Optimal robust bounds for variance options

Alexander M. G. Cox* Jiajie Wang†

August 19, 2013

Abstract

Robust, or model-independent properties of the variance swap are well-known, and date back to Dupire [19] and Neuberger [37], who showed that, given the price of co-terminal call options, the price of a variance swap was exactly specified under the assumption that the price process is continuous. In Cox and Wang [11] we showed that a lower bound on the price of a variance call could be established using a solution to the Skorokhod embedding problem due to Root [45]. In this paper, we provide a construction, and a proof of optimality of the upper bound, using results of Rost [46] and Chacon [9], and show how this proof can be used to determine a super-hedging strategy which is model-independent. In addition, we outline how the hedging strategy may be computed numerically. Using these methods, we also show that the Heston-Nandi model is 'asymptotically extreme' in the sense that, for large maturities, the Heston-Nandi model gives prices for variance call options which are approximately the lowest values consistent with the same call price data.

1 Introduction

The classical approach to derivative pricing problems is to hypothesise a certain model for the underlying asset, and to invoke the Fundamental Theorem of Asset Pricing to identify arbitrage-free prices with discounted expectations under risk-neutral measures. In this setting, market information, for example in the form of traded 'vanilla' options, may be incorporated by choosing a parametrised class of models, and determining the parameters by finding the best fit to the observed prices.

An alternative approach to incorporating market information is to use the traded options as part of a hedging strategy. This approach can be particularly beneficial in the presence of model-risk, since, if carefully chosen, the hedging properties of the strategy may still hold under a wide class of models. The archetypal example of this is the hedging of a variance swap using a log-contract due to Dupire [19] and Neuberger [37]. Suppose a (discounted) asset price has dynamics under the risk-neutral measure:

$$\frac{\mathrm{d}S_t}{S_t} = \sigma_t \ \mathrm{d}W_t, \tag{1.1}$$

*Department of Mathematical Sciences, University of Bath, Bath, U. K.. e-mail: a.m.g.cox@bath.ac.uk, web: http://www.maths.bath.ac.uk/∼mapamgc/

†Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza, Università di Roma 'La Sapienza', Italy. e-mail: jiajie.wang@uniroma1.it

where the process σ_t is not necessarily known. A variance swap is a contract where the payoff depends on the realised quadratic variation of the log-price process: $\langle \ln S \rangle_T = \int_0^T \sigma_t^2 \, \mathrm{d}t$. Dupire and Neuberger observed that

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

from which we conclude that

$$\int_0^T \sigma_t^2 \, \mathrm{d}t = 2\ln(S_0) - 2\ln(S_T) + 2\int \frac{1}{S_t} \, \mathrm{d}S_t.$$

It follows that one can replicate (up to a constant) the payoff of a variance swap by shorting two log-contracts (that is, contracts which pay $\ln(S_T)$ at maturity), and dynamically trading in the asset so as to always hold $2/S_t$ units of the asset. The resulting portfolio will hedge the variance swap under essentially *any* model of the form (1.1) (subject to very mild measurability and integrability conditions on σ_t).

It follows that the price of a variance swap is essentially determined once one observes the price of a log-contract. In practice, call options are more liquidly traded, but the log contract can be statically replicated given a continuum of call options, and so this information is more commonly used to determine the price.

In this paper, we consider what can be said about the price of options which pay the holder a function of the realised variance in the presence of co-terminal call options. Two important examples of these contracts are the variance call, which has payoff $(\langle \ln S \rangle_T - K)_+$ and the volatility swap, which has payoff $\sqrt{\langle \ln S \rangle_T} - K$. Unlike in the case of the variance swap, it is no longer possible to provide a trading strategy which exactly replicates the payoff in any model, however, we are able to provide both super- and sub-hedging strategies which work for a large class of models, and therefore provide model-independent bounds on the prices of such options. Our methods rely on techniques from the theory of Skorokhod embeddings, and in particular, we need a novel proof of optimality of some existing constructions. In one direction, the bounds have been established in a preceding paper, Cox and Wang [11], and some of the methods used in this paper for the other direction follow similar approaches, although there remain substantial technical differences.

In addition to establishing theoretical bounds on the prices of these options, we provide some numerical investigation of the bounds we obtain, and a justification of our numerical techniques. Finally, we are also able to establish, using the optimality techniques of [11], a result relating to the extremality of the Heston-Nandi model (the classical stochastic volatility model of Heston, where the volatility and asset processes are perfectly anti-correlated). Essentially, we show that for large maturities, under the Heston-Nandi model, the price of a variance call option will closely approximate the lowest price possible in the class of models which produce identical call prices. Given that the Heston model is commonly used for pricing options on variance, and that calibration can often lead to values of the correlation parameter close to -1, this suggests that using this model for pricing variance options amounts to taking a strong 'bet' on which model most accurately reflects reality.

The theme of model-independent, or robust, pricing is one that has received a great deal of attention in recent years. The approach we take in this paper can be traced back to Hobson [28], and more recent work, closely related to variance options, includes Dupire [18], Carr and Lee [8], Davis, Obłój, and Raval [16], Cox and Wang [11] and Oberhauser and dos Reis [39]. In addition, Hobson and Klimmek [29] consider variance swaps where the model may include jumps — a case which we exclude. An alternative, related approach is based on the *uncertain volatility* models of Avellaneda, Levy, and Parás [1]. Recent papers which take this approach include Galichon, Henry-Labordere, and Touzi [23], Possamaï, Royer, and Touzi [42] and Neufeld and Nutz [38]. We explain how our results may be interpreted in this framework in Remark 4.8. Other recent connected work in this direction includes Beiglböck, Henry-Labordère, and Penkner [5] and Dolinsky and Soner [17], where connections with optimal transport are established.

We proceed as follows: in Section 2 we introduce our financial setup, and explain why the financial problem of interest can be related to the Skorokhod embedding problem. This motivates Section 3, where we provide a characterisation of the solution of Rost [46] and Chacon [9] to the Skorokhod embedding problem. In Section 4 we give a novel proof of the optimality of these barriers, and explain how these constructions may be used to derive superhedges for certain options. In Section 5 we show how the solutions may be computed numerically, and provide some graphical evidence of the behaviour of the hedging strategies. In Section 6 we prove our optimality result for the Heston-Nandi model.

2 Financial Motivation

To motivate our financial models, we begin with a fairly classical setup: we suppose that there is a market which consists of a traded asset, with price S_t defined on a probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, satisfying the usual conditions, with:

$$\frac{dS_t}{S_t} = r_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t \tag{2.1}$$

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, progressively measurable processes so that the integral in (2.1) 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, at least initially, suppose:

Assumption 2.1. *The asset price process, under some probability measure $\mathbb{Q} \sim \mathbb{P}$, is the solution to* (2.1)*, where r_t and σ_t are locally bounded, predictable processes.*

We shall later see that some relaxation of this condition is possible, using pathwise approaches to stochastic integration.

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

Assumption 2.2. *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 [14]). That there are plausible situations in which these assumptions do not hold can be seen by considering models with bubbles (e.g. [13]), 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. Since (classically) 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 [7]:

$$\mu((K, \infty)) = \mathbb{Q}^*(B_T^{-1} S_T \in (K, \infty)) = -B_T C'_+(B_T K). \tag{2.2}$$

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 $\mathbb{Q}$. 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. $\mu((0, \infty)) = 1$. In general, the equality here could be replaced with an inequality ($C'_+(0) \geq -B_T^{-1}$) if one wished to consider models with a positive probability of the asset being worthless at time T. 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. 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 or upper bound on the price of an option which has payoff

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

The term $\int_0^T \sigma_t^2 \,\mathrm{d}t$ is commonly referred to as the *realised variance*. There are a number of pertinent examples which motivate us: the most common case, and where the answer is well known, is the case of a variance swap, where the payoff of the option is $(\langle \ln S \rangle_T - K)$. An obvious modification of this is the *variance call* which has payoff $(\langle \ln S \rangle_T - K)_+$, and the corresponding variance put. In addition, *volatility swaps* are traded, where the payoff is $\sqrt{\langle \ln S \rangle_T} - K$. As well as options written on the realised variance, there are classes of options which trade on various forms of weighted realised variance: define

$$RV_T^w = \int_0^T w(S_t) \,\mathrm{d} \langle \ln S \rangle_t = \int_0^T w(S_t) \sigma_t^2 \,\mathrm{d}t$$

then many of the above options can be recast in terms of their weighted versions. Common examples of these include options on corridor variance, where $w(x) = \mathbf{1}_{\{x \in [a,b]\}}$, and the gamma swap [33], where $w(x) = x$. In fact, for a simplified exposition, we will assume that the weight depends not on the spot price, but rather on the discounted spot price (or equivalently, the forward price). In the case of most interest, where $w(x) = 1$, this makes no difference, as it would, for example in an equity setting where the dividend yield and the interest rate were the same; we also refer to Lee [34], where it is indicated how such an approximation may be accounted for using a model-independent hedge involving calls of *all* maturities, although this is beyond the general methodology described in the article, where we will generally assume that only calls of one maturity are observed.

Note we assume that our underlying price process, S_t, has continuous paths. This is an important assumption, and our conclusions will not generally hold otherwise. Hobson and Klimmek [29] consider related questions in the case where the underlying asset may jump. We also assume that the payoff of the option is exactly the realised quadratic variation, wheras in reality, financial contracts will be written on a discretised version of the quadratic variation (for example, the sum of squared daily log-returns); the effects of this approximation are also considered by, for example, Hobson and Klimmek [29].

Our approach is motivated by the following heuristics. 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 2.1, X_t satisfies the SDE:

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

Let $\lambda(x)$ be a strictly positive, continuous function, and define a time change $\tau_t = \int_0^t \lambda(X_s)\sigma_s^2 \, \mathrm{d}s$. Writing A_t for the right-continuous inverse, so that $\tau_{A_t} = t$, we note that $\widetilde{W}_t = \int_0^{A_t} \sigma_s \lambda(X_s)^{1/2} \, \mathrm{d}W_s$ 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:

$$\mathrm{d}\widetilde{X}_t = \widetilde{X}_t \lambda(\widetilde{X}_t)^{-1/2} \, \mathrm{d}\widetilde{W}_t. \tag{2.3}$$

In particular, under mild assumptions on λ, $\widetilde{X}_t$ is now a diffusion on natural scale, and we note also that $\widetilde{X}_0 = S_0$ and $\widetilde{X}_{\tau_T} = X_T = B_T^{-1} S_T$. It follows that $(\widetilde{X}_{\tau_T}, \tau_T) = (B_T^{-1} S_T, \int_0^T \lambda(X_s)\sigma_s^2 \, \mathrm{d}s)$, and therefore that (2.2), which implies knowledge of the law of $B_T^{-1} S_T$, also tells us the law of $\widetilde{X}_{\tau_T}$. The key observation is that there is now a correspondence between the possible joint laws of a stopped diffusion and its stopping time, and the joint laws of the (discounted) asset price at a fixed time, and the weighted realised variance at that time. Since we wish to find the extremal possible prices of options whose payoff is $F\left(\int_0^T w(X_t)\sigma_t^2 \, \mathrm{d}t\right)$, if we take $\lambda(x) \equiv w(x)$, the problem would appear to be equivalent to that of finding a stopping time which maximises or minimises $\mathbb{E}F(\tau)$ subject to $\mathcal{L}(\widetilde{X}_\tau) = \mu$, where μ is the law of $B_T^{-1} S_T$ inferred by the market call prices. The general problem of finding a stopping time for a process which has a given distribution is known as the Skorokhod Embedding problem, and solutions with given optimality properties have been well studied in recent years [2, 28, 12, 10] — for a survey of these results, we refer the reader to Hobson [27]. In Cox and Wang [11], we established that a construction of a Skorokhod embedding due to Root [45] corresponded to minimising payoffs of the above form where $F(\cdot)$ is a convex increasing function. In this paper, we show that a related construction, which can be traced back to work of Rost [46], and Chacon

[9], maximises such payoffs. In addition, we shall show how these constructions may be used to derive trading strategies, which will super- or sub-hedge in any of the models under consideration, and are a hedge in the extremal model.

Finally, we will also briefly consider a similar problem where the option is *forward starting* — so the payoff depends on the realised variance accumulated between future dates T_0 and T_1. In this case, we assume that we observe call prices at times T_0 and T_1, which correspond to distributions at times T_0 and T_1. The implied distribution at time T_0, ν say, can be interpreted as the law of X_{T_0}. Taking $\widetilde{X}_t = X_{T_0+\tau_t}$, we get a problem similar to above, but with $\widetilde{X}_0 \sim \nu$ for some new probability measure ν. This setting will be considered further in Remark 4.6.

3 Construction of Rost's Barrier

3.1 Background

In this section, we recall some important results due to Rost and Chacon concerning the construction of solutions to the Skorokhod embedding problem. These results are established under fairly general assumptions on the underlying process $\widetilde{X}_t$. One of the main goals of this section is to provide conditions under which we can apply their results in our setting. We begin by recalling the notion of a reversed barrier.

Definition 3.1 (Reversed Barrier ([40])). A closed subset B of $[-\infty, +\infty] \times [0, +\infty]$ is a *reversed barrier* if

(i). $(x, 0) \in B$ for all $x \in [-\infty, +\infty]$;

(ii). $(\pm\infty, t) \in B$ for all $t \in [0, \infty]$;

(iii). if $(x, t) \in B$ then $(x, s) \in B$ whenever $s < t$.

Given a reversed barrier, we can construct a stopping time of a process $\widetilde{X}_t$ as $\tau = \inf\{t > 0 : (\widetilde{X}_t, t) \in B\}$. Then it is known that, given a measure μ satisfying certain conditions, there exists a reversed barrier which embeds the law μ, that is, such that $\widetilde{X}_\tau \sim \mu$. Moreover, in the case where $\widetilde{X}_t$ is a diffusion, the reversed barrier has the property that the corresponding stopping time minimises the *capped expectation* $\mathbb{E}[\tau \wedge t]$ over solutions to the Skorokhod embedding problem, for all $t \geq 0$. Using the observation that

$$F(\tau) = F(0) + \tau F'(0) + \int_0^\infty F''(t)\,(\tau - t)_+ \,\mathrm{d}t$$

and $(\tau - t)_+ = \tau - \tau \wedge t$, and the fact that $\mathbb{E}\sigma$ depends only on the law of X_σ for 'nice' stopping times (a point we will elaborate on shortly), then it is immediate that $\mathbb{E}F(\tau)$ is maximised over such stopping times for all convex functions with $F'(0) = 0$.

These observations are essentially due to work of Rost and Chacon. In Rost [46], the notion of a *filling scheme* stopping time was introduced for a general Markov process, and this was shown to embed and have the optimality property described. Later, Chacon [9][1] proved that in many cases, the filling scheme

[1] We note that Chacon calls what we refer to a reversed barrier as a barrier. We follow the terminology established in Obłój [40].

stopping time would actually be almost-surely equal to a stopping time generated by a reversed barrier. More recent results concerning these constructions can be found in [15].

Before proving our results characterising the reversed barriers, we recall some important background. Given a probability distribution μ and a Markov process $\widetilde{X}$, the Skorokhod embedding problem is to find a stopping time τ such that $\widetilde{X}_\tau \sim \mu$. Motivated by the financial setting, we consider the case that

$$\mu \text{ is a probability distribution with } \mu((0,\infty)) = 1, \tag{3.1}$$

and $\widetilde{X}$ is a regular (see Rogers and Williams [44] for terminology relating to one-dimensional diffusions) diffusion on $I = (0,\infty)$, which is a solution to (2.3), with initial distribution $\widetilde{X}_0 \sim \nu$, for some given distribution ν, and a continuous function λ on I which is strictly positive. Since $0 \notin I$, 0 is inaccessible for $\widetilde{X}_t$.

Recall that we wish to include the case of forward-starting options, in which case $\widetilde{X}_0$ is assumed to have the law inferred from the call prices at the start date of the contract. In the theory of Skorokhod embeddings, it is usually natural to restrict to the class of *minimal* stopping times, however since $\widetilde{X}_t$ is transient, any embedding will be minimal. Moreover, for example by considering $\widetilde{X}_t$ as a time change of a Brownian motion stopped on hitting 0, if we restrict ourselves to laws μ and ν which have the same mean, then any embedding of μ must have $(\widetilde{X}_{t\wedge\tau})_{t\geq 0}$ a uniformly integrable (UI) process, and moreover a necessary and sufficient condition for the existence of an embedding is that

$$\mathrm{U}_\nu(x) := -\int_{\mathbb{R}} |y-x|\,\nu(\mathrm{d}y) \geq -\int_{\mathbb{R}} |y-x|\,\mu(\mathrm{d}y) =: \mathrm{U}_\mu(x) > -\infty, \tag{3.2}$$

for all $x \in \mathbb{R}$. By Jensen's inequality, such a constraint is clearly necessary for the existence of an embedding; further, using time-change arguments, and reducing to the Brownian case, it is the only restriction that is required. To understand this notion in the financial setting, note that from (2.2) 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.

For any given reversed barrier, we define $D := (\mathbb{R}\times I)\backslash B$, and then one can show that there exists a unique upper semi-continuous function $R : I \to [0,\infty]$ such that

$$D = \{(x,t) : t > R(x)\} \quad \text{and} \quad B = \{(x,t) : 0 \leq t \leq R(x)\}.$$

Then the stopping time of interest is:

$$\tau_D = \inf\{t > 0 :\, (\widetilde{X}_t, t) \notin D\} = \inf\{t > 0 : t \leq R(\widetilde{X}_t)\},$$

and we will call such an embedding a *Rost stopping time*, or a *Rost embedding*. We may also refer to D as a reversed barrier, with the meaning intended to be inferred from the notation. Note that multiple barriers may solve the same stopping problem: for example, if the target distribution contains atoms, the barrier between atoms may be unspecified away from the starting point, provided it is never beyond the 'spikes' of the atoms.

We can put together the work of Rost and Chacon in the specific case that the underlying process is a regular diffusion on I, satisfying a certain regularity condition needed to ensure Chacon's result holds. Specifically, we introduce the set:

$$\mathcal{D} = \left\{ \lambda \in C(I;\mathbb{R}) : \begin{array}{l} \lambda \text{ is strictly positive and the solution to (2.3) de-} \\ \text{fines a regular diffusion on } I, \text{ with transition den-} \\ \text{sity } p(t,x,y) \text{ with respect to Lebesgue such that,} \\ \text{for any } x_0 \in I,\ c > 0, \text{ open set } A \text{ contain-} \\ \text{ing } x_0 \text{ and } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that} \\ |\left(p(x,x_0,t) - p(x,x_0,s)\right) x_0^2 \lambda(x_0)^{-1}| < \varepsilon \text{ whenever} \\ |s-t| < \delta \text{ and either } x_0 \notin A \text{ or } t > c. \end{array} \right\} \tag{3.3}$$

Then we can prove the following result:

Theorem 3.2. *Suppose μ and ν are probability measures on I such that* (3.2) *holds, and suppose $\widetilde{X}_t$ solves* (2.3) *for some $\lambda \in \mathcal{D}$ with $\widetilde{X}_0 \sim \nu$:*

(i). if μ and ν have disjoint support, then there exists a reversed barrier D such that $\widetilde{X}_{\tau_D} \sim \mu$;

(ii). if μ and ν do not have disjoint support, then (on a possibly enlarged probability space) there exists a random variable $S \in \{0,\infty\}$, and reversed barrier D, such that $\widetilde{X}_{\tau_D \wedge S} \sim \mu$.

Moreover, in both cases, the resulting embedding maximises $\mathbb{E}F(\sigma)$ over all stopping times σ with $\widetilde{X}_\sigma \sim \mu$ and $\mathbb{E}\sigma = \mathbb{E}\tau_D < \infty$, for any convex function F on $[0,\infty)$.

That the condition $\mathbb{E}\sigma = \mathbb{E}\tau_D$ is reasonable can be seen by considering

$$Q(\widetilde{X}_t) := \int_{x_0}^{\widetilde{X}_t} \int_{x_0}^{y} \lambda(z) z^{-2}\, \mathrm{d}z \mathrm{d}y = \int_0^t \int_{x_0}^{x_s} \lambda(y) y^{-2}\, \mathrm{d}y \mathrm{d}\widetilde{X}_s + \frac{1}{2} t + Q(\widetilde{X}_0). \tag{3.4}$$

Noting that $Q(x)$ is convex, we can take expectations along a localising sequence and apply Fatou's Lemma to see that, for any stopping time τ, $\mathbb{E}\tau \geq \mathbb{E}Q(\widetilde{X}_\tau)$, the second term depending only on the law of $\widetilde{X}_\tau$. For well behaved stopping times (specifically, where $\mathbb{E}Q(\widetilde{X}_{\tau\wedge N}) \to \mathbb{E}Q(\widetilde{X}_\tau)$), we would in fact expect equality here.

We also observe that there is a trivial extension of this result when $F(\cdot)$ is a concave function, then $-F(\cdot)$ is a convex function, and so the resulting stopping time minimises $\mathbb{E}F(\sigma)$ over the same class of stopping times.

Proof of Theorem 3.2. We first show that, under the conditions above, there exists a filling scheme stopping time.

Standard time-change arguments, and reduction to the Brownian case show that when μ and ν have the same mean, then (3.2) is both necessary and sufficient for the existence of a Skorokhod embedding which solves the problem. In addition, by Rost [46, Theorem 4], this is sufficient to deduce the existence of a filling scheme stopping time: from the proof of this result, it is clear that whenever an embedding exists, it can be taken as a filling scheme stopping time.

Now consider the case where μ and ν have disjoint support. Then Chacon [9, Theorem 3.24] states that a filling scheme stopping time is a reversed barrier stopping time provided:

(i). $\widetilde{X}_t$ is a standard Markov process, in duality with a standard Markov process $\hat{X}_t$ (we refer the reader to e.g. Blumenthal and Getoor [6, Definition VI.1.2]);

(ii). the transition measures relating to $\widetilde{X}_t$ and $\hat{X}_t$ have densities;

(iii). the transition density $p(x, y, t)$ for $\widetilde{X}_t$ satisfies an equicontinuity property: for any $x_0 \in \mathbb{R}$, $c > 0$ and open set A containing x_0, then given $\varepsilon > 0$, there exists $\delta > 0$ such that $|p(x, x_0, t) - p(x, x_0, s)| < \varepsilon$ whenever $|s - t| < \delta$ and either $x_0 \notin A$ or $t > c$.

Since $\widetilde{X}_t$ is a regular diffusion with inaccessible endpoints, it is its own dual process, with respect to the speed measure (which is $\lambda(x)x^{-2}\mathrm{d}x$), see Fitzsimmons [20, Remark 1.15]. Moreover, under these conditions, the transition density exists (e.g. Rogers and Williams [44, Theorem V.50.11]), and (3.3) guarantees that *(iii)* holds.

Finally, suppose μ and ν are not disjointly supported. Then, using the Hahn-Jordan decomposition, we can find disjoint, non-negative measures ν_0 and μ_0, such that $\nu_0(A) \leq \nu(A)$, $\mu_0(A) \leq \mu(A)$, for all $A \in \mathcal{B}(0, \infty)$, and $\mu - \nu = \mu_0 - \nu_0$. Since μ and ν are not disjoint, μ_0 and ν_0 are non-trivial. We write in addition $\nu \wedge \mu = \nu - \nu_0$, observing that also then $\nu \wedge \mu = \mu - \mu_0$. By enlarging the probability space if necessary, let Z be a uniform random variable on $(0, 1)$, independent of the process, and define the Radon-Nikodym derivative $f = \mathrm{d}(\nu \wedge \mu)/\,\mathrm{d}\nu$. Then if we set

$$S = \begin{cases} 0, & \text{if } Z \leq f(\widetilde{X}_0); \\ \infty, & \text{if } Z > f(\widetilde{X}_0), \end{cases}$$

it follows that $\mathbb{P}(X \in A, S = 0) = (\nu \wedge \mu)(A)$. Now define the normalised measures, $\nu^*(A) = \nu_0(A)/\nu_0((0, \infty))$ and $\mu^*(A) = \mu_0(A)/\mu_0((0, \infty))$, and construct the reversed barriers for the initial distribution ν^* and target distribution μ^*. (It is straightforward to check that $U_{\nu_0}(x) - U_{\mu_0}(x) = U_\nu(x) - U_\mu(x)$, and therefore that the construction is possible.) However, it is now clear that this embeds and is exactly the stopping time described in the statement of the theorem. Moreover, this description of the first step of the construction is exactly the first step described in the construction of the filling scheme by Rost, so it follows that this stopping time is a filling scheme stopping time.

The final statement is Chacon [9, Proposition 2.2]. $\square$

It is clear that the above conditions include the main case of interest — the case where $\lambda(x) = 1$, which is the case corresponding to options on realised variance. For the point 0 to be inaccessible, we require $\int_{0+} \lambda(x)x^{-1}\,\mathrm{d}x = \infty$ [44, Theorem V.51.2]. In principle, this would exclude, for example, the case where $\lambda(x) = x$, the Gamma swap, however in practice, this case could be approximated by taking $\lambda(x) = x \vee \varepsilon$ for some small $\varepsilon > 0$. Note however that in the case where $\lambda(x) = x$, $\widetilde{X}_t$ is a Bessel process of dimension 0, and so in particular, the process will hit zero with positive probability; this is not crucial, since we assume our target measure has no atom of mass at 0, and so we would expect the reversed barrier to stop the process before hitting zero, and therefore the exact behaviour near zero should not affect the barrier substantially.

3.2 Construction of reversed barriers

In this section, we show how the reversed barrier determined in Theorem 3.2 can be constructed. As above, we suppose that we have a time-homogeneous diffusion, $\widetilde{X}_t$, such that:

$$\begin{aligned} \mathrm{d}\widetilde{X}_t &= \sigma(\widetilde{X}_t)\mathrm{d}W_t, \\ \widetilde{X}_0 &\sim \nu, \end{aligned} \tag{3.5}$$

so $\sigma(x) = x\lambda(x)^{-1/2}$. In addition, we suppose that the diffusion coefficient, $\sigma : I \to (0, \infty)$ is a continuously differentiable function such that:

$$x^2\sigma(x)^{-2} \in \mathcal{D}, \quad |\sigma(x)x^{-1}| \text{ and } |\sigma'(x)\sigma(x)x^{-1}| \text{ are bounded on } (0,\infty), \tag{3.6}$$

or equivalently, that $\lambda : I \to (0, \infty)$ is continuously differentiable, and

$$\lambda(x) \in \mathcal{D}, \quad |\lambda(x)^{-1}| \text{ and } |\lambda'(x)\lambda(x)^{-2}x| \text{ are bounded on } (0,\infty). \tag{3.7}$$

Then our general problem is:

SEP$^*(\sigma, \nu, \mu)$: Find an upper-semicontinuous function $R(x)$ such that the domain $D = \{(x,t) : t > R(x)\}$ has $\widetilde{X}_{\tau_D \wedge S} \sim \mu$, where $\widetilde{X}_t$ is given by (3.5), $\tau_D = \inf\{t > 0 : (\widetilde{X}_t, t) \notin D\} = \inf\{t > 0 : t \le R(\widetilde{X}_t)\}$, and $S \in \{0, \infty\}$ is an $\mathcal{F}_0$-measurable random variable such that $\widetilde{X}_0 \sim \nu \wedge \mu$ on $\{S = 0\}$.

Here, the measure $\nu \wedge \mu$ is as defined in the proof of Theorem 3.2. We restrict ourselves to the case where $\mathrm{U}_\nu(x) \ge \mathrm{U}_\mu(x)$. We will also introduce the notation $\bar{\tau}_D = \tau_D \wedge S$.

Then we have the following result:

Theorem 3.3. *Suppose* (3.5) *and* (3.6) *hold. Assume D solves* **SEP**$^*(\sigma,\nu,\mu)$. *Then* $u(x,t) = \mathrm{U}_\mu(x) + \mathbb{E}^\nu \left| x - \widetilde{X}_{t\wedge\bar{\tau}_D} \right|$ *is the unique bounded viscosity solution to:*

$$\frac{\partial u}{\partial t}(x,t) = \left(\frac{\sigma(x)^2}{2} \frac{\partial^2 u}{\partial x^2}(x,t) \right)_+ \tag{3.8a}$$

$$u(0,x) = \mathrm{U}_\mu(x) - \mathrm{U}_\nu(x). \tag{3.8b}$$

Moreover, given the solution u to (3.8)*, a reversed barrier D which solves* **SEP**$^*(\sigma,\nu,\mu)$ *can be recovered by* $D = \{(x,t) : u(x,t) > u(0,t)\}$.

Remark 3.4. In a recent paper, Oberhauser and dos Reis [39] make a very similar observation, and they also use a viscosity solution approach to derive existence and uniqueness in the Rost setting (unlike [11] where a variational inequality-based approach is taken). They also work in the slightly more general setting where the diffusion coefficient σ may depend both on time and space; it seems very likely that the results should extend to this setting, but we observe that the optimality of such a construction is no longer easily determined; given that we are interested in the optimality properties of such processes, we restrict ourselves to the time-homogenous case.

We wish to use standard results on viscosity solutions from Fleming and Soner [21]. The equation (3.8) is really a forward equation, rather than a backward equation, which is the setting in [21]; however we can apply their results to the function $v(x, T-t) = -u(x,t)$, for a fixed $T > 0$. Since T is arbitrary, the extension to an infinite horizon is straightforward.

Proof. We first note that since μ and ν are integrable, the function u is bounded, both $-\mathrm{U}_\mu(x)$ and $h(x,t) = \mathbb{E}^\nu|x - \widetilde{X}_{t\wedge\bar{\tau}_D}|$ are convex, continuous functions, and so their second derivatives (in x) exist as positive measures. Moreover, $-\mathrm{U}_\mu''(x) = 2\mu(\mathrm{d}x)$, and we can also decompose $h_{xx}(x,t)/2$ into two measures μ_t^1 and μ_t^2 defined by:

$$\int f(x)\,\mu_t^1(\mathrm{d}x) = \mathbb{E}^\nu\left[f(\widetilde{X}_t); t < \bar{\tau}_D\right]$$
$$\int f(x)\,\mu_t^2(\mathrm{d}x) = \mathbb{E}^\nu\left[f(\widetilde{X}_{\bar{\tau}_D}); t \geq \bar{\tau}_D\right].$$

In particular, since $\bar{\tau}_D$ embeds μ, we must have $\mu_t^2(A) \leq \mu(A)$ for all A, and $\mu_t^2(A) = \mu(A)$ for all $A \subseteq \{x | t \geq R(x)\}$. In addition, μ_t^1 will be dominated by the transition density of a diffusion (which exists by (3.3)) started with distribution ν, and so will have a density $f^1(x,t)$ with respect to Lebesgue for all $t > 0$, and since D is open, by a slight modification of Cox and Wang [11, Lemma 3.3], it is easily checked that $u(x,t)$ satisfies $\frac{\partial u}{\partial t}(x,t) = \frac{\sigma(x)^2}{2}\frac{\partial^2 u}{\partial x^2}(x,t)$ in D. We observe also that $f^1(x_n,t_n) \to 0$ whenever $(x_n,t_n) \to (x,t) \in B$, since f^1 is dominated by the density of a diffusion with initial law ν, killed if it hits x before t, which also has this property.

The result now follows from Fleming and Soner [21, Proposition V.4.1], when we observe that for $t \leq R(x)$, any smooth function w such that $w - u$ is a local minimum at (x,t) must have $\frac{\partial w}{\partial t} = 0$. For $t < R(x)$ this follows from observing that $\mathbb{E}^\nu\left|x - \widetilde{X}_{t\wedge\bar{\tau}_D}\right|$ is constant in t whenever $t < R(x)$. For $t = R(x)$, we observe that $f^1(x,t)$ can be made arbitrarily small in a ball near (x,t), and $\frac{\partial^2 u}{\partial x^2} = f^1(x,t)$ in D. Hence, for such w, we have $-w_t + \left(\sigma^2(x)w_{xx}\right)_+ \geq 0$. For smooth w such that $w - u$ has a local maximum at (x,t), we first observe that for $t \leq R(x)$, the argument above implies that $u_{xx}(x,t) \leq 0$, and so $w_{xx}(x,t) \leq 0$ also. In addition, by Jensen's inequality, $u(x,t)$ is non-decreasing, so $w_t(x,t) \geq 0$. It follows that $-w_t + \left(\sigma^2(x)w_{xx}\right)_+ \leq 0$.

It follows that u is indeed a bounded viscosity solution. To see that it is unique, we apply Fleming and Soner [21, Theorem V.9.1] to $u(t, \mathrm{e}^y)$ (noting also the comment immediately preceding this proof). It is now routine to check that (3.6) is sufficient to ensure that there is a unique bounded viscosity solution to (3.8).

To conclude that a reversed barrier can be recovered from a solution u to (3.8), we observe that Theorem 3.2 guarantees the existence of a reversed barrier D, and by Jensen's inequality, we see that $D^* = \{(x,t) : u(x,t) > u(t,0)\}$ does indeed define a reversed barrier. From the arguments above, we also conclude that $\mathbb{P}(\bar{\tau}_D = \bar{\tau}_{D^*}) = 1$, since $\widetilde{X}_{\bar{\tau}_D}$ is supported on the set where $u(x,t) = u(x,0)$, and $f(x,t) = 0$ on this set also. $\square$

Remark 3.5. We observe also that the solution $u(x,t)$ has an interpretation in terms of an optimal stopping problem [c.f. 11, Remark 4.4]. Fix $T > 0$ and set

$$v(x,t) = \sup_{\tau \in [t,T]} \mathbb{E}^{(x,t)} \left[\mathrm{U}_\mu(\widetilde{X}_\tau) - \mathrm{U}_\nu(\widetilde{X}_\tau) \right],$$

where the supremum is taken over stopping times τ, and the expectation is taken conditional on $\widetilde{X}_t = x$. Then standard results for optimal stopping problems suggest that $v(x,t)$ is the solution to the viscosity equation:

$$\begin{aligned} &\max\left\{ \frac{1}{2}\sigma(x)^2 \frac{\partial^2 v}{\partial x^2}(x,t) + \frac{\partial v}{\partial t}(x,t), v(x,t) - (\mathrm{U}_\mu(x) - \mathrm{U}_\nu(x)) \right\} = 0, \\ &v(x,T) = \mathrm{U}_\mu(x) - \mathrm{U}_\nu(x). \end{aligned} \tag{3.9}$$

Now observe from the problem formulation that $v(x,t)$ is certainly decreasing in t (a stopping time which is feasible for t_1 is also feasible for $t_0 < t_1$, with the same reward). Using the fact that both the solution to (3.8) and (3.9) are monotone in t, it is possible to deduce that $v(x, T-t)$ solves (3.8), and $u(x, T-t)$ solves (3.9), so that they must be the same function.

4 Optimality of Rost's Barrier, and superhedging strategies

4.1 Optimality via pathwise inequalities

For a given distribution μ, Theorem 3.2 says that Rost's solution is the "maximal variance" embedding. A slight generalisation of this result leads us to consider the following problem:

OPT$^*(\sigma, \nu, \mu)$: Suppose (3.1) and (3.2) hold, and $\widetilde{X}$ solves (3.5). Find a stopping time τ such that $\widetilde{X}_\tau \sim \mu$ and, for a given increasing convex function F with $F(0) = 0$,

$$\mathbb{E}\big[F(\tau)\big] \;=\; \sup_{\rho : \widetilde{X}_\rho \sim \mu} \mathbb{E}\big[F(\rho)\big].$$

Our aim in this section is to find the super-replicating hedging strategy for call-type payoffs on variance options, however it is not immediately obvious how to recover such an identity directly from the proofs of the optimality criterion given in Chacon [9]. Rather, we shall provide a 'pathwise inequality' which encodes the optimality in the sense that we can find a supermartingale G_t, and a function $H(x)$ such that

$$F(t) \le G_t + H(\widetilde{X}_t) \tag{4.1}$$

and such that, for Rost's embedding $\bar{\tau}_D$, equality holds in (4.1) and $G_{t\wedge\bar{\tau}_D}$ is a UI martingale. It then follows that $\bar{\tau}_D$ does indeed maximise $\mathbb{E}F(\tau)$ among all solutions to the Skorokhod embedding problem, and further, using (4.1), we can super-replicate call-type payoffs on variance options by a dynamic trading strategy.

Suppose that we have already found the solution to $\mathbf{SEP}^*(\sigma,\nu,\mu)$, $\bar{\tau}_D$. Define the function

$$M(x,t) = \mathbb{E}^{(x,t)}\left[f(\bar{\tau}_D)\right], \tag{4.2}$$

where f is the left derivative of F and $\bar{\tau}_D$ is the corresponding hitting time of B. Specifically, observe that $\widetilde{X}_t$ is a Markov process, and we interpret the expectation as the average of $f(\bar{\tau}_D)$ given that we start at $\widetilde{X}_t = x$, with $\bar{\tau}_D \geq t$. At zero, given the possibility of stopping at time 0, it is not immediately clear how to interpret the conditioning — it will turn out not to matter, but a natural choice would be to replace $\bar{\tau}_D$ with τ_D. In the following, we shall assume:

$$M(x,t) \text{ is locally bounded on } \mathbb{R}\times\mathbb{R}_+. \tag{4.3}$$

Obviously, $M(x,t) = f(t)$ whenever $0 \leq t < R(x)$. Now given a fixed time $T > 0$, and choosing $S_0^* \in (\inf \operatorname{supp}(\mu), \sup \operatorname{supp}(\mu))$ (which is non-empty if μ is non-trivial), we define

$$Z_T(x) \;=\; 2\int_{S_0^*}^{x}\int_{S_0^*}^{y} \frac{M(z,T)}{\sigma(z)^2}\,\mathrm{d}z\,\mathrm{d}y,$$

and in particular, $Z_T''(x) = 2M(x,T)/\sigma^2(x)$, and Z_T is a convex function. Define also

$$\begin{aligned}
G_T(x,t) &= F(T) - \int_t^T M(x,s)\,\mathrm{d}s - Z_T(x)\\
H_T(x) &= \int_{R(x)}^{T}\Big[M(x,s) - f(s)\Big]\,\mathrm{d}s + Z_T(x)\\
&= \int_{R(x)\wedge T}^{T}\Big[M(x,s) - f(s)\Big]\,\mathrm{d}s + Z_T(x).\\
Q(x) &= \int_{S_0^*}^{x}\int_{S_0^*}^{y}\frac{2}{\sigma(z)^2}\,\mathrm{d}z\,\mathrm{d}y.
\end{aligned} \tag{4.4}$$

Then we have the following results

Proposition 4.1. *For all $(x,t,T) \in \mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}_+$, we have,*

$$\begin{cases} G_T(x,t) \;+\; H_T(x) \;\geq\; F(t), & \text{if } t > R(x)\,;\\ G_T(x,t) \;+\; H_T(x) \;=\; F(t), & \text{if } t \leq R(x)\,. \end{cases} \tag{4.5}$$

Lemma 4.2. *Under the setting (3.1) – (3.2), suppose that the stopping time $\bar{\tau}_D$ is the solution to $\mathbf{SEP}^*(\sigma,\nu,\mu)$. Moreover, assume f is bounded and*

$$\textit{for any } T>0,\quad \Big(Q(\widetilde{X}_t);\, 0\leq t\leq T\Big)\quad \textit{is a uniformly integrable family.} \tag{4.6}$$

Then for any $T > 0$, the process

$$\Big(G_T(\widetilde{X}_{t\wedge\bar{\tau}_D}, t\wedge\bar{\tau}_D);\, 0\leq t\leq T\Big)\quad \textit{is a martingale,} \tag{4.7}$$

and

$$\Big(G_T(\widetilde{X}_t, t);\, 0\leq t\leq T\Big)\quad \textit{is a supermartingale.} \tag{4.8}$$

We note that when $\sigma(x) = x$, so $\widetilde{X}_t$ is geometric Brownian motion, then it is straightforward to check that, for all $T > 0$, $\sup_{t \le T} \mathbb{E}[Q(\widetilde{X}_t)^2] < \infty$, and so (4.6) is trivially satisfied provided $\mathbb{E}[Q(\widetilde{X}_0)^2] < \infty$. Note also that since $\widetilde{X}_t$ is a local martingale bounded below, for any embedding τ which embeds μ we have $\mathbb{E}[\widetilde{X}_\tau] = \mathbb{E}[\widetilde{X}_{\bar{\tau}_D}]$. It follows that if $\mathbb{E}[\widetilde{X}_{\bar{\tau}_D}] = \mathbb{E}[\widetilde{X}_0]$, any embedding of μ is a martingale, and not just a local-martingale.

Then the main result of this section follows.

Theorem 4.3. *Suppose that $\bar{\tau}_D$ is the solution to* $\mathbf{SEP}^*(\sigma, \nu, \mu)$*, and* (4.6) *holds, then $\bar{\tau}_D$ solves* $\mathbf{OPT}^*(\sigma, \nu, \mu)$*.*

Proof. We first consider the case where $\mathbb{E}[\bar{\tau}_D] = \infty$. Since $F(t) \ge \alpha + \beta t$ for some constants $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$, we must have $\mathbb{E}[F(\bar{\tau}_D)] = \infty$. The result is trivial. So we always assume $\mathbb{E}[\bar{\tau}_D] < \infty$.

Under the assumption $\mathbb{E}[\bar{\tau}_D] < \infty$, consider $Q(\cdot)$ given by (4.4). We have (recall (3.4)) $\mathbb{E}[Q(\widetilde{X}_{\bar{\tau}_D})] = \mathbb{E}[\bar{\tau}_D] + \mathbb{E}[Q(\widetilde{X}_0)] < \infty$. Therefore, for all τ embedding μ, $\mathbb{E}[\tau] = \mathbb{E}[Q(\widetilde{X}_\tau)] = \mathbb{E}[Q(\widetilde{X}_{\bar{\tau}_D})] = \mathbb{E}[\bar{\tau}_D] + \mathbb{E}[Q(\widetilde{X}_0)] < \infty$. In the remainder of this proof, we always assume $\mathbb{E}[\tau] = \mathbb{E}[\bar{\tau}_D] < \infty$.

We first assume f is bounded, since f is increasing,

$$\text{there exists } C < \infty, \text{ such that } \lim_{t\to\infty} f(t) = C. \tag{4.9}$$

For $T > 0$, since $M(\cdot, T)$ is also bounded by C, then

$$\mathbb{E}[Z_T(\widetilde{X}_{t\wedge\tau})] \le C\mathbb{E}[Q(\widetilde{X}_{t\wedge\tau})] = C(\mathbb{E}[t \wedge \tau] + \mathbb{E}[Q(\widetilde{X}_0)]) < \infty,$$

and the same argument implies $\mathbb{E}[Z_T(\widetilde{X}_\tau)] < \infty$. So $\mathbb{E}[Z_T(\widetilde{X}_\tau)|\mathcal{F}_t]$ is a uniformly integrable martingale, and by convexity, $Z_T(\widetilde{X}_{t\wedge\tau}) \le \mathbb{E}[Z_T(\widetilde{X}_\tau)|\mathcal{F}_t]$. Therefore,

$$-C|T - (t \wedge \tau)| \;\le\; F(T) - G_T(\widetilde{X}_{t\wedge\tau}, t \wedge \tau) \;\le\; C|T - (t \wedge \tau)| + \mathbb{E}[Z_T(\widetilde{X}_\tau)|\mathcal{F}_t].$$

It follows that $\mathbb{E}[G_T(\widetilde{X}_{t\wedge\tau}, t \wedge \tau)] \to \mathbb{E}[G_T(\widetilde{X}_\tau, \tau)]$ as $t \to \infty$. On the other hand,

$$\mathbb{E}\left[H_T(\widetilde{X}_\tau)\right] \;=\; \mathbb{E}\left[\int_{T\wedge R(\widetilde{X}_\tau)}^{T} \left[M(\widetilde{X}_\tau, s) - f(s)\right] \mathrm{d}s\right] + \mathbb{E}\left[Z_T(\widetilde{X}_\tau)\right] \;<\; \infty.$$

The same arguments hold when τ is replaced by $\bar{\tau}_D$, and then we have

$$\mathbb{E}\left[H_T(\widetilde{X}_\tau)\right] \;=\; \mathbb{E}\left[H_T(\widetilde{X}_{\bar{\tau}_D})\right] \qquad \text{and} \qquad \mathbb{E}\left[Z_T(\widetilde{X}_\tau)\right] \;=\; \mathbb{E}\left[Z_T(\widetilde{X}_{\bar{\tau}_D})\right].$$

In addition, by Lemma 4.2, we have,

$$\mathbb{E}\left[G_T(\widetilde{X}_{T\wedge\bar{\tau}_D}, T \wedge \bar{\tau}_D)\right] \ge\; \mathbb{E}\left[G_T(\widetilde{X}_{T\wedge\tau}, T \wedge \tau)\right].$$

Combining the results above with (4.5), we have

$$\begin{aligned}
\mathbb{E}F(\tau) &\le \mathbb{E}\big[G_T(\widetilde{X}_\tau,\tau)+H_T(\widetilde{X}_\tau)\big]\\
&= \mathbb{E}\big[G_T(\widetilde{X}_{T\wedge\tau},T\wedge\tau)+H_T(\widetilde{X}_\tau)\big]+\mathbb{E}\big[G_T(\widetilde{X}_\tau,\tau)-G_T(\widetilde{X}_{T\wedge\tau},T\wedge\tau)\big]\\
&\le \mathbb{E}\big[G_T(\widetilde{X}_{T\wedge\bar{\tau}_D},T\wedge\bar{\tau}_D)+H_T(\widetilde{X}_{\bar{\tau}_D})\big]\\
&\qquad +\mathbb{E}\big[G_T(\widetilde{X}_\tau,\tau)-G_T(\widetilde{X}_{T\wedge\tau},T\wedge\tau)\big]\\
&= \mathbb{E}\big[G_T(\widetilde{X}_{\bar{\tau}_D},\bar{\tau}_D)+H_T(\widetilde{X}_{\bar{\tau}_D})\big]+\mathbb{E}\big[G_T(\widetilde{X}_\tau,\tau)-G_T(\widetilde{X}_{T\wedge\tau},T\wedge\tau)\big]\\
&\qquad -\mathbb{E}\big[G_T(\widetilde{X}_{\bar{\tau}_D},\bar{\tau}_D)-G_T(\widetilde{X}_{T\wedge\bar{\tau}_D},T\wedge\bar{\tau}_D)\big]\\
&= \mathbb{E}\big[F(\bar{\tau}_D)\big]+\mathbb{E}\left[\int_{T\wedge\tau}^{T}M(\widetilde{X}_{T\wedge\tau},s)\,\mathrm{d}s-\int_{\tau}^{T}M(\widetilde{X}_\tau,s)\,\mathrm{d}s+Z_T(\widetilde{X}_{T\wedge\tau})\right]\\
&\qquad -\mathbb{E}\left[\int_{T\wedge\bar{\tau}_D}^{T}M(\widetilde{X}_{T\wedge\bar{\tau}_D},s)\,\mathrm{d}s-\int_{\bar{\tau}_D}^{T}M(\widetilde{X}_{\bar{\tau}_D},s)\,\mathrm{d}s+Z_T(\widetilde{X}_{T\wedge\bar{\tau}_D})\right]\\
&= \mathbb{E}\big[F(\bar{\tau}_D)\big]+\mathbb{E}\left[\mathbf{1}_{[\tau>T]}\int_T^{\tau}M(\widetilde{X}_\tau,s)\,\mathrm{d}s\right]\\
&\qquad -\mathbb{E}\left[\mathbf{1}_{[\bar{\tau}_D>T]}\int_T^{\bar{\tau}_D}M(\widetilde{X}_{\bar{\tau}_D},s)\,\mathrm{d}s\right]+\mathbb{E}\Big[Z_T(\widetilde{X}_{T\wedge\tau})-Z_T(\widetilde{X}_{T\wedge\bar{\tau}_D})\Big].
\end{aligned} \tag{4.10}$$

Since $f\le C$, we have

$$\begin{aligned}
0\le \mathbb{E}\left[\mathbf{1}_{[\tau>T]}\int_T^{\tau}M(\widetilde{X}_\tau,s)\,\mathrm{d}s\right] &\le C\mathbb{E}\big[\mathbf{1}_{[\tau>T]}(\tau-T)\big]\\
&= C\mathbb{E}\big[\tau-T\wedge\tau\big] \;\longrightarrow\; 0, \quad \text{as } T\to\infty.
\end{aligned}$$

Similarly,

$$\lim_{T\to\infty}\mathbb{E}\left[\mathbf{1}_{[\bar{\tau}_D>T]}\int_T^{\bar{\tau}_D}M(\widetilde{X}_{\bar{\tau}_D},s)\,\mathrm{d}s\right] \;=\; 0.$$

Now, by the fact that $\mathbb{E}[Q(\widetilde{X}_{T\wedge\tau})]=\mathbb{E}[T\wedge\tau]+\mathbb{E}[Q(\widetilde{X}_0)]$ and the convexity of Q, $Q(\widetilde{X}_{T\wedge\tau})\le\mathbb{E}[Q(\widetilde{X}_\tau)|\mathcal{F}_T]$, hence, $Q(\widetilde{X}_{t\wedge\tau})\to Q(\widetilde{X}_\tau)$ in L^1. Noting that $Z_T(\widetilde{X}_{T\wedge\tau})\le CQ(\widetilde{X}_{T\wedge\tau})$ and $Z_T(\widetilde{X}_{T\wedge\tau})\to CQ(\widetilde{X}_\tau)$ a.s. as $T\to\infty$, we have

$$\lim_{T\to\infty}\mathbb{E}\left[Z_T(\widetilde{X}_{T\wedge\tau})\right] \;=\; C\mathbb{E}\left[Q(\widetilde{X}_\tau)\right] \;<\infty.$$

The same arguments hold when τ is replaced by $\bar{\tau}_D$, and moreover, $\mathbb{E}[Q(\widetilde{X}_\tau)]=\mathbb{E}[Q(\widetilde{X}_{\bar{\tau}_D})]$. Now, let T go to infinity in (4.10), and we have

$$\mathbb{E}\big[F(\tau)\big] \;\le\; \mathbb{E}\big[F(\bar{\tau}_D)\big].$$

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 $F_N(t)=\int_0^t f(s)\wedge N\,\mathrm{d}s$ to get $\mathbb{E}\left[F_N(\bar{\tau}_D)\right]\ge\mathbb{E}\left[F_N(\tau)\right]$, and the conclusion follows on letting $N\to\infty$. $\square$

Now we turn to the proofs of Proposition 4.1 and Lemma 4.2.

Proof of Proposition 4.1. If $(x,t) \in D$, i.e. $t > R(x)$,

$$\begin{aligned} G_T(x,t) + H_T(x) &= \int_{R(x)}^{t} M(x,s)\,\mathrm{d}s + F(R(x)) \\ &\geq \int_{R(x)}^{t} f(s)\,\mathrm{d}s + F(R(x)) = F(t). \end{aligned}$$

If $(x,t) \notin D$, i.e. $t \leq R(x)$,

$$\begin{aligned} G_T(x,t) + H_T(x) &= -\int_{t}^{R(x)} M(x,s)\,\mathrm{d}s + F(R(x)) \\ &= -\int_{t}^{R(x)} f(s)\,\mathrm{d}s + F(R(x)) = F(t). \end{aligned}$$

□

Proof of Lemma 4.2. For $s \leq t \leq T$, by (4.6), the Meyer-Itô formula gives,

$$Z_T(\widetilde{X}_t) - Z_T(\widetilde{X}_s) \;=\; \int_s^t Z_T'(\widetilde{X}_u)\,\mathrm{d}\widetilde{X}_u + \int_s^t M(\widetilde{X}_u, T)\,\mathrm{d}u.$$

By (4.6) and the fact f is bounded, it is easy to see that the family $(Z_T(\widetilde{X}_t); 0 \leq t \leq T)$ is uniformly integrable. By the Doob-Meyer decomposition theorem (e.g. Karatzas and Shreve [32, Theorem 4.10, Chapter 1]), the first term on the right-hand side is a uniformly integrable martingale,

$$\mathbb{E}\big[Z_T(\widetilde{X}_t) - Z_T(\widetilde{X}_s)\,\big|\,\mathcal{F}_s\big] \;=\; \int_s^t \mathbb{E}\big[M(\widetilde{X}_u, T)\,\big|\,\mathcal{F}_s\big]\,\mathrm{d}u.$$

Then we have,

$$\begin{aligned} &G_T(\widetilde{X}_s, s) - \mathbb{E}\left[G_T(\widetilde{X}_t, t)\big|\mathcal{F}_s\right] \\ &= \int_t^T \mathbb{E}\big[M(\widetilde{X}_t, u)\big|\mathcal{F}_s\big]\,\mathrm{d}u + \int_s^t \mathbb{E}\big[M(\widetilde{X}_u, T)\,\big|\,\mathcal{F}_s\big]\,\mathrm{d}u - \int_s^T M(\widetilde{X}_s, u)\,\mathrm{d}u \\ &= \int_t^T \mathbb{E}\big[M(\widetilde{X}_t, u)\big|\mathcal{F}_s\big]\,\mathrm{d}u - \int_s^{T-t+s} M(\widetilde{X}_s, u)\,\mathrm{d}u \\ &\qquad\qquad + \int_s^t \mathbb{E}\big[M(\widetilde{X}_u, T)\,\big|\,\mathcal{F}_s\big]\,\mathrm{d}u - \int_{T-t+s}^T M(\widetilde{X}_s, u)\,\mathrm{d}u \\ &= \int_t^T \Big\{\mathbb{E}\big[M(\widetilde{X}_t, u)\,\big|\,\mathcal{F}_s\big] - M(\widetilde{X}_s, u-(t-s))\Big\}\,\mathrm{d}u \\ &\qquad\qquad + \int_s^t \Big\{\mathbb{E}\big[M(\widetilde{X}_u, T)\,\big|\,\mathcal{F}_s\big] - M(\widetilde{X}_s, T-(t-u))\Big\}\,\mathrm{d}u. \end{aligned}$$

Now, observe that $\widetilde{X}_t$ is a Markov process, so we can write Y_t for an independent copy of $\widetilde{X}$, and $\bar{\sigma}_D$ for the corresponding hitting time of the reversed barrier, and write $\widetilde{X}^x$ for $\widetilde{X}$ started at x. We have, for $u \in (t,T]$:

$$\begin{aligned} \mathbb{E}^{(x,u-(t-s))}\big[f(\bar{\tau}_D)\big|\mathcal{F}_u\big] \leq &\mathbf{1}_{[\bar{\tau}_D \leq u]} f(u) + \mathbf{1}_{[\bar{\tau}_D > u]} \mathbb{E}^{(x,u-(t-s))}\big[f(\bar{\tau}_D)\big|\mathcal{F}_u\big] \\ = &\mathbf{1}_{[\bar{\tau}_D \leq u]} f(u) + \mathbf{1}_{[\bar{\tau}_D > u]} \mathbb{E}^{(\widetilde{X}^x_{t-s}, u)}\big[f(\bar{\sigma}_D)\big] \\ \leq &M(\widetilde{X}^x_{t-s}, u). \end{aligned} \tag{4.11}$$

Hence,

$$\begin{aligned}\mathbb{E}\big[M(\widetilde{X}_t,u)\,\big|\,\mathcal{F}_s\big] \;&=\; \mathbb{E}^{\widetilde{X}_s}M(\widetilde{X}_{t-s},u) \\ &\geq\; \mathbb{E}^{(\widetilde{X}_s,u-(t-s))}\big[f(\widetilde{\tau}_D)\big] \;=\; M(\widetilde{X}_s,u-(t-s)).\end{aligned} \tag{4.12}$$

For $u \in (s,T)$, replacing u by T and t by u in (4.11) gives that

$$\mathbb{E}^{(x,T-(u-s))}\big[f(\bar{\tau}_D)\,|\,\mathcal{F}_T\big] \;\leq\; M(\widetilde{X}^x_{u-s},T),$$

and hence,

$$\mathbb{E}\big[M(\widetilde{X}_u,T)\,\big|\,\mathcal{F}_s\big] \;\geq\; M(\widetilde{X}_s,T-(u-s)). \tag{4.13}$$

It follows that

$$\begin{aligned}&\int_s^t \Big\{\mathbb{E}\big[M(\widetilde{X}_u,T)\,\big|\,\mathcal{F}_s\big] - M(\widetilde{X}_s,T-(t-u))\Big\}\,\mathrm{d}u \\ &\qquad = \int_s^t \Big\{\mathbb{E}\big[M(\widetilde{X}_u,T)\,\big|\,\mathcal{F}_s\big] - M(\widetilde{X}_s,T-(u-s))\Big\}\,\mathrm{d}u \;\geq 0.\end{aligned}$$

Therefore,

$$G_T(\widetilde{X}_s,s) - \mathbb{E}\left[G_T(\widetilde{X}_t,t)|\mathcal{F}_s\right] \;\geq\; 0,$$

which implies (4.8).

On the other hand, as a part of (4.11),

$$\mathbf{1}_{[\bar{\tau}_D>u]}\mathbb{E}^{(x,u-(t-s))}\big[f(\bar{\tau}_D)\,|\,\mathcal{F}_u\big] \;=\; \mathbf{1}_{[\bar{\tau}_D>u]}\mathbb{E}^{(\widetilde{X}^x_{t-s},u)}\big[f(\bar{\sigma}_D)\big],$$

and on $\{u < \bar{\tau}_D\}$ equality holds in the inequalities (4.12) and (4.13). Thus, (4.7) follows. $\square$

For bounded f, although the pathwise inequality in this section $G_T(\widetilde{X}_t,t) + H_T(\widetilde{X}_t) \geq F(t)$ holds for all $T,t>0$, $G_T(\widetilde{X}_t,t)$ is a supermartingale only on $[0,T]$. For hedging purposes, we would really like to know: can we find a *global* pathwise inequality $G^*_t + H^*(\widetilde{X}_t) \geq F(t)$, such that G^*_t is a supermartingale on $[0,\infty]$ and a martingale on $[0,\bar{\tau}_D]$? We now provide conditions where we can find such G^* and H^*.

We replace (4.9) by a stronger assumption: there exists some $\alpha > 1$, such that

$$\text{for } t \text{ sufficiently large,}\;\; C \;\geq\; f(t) \;\geq\; C - O(t^{-\alpha}). \tag{4.14}$$

Under this assumption, it is easy to check there exists a $J(x,t)$ such that

$$J(x,t) \;=\; \lim_{T\to\infty}\int_t^T [M(x,s)-f(s)]\,\mathrm{d}s, \tag{4.15}$$

then we define

$$\begin{cases} G(x,t) \;=\; \lim\limits_{T\to\infty} G_T(x,t) \;=\; F(t) - J(x,t) - CQ(x); \\ H(x) \;=\; \lim\limits_{T\to\infty} H_T(x) \;=\; J(x,R(x)) + CQ(x). \end{cases} \tag{4.16}$$

Letting $T \to \infty$ in (4.5),

$$\begin{cases} G(x,t) + H(x) \ > \ F(t), & \text{if } \ t > R(x); \\ G(x,t) + H(x) \ = \ F(t), & \text{if } \ t \le R(x). \end{cases} \tag{4.17}$$

By the monotone convergence theorem, for all $t > 0$, $\mathbb{E}[\int_t^T [M(\widetilde{X}_t, s) - f(s)] \, \mathrm{d}s] \to \mathbb{E}[J(\widetilde{X}_t, t)]$ as $T \to \infty$, and then by Scheffé's Lemma, $\int_t^T [M(\widetilde{X}_t, s) - f(s)] \, \mathrm{d}s \to J(\widetilde{X}_t, t)$ in L^1. On the other hand, since $Z_T(\widetilde{X}_t) \to CQ(\widetilde{X}_t)$ in L^1,

$$G_T(\widetilde{X}_t, t) \ \xrightarrow{L^1} \ G(\widetilde{X}_t, t) \ \text{ and } \ H_T(\widetilde{X}_t) \ \xrightarrow{L^1} \ H(\widetilde{X}_t).$$

It follows that the process $\left(G(\widetilde{X}_t, t); t \ge 0\right)$ is a supermartingale and the process $\left(G(\widetilde{X}_{t \wedge \bar{\tau}_D}, t \wedge \bar{\tau}_D); t \ge 0\right)$ is a martingale (since the conditional expectation, as an operator, is continuous in L^p for $p \ge 1$). We then can show as before that (if τ, $\bar{\tau}_D$ are integrable),

$$\begin{aligned} \mathbb{E}\left[F(\tau)\right] \ &\le \ \mathbb{E}\left[G(\widetilde{X}_\tau, \tau) + H(\widetilde{X}_\tau)\right] \\ &\le \ \mathbb{E}\left[G(\widetilde{X}_{\bar{\tau}_D}, \bar{\tau}_D) + H(\widetilde{X}_{\bar{\tau}_D})\right] \ = \ \mathbb{E}\left[F(\bar{\tau}_D)\right]. \end{aligned}$$

An example where (4.14) holds is the call-type payoff: $F(t) = (t - K)_+$. We see that for $t > K$, the left derivative $f(t) = 1$, and hence

$$J(x,t) = \int_t^K [M(x,s) - f(s)] \, \mathrm{d}s,$$

we then repeat all arguments above to obtain the pathwise inequality and the optimality result. On the other hand, the condition will fail if e.g. $F(t) = t^2$. Recall that a volatility swap corresponds to the choice of $F(t) = \sqrt{t}$, and that we can consider concave functions by taking $-F(t)$. This causes difficulties since $-F$ is not increasing, nor can we make it increasing by considering $-F(t) + \alpha t$, for some $\alpha > 0$, since $F'(t) \to \infty$ as $t \to \infty$. However $F(t)$ can be approximated from both above and below by functions $F^1(t)$ and $F^2(t)$ which are concave, and have bounded derivatives. (In the case of the upper approximation, we have $F^1(0) > 0$, but this can be made arbitrarily small). An optimality result will then follow in this setting — note also that the condition (4.14) will be satisfied in this case.

4.2 Superhedging options on weighted realised variance

We now return to the financial context described in Section 2. Our aim is to use the construction we produced for the proof of optimality in Section 4.1 to provide a model-independent hedging strategy for derivatives which are convex functions of weighted realised variance. We will suppose initially that our options are not forward starting, so $\nu = \delta_{S_0}$.

We now define $\bar{\tau}_D$ as the embedding of μ for the diffusion $\widetilde{X}$, and define functions: G, H, J, and Q as in the previous section (so (4.14) holds). Our aim is to use (4.17), which now reads:

$$G(X_{A_t},t)+H(X_{A_t}) \;=\; G(\widetilde{X}_t,t)+H(\widetilde{X}_t) \;\geq\; F(t) \;=\; F\left(\int_0^{A_t} w(X_s)\sigma_s^2\,\mathrm{d}s\right), \tag{4.18}$$

to construct a super-replicating portfolio. We shall first show that we can construct a trading strategy that super-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 $(G(\widetilde{X}_t,t))_{t\geq 0}$ is a supermartingale, 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 $(\widetilde{\phi}_t)_{t\geq 0}$ such that:

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

and such that there is equality at $t=\bar{\tau}_D$. Moreover, since $(G(\widetilde{X}_{t\wedge\bar{\tau}_D},t\wedge\bar{\tau}_D))_{t\geq 0}$ is a martingale, and G is of $\mathbb{C}^{2,1}$ class in D (since $M(x,t)$ is), we have:

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

More generally, we would not expect $\partial G/\partial x$ 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.

It follows then that we can identify a process $\left(\widetilde{\phi}_t;t\geq 0\right)$ with

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

where we have used e.g. Revuz and Yor [43, Proposition V.1.4]. Finally, writing $\phi_t=\widetilde{\phi}_{\tau_t}$, then

$$G(X_t,\tau_t) \;\leq\; G(X_0,0)+\int_0^t\phi_s\,\mathrm{d}X_s \;=\; G(X_0,0)+\int_0^t\phi_s\,\mathrm{d}\left(B_s^{-1}S_s\right). \tag{4.19}$$

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 $V_0 \;=\; G(X_0,0)B_T^{-1}$, and therefore

$$V_T \;=\; B_T\left(V_0B_0^{-1}+\int_0^T B_T^{-1}\phi_s\,\mathrm{d}\left(B_s^{-1}S_s\right)\right) \;=\; G(X_0,0)+\int_0^T\phi_s\,\mathrm{d}X_s.$$

We now turn to the $H(X_T)$ component in (4.18). 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:

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

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

$$\begin{aligned} &B_T^{-1}\big[H(S_0) - S_0 H'_+(S_0)\big] \quad \text{in cash;} \\ &B_T^{-1} H'_+(S_0) \quad \text{units of the asset;} \\ &B_T^{-1} H''(\mathrm{d}K) \quad \text{units of the call with strike } B_T K \text{ for } K \in (S_0, \infty); \\ &B_T^{-1} H''(\mathrm{d}K) \quad \text{units of the put with strike } B_T K \text{ for } K \in (0, S_0], \end{aligned} \tag{4.20}$$

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 the integrability condition:

$$\int_{(0,S_0]} P(B_T K)|H''|(\mathrm{d}K) + \int_{(S_0,\infty)} C(B_T K)|H''|(\mathrm{d}K) \ < \ \infty, \tag{4.21}$$

holds, 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 4.4. *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 4.5. *Suppose Assumptions 2.1, 2.2 and 4.4 hold, and suppose $F(\cdot)$ is a convex, increasing function with $F(0) = 0$ and the left derivative $f(t) := F'_-(t)$ satisfies* (4.14)*. Let $M(x,t)$ and $J(x,t)$ be given by* (4.2) *and* (4.15)*, and are determined by the solution to $\mathbf{SEP}^*(xw(x)^{-1/2}, \delta_{S_0}, \mu)$, where μ is determined by* (2.2) *and $w \in \mathcal{D}$. We also define Q after* (4.4)*, such that* (4.6) *holds and then the functions G and H are given by* (4.16).

Then there exists an arbitrage if the price of an option with payoff $F(RV_T^w)$ is strictly greater than

$$B_T^{-1}\Big[G(S_0, 0) + H(S_0) + \int_{(S_0,\infty)} C(B_T K) H''(\mathrm{d}K) \\ + \int_{(0,S_0]} P(B_T K) H''(\mathrm{d}K)\Big]. \tag{4.22}$$

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, and the arbitrage strategy can be chosen independent of the model.

Proof. According to the arguments above, our superhedge of the variance option can be described as the combination of a static portfolio (4.20) and a self-financing dynamic portfolio which consists of an additional $B_T^{-1}\psi_t$ units of the risky asset and an additional initial cash holding of $B_T^{-1}(G(S_0, 0) - \psi_0 S_0)$. In the case where $G(x, t)$ is sufficiently differentiable, we can identify the process $\psi_t = \widetilde{\psi}_{\tau_t}$ by

$$\psi_t = \frac{\partial G}{\partial x}(X_t, \tau_t).$$

We observe that this strategy is independent of the true model. It is easy to see that the total initial investment of this superhedge is given be (4.22).

In the case where G is not sufficiently differentiable, we first observe that $G(x, t)$ is continuous: note that $Q(x)$ and $F(t)$ are trivially so by (4.16), and $M(x, t)$ is continuous in D, and in $D^{\complement}$, and additionally at jumps of $R(x)$: it follows that $M(x_n, t) \to M(x, t)$ as $x_n \to x$ except possibly at a set of Lebesgue measure zero, and hence $G(x, t)$ is continuous.

Now consider $G(x, t)$ on a bounded open set of the form $\mathcal{O} = (y_0, y_1) \times (t_0, t_1)$. By continuity, G can be approximated uniformly on the boundary $\partial\mathcal{O}$ (or more relevantly, on the boundary where $t > t_0$, $\partial\mathcal{O}_+$) by a smooth function. Specifically, for fixed $\varepsilon > 0$, there exists a function $G_\varepsilon(x, t)$ such that $G(x, t) + \varepsilon \geq G_\varepsilon(x, t) \geq G(x, t)$ on $\partial\mathcal{O}_+$. Moreover, the function

$$G_\varepsilon(x, t) = \mathbb{E}^{(x,t)}\left[G_\varepsilon(\widetilde{X}_{\tau_{\partial\mathcal{O}_+}}, \tau_{\partial\mathcal{O}_+})\right] \tag{4.23}$$

is $C^{2,1}$ and a martingale on $\bar{\mathcal{O}}$, and so

$$G_\varepsilon(\widetilde{X}_{\tau_{\partial\mathcal{O}_+}}, \tau_{\partial\mathcal{O}_+}) = G_\varepsilon(\widetilde{X}_0, t_0) + \int_{t_0}^{\tau_{\partial\mathcal{O}_+}} \frac{\partial G_\varepsilon}{\partial x}\, \mathrm{d}\widetilde{X}_s.$$

Since G is a supermartingale, for $(x, t) \in \mathcal{O}$, from (4.23) we have $G(x, t) \geq G_\varepsilon(x, t) - \varepsilon$.

Now observe that we can choose a countable sequence of such sets $\mathcal{O}_1, \mathcal{O}_2, \ldots$ with each set centred at the exit point of the previous set, and such that any continuous path is guaranteed to pass through only finitely many such sets on a finite time interval. For any fixed $\delta > 0$, we can take a sequence of strictly positive $\varepsilon_1, \varepsilon_2, \ldots$ such that $\sum_i \varepsilon_i = \delta$, and apply the arguments above to generate a sequence of functions $G_{\varepsilon_i}(x, t)$ on $\mathcal{O}_i$. It follows that, given $\delta > 0$, we can always find a function $\widetilde{\psi}_t$ such that

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

Since δ was arbitrary, whenever the price of an option is strictly greater than (4.22) , we can choose δ sufficiently small that the arbitrage still works. Finally, we observe that at any time $t \in [0, T]$, the arbitrage strategy is worth at least $F(\tau_t) \geq F(0)$, so the strategy is bounded below, and hence admissible.

To see that this is the best possible upper bound, we need to show that there is a model which satisfies Assumption 2.1, has law μ under $\mathbb{Q}$ at time T, and such that the superhedge is actually a hedge. But consider the stopping time $\bar{\tau}_D$ for the process $\widetilde{X}_t$. Define the process $(X_t; 0 \leq t \leq T)$ by

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

and then X_t satisfies the stochastic differential equation

$$\mathrm{d}X_s = \hat{\sigma}_s X_s w(X_s)^{-1/2} \, \mathrm{d}W_s = \sigma_s^2 X_s \, \mathrm{d}W_s$$

with the choice of

$$\hat{\sigma}_s^2 = \frac{T}{(T-s)^2} \mathbf{1}_{[\frac{s}{T-s} < \bar{\tau}_D]}, \quad \sigma_s^2 = \hat{\sigma}_s^2 w(X_s)^{-1/2}. \tag{4.24}$$

Since $\bar{\tau}_D < \infty$, a.s., then $X_T = \widetilde{X}_{\bar{\tau}_D}$, and

$$\tau_T = \int_0^T w(X_s) \sigma_s^2 \mathrm{d}s = \int_0^T \frac{T}{(T-s)^2} \mathbf{1}_{[\frac{s}{T-s} < \bar{\tau}_D]} = \bar{\tau}_D.$$

Hence $S_t = X_t B_t$ is a price process satisfying Assumption 2.1 with

$$F\left(\int_0^T w(X_s) \sigma_t^2 \, \mathrm{d}t \right) \ = \ F(\bar{\tau}_D).$$

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

Remark 4.6. The above result assumes that the option payoff depends on the realised weighted variation computed between time 0 and a fixed time T. In some situations, *forward-starting* versions of these derivatives may be traded. Here, one is interested in the payoff of an option written on the variation observed between a fixed time $T_0 > 0$ and the maturity date T_1: $\int_{T_0}^{T_1} w(X_t) \sigma_t^2 \, \mathrm{d}t$. If one observes traded options at both T_0 and T_1, these again imply the (hypothesised, risk-neutral) distributions at times T_0, and T_1, and it is reasonable to suppose that the upper bound on the price of an option (for suitable, convex $F(\cdot)$) should correspond to the solution of $\mathbf{SEP}^*(\sigma, \nu, \mu)$ determined above. Let G and H be the functions derived above. The question remains as to how one includes the additional information at time T_0 in the hedging strategy. (For clarity, we suppose $B_t = 1$ for all $t \geq 0$.)

In order to have the correct hedge for $t \in [T_0, T_1]$, we need a portfolio of call options maturing at time T_1 with payoff $H(X_{T_1})$. In addition to the payoff at maturity, we need a dynamic portfolio worth (at least) $G(X_t, \tau_t)$, where now $\tau_t = \int_{T_0}^t \lambda(X_s) \sigma_s^2 \, \mathrm{d}s$ — specifically, recalling $F(0) = 0$, and Proposition 4.1, we should have $G(X_{T_0}, 0) + H(X_{T_0}) = 0$. This implies that we need a portfolio of call options with maturity T_0 and with payoff $-H(X_{T_0})$. Under a similar assumption to Assumption 4.4, this is possible, and the resulting strategy will give a superhedge which is a hedge under the optimal model corresponding to the Rost embedding.

Strictly speaking, Theorem 4.5 is model-dependent: our arbitrage strategy is specified in a way that is independent of the exact model, but some of the underlying concepts — specifically the quadratic variation in the option payoff, and the stochastic integral term that is implemented in the hedge both depend on an underlying probability space, and it could therefore be argued that the strategies are not truly model-independent. In the following remarks, we briefly outline how one might relax this assumption.

Remark 4.7. In a similar manner to recent work of Davis, Obłój, and Raval [16], we can formulate this result without any need for a probabilistic framework. The difficulty in treating the previous arguments on a purely pathwise basis is that we need to make sense of the stochastic integral term in (4.19), and the quadratic variation in the option payoff. However, under mild assumptions on the paths of $S_t = B_t^{-1} X_t$, and a stronger assumption on G (specifically, that G is $\mathbb{C}^{2,1})^2$, we can recover a pathwise result, based on a version of Itô's formula due to Föllmer [22].

Suppose we fix a sequence of partitions $\pi_n = \{0 = t_0^n \leq t_1^n \leq t_2^n \leq \cdots \leq t_n^n = T\}$ of $[0,T]$, such that $\sup_{i \leq n,} |t_i^n - t_{i-1}^n| \to 0$ as $n \to \infty$. Then we define the class $\mathcal{QV}$ of continuous, strictly positive paths X_t such that

$$\sum_{i=1}^{n} \left(\frac{X_{t_i^n} - X_{t_{i-1}}^n}{X_{t_{i-1}}^n} \right)^2 \delta_{t_{i-1}^n} = \mu^n \to \mu \text{ where } \mu([0,t]) = \int_0^t \sigma_s^2 \, \mathrm{d}s \tag{4.25}$$

for some bounded measurable function $\sigma_s : [0,T] \to \mathbb{R}_+$. Here δ_t is the Dirac measure at t, and the convergence is in the sense of weak convergence of measures as $n \to \infty$, possibly down a subsequence.

Then, following the proof of the main theorem in Föllmer [22], an application of Taylor's Theorem to the terms $G(X_{t_i^n}, \tau_{t_i^n}) - G(X_{t_{i-1}^n}, \tau_{t_{i-1}^n})$, where $\tau_t = \int_0^t \lambda(X_s)\sigma_s^2 \, \mathrm{d}s$, gives

$$\begin{aligned} G(X_T, \tau_T) - G(X_0, 0) = & \sum_{i=1}^{n} \frac{\partial G}{\partial x} \left(X_{t_i^n} - X_{t_{i-1}^n} \right) + \sum_{i=1}^{n} \frac{\partial G}{\partial t} \left(\tau_{t_i^n} - \tau_{t_{i-1}^n} \right) \\ & + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 G}{\partial x^2} \left(X_{t_i^n} - X_{t_{i-1}^n} \right)^2 . \end{aligned}$$

It follows that, whenever X_t is a path in $\mathcal{QV}$, then:

$$\begin{aligned} \sum_{i=1}^{n} \frac{\partial G}{\partial x} \left(X_{t_i^n} - X_{t_{i-1}^n} \right) \to & \, G(X_T, \tau_T) - G(X_0, 0) \\ & - \int_0^T \sigma_s^2 \left(\lambda(X_s) \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} X_s^2 \right) \mathrm{d}s. \end{aligned}$$

Recall that $\sigma(x) = x\lambda^{-1/2}(x)$, and since $G \in \mathbb{C}^{2,1}$ and a supermartingale for X_t where X_t solves (3.5), the final integrand will be negative. We conclude that, in the limit as we trade more often, *for any* $X_t \in \mathcal{QV}$, we will have a portfolio which superhedges. One could then recover the statement in the probabilistic setting by observing that, almost surely, a path from a model of the form described by Assumption 2.1 lies in $\mathcal{QV}$.

[2] This would appear to be a very strong assumption on G. However, along the lines of [16], it seems reasonable that the conclusions would hold in a milder sense; what seems harder is to both provide a set of conditions under which these conclusions hold, and which can be verified under relatively natural constraints on our modelling setup.

Remark 4.8. An alternative approach that is still within a more general, model-independent setting, but where we do not need to assume strong differentiability conditions on G, can be constructed using the uncertain volatility approach, originally introduced by Avellaneda, Levy, and Parás [1]. We base our presentation on the paper of Possamaï, Royer, and Touzi [42].

Let $\Omega = \{\omega \in C([0,T];(0,\infty)), \omega(0) = X_0\}$ be a path space, equipped with the uniform norm, $||\omega|| = \sup_{t\in[0,T]} |\omega(t)|$, and let $X_t(\omega) = \omega(t)$ be the canonical process. Let $\mathbb{P}_0$ be the probability measure on Ω such that X_t is a standard geometric Brownian motion (i.e. $\log X_t$ has quadratic variation t). Let $\mathbb{F}$ be the filtration generated by X.

Let $\mathbb{H}^e_{loc}(\mathbb{P}_0, \mathbb{F})$ be the set of non-negative, $\mathbb{F}$-progressively measurable processes α_t such that $\exp\left\{\frac{1}{2}\int_0^\cdot \alpha_s \,\mathrm{d}s\right\}$ is $\mathbb{P}_0$-locally integrable. Then for $\alpha \in \mathbb{H}^e_{loc}(\mathbb{P}_0, \mathbb{F})$ we can define

$$X_t^\alpha = \exp\left\{\int_0^t \alpha_s^{1/2} \,\mathrm{d}\log(X_s) - \frac{1}{2}\int_0^t \alpha_s \,\mathrm{d}s\right\}.$$

In particular, under $\mathbb{P}_0$, X_t^α has $\langle \log X\rangle_t = \int_0^t \alpha_s \,\mathrm{d}s$. Then we can define a probability measure on Ω by $\mathbb{P}^\alpha(X_t \in A) = \mathbb{P}_0(X_t^\alpha \in A)$, or equivalently, $\mathbb{P}^\alpha = \mathbb{P}_0 \circ (X_\cdot^\alpha)^{-1}$. It follows that there is a class of probability measures $\mathcal{P} = \{\mathbb{P}^\alpha : \alpha \in \mathbb{H}^e_{loc}(\mathbb{P}_0, \mathbb{F})\}$ on the space $(\Omega, \mathbb{F})$. We aim to produce conclusions which hold for all $\mathbb{P}^\alpha$, and we say that something holds $\mathcal{P}$-quasi surely (q.s.) if it holds $\mathbb{P}$-a.s. for all $\mathbb{P} \in \mathcal{P}$.

We now have a filtered space $(\Omega, \mathbb{F})$, and a class of (non-dominated) probability measures $\mathcal{P}$ under which we can discuss trading strategies simultaneously. Observe that the variance process $\langle \log X\rangle_t$ can be defined pathwise on Ω using the results of Karandikar [31]: set $a_0^n = 0$, and $a_{i+1}^n = \inf\{t \geq a_i^n : |\log(\omega(t)) - \log(\omega(a_i^n))| \geq 2^{-n}\}$, and consider the process $V_t(\omega)$ defined by

$$V_t(\omega) = \lim_{n\to\infty} \sum_{i=0}^n \big(\log(\omega(a_{i+1}^n \wedge t)) - \log(\omega(a_i^n \wedge t))\big) \tag{4.26}$$

if the limit exists, where the limit is taken in the sense of uniform convergence on $[0,T]$ and defined to be zero otherwise. The limit exists $\mathbb{P}^\alpha$-a.s. for each α, and when the limit exists, the limit is $\mathcal{F}_t$ measurable and $\mathbb{P}^\alpha$-a.s. equal to $\langle \log X\rangle_t$. If we write

$$\mathcal{N}^{\mathcal{P}} = \{E \subset \Omega : \exists \widetilde{E} \in \mathcal{F} \ s.t. \ E \subseteq \widetilde{E}, \ \mathbb{P}(E) = 0 \ \forall \ \mathbb{P} \in \mathcal{P}\},$$

then the set of ω for which the limit in (4.26) fails is an element of $\mathcal{N}^{\mathcal{P}}$. As a result, we can make the process V_t adapted by considering the augmented filtration:

$$\mathcal{F}'_t = \mathcal{F}_t \vee \mathcal{N}^{\mathcal{P}}, \quad \mathbb{F}' = \{\mathcal{F}'_t, t \geq 0\}.$$

Under this augmented filtration, the process V_t remains $\mathbb{F}'$-progressively measurable, and indeed is continuous; it follows that the trading strategy described in the proof of Theorem 4.5 can be constructed, giving a càdlàg process ψ_t which is continuous except at the times when the process $(\omega(t), V_t(\omega))$ exits the sets $\mathcal{O}_1, \mathcal{O}_2, \ldots$. Using the construction of Karandikar [31], we can again define pathwise a process I_t, which agrees $\mathbb{P}^\alpha$-a.s. with the classical stochastic integral $J_t^{\mathbb{P}} = \int_0^t \psi_s \,\mathrm{d}X_s$. (Observe however that we may need to work in a $\mathbb{P}^\alpha$-augmented filtration for this latter object to be defined). Since by Theorem 4.5 we have $J_T^{\mathbb{P}} \geq G(X_T, \langle \log X \rangle_T) - \varepsilon$, it follows that $I_T \geq G(X_T, V_T) - \varepsilon$ $\mathcal{P}$-q.s., and therefore the strategy we describe makes sense in the uncertain volatility setting.

The fact that we have a concrete characterisation of ψ_t enables us to avoid much of the technical difficulties that arise in [42] and related papers. However, our results are in one sense also not quite so strong: we only obtain a strategy which superhedges our payoff less some $\varepsilon > 0$. The results in [42] suggest that this is unnecessary. However, our results are stronger in another direction: we do not require any integrability restriction on the payoff of the option under the class of models we consider — this constraint is already embedded in our restriction to non-negative price processes.

5 Numerical Results

5.1 Numerical solution of the viscosity equation

An important goal is to use the results of the previous sections to find numerical bounds, and their associated option prices and hedging strategies, corresponding to the solutions of Rost and Root. The hardest aspect of this is finding the numerical solution to the viscosity equation (3.8), and its equivalent for the Root solution. The solution to the Rost viscosity equation is roughly equivalent to solving a parabolic PDE inside the continuation region, while outside this region we know the solution will be equal to the initial boundary condition.

The numerical solution is made harder by the fact that, particularly in the case of the Rost solution, we expect the behaviour of the barrier near the initial starting point to be very sensitive to any discretisation: in the case where the starting measure is a point mass at X_0, and the target measure also places mass continuously (say) near X_0, then we are looking for a barrier function $R(x)$ with $R(X_0) = 0$, and a positive, but non-zero probability that $R(X_t) > t$ for some small time t. According to the law of the iterated logarithm, the behaviour of the stopped process will be very sensitive to small changes in $R(\cdot)$. As a result, a numerical method that can concentrate on this initial region would be beneficial. On the other hand, the behaviour of the barrier at large times is also of interest, although here we expect the numerics are likely to be less sensitive to discretisation.

A second question concerns the convergence and stability of our numerical methods. The theory behind the numerical approximation of viscosity equations is fairly well understood — dating back to the methods of Barles and Souganidis [3]. In this paper, we use the results of Barles and Jakobsen [4], which are suited to our purposes. Since we wish to use a large range of time steps, and we look to have non-equal grid point spacings, we will look to use an implicit method, in order to provide unconditional stability of the numerical regime. The results of [4] provide us with the necessary justification.

To outline the numerical method used, we consider a standard numerical scheme, with $\mathbf{u}^n$ the (vector valued) approximation to $u(x,t)$ evaluated at $t = t^n$, and at spatial positions $\mathbf{x}$. We approximate $\mathcal{L}u = \frac{\sigma(x)^2}{2}\frac{\partial^2 u}{\partial x^2}(x,t)$ using the Kushner approximation described by [4], which ensures that the finite difference operator L can be written in the form $(L\mathbf{u})_i = \sum_j c_j(t^n, x_{i+j})(u_{i+j} - u_i)$, where the c_j's are non-negative and zero except on some finite subset of $\mathbb{Z} \setminus \{0\}$. We also need to assume that the measures μ and ν both have compact support and the same mean, in which case $\mathbf{u}$ is constant and zero at the endpoints of $\mathbf{x}$.

Then an implicit numerical scheme to solve (3.8) will take $\mathbf{u}_0 = u(\mathbf{x}, 0)$, and solve iteratively

$$\frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{t^{n+1} - t^n} = \max\{L\mathbf{u}_{n+1}, 0\}. \tag{5.1}$$

The difficulty here arising from the fact that the maximisation depends on the unknown $\mathbf{u}_{n+1}$. We can rearrange this expression, writing $\mathbf{z} = \mathbf{u}_{n+1} - \mathbf{u}_n$, and $\Delta t^n = t^{n+1} - t^n$, to see that this is equivalent to the problem of finding $\mathbf{z} \geq 0$ such that: $(I - \Delta t^n L)\mathbf{z} - \Delta t^n L\mathbf{u}^n \geq 0$, and $\mathbf{z}^\intercal\left((I - \Delta t^n L)\mathbf{z} - \Delta t^n L\mathbf{u}^n\right) = 0$. This is a classical linear complementarity problem (LCP), and may be hard to solve (or at least, may involve many evaluations of the matrix multiplication inside the maximisation), however, at this point we can exploit the fact that the structure of the solution implies that $\mathbf{z}$ will be zero at exactly the points where we are in the barrier. Since the barrier should generally change relatively slowly, as an initial supposition, it is likely that the spatial values where $\mathbf{z} = 0$ for the previous time-step are likely to be roughly the same at the next step. It follows that a numerical scheme for solving LCPs which involves pivoting on a set of basis variables may be very efficient at solving (5.1). The algorithm we will use for this purpose is the Complementary Pivot (or Lemke's) algorithm. We refer to Murty [36] for details on the numerical implementation of the Complementary Pivot algorithm. We note also that a similar method can be used to justify implicit methods for the Root solution (the case of explicit solutions being justified directly by the results of [39]).

5.2 Analysis of numerical evidence

Using the methods outlined above, we can analyse the solutions of Root and Rost numerically. In general, we consider $\nu = \delta_{S_0}$ and μ will be determined by assuming that we observe prices of call options which are consistent with a Heston market model. In general we will consider features of barriers under Heston models since they permit relatively straightforward computation of both call prices, and prices of variance options. In what follows, we take our given prices to come from a Heston model with parameters: $\rho = -0.65, v_0 = 0.04, \theta = 0.035, \kappa = 1.2, r = 0, \xi = 0.5, S_0 = 2$ (see (6.1) for the meaning of the parameters).

Figure 1 shows the functions $u(x, t)$ in both the Rost and Root solutions, and their corresponding barrier functions. We can confirm that these functions do indeed embed the correct distributions by simulation: we compute the distribution of a process stopped on exit from the barrier and compute the corresponding call prices empirically. In fact, it is more informative to plot the implied volatility of the empirically obtained call prices. This is done in Figure 2.

One can also consider the behaviour of the barriers in time. In Figure 3 we plot the barriers for a sequence of call prices with increasing maturity.

Of course, our interest lies in the implied bounds of options on variance. We first consider the case of a variance call. In Figure 4 we display the upper bound on the price of a variance call derived in Section 4. As might be expected,

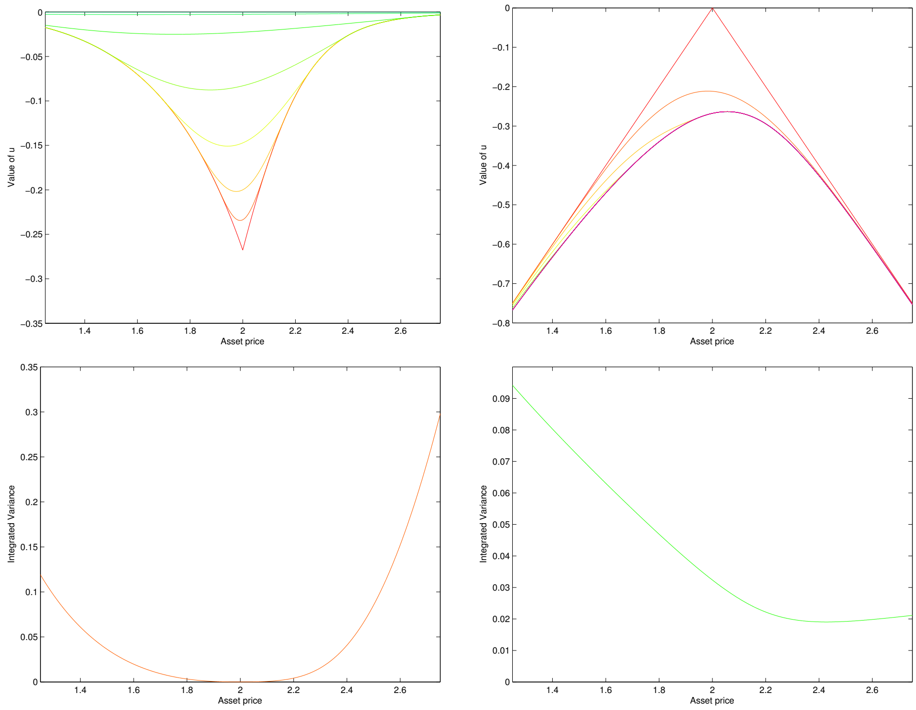

Figure 1: Plots of the function $u(x,t)$ (top) and the corresponding barriers (bottom) for the Rost (left) and Root (right) barriers.

there is a substantial difference between the upper bound and the model-implied price.

To see how the hedges constructed in perform in a given realisation, we can simulate a path, and compute the values of the super- and sub-hedging strategies along the realisation. In this example, we consider an option on variance with payoff $F(\langle \ln S\rangle_T)$, where $F(t) = t(t \wedge v_K)$. This is shown in Figure 5. The main attraction of these hedging portfolios is that they remain super/sub-hedges under a different model. For example, in Figure 6 we show how these hedges behave if the path realisation comes from a Heston model with different parameters. Here we set: $\rho' = 0.5, \theta' = 0.07$ and $\kappa' = 2.4$. To conclude, we show that the sub- and super-hedges provide good model-robustness by computing (empirically) the difference between the payoff of an option on variance, and the corresponding super- or sub-hedge. This is shown in Figure 7, which also shows the effect of model-misspecification on the distribution of the hedging error.

6 Extremality and the Heston-Nandi model

In this section, we consider a particular, commonly used model for asset prices — the Heston-Nandi model — and show that it can have particularly bad implications for the pricing of variance options.

The Heston-Nandi model [26] is the common Heston stochastic volatility model [25], where the correlation ρ between the Brownian motions driving the

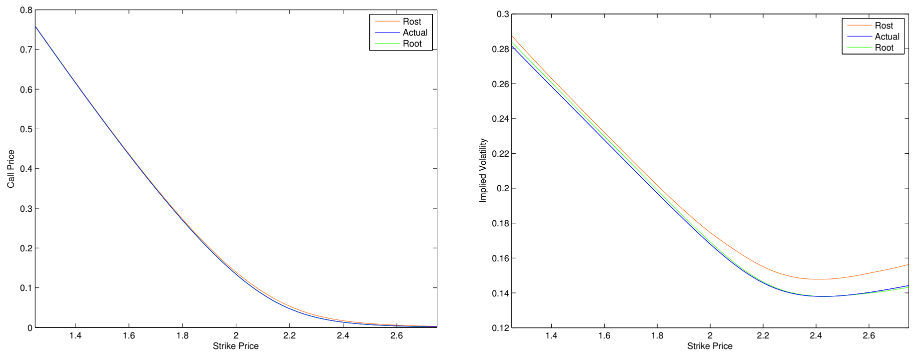

Figure 2: The original call prices from which we obtained our barrier, and the empirical call prices obtained by simulation for the Rost and Root barriers (left); The implied volatility of the call prices (right). Note that numerically the Rost barrier proves harder to correctly compute/simulate.

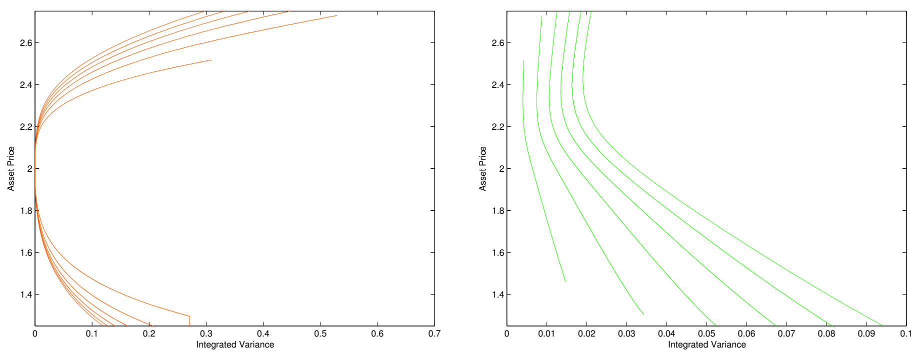

Figure 3: We compare the barriers for multiple maturities. In this figure we compute the barriers at equal spaced maturities of the underlying Heston model (the last barrier corresponding to $T = 1$) for the Rost (left) and Root (right) cases.

asset and the volatility processes is taken to be -1. Since in some asset classes, $\rho \approx -1$ is not abnormal (e.g. a Heston model calibrated to the S&P 500 typically has $\rho \approx -0.9 \pm 0.1$, for example, see Guillaume and Schoutens [24]), and the pricing of options on variance in Heston models is also common practice [30], this may substantially bias the prices towards the extreme models.

The Heston model is given (under the risk-neutral measure) by:

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sqrt{v_t}S_t\,\mathrm{d}B_t, \\ \mathrm{d}v_t &= \kappa(\theta - v_t)\,\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}\widetilde{B}_t, \end{aligned} \tag{6.1}$$

where B_t and $\widetilde{B}_t$ are Brownian motions with correlation ρ. The Heston-Nandi model is the restricted case where $\rho = -1$, and so $\widetilde{B}_t = -B_t$. Note that $v_t = \sigma_t^2$ in our previous notation, so we are interested in options on $\int_0^T v_t\,\mathrm{d}t$

The simplification $\rho = -1$ allows for the following observation: using Itô's

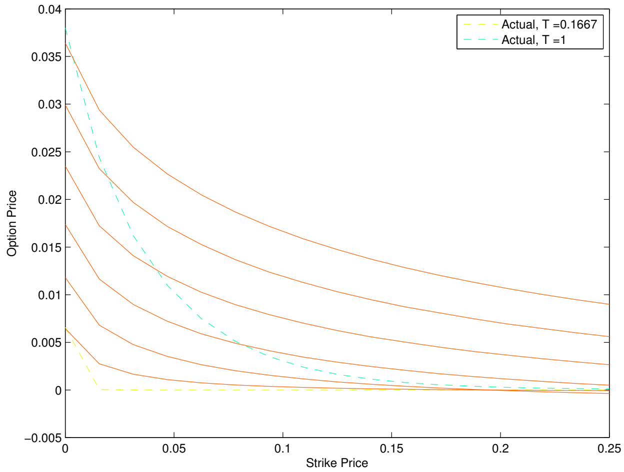

Figure 4: A plot of the upper bound on the price of a Variance call for different maturities (equally spaced, up to $T = 1$). For comparison, we also plot the actual prices of the variance calls under the Heston model corresponding to the shortest and longest maturities.

Lemma, we know

$$\begin{aligned} \mathrm{d}(\log(\mathrm{e}^{-rt}S_t)) &= -\frac{1}{2}v_t\,\mathrm{d}t + \sqrt{v_t}\,\mathrm{d}B_t \\ &= \left(\frac{\kappa\theta}{\xi} - \left(\frac{\kappa}{\xi}+\frac{1}{2}\right)v_t\right)\,\mathrm{d}t - \frac{1}{\xi}\mathrm{d}v_t. \end{aligned}$$

Solving, we see that

$$\log\left(\frac{\mathrm{e}^{-rT}S_T}{S_0}\right) = \frac{1}{\xi}(v_0 - v_T) + \frac{\kappa\theta}{\xi}T - \left(\frac{\kappa}{\xi}+\frac{1}{2}\right)\int_0^T v_t\,\mathrm{d}t. \tag{6.2}$$

If we assume that the maturity time of our option, T, is sufficiently large, since v_T is mean reverting, $(v_T - v_0) \approx (\theta - v_0)$ will be small in relation to the other terms on the right-hand side.

If we temporarily ignore the $v_T - v_0$ term, (6.2) tells us that, at time T, we have

$$\int_0^T v_t\,\mathrm{d}t \approx \left(\frac{\kappa}{\xi}+\frac{1}{2}\right)^{-1}\left[\log\left(\frac{S_0}{S_T}\right) + \frac{\kappa\theta}{\xi}T\right].$$

Writing

$$R_T(x) = \left(\frac{\kappa}{\xi}+\frac{1}{2}\right)^{-1}\left[\log\left(\frac{S_0}{x}\right) + \frac{\kappa\theta}{\xi}T\right], \tag{6.3}$$

then we have

$$T \approx \inf\left\{s \geq 0 : \int_0^s v_t\,\mathrm{d}t \geq R_T(X_s)\right\}.$$

This describes a barrier stopping time, corresponding to a Root stopping time, with

$$D_T = \left\{(x,t) \in \mathbb{R}\times\mathbb{R}_+ : t < R_T(x)\right\}. \tag{6.4}$$

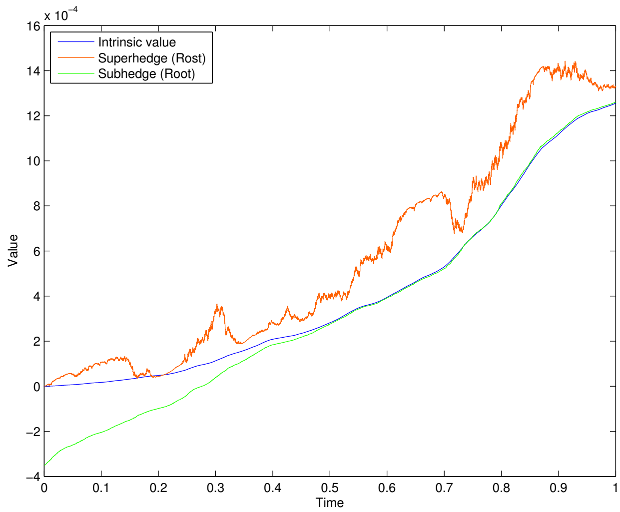

Figure 5: A plot of a realisation of the intrinsic value of the option $(F(\langle \ln S \rangle_t)$, and the value of the super-hedging and sub-hedging portfolios at time t, for $t \in [0, T]$. Here $F(t) = t(t - v_K)$, where $v_K = 0.01875$.

So, ignoring the term in v_T, we might conjecture that the corresponding model minimises the value of a derivative which is a convex, increasing function of v_T over all models with the same law at time T.

This leads to the following result:

Theorem 6.1. *Let $M > 0$ and suppose $\xi, \theta, \kappa, r > 0, \xi \neq 2\kappa$ are given parameters of a Heston-Nandi model, $\mathbb{Q}^{HN}$. Suppose $\mathcal{Q}_T$ is the class of models $\mathbb{Q}$ satisfying Assumption 2.1 and $\mathbb{E}^{\mathbb{Q}^{HN}}(S_T - K)_+ = \mathbb{E}^{\mathbb{Q}}(S_T - K)_+$ for all $K \geq 0$.*

Then there exists a constant κ, depending only on M and the parameters of the Heston-Nandi model, such that for all convex, increasing functions $F(t)$ with suitably smooth derivative $f(t) = F'(t)$ such that $f(t), f'(t) \leq M^*$, and for all $T \geq 0$

$$\mathbb{E}^{\mathbb{Q}^{HN}} F\left(\langle \log S \rangle_T\right) \leq \inf_{\mathbb{Q} \in \mathcal{Q}_T} \mathbb{E}^{\mathbb{Q}} F\left(\langle \log S \rangle_T\right) + \kappa. \tag{6.5}$$

Note that the strength of the result depends on the fact that the constant κ is independent of both T and F. In particular, $\langle \log S \rangle_T$ should be both growing in T and increasing in variance as T increases. That this does not appear in the bound leads us to claim that $\mathbb{Q}^{HN}$ is *asymptotically optimal.*

In fact, the continuity assumptions can be trivially relaxed, and this leads to the simple corollary:

Corollary 6.2. *The conclusions of Theorem 6.1 hold where the class of functions F considered is the set of variance call payoffs: $F_K(t) = (t - K)_+$ for all $K \in \mathbb{R}^+$ and all maturity dates $T > 0$.*

Proof. For fixed M^*, the function $F_K(t)$ can be approximated uniformly from above and below by a suitably smooth function satisfying the conditions of Theorem 6.1, independent of K. The result follows. $\square$

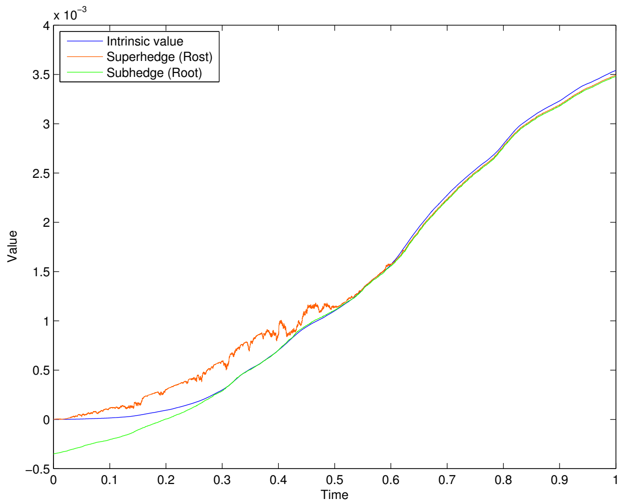

Figure 6: A plot of a realisation of the intrinsic value of the option, as in Figure 5, and the corresponding sub- and super-hedges, where the realisation is taken from a different model to that under which the original hedge was constructed.

The above result demonstrates that the seemingly strong assumptions on the function $F(t)$ required in Theorem 6.1 are not a big restriction: by allowing a slightly larger constant, we can consider the class of functions which can be approximated by such functions uniformly. As we will see, the exact smoothness requirements on $f(t)$ are that $f(t)$ has a Hölder continuous second derivative (although we believe that this assumption could be relaxed).

Our arguments rely on the construction of a barrier, and the proof of optimality described in [11]. We recall some important definitions here. We suppose that we are given a barrier function $R_T(x)$ as defined in (6.3), and consider the geometric Brownian motion $(\widetilde{X}_t)$ on this domain, with corresponding hitting time τ_{D_T}, where D_T is as defined in (6.4). Then we define the function

$$M(x,t) = \mathbb{E}^{(x,t)} f(\tau_{D_T}) \tag{6.6}$$

and observe that (under the assumptions of Theorem 6.1), we have $M(x,t)$ bounded. Since we consider the case where $\widetilde{X}$ is geometric Brownian motion, we can assume $\sigma(x) = x$ in the formulae from [11].

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

$$Z(x) = 2\int_{S_0}^{x}\int_{S_0}^{y} \frac{M(z,0)}{z^2}\,\mathrm{d}z\,\mathrm{d}y. \tag{6.7}$$

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

$$G(x,t) = \int_0^t M(x,s)\,\mathrm{d}s - Z(x), \tag{6.8}$$

and

$$H(x) = \int_0^{R_T(x)} (f(s) - M(x,s))\,\mathrm{d}s + Z(x). \tag{6.9}$$

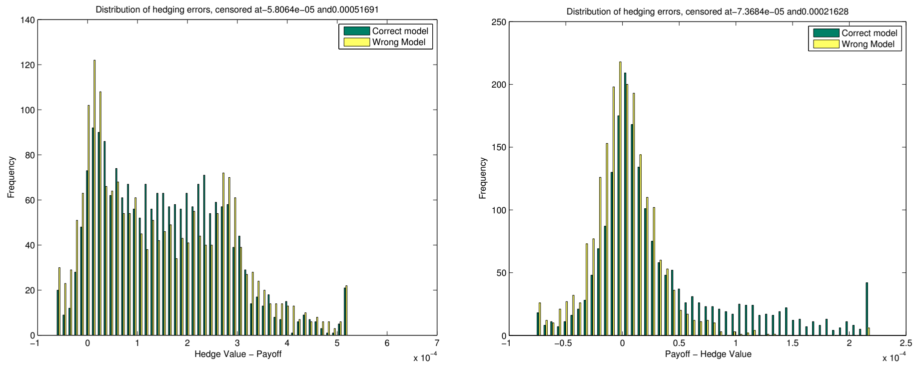

Figure 7: The (empirically computed) distribution of the hedging error for the super-hedge (left) and sub-hedge (right). In the former case, a positive error represents a surplus in the hedge, while in the latter case, a positive error represents an underhedge. We also compare the tracking error with the tracking error for the same strategy when our realised path is determined by the mis-specified model of Figure 6.

Then (Cox and Wang [11, Proposition 5.1]) for all $(x,t) \in \mathbb{R}_+ \times \mathbb{R}_+$:

$$G(x,t) + H(x) \leq F(t) \tag{6.10}$$

with equality when $t = R_T(x)$. In addition, if for any $T > 0$:

$$\mathbb{E}\left[\int_0^T Z'(\widetilde{X}_s)^2 \sigma(\widetilde{X}_s)^2 \,\mathrm{d}s\right] < \infty, \quad \mathbb{E}Z(\widetilde{X}_0) < \infty, \tag{6.11}$$

then the process

$$G(\widetilde{X}_{t\wedge\tau_{D_T}}, t\wedge\tau_{D_T}) \text{ is a martingale,} \tag{6.12}$$

and

$$G(\widetilde{X}_t, t) \text{ is a submartingale.} \tag{6.13}$$

We collect some useful properties of these functions in the following lemma:

Lemma 6.3. *Under the assumptions of Theorem 6.1, the functions $Z(x), H(x)$ and $G(x,t)$ as defined above have the following properties:*

(i). $|\frac{\partial G}{\partial t}(x,t)| \leq M^*$ *for all* $(x,t) \in \mathbb{R}_+ \times \mathbb{R}_+$.

(ii). $G(\widetilde{X}_t, t)$ *is a submartingale, with decomposition:*

$$G(\widetilde{X}_t, t) = G(\widetilde{X}_0, 0) + \int_0^t \widetilde{X}_s \left(\int_0^{R_T(\widetilde{X}_s)\wedge t} \frac{\partial M}{\partial x}(\widetilde{X}_s, r)\,\mathrm{d}r - Z'(\widetilde{X}_s)\right) \mathrm{d}\widetilde{X}_s$$
$$- \int_0^t \gamma(\widetilde{X}_s)\mathbf{1}_{\{s>R_T(\widetilde{X}_s)\}}\,\mathrm{d}s,$$

where

$$\gamma(x) = f(R_T(x)) - \frac{1}{2}x^2\frac{\partial M}{\partial x}(x, R_T(x)-)R'_T(x) \geq 0$$

is a bounded function.

Proof of Theorem 6.1. Let v_t be the squared volatility process for the Heston-Nandi price process S_t, and suppose we fix $T > 0$ (although we will want our constants to be independent of T). Define the time-change process $\tau_T = \int_0^T v_t \,\mathrm{d}t$, and let A_t be the right-inverse of τ_t. In particular, if we define as usual $\widetilde{X}_t = \mathrm{e}^{-rA_t} S_{A_t}$, then $\widetilde{X}_t$ is a geometric Brownian motion with fixed law μ_T at time τ_T. Using (6.2) in (6.3) we get:

$$R_T(\widetilde{X}_{\tau_T}) = \int_0^T v_s \,\mathrm{d}s + \frac{1}{\xi}\left(\frac{\kappa}{\xi} + \frac{1}{2}\right)^{-1} (v_T - v_0). \tag{6.14}$$

Since the variance process v_s is mean reverting $\mathbb{E}|v_T - v_0|$, can be bounded uniformly for all T by some constant depending only on the parameters of the model and so in particular, there exists a constant κ_1 such that

$$\mathbb{E}|R_T(\widetilde{X}_{\tau_T}) - \tau_T| < \kappa_1.$$

From the bound on $f(t)$, it then follows that:

$$\mathbb{E}[|F(\tau_T) - F(R_T(\widetilde{X}_{\tau_T}))|] \leq M^* \kappa_1. \tag{6.15}$$

Similarly, using Lemma 6.3.*(i)*, we get

$$\mathbb{E}[|G(\widetilde{X}_{\tau_T}, \tau_T) - G(\widetilde{X}_{\tau_T}, R_T(\widetilde{X}_{\tau_T}))|] \leq M^* \kappa_1.$$

In addition, using the decomposition from Lemma 6.3.*(ii)*, and noting that $\gamma(x)$ is bounded above by a constant, κ_2 say, we have:

$$\mathbb{E}G(\widetilde{X}_{\tau_T}, \tau_T) \leq G(\widetilde{X}_0, 0) + \kappa_2 \mathbb{E}\left[\int_0^{\tau_T} \mathbf{1}_{\{s > R_T(\widetilde{X}_s)\}} \,\mathrm{d}s\right].$$

Observe from the definition of τ_t, (6.3) and (6.2) evaluated at a general time $t = A_s$:

$$\begin{aligned} \{R_T(\widetilde{X}_s) < s\} &= \left\{\left(\frac{\kappa}{\xi} + \frac{1}{2}\right)^{-1}\left[\frac{\kappa\theta}{\xi}(T - A_s) + \frac{1}{\xi}(v_{A_s} - v_0)\right] \leq 0\right\} \\ &= \left\{A_s \geq \frac{\kappa\theta T + v_{A_s} - v_0}{\kappa\theta}\right\} \subseteq \left\{A_s \geq \frac{\kappa\theta T - v_0}{\kappa\theta}\right\}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}\left[\int_0^{\tau_T} \mathbf{1}_{\{s > R_T(\widetilde{X}_s)\}} \,\mathrm{d}s\right] &\leq \mathbb{E}\left[\int_0^{\tau_T} \mathbf{1}_{\{A_s \geq T - \frac{v_0}{\kappa\theta}\}} \,\mathrm{d}s\right] = \mathbb{E}\left[\int_{\tau_{T - \frac{v_0}{\kappa\theta}}}^{\tau_T} \,\mathrm{d}s\right] \\ &= \mathbb{E}\left[\tau_T - \tau_{T - \frac{v_0}{\kappa\theta}}\right] = \mathbb{E}\left[\int_{T - \frac{v_0}{\kappa\theta}}^{T} v_s \,\mathrm{d}s\right]. \end{aligned}$$

Again, since v_s is mean reverting, the right-hand-side can be bounded independently of T, and so

$$\mathbb{E}\left[G(\widetilde{X}_{\tau_T}, \tau_T)\right] \leq G(\widetilde{X}_0, 0) + \kappa_2 \tag{6.16}$$

for some constant κ_2.

Now, using (6.15) and the fact that (6.10) holds with equality when $t = R_T(x)$, we have:

$$\begin{aligned}
\mathbb{E}\left[F(\tau_T)\right] &\leq \mathbb{E}\left[F(R_T(\widetilde{X}_{\tau_T}))\right] + M^*\kappa_1 \\
&= \mathbb{E}\left[G(\widetilde{X}_{\tau_T}, R_T(\widetilde{X}_{\tau_T}))\right] + \mathbb{E}\left[H(\widetilde{X}_{\tau_T})\right] + M^*\kappa_1 \\
&= \mathbb{E}\left[G(\widetilde{X}_{\tau_T}, \tau_T)\right] + \mathbb{E}\left[H(\widetilde{X}_{\tau_T})\right] + 2M^*\kappa_1 \\
&\leq G(\widetilde{X}_0, 0) + \mathbb{E}\left[H(\widetilde{X}_{\tau_T})\right] + 2M^*\kappa_1 + \kappa_2
\end{aligned}$$

It remains for us to show that $\mathbb{E}\left[F(\widetilde{X}_\sigma)\right] \geq G(\widetilde{X}_0, 0) + \mathbb{E}\left[H(\widetilde{X}_{\tau_T})\right] = G(\widetilde{X}_0, 0) + \mathbb{E}\left[H(\widetilde{X}_\sigma)\right]$ for any stopping time σ with $\widetilde{X}_\sigma \sim \widetilde{X}_{\tau_T}$.

We consider a localising sequence, $\sigma_N \uparrow \sigma$, and note that we then have: $\mathbb{E}\left[G(\widetilde{X}_{\sigma_N}, \sigma_N)\right] \geq G(\widetilde{X}_0, 0)$, since G is a submartingale, and in addition, $\mathbb{E}\left[F(\sigma_N)\right] \uparrow \mathbb{E}\left[F(\sigma)\right]$ since $F(\cdot)$ is increasing. On account of (6.10), it remains only to show $\mathbb{E}\left[H(\widetilde{X}_\sigma)\right] \leq \mathbb{E}\left[H(\widetilde{X}_{\sigma_N})\right]$. We first observe that f is an increasing and bounded function, and if $f(t) = f(\infty)$ for all $t \geq t_0$, for some $t_0 \in \mathbb{R}_+$, $M(x,t) = f(t)$ for all $t \geq t_0$. Since also $Z(x) \geq 0$ ($Z(x)$ is convex with $Z(S_0) = Z'(S_0) = 0$), we must have $H(x)$ bounded below. We can therefore apply Fatou's Lemma to deduce $\mathbb{E}\left[H(\widetilde{X}_\sigma)\right] \leq \mathbb{E}\left[H(\widetilde{X}_{\sigma_N})\right]$. To remove the assumption on $f(t)$, we observe that, by smoothly truncating f, we can approximate F from below by an increasing sequence F_N of functions which each have constant derivative, and such that $\mathbb{E}\left[F_N(\widetilde{X}_{\tau_T})\right] \uparrow \mathbb{E}\left[F(\widetilde{X}_{\tau_T})\right]$. Since each approximation satisfies the bound, the same must be true in the limit.

□

Proof of Lemma 6.3. We first show *(i)*. Observe that $M(x, R_T(x)) = f(R_T(x))$, so in particular, M is continuous, and $f(t) \leq M(x,t) \leq M^*$ since $M(x,t)$ is increasing in t. It follows immediately that $\frac{\partial G}{\partial t}(x,t) = M(x,t)$ is continuous and bounded, and in fact, is non-negative.

For *(ii)* we aim to use Peskir [41, Theorem 3.1]. We note that $M(x,t)$ is $C^{2,1}$ in D_T since it is a martingale (in particular, M is the (unique bounded) solution to a parabolic initial-value boundary problem). In fact, by Lieberman [35, Theorem 5.14], if we assume that $f''(t)$ is bounded in a Hölder norm, it follows that $M(x,t)$ has Hölder-bounded first and second spatial derivatives, and first time derivative. It is easy to check that $G(x,t) = \int_0^t M(x,s)\,\mathrm{d}s - Z(x)$ is also $C^{2,1}$ in D_T as a consequence. Moreover, computing explicitly, we see that $\frac{\partial G}{\partial x}(x,t) = \int_0^t \frac{\partial M}{\partial x}(x,s)\,\mathrm{d}s - Z'(x)$ is also continuous on $\bar{D}_T$. If we write $C = \{(x,t) : R_T(x) < t\}$, so $\bar{C} \cup \bar{D}_T = \mathbb{R}_+ \times \mathbb{R}_+$ with boundary $R_T(x) = t$, we see that on C,

$$G(x,t) = \int_0^{R_T(x)} M(x,s)\,\mathrm{d}s + \int_{R_T(x)}^t f(s)\,\mathrm{d}s - Z(x)$$

and again, G is $C^{2,1}$ in C, and $\frac{\partial G}{\partial x}(x,t)$ is continuous on $\bar{C}$.

Considering the conditions required for Theorem 3.1 of [41], we observe that (3.18), (3.19), (3.26), (3.30) and (3.33) of [41] have now been shown, and so the theorem holds. Moreover, since the first spatial derivative of G is continuous across the boundary, we do not get a local-time term on the boundary. Computing $\frac{\partial G}{\partial t}(x,t) + \frac{1}{2}\sigma(x)^2 \frac{\partial^2 G}{\partial x^2}(x,t)$ results in the expression for $\gamma(x)$ stated, and we observe that the boundary function $R_T(x)$ is a decreasing function of x, which implies in turn that $\frac{\partial M}{\partial x}(x,t)$ is positive at the boundary (since f is increasing), so $\gamma(x) \geq 0$.

We finally show that $\gamma(x)$ is bounded. By assumption, f is bounded, so we need only consider the second term. We have $M(x, R_T(x)) = f(R_T(x))$, and differentiating (recall that the derivatives of $M(x,t)$ on D_T are Hölder continuous, and so extend continuously to the boundary) and rearranging we get:

$$\frac{\partial M}{\partial x}(x, R_T(x)) = R_T'(x) \left(f'(R_T(x)) - \frac{\partial M}{\partial t}(x, R_T(x)) \right)$$

Observing that (via a standard coupling argument)

$$\mathbb{P}^{(x,t)}(\tau_{D_T} > t) \geq \mathbb{P}^{(x,t+\delta t)}(\tau_{D_T} > t + \delta t)$$

whenever $(x,t) \in D_T$ (the later path sees a 'bigger' stopping region). It follows that $\mathbb{E}^{(x,t+\delta t)}[f(\tau_{D_T})] \leq \mathbb{E}^{(x,t)}[f(\tau_{D_T} + \delta t)]$, and therefore that $\frac{\partial M}{\partial t}(x,t) \leq \sup_{t\in\mathbb{R}_+} f'(t) \leq M^*$.

Recalling finally that $\sigma(x) = x$, and observing that $R_T'(x) = -(\kappa/\xi + 1/2)^{-1}x^{-1}$, we conclude that $\gamma(x)$ is indeed bounded. $\square$

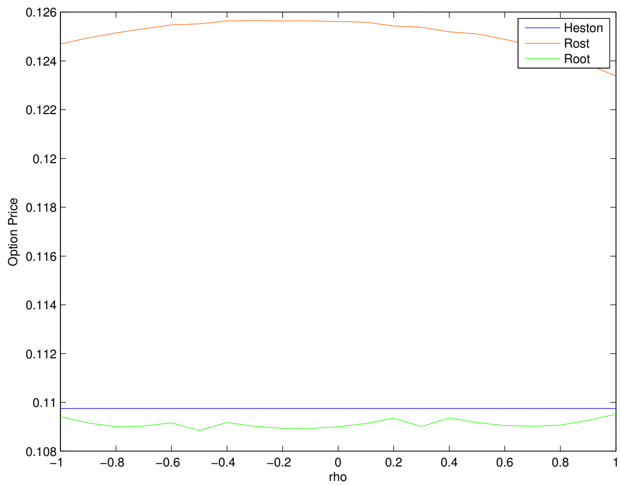

Figure 8: The upper and lower model-independent bounds on the price of a variance call, plotted as functions of the correlation between the asset and volatility processes. The constant line represents the price of the variance call under the Heston model — this is constant, since the price is unaffected by the choice of ρ.

We finish with some numerical evidence to support our conjecture. Figure 8 shows the upper and lower bounds on the price of a variance call, as determined using the numerical methods of Section 5, seen as a function of the parameter ρ. In this example, we use the same parameters as before, but with $T = 4$. It is notable that the lower bound and the price arising from the Heston model are certainly close. It is also interesting to observe that it is not only the Heston-Nandi model that seems to be close to extremal; rather this seems to be a more general property of the Heston model. A good explanation of this fact eludes us, but better understanding of this behaviour would appear to be both practically relevant, and theoretically interesting.

References

[1] M. Avellaneda, A. Levy, and A. Parás. "Pricing and hedging derivative securities in markets with uncertain volatilities". *Applied Mathematical Finance* 2.2 (1995), pp. 73–88. DOI: 10.1080/13504869500000005.

[2] J. Azéma and M. Yor. "Une solution simple au problème de Skorokhod". *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78).* Vol. 721. 1979, pp. 90–115.

[3] G. Barles and P. Souganidis. "Convergence of approximation schemes for fully nonlinear second order equations". *Asymptotic Analysis* 4.3 (1991), pp. 271–283. DOI: 10.3233/ASY-1991-4305.

[4] G. Barles and E. Jakobsen. "Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations". *Mathematics of Computation* 76.260 (2007), pp. 1861–1893. DOI: 10.1090/S0025-5718-07-02000-5.

[5] M. Beiglböck, P. Henry-Labordère, and F. Penkner. "Model-independent bounds for option prices—a mass transport approach". *Finance and Stochastics* 17.3 (2013), pp. 477–501. DOI: 10.1007/s00780-013-0205-8.

[6] R. M. Blumenthal and R. K. Getoor. *Markov processes and potential theory.* Academic Press, 1968, pp. x+313.

[7] D. T. Breeden and R. H. Litzenberger. "Prices of State-contingent Claims Implicit in Option Prices". *Journal of Business* 51.4 (1978), pp. 621–651.

[8] P. Carr and R. Lee. "Hedging variance options on continuous semimartingales". *Finance and Stochastics* 14.2 (2010), pp. 179–207. DOI: 10.1007/s00780-009-0110-3.

[9] R. M. Chacon. "Barrier stopping times and the filling scheme". PhD thesis. University of Washington, 1985, p. 58.

[10] A. M. G. Cox and J. Obłój. "Robust pricing and hedging of double no-touch options". *Finance and Stochastics* 15.3 (2011), pp. 573–605. DOI: 10.1007/s00780-011-0154-z.

[11] A. M. G. Cox and J. Wang. "Root's barrier: Construction, optimality and applications to variance options". *The Annals of Applied Probability* 23.3 (2013), pp. 859–894. DOI: 10.1214/12-AAP857.

[12] A. M. G. Cox, D. Hobson, and J. Obłój. "Pathwise inequalities for local time: Applications to Skorokhod embeddings and optimal stopping". *The Annals of Applied Probability* 18.5 (2008), pp. 1870–1896. DOI: 10.1214/07-AAP507.

[13] A. M. G. Cox and D. G. Hobson. "Local martingales, bubbles and option prices". *Finance and Stochastics* 9.4 (2005), pp. 477–492. DOI: 10.1007/s00780-005-0162-y.

[14] A. M. G. Cox and J. Obłój. "Robust Hedging of Double Touch Barrier Options". *SIAM Journal on Financial Mathematics* 2.1 (2011), pp. 141–182. DOI: 10.1137/090777487.

[15] A. M. G. Cox and G. Peskir. "Embedding Laws in Diffusions by Functions of Time". 2012. URL: http://arxiv.org/abs/1201.5321.

[16] M. Davis, J. Obłój, and V. Raval. "Arbitrage Bounds for Prices of Weighted Variance Swaps". *Mathematical Finance* (2013). DOI: 10.1111/mafi.12021.

[17] Y. Dolinsky and H. M. Soner. "Martingale Optimal Transport and Robust Hedging in Continuous Time". SSRN Scholarly Paper ID 2245481. 2013. URL: http://papers.ssrn.com/abstract=2245481.

[18] B. Dupire. "Arbitrage Bounds for Volatility Derivatives as Free Boundary Problem". Presentation at 'PDE and Mathematical Finance', KTH, Stockholm. 2005.

[19] B. Dupire. "Model art". *Risk* 6.9 (1993), pp. 118–120.

[20] P. J. Fitzsimmons. "Markov Processes with Identical Bridges". *Electronic Journal of Probability* 3 (1998). DOI: 10.1214/EJP.v3-34.

[21] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Second. Vol. 25. Springer, 2006, pp. xviii+429.

[22] H. Föllmer. "Calcul d'Itô sans probabilités". *Séminaire de Probabilités XV 1979/80*. Vol. 850. 1981, pp. 143–150.

[23] A. Galichon, P. Henry-Labordere, and N. Touzi. "A Stochastic Control Approach to No-Arbitrage Bounds Given Marginals, with an Application to Lookback Options". To Appear, *The Annals of Applied Probability*. 2011. URL: http://ssrn.com/abstract=1912477.

[24] F. Guillaume and W. Schoutens. "Calibration risk: Illustrating the impact of calibration risk under the Heston model". *Review of Derivatives Research* 15 (1 2012). 10.1007/s11147-011-9069-2, pp. 57–79.

[25] S. L. Heston. "A closed-form solution for options with stochastic volatility with applications to bond and currency options". *The Review of Financial Studies* 6.2 (1993), pp. 327–343.

[26] S. Heston and S. Nandi. "Preference-Free Option Pricing with Path-Dependent Volatility: A Closed-Form Approach". Federal Reserve Bank of Atlanta Working Paper, 98-20. 1998.

[27] D. Hobson. "The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices". *Paris-Princeton Lectures on Mathematical Finance 2010*. Vol. 2003. 2010, pp. 267–318.

[28] D. G. Hobson. "Robust hedging of the lookback option". *Finance and Stochastics* 2.4 (1998), pp. 329–347. DOI: 10.1007/s007800050044.

[29] D. Hobson and M. Klimmek. "Model-independent hedging strategies for variance swaps". *Finance and Stochastics* 16.4 (2012), pp. 611–649. DOI: 10.1007/s00780-012-0190-3.

[30] A. Javaheri, P. Wilmott, and E. G. Haug. “GARCH and volatility swaps”. *Quantitative Finance* 4.5 (2004), pp. 589–595. DOI: 10.1080/14697680400008700.

[31] R. L. Karandikar. “On pathwise stochastic integration”. *Stochastic Processes and their Applications* 57.1 (1995), pp. 11–18. DOI: 10.1016/0304-4149(95)00002-O.

[32] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Second. Vol. 113. Springer-Verlag, 1991, pp. xxiv+470.

[33] R. Lee. “Gamma Swap”. *Encyclopedia of Quantitative Finance*. 2010.

[34] R. Lee. “Weighted Variance Swap”. *Encyclopedia of Quantitative Finance*. 2010.

[35] G. M. Lieberman. *Second Order Parabolic Differential Equations*. World Scientific, 1996.

[36] K. G. Murty. *Linear Complementarity, Linear and Non Linear Programming*. Heldermann Verlag, 1988.

[37] A. Neuberger. “The Log Contract”. *The Journal of Portfolio Management* 20.2 (1994), pp. 74–80. DOI: 10.3905/jpm.1994.409478.

[38] A. Neufeld and M. Nutz. “Superreplication under volatility uncertainty for measurable claims”. *Electronic Journal of Probability* 18 (2013). DOI: 10.1214/EJP.v18-2358.

[39] H. Oberhauser and G. dos Reis. “Root’s barrier, viscosity solutions of obstacle problems and reflected FBSDEs”. 2013. URL: http://arxiv.org/abs/1301.3798.

[40] J. Obłój. “The Skorokhod embedding problem and its offspring”. *Probability Surveys* 1 (2004), 321–390 (electronic). DOI: 10.1214/154957804100000060.

[41] G. Peskir. “A Change-of-Variable Formula with Local Time on Curves”. *Journal of Theoretical Probability* 18.3 (2005), pp. 499–535. DOI: 10.1007/s10959-005-3517-6.

[42] D. Possamaï, G. Royer, and N. Touzi. “On the Robust superhedging of measurable claims”. 2013. URL: http://arxiv.org/abs/1302.1850.

[43] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Third. Vol. 293. Springer-Verlag, 1999, pp. xiv+602.

[44] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Itô calculus, Reprint of the second (1994) edition. Cambridge University Press, 2000, pp. xiv+480.

[45] D. H. Root. “The existence of certain stopping times on Brownian motion”. *Annals of Mathematical Statistics* 40 (1969), pp. 715–718.

[46] H. Rost. “The stopping distributions of a Markov Process”. *Inventiones Mathematicae* 14 (1971), pp. 1–16.