb{Q}$. From (6.2) we deduce that $\mathrm{U}_\mu(x) = S_0 - 2C(B_T x) - x$, giving an affine mapping between the function $\mathrm{U}_\mu(x)$ and the call prices. We do not impose the condition that the law of $B_T^{-1} S_T$ under $\mathbb{Q}$ is μ, we merely note that this is the law implied by the traded options. We also do not assume anything about the price paths of the call options: our only assumptions are their initial prices, and that they return the usual payoff at maturity. It can now also be seen that the assumption that $C'_+(0) = -B_T^{-1}$ is equivalent to assuming that there is no atom at 0 — i.e. μ is supported on $(0,\infty)$. Finally, it follows from the assumptions that μ is an integrable measure with mean S_0.

Our goal is to now to use the knowledge of the call prices to find a lower bound on the price of an option which has payoff

$$F\left(\int_0^T \sigma_t^2 \, \mathrm{d}t\right) = F\left(\langle \ln S \rangle_T\right).$$

Consider the discounted stock price:

$$X_t = e^{-\int_0^t r_s \, \mathrm{d}s} S_t = B_t^{-1} S_t.$$

Under Assumption 6.1, X_t satisfies the SDE:

$$\mathrm{d}X_t = X_t \sigma_t \, \mathrm{d}W_t.$$

Defining a time change $\tau_t = \int_0^t \sigma_s^2 \, \mathrm{d}s$, and writing A_t for the right-continuous inverse, so that $\tau_{A_t} = t$, we note that $\widetilde{W}_t = \int_0^{A_t} \sigma_s \, \mathrm{d}W_s$ is a Brownian motion with respect to the filtration $\widetilde{\mathcal{F}}_t = \mathcal{F}_{A_t}$, and if we set $\widetilde{X}_t = X_{A_t}$, we have:

$$\mathrm{d}\widetilde{X}_t = \widetilde{X}_t \, \mathrm{d}\widetilde{W}_t.$$

In particular, $\widetilde{X}_t$ is now of a form where we may apply our earlier results, using the target distribution arising from (6.2), and noting also that $\widetilde{X}_0 = S_0$ and $\widetilde{X}_{\tau_T} = X_T = B_T^{-1} S_T$.

We now define functions as in Section 5, so that $f(t) = F'_+(t)$ and (5.2)–(5.6) hold. Our aim is to use (5.7), which now reads:

$$G(X_{A_t}, t) + H(X_{A_t}) = G(\widetilde{X}_t, t) + H(\widetilde{X}_t) \le F(t) = F\left(\int_0^{A_t} \sigma_s^2 \, \mathrm{d}s\right), \tag{6.3}$$

to construct a sub-replicating portfolio. We shall first show that we can construct a trading strategy that sub-replicates the $G(\widetilde{X}_t, t)$ portion of the portfolio.

Then we argue that we are able, using a portfolio of calls, puts, cash and the underlying, to replicate the payoff $H(X_T)$.

Since $G(\widetilde{X}_t, t)$ is a submartingale, we do not expect to be able to replicate this in a completely self-financing manner. However, by the Doob-Meyer Decomposition Theorem, and the Martingale Representation Theorem, we can certainly find some process $\widetilde{\phi}_t$ such that:

$$G(\widetilde{X}_t, t) \geq G(\widetilde{X}_0, 0) + \int_0^t \widetilde{\phi}_s \, \mathrm{d}\widetilde{X}_s$$

and such that there is equality at $t = \tau_D$. Moreover, since $G(\widetilde{X}_{\tau_D \wedge t}, \tau_D \wedge t)$ is a martingale, and G is $C^{2,1}$ in D, we have:

$$G(\widetilde{X}_{\tau_D \wedge t}, \tau_D \wedge t) = G(\widetilde{X}_0, 0) + \int_0^{\tau_D \wedge t} \frac{\partial G}{\partial x}(\widetilde{X}_{\tau_D \wedge s}, \tau_D \wedge s) \, \mathrm{d}\widetilde{X}_s.$$

More generally, we would not expect $\frac{\partial G}{\partial x}$ to exist everywhere in $D^\complement$, however, if for example left and right derivatives exist, then we could choose $\widetilde{\phi}_t \in \left[\frac{\partial G}{\partial x}(x-, t), \frac{\partial G}{\partial x}(x+, t)\right]$ as our holding of the risky asset (or alternatively, but less explicitly, take $\widetilde{\phi}_t = \partial/\partial x \left[\mathbb{E}^{x,t} G(\widetilde{X}_{t+\delta}, t_0 + \delta)\right]$, for $t \in [t_0, t_0 + \delta]$).

It follows that we can identify a process $\widetilde{\phi}_t$ with

$$G(\widetilde{X}_{\tau_t}, \tau_t) \geq G(\widetilde{X}_0, 0) + \int_0^{\tau_t} \widetilde{\phi}_s \, \mathrm{d}\widetilde{X}_s = G(X_0, 0) + \int_0^t \widetilde{\phi}_{\tau_s} \, \mathrm{d}X_s$$

where we have used e.g. Revuz and Yor (1999, Proposition V.1.4). Finally, writing $\phi_s = \widetilde{\phi}_{\tau_s}$, we have:

$$G(X_t, \tau_t) \geq G(X_0, 0) + \int_0^t \phi_s \, \mathrm{d}X_s = G(X_0, 0) + \int_0^t \phi_s \, \mathrm{d}\left(B_s^{-1} S_s\right).$$

If we consider the self-financing portfolio which consists of holding $\phi_s B_T^{-1}$ units of the risky asset, and an initial investment of $G(X_0, 0) B_T^{-1} - \phi_0 S_0 B_T^{-1}$ in the risk-free asset, this has value V_t at time t, where

$$\mathrm{d}\left(B_t^{-1} V_t\right) = B_T^{-1} \phi_t \, \mathrm{d}\left(B_t^{-1} S_t\right),$$

and therefore

$$V_T = B_T \left(V_0 B_0^{-1} + \int_0^T B_T^{-1} \phi_s \, \mathrm{d}\left(B_s^{-1} S_s\right) \right) = G(X_0, 0) + \int_0^T \phi_s \, \mathrm{d}X_s.$$

We now turn to the $H(X_T)$ component in (6.3). If $H(x)$ can be written as the difference of two convex functions (so in particular, $H''(\mathrm{d}K)$ is a well defined signed measure) we can write:

$$H(x) = H(S_0) + H'_+(S_0)(x - S_0) + \int_{(S_0, \infty)} (x - K)_+ \, H''(\mathrm{d}K) + \int_{(0, S_0]} (K - x)_+ \, H''(\mathrm{d}K).$$

Taking $x = X_T = B_T^{-1} S_T$ we get:

$$H(X_T) = H(S_0) + H'_+(S_0)\left(B_T^{-1} S_T - S_0\right) + B_T^{-1} \int_{(S_0,\infty)} (S_T - B_T K)_+ \, H''(\mathrm{d}K) + B_T^{-1} \int_{(0,S_0]} (B_T K - S_T)_+ \, H''(\mathrm{d}K).$$

This implies that the payoff $H(X_T)$ can be replicated at time T by 'holding' a portfolio of:

$$\begin{aligned} &B_T^{-1}\left(H(S_0) - H'_+(S_0) S_0\right) \text{ in cash;} \\ &B_T^{-1} H'_+(S_0) \text{ units of the asset;} \\ &B_T^{-1}\, H''(\mathrm{d}K) \text{ units of the call with strike } B_T K \text{ for } K \in (S_0, \infty); \\ &B_T^{-1}\, H''(\mathrm{d}K) \text{ units of the put with strike } B_T K \text{ for } K \in (0, S_0]; \end{aligned} \tag{6.4}$$

where the final two terms should be interpreted appropriately. In practice, the function $H(\cdot)$ can typically be approximated by a piecewise linear function, where the 'kinks' in the function correspond to traded strikes of calls or puts, in which case the number of units of each option to hold is determined by the change in the gradient at the relevant strike. The initial cost of setting up such a portfolio is well defined provided

$$\int_{(0,S_0]} P(B_T K)\, |H''|(\mathrm{d}K) + \int_{(S_0,\infty)} C(B_T K)\, |H''|(\mathrm{d}K) < \infty, \tag{6.5}$$

where $|H''|(\mathrm{d}K)$ is the total variation of the signed measure $H''(\mathrm{d}K)$. We therefore shall make the following assumption:

Assumption 6.3. *The payoff $H(X_T)$ can be replicated using a suitable portfolio of call and put options, cash and the underlying, with a finite price at time 0.*

We can therefore combine these to get the following theorem:

Theorem 6.4. *Suppose that Assumptions 6.1, 6.2 and 6.3 hold, and suppose $F(\cdot)$ is a convex, increasing function with $F(0) = 0$ and right derivative $f(t) = F'_+(t)$ which is bounded. Then there exists an arbitrage if the price of an option with payoff $F(\langle \ln S \rangle_T)$ is less than:*

$$B_T^{-1} G(S_0, 0) + B_T^{-1} H(S_0) + B_T^{-1} \int_{(S_0,\infty)} C(B_T K)\, H''(\mathrm{d}K) + B_T^{-1} \int_{(0,S_0]} P(B_T K)\, H''(\mathrm{d}K), \tag{6.6}$$

where the functions G and H are as defined in (5.5) *and* (5.6)*, and are determined by the solution τ_D to $\mathbf{SEP}(\sigma, \delta_{S_0}, \mu)$ for $\sigma(x) = x$, and where μ is determined by* (6.2)*.*

Moreover, this bound is optimal in the sense that there exists a model which is free of arbitrage, under which the bound can be attained.

Proof. It follows from Theorem 4.6 that, given μ, we can find a domain D and corresponding stopping time τ_D which solves $\mathbf{SEP}(\sigma, \delta_{S_0}, \mu)$. Applying

Proposition 5.1 (and bearing in mind Remark 5.5), we conclude that the strategy described above will indeed sub-replicate, and we can therefore produce an arbitrage by purchasing the option, and selling short the portfolio of calls, puts and the underlying given in (6.4), and in addition, holding the dynamic portfolio with $-\phi_t B_T^{-1}$ units of the underlying at time t. It is not hard to check, given that f is bounded (and choosing the lower limits in (5.4) to be S_0 rather than 0) that $(Z'(\widetilde{X}_s)\sigma(\widetilde{X}_s))^2 \leq (\widetilde{X}_s/\widetilde{X}_0 - 1)^2$, and hence that (5.8) holds. The condition (5.3) also clearly holds. As a consequence, we do indeed have a subhedge.

To see that this is the best possible bound, we need to show that there is a model which satisfies Assumption 6.1, has law μ under $\mathbb{Q}$ at time T, and such that the subhedge is actually a hedge. But consider the stopping time τ_D for the process $\widetilde{X}_t$. Define the process

$$X_t = \widetilde{X}_{\frac{t}{T-t} \wedge \tau_D} \quad \text{for } t \in [0, T]$$

which corresponds to the choice of $\sigma_s^2 = \frac{T-s+1}{(T-t)^2} \mathbf{1}_{\{\frac{s}{T-s} < \tau_D\}}$. Since $\tau_D < \infty$ a.s., then $X_T = \widetilde{X}_{\tau_D}$, $\tau_T = \tau_D$ and $S_t = X_t B_t$ is a price process satisfying Assumption 6.1 with

$$F\left(\int_0^T \sigma_t^2 \, \mathrm{d}t\right) = F(\tau_D).$$

Finally, it follows from (5.12) that at time T, the value of the hedging portfolio exactly equals the payoff of the option. □

Remark 6.5. The above results are given in the context of an increasing, convex function, but there is also a similar result concerning increasing, concave functions which can be derived. Consider a bounded, increasing function f as before, and define the function

$$L(t) = \int_0^t (f(\infty) - f(s)) \, \mathrm{d}s = f(\infty)t - F(t).$$

Using Theorem 6.4 and (1.2), it is easy to see that the price of a contract with payoff $L(\langle \ln S \rangle_T)$ must be bounded above by:

$$\begin{aligned} 2f(\infty)Q - &2f(\infty)B_T^{-1}\log(S_0) - B_T^{-1}G(S_0, 0) - B_T^{-1}H(S_0) \\ &- B_T^{-1}\int_{(S_0,\infty)} C(B_T K)\, H''(\mathrm{d}K) - B_T^{-1}\int_{(0,S_0]} P(B_T K)\, H''(\mathrm{d}K), \end{aligned}$$

where Q is the price of a log-contract (that is, an option with payoff $\ln(S_T)$). As before, this upper bound is the best possible, under a similar set of assumptions.

Remark 6.6. An analogous result can be shown for *forward start* options. Suppose that the option has payoff

$$F\left(\int_S^T \sigma_t^2 \, \mathrm{d}t\right) = F\left(\langle S \rangle_T - \langle S \rangle_S\right)$$

for fixed times $0 < S < T$. Then we can use the previous results for general starting distributions to deduce a similar result to Theorem 6.4 for forward start options, provided we assume that there are calls traded at both S and T.

We use essentially the same idea as above: we aim to hold a portfolio which (sub-)replicates $G(X_t, \tau_t)$ for $t \in [S,T]$, and hold the payoff $H(X_T)$ as a portfolio of calls. However, we now have $\tau_t = \int_S^t \sigma_s^2 \, \mathrm{d}s$, and so $\widetilde{X}_t = X_{A_t}$, gives $\widetilde{X}_0 = X_S$ (recall that A_t was assumed right-continuous). The procedure is much as above, except that we need to use the solution to Theorem 5.3 with a general target distribution, and the amount $G(\widetilde{X}_0, 0)$ will be a $\mathcal{F}_S$-random variable. The initial distribution ν can be derived using the Breeden-Litzenberger formula (6.2) at time S. To ensure that we hold the amount $G(\widetilde{X}_0, 0)$ at time S, we observe that $G(\widetilde{X}_0, 0) = G(X_S, 0)$. Hence, if e.g. $G(x, 0)$ can be written as the difference of two convex functions, we can replicate this amount by holding a portfolio of calls and puts with maturity S in a similar manner to (6.4). The remaining details follow as in the hedge described in Theorem 6.4

Remark 6.7. We can also consider modifications to the realised variance. Consider a slightly different time-change: suppose we set

$$\tau_t = \int_0^t \sigma_s^2 \lambda(X_s) \, \mathrm{d}s,$$

for some 'nice' function $\lambda(x)$, which in particular we suppose is bounded above and below by positive constants. Then following the computations above, we see that

$$\widetilde{X}_t = X_{A_t} = \int_0^{A_t} X_s \lambda(X_s)^{-1/2} \left(\sigma_s \lambda(X_s)^{1/2} \, \mathrm{d}W_s \right) = \int_0^t X_{A_s} \lambda(X_{A_s})^{-1/2} \, \mathrm{d}\widetilde{W}_s,$$

and therefore $\mathrm{d}\widetilde{X}_t = \sigma(\widetilde{X}_t) \, \mathrm{d}\widetilde{W}_t$, where $\sigma(x) = x\lambda(x)^{-1/2}$. It seems feasible (Theorem 4.6 would need to be extended, but for 'nice' λ, this should be possible) that the above arguments could then be extended to provide robust hedges on convex payoffs of the form:

$$F\left(\int_0^T \sigma_s^2 \lambda(X_s) \, \mathrm{d}s \right).$$

An interesting special case of this would then be to give robust bounds on the price of an option on corridor variance:

$$F\left(\int_0^T \sigma_s^2 \mathbf{1}_{\{S_s \in [a,b]\}} \, \mathrm{d}s \right), \tag{6.7}$$

by considering $\lambda(x) = \mathbf{1}_{\{x \in [a,b]\}}$, however this would only work in the case where there are no discount rates (i.e. $B_t = 1$). In general, we can only give a tight lower bound for options on:

$$F\left(\int_0^T \sigma_s^2 \mathbf{1}_{\{X_s \in [\widetilde{a},\widetilde{b}]\}} \, \mathrm{d}s \right),$$

although this does provide a lower bound for (6.7) by considering the case where $\widetilde{a} = a$ and $\widetilde{b} = B_T b$.

7 Conclusions

We conclude by summarising the results, and describing some interesting questions for future work. In this paper, we have given a variational inequality representation of Root's solution to the Skorokhod embedding problem, and provided a novel proof of optimality, which allows us to construct a model-independent subhedge for options on variance. We believe that our results provide interesting insights into all three aspects of the work: the construction of solutions to the Skorokhod embedding problem, proving optimality results for the same, and finally the connections with model-independent hedging.

We also believe that there are interesting lines of research that now arise. The construction opens up a number of questions regarding Root's solution to the Skorokhod embedding problem: for example, what can be said about the shape of the boundary? Under what conditions on μ will the boundary be smooth? When does $R(x) \to 0$ as $x \to \pm\infty$? When is $R(x)$ bounded? Properties of free boundaries are well-studied in the analytic literature, and may be useful in answering these questions. The connection to minimality and non-centred target distributions raised in Remark 4.5, and the question asked at the end of this remark would also be interesting lines for research.

The connection with optimal stopping noted in Remark 4.4 is interesting, and obtaining a deeper understanding between optimal stopping problems and optimal Skorokhod embeddings seems to be an interesting area of research.

Another natural question concerns the upper bound/super-hedging strategy. It has been remarked by Obłój (2004), and Carr and Lee (2010) that a related construction of Rost should provide a suitable upper bound, but similar questions to those answered here remain (although we hope to be able to provide some answers in subsequent work). We note however that numerical evidence seems to suggest that the Root bounds may be more appropriate in the financial applications. It would also be of interest to see to what extent these model-independent bounds may be useful in practice. In Cox and Obłój (2011b), an analysis of the use of model-independent bounds as a hedging strategy for barrier options was performed. A similar analysis of the strategies derived in this work would also be of interest.

Other questions that arise from the practical standpoint include how to incorporate additional market information (e.g. calls at an intermediate time (Brown et al., 2001b)), and how to adjust for the fact that there will generally only be a finite set of quoted calls (see (Davis et al., 2010) for a related question). Remark 6.7 also suggests open questions regarding more general choices of $\sigma(x)$.

Acknowledgements

We are grateful to Sam Howison for a helpful discussion which has much improved the material in Section 3 and 4.

References

A. Bensoussan and J.-L. Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1982.

D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978.

M. Broadie and A. Jain. Pricing and hedging volatility derivatives. *The Journal of Derivatives*, 15(3):7–24, 2008.

H. Brown, D. Hobson, and L. C. G. Rogers. Robust hedging of barrier options. *Math. Finance*, 11(3):285–314, 2001a.

H. Brown, D. Hobson, and L. C. G. Rogers. The maximum maximum of a martingale constrained by an intermediate law. *Probab. Theory Related Fields*, 119(4):558–578, 2001b.

P. Carr and R. Lee. Hedging variance options on continuous semimartingales. *Finance and Stochastics*, 14:179–207, 2010.

P. Carr, R. Lee, and L. Wu. Variance swaps on time-changed lévy processes. *Finance and Stochastics*, 2011. doi: 10.1007/s00780-011-0157-9. *To Appear.*

A. M. G. Cox. Extending Chacon-Walsh: minimality and generalised starting distributions. In C. Donati-Martin et al., editor, *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 233–264. Springer, Berlin, 2008.

A. M. G. Cox and D. G. Hobson. Local martingales, bubbles and option prices. *Finance Stoch.*, 9(4):477–492, 2005.

A. M. G. Cox and J. Obłój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 2011a. doi: 10.1007/s00780-011-0154-z. *To Appear.*

A. M. G. Cox and J. Obłój. Robust hedging of double touch barrier options. *SIAM Journal on Financial Mathematics*, 2:141–182, 2011b.

A. M. G. Cox, D. G. Hobson, and J. Obloj. Pathwise inequalities for local time: applications to Skorokhod embeddings and optimal stopping. *Ann. Appl. Probab.*, 18(5):1870–1896, 2008.

M. H. A. Davis, J. Obłój, and V. Raval. Arbitrage bounds for weighted variance swap prices. `http://arxiv.org/abs/1001.2678`, 2010.

H. Dinges. Stopping sequences. In *Séminaire de Probabilitiés, VIII (Univ. Strasbourg, année universitaire 1972–1973)*, pages 27–36. Lecture Notes in Math., Vol. 381. Springer, Berlin, 1974. Journées de la Sociéte Mathématique de France de Probabilités, Strasbourg, 25 Mai 1973.

B. Dupire. Model art. *Risk*, 6(9):118–120, 1993.

B. Dupire. Arbitrage bounds for volatility derivatives as free boundary problem. Presentation at 'PDE and Mathematical Finance', KTH, Stockholm, 2005.

A. Friedman. *Generalized functions and partial differential equations*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1963.

D. Hobson. The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices. In R.A. Carmona, E. Çinlar, I. Ekeland, E. Jouini, J.A. Scheinkman, and N. Touzi, editors, *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, 2010.

S. Howison, A. Rafailidis, and H. Rasmussen. On the pricing and hedging of volatility derivatives. *Applied Mathematical Finance*, 11(4):317, 2004.

J. Kallsen, J. Muhle-Karbe, and M. Voß. Pricing options on variance in affine stochastic volatility models. *Mathematical Finance*, 2010. doi: 10.1111/j.1467-9965.2010.00447.x. *To Appear.*

M. Keller-Ressel. Convex order properties of discrete realized variance and applications to variance options. `http://arxiv.org/abs/1103.2310`, 2011.

M. Keller-Ressel and J. Muhle-Karbe. Asymptotic and Exact Pricing of Options on Variance. `http://arxiv.org/abs/1003.5514`, 2010.

R. M. Loynes. Stopping times on Brownian motion: Some properties of Root's construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16:211–218, 1970.

I Monroe. On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.*, 43:1293–1311, 1972.

A. Neuberger. The log contract. *The Journal of Portfolio Management*, 20(2): 74–80, 1994.

J. Obłój. The maximality principle revisited: on certain optimal stopping problems. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 309–328. Springer, Berlin, 2007.

J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390 (electronic), 2004.

G. Peskir. Optimal stopping of the maximum process: the maximality principle. *Ann. Probab.*, 26(4):1614–1640, 1998.

P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, third edition, 1999.

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

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

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

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

D. W. Stroock. *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.