# The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices

David Hobson

July 20, 2009

### Abstract

This set of lecture notes is concerned with the following pair of ideas and concepts:

1) The Skorokhod Embedding problem (SEP) is, given a stochastic process $X = (X_t)_{t \geq 0}$ and a measure $\mu$ on the state space of $X$, to find a stopping time $\tau$ such that the stopped process $X_\tau$ has law $\mu$. Most often we take the process $X$ to be Brownian motion, and $\mu$ to be a centred probability measure.

2) The standard approach for the pricing of financial options is to postulate a model and then to calculate the price of a contingent claim as the suitably discounted, risk-neutral expectation of the payoff under that model. In practice we can observe traded option prices, but know little or nothing about the model. Hence the question arises, if we know vanilla option prices, what can we infer about the underlying model?

If we know a single call price, then we can calibrate the volatility of the Black-Scholes model (but if we know the prices of more than one call then together they will typically be inconsistent with the Black-Scholes model). At the other extreme, if we know the prices of call options for all strikes and maturities, then we can find a unique martingale diffusion consistent with those prices.

If we know call prices of all strikes for a single maturity, then we know the marginal distribution of the asset price, but there may be many martingales with the same marginal at a single fixed time. Any martingale with the given marginal is a candidate price process. On the other hand, after a time change it becomes a Brownian motion with a given distribution at a random time. Hence there is a 1-1 correspondence between candidate price processes which are consistent with observed prices, and solutions of the Skorokhod embedding problem.

These notes are about this correspondence, and the idea that extremal solutions of the Skorokhod embedding problem lead to robust, model independent prices and hedges for exotic options.

## 1 Motivation

Consider the problem of pricing and hedging a one-touch digital option on an underlying. The standard approach to such a problem is to postulate a stochastic

model for the underlying, perhaps the Osborne-Samuelson-Black-Scholes (exponential Brownian motion) model, and to price the option as the discounted expectation under the risk-neutral measure. In perfect frictionless markets this approach is justified by the theory of replication, and allows either counter-party in a transaction to eliminate the market risk.

However, the success of a replicating strategy is predicated on the fundamental truth of the model on which it is based. Although market risk (the known unknown) is eliminated, model risk (the unknown unknown) remains. The classical hedging strategies remove hedging risk, but leave agents exposed to Knightian uncertainty.

If the problem at issue is to price and hedge the exotic digital option then it seems reasonable to assume that simpler, vanilla options (such as call options) would also be traded. Then, at the very least, the volatility in the exponential Brownian motion can be calibrated with reference to the price of a traded call. However, this leads to a potential inconsistency, since many calls may be traded, each with their distinct volatilities. Ideally we should use a model which calibrates perfectly to the full spectrum of traded calls. However, in principle there are many such models, and associated with each model which is consistent with the market prices of (liquidly) traded options, there may be a different price for the exotic. Instead, one might attempt to characterise the class of models which are consistent with the market prices of options. This is a very challenging problem, and a less ambitious target is to characterise the extremal elements of this set, and especially those models for which the price of the exotic is maximised or minimised.

Suppose the one-touch digital option is written on a forward price $(S_t)_{0 \leq t \leq T}$, and that the payoff is given by $\tilde{F} = I_A$ where $I$ is the indicator function and $A = \{S_t \geq B, \text{ for some } t \in [0, T]\}$. Here the payoff is made at time $T$, we take 0 to be the current time and we assume that the barrier $B$ is above the initial price $B > S_0$, and that $S$ is right continuous. If we write $H_B$ for the first time the underlying reaches the barrier then we have $H_B = \inf\{u > 0 : S_u \geq B\}$, and $\tilde{F} = I_{\{H_B \leq T\}}$.

The key observation is contained in the following inequality which is valid for any $K < B$:

$$I_{\{H_B \leq T\}} \leq \frac{(S_T - K)^+}{B - K} + \frac{(S_{H_B} - S_T)}{B - K} I_{\{H_B \leq T\}}. \tag{1}$$

Note that (1) is a path-wise inequality. The left-hand-side is the payoff of the option. The two terms on the right-hand-side have simple financial interpretations as the payoff from $1/(B - K)$ call options with strike $K$ (with maturity taken to match that of the exotic), and the gains from trade from a forward transaction in $S$, struck the first time, if ever, that the underlying crosses the barrier. Provided the call with strike $K$ is traded, and provided it is possible to invest forward in $S$, then the right-hand-side of (1) describes a super-replicating strategy for the one-touch-digital; furthermore by no-arbitrage it determines an upper bound on the price of the one-touch-digital given by $C(K)/(B - K)$, where $C(k)$ is the traded price of a call option on the forward strike with $k$. (By definition, the forward transaction is costless).

Since the strike $K$ in (1) is arbitrary we can optimise over $K$. If we write the

arbitrage-free price of the one-touch digital as $\mathcal{P}(\tilde{F})$ then

$$\mathcal{P}(\tilde{F}) \leq \inf \frac{C(K)}{B-K} \tag{2}$$

where the infimum is taken over traded strikes $K < B$.

In deriving this bound the only assumption that has been used is that it is possible to generate constant multiples of the payoff $(S_{H_B} - S_T)I_{\{H_B \leq T\}}$ at zero cost, which we interpret as an investment in the forward market. Note that we have imposed no probabilistic structure, we have not written down a stochastic model $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ nor have we postulated the existence of a martingale measure $\mathbb{Q}$. Instead, the bound is based on the ability to hedge using simple strategies in the forward market, and to take a static initial positions in vanilla securities at their prevailing time-0 market price. In this sense the resulting bound is model-independent. The explicit modelling of the dynamics of $S$ has been replaced by the requirement that the model is consistent with observed call prices, (which places implicit conditions on the stochastic properties of the underlying price process) and $\delta$-hedging is replaced by super-hedging with a portfolio of call options.

The quantity $\overline{D} = \inf_{K<B} \frac{C(K)}{B-K}$ gives a model-independent bound on the price of a digital option, in the sense that if we introduce any model for the underlying for which call prices are given by the pricing function $C(k)$, then the discounted expected payoff of the option under (any) martingale measure is less than $\overline{D}$.

The question immediately arises: is this bound best possible? We will show in Section 2.7 below, by demonstrating a model which matches call prices and for which the price of the one-touch digital is equal to the bound, that this is indeed the case, and the bound is tight. (The model we give may not be realistic, but unless it can be ruled out, the bound $\overline{D}$ cannot be refined.)

Return to the characterisation problem of describing the models which are consistent with observed option prices. It turns out that knowledge of the prices of puts and calls is equivalent to knowledge of the marginal distribution of the underlying asset under the measure used by the market for pricing. Further, if we assume temporarily that the market uses a pricing measure under which the forward price is a martingale, then the problem is to find a martingale with a given law at time $T$. However, any martingale is a time-change of Brownian motion. Hence the characterisation problem is reduced to the following: find stopping times such that the stopped Brownian motion has the given law. This is the classical Skorokhod embedding problem (SEP), first introduced (and solved) by Skorokhod [54].

The idea of relating the range of no-arbitrage prices of exotics to the prices of vanilla puts and calls via solutions of the Skorokhod embedding problem was first developed by Hobson [31] in the context of lookback options. The example above of a one-touch barrier option, which allows perhaps for the simplest exposition of the theory, is from Brown et al [10]. Recent applications include Cox and Oblój [18, 19] to two-sided barrier options, Hobson and Neuberger [35] to forward-starting straddles and Dupire [24] and Carr and Lee [12] to volatility options, but, in principle, the method applies to a wide range of path-dependent derivatives. The connection between model independent option price bounds and the Skorokhod embedding

3

problem can be exploited in both directions. Existing solutions to the SEP (see Oblój [43] for a comprehensive survey) can be used to give bounds on the prices of financial options; conversely questions about the range of no-arbitrage prices for path-dependent options motivates the search for further solutions of the SEP.

# 2 Model Independent Option Pricing

## 2.1 The classical approach

The standard approach in derivative pricing is to begin by postulating a model (or a parametric family of models) for the price process $(P_t)_{t \geq 0}$, supported on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The price of a contingent claim with payoff $F_T = F(P_t; 0 \leq t \leq T)$, payable at $T$, is calculated as a discounted expectation:

$$\mathbb{E}^{\mathbb{Q}}[e^{-rT} F_T] = \mathbb{E}^{\mathbb{Q}}[e^{-rT} F(P_t; 0 \leq t \leq T)].$$

Here $\mathbb{Q}$ is an equivalent (local) martingale measure. In a complete market, the rationale for this pricing rule is supported by the replication (martingale representation)

$$e^{-rT} F_T = \mathbb{E}^{\mathbb{Q}}[e^{-rT} F_T] + \int_0^T \theta_t (dP_t - r P_t dt), \qquad \mathbb{Q} \text{ a.s.}$$

(Note that since $\mathbb{Q}$ is equivalent to $\mathbb{P}$, any statement which holds $\mathbb{Q}$ almost surely, also holds $\mathbb{P}$ almost surely.) The quantity $\theta_t$ determines the investment or hedge in the risky asset.

For example, in the exponential Brownian motion parametric family of models we have that the asset price $P$ follows

$$\frac{dP}{P} = \sigma dW + \gamma dt$$

and then for a call option with strike $K$, maturity $T$ and payoff $F_T = (P_T - K)^+$ we have the price $\mathcal{C}(K, T; P_0; r, \gamma, \sigma)$, as given by the Black-Scholes formula.

Here the quantities $K$ and $T$ are characteristics of the derivative, $P_0$ is the observed initial value of the asset, $r$ is assumed known (or observed from bond prices), the drift $\gamma$ is irrelevant, since we price under the risk-neutral measure, and the volatility $\sigma$ is estimated from historical data.

When the derivative has a payoff which is convex in the the price process $P_t$ (for example, a call option) then the model price $\mathcal{C} = \mathcal{C}(\sigma)$ is increasing in $\sigma$. Hence in practice, the volatility is selected via a calibration process in which the price of a liquidly traded option $\mathcal{C}^{\text{traded}}$ is used to determine the implied volatility $\sigma_I$ as the solution to $\mathcal{C}^{\text{traded}} = \mathcal{C}(\sigma_I)$. In this sense the real purpose of the pricing function $\mathcal{C}(K, T)$ is as an extrapolation device. See Figlewski [26] for a discussion of other choices of extrapolation functions.

## 2.2 Model misspecification and incomplete markets

In a complete market the price $\mathbb{E}^{\mathbb{Q}}[e^{-rT} F_T]$ and the hedge (or replication strategy) $\theta_t$ are uniquely determined, and replication is perfect provided that the model provides

an exact description of reality.

Clearly, this last provision never holds. However, we have the following reassuring result, see Bergman et al [6], El Karoui et al [25], Hobson [32] and Janson and Tysk [37].

**Theorem 2.1** *Suppose that the goal is to price and hedge a claim on $P_T$ with convex payoff function $F_T = F(P_T)$. Suppose that the claim is priced and hedged under a diffusion model $dP_t = P_t \hat{\sigma}(P_t) dB_t + rP_t dt$ and that the model is sufficiently regular that the solution, denoted $\hat{P}$, is such that $(e^{-rt} \hat{P}_t)_{t \geq 0}$, is a true martingale. The model is Markovian, so that we may write the replicating hedge of the claim $F$ as $\hat{\theta}_t = \hat{\theta}(P_t, t)$.*

*Now suppose that the real-world dynamics are such that $dP_t = P_t \sigma_t dB_t + rP_t dt$ and that $\sigma_t \leq \hat{\sigma}(P_t)$. Then*

$$\hat{\mathbb{E}}^{\mathbb{Q}}[e^{-rT} F_T] + \int_0^T \hat{\theta}(P_t, t)(dP_t - rP_t dt) \geq e^{-rT} F(P_T), \qquad \mathbb{Q} \ a.s.,$$

*where $\hat{\mathbb{E}}^{\mathbb{Q}}$ denotes the fact that expectation has been calculated under a martingale measure for the model with volatility $\hat{\sigma}$.*

The content of the theorem is that if an agent uses a model which overestimates volatility, then provided she uses both the model price and the model hedge, then she will still succeed in super-replicating the option on a path-wise basis. We will extend this idea of path-wise super-replication so that it does not rely on an assumption that the true volatility is bounded by a model volatility, but rather so that super-replication holds whatever the behaviour of the underlying.

In general, in incomplete markets there are several alternative criteria which can be used for option pricing. One alternative is to select arbitrarily (or otherwise) a martingale measure, perhaps by fixing the market prices of risk of those Brownian motions orthogonal to those driving the traded assets to be zero. Alternatively, utility indifference pricing incorporates the agents' attitudes to risk to find a price at which an option seller will be prepared to enter into a contract for sale of a derivative. The extreme version of utility indifference pricing is super-replication, whereby the agent is not willing to accept any risk.

These approaches are all conditional on an assumption that the dynamics of $P$ follow a given model, or family of models. Often, (e.g. in the case of utility indifference pricing) the investor needs knowledge of the drift parameter $\gamma$ (which is notoriously difficult to estimate) as well as the volatility $\sigma$.

## 2.3 A reverse approach: recovering a model from prices

The starting point of our philosophy is to take option prices (at least those of liquidly traded options) as exogenously given by the market, and to use those prices to learn about the stochastic properties of the underlying.

Since market prices are determined under the measure used by the market for pricing, the conclusions we draw are already phrased in a fashion whereby they can be used for pricing other derivatives.
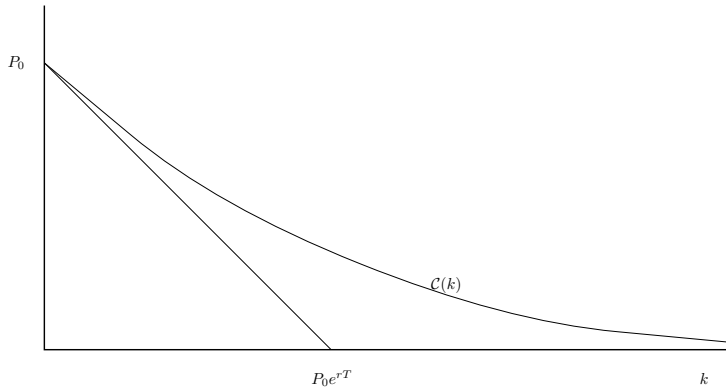
Figure 1: No arbitrage considerations force that the call price curve $\mathcal{C}$ is a decreasing convex function of $k$ (for maturity $T$ fixed) with $\mathcal{C}(0) = P_0$ and $\mathcal{C}(k) \geq (P_0 - Ke^{-rT})^+$.

**Lemma 2.2 (Breeden and Litzenberger [8])** *Fix $T \in (0, \infty)$. Suppose that call prices with maturity $T$ are known for every $K \in (0, \infty)$. Then assuming call prices are calculated as the discounted expected payoff under a model $\mathbb{Q}$, so that*

$$\mathcal{C}(K, T) = \mathbb{E}^{\mathbb{Q}}[e^{-rT}(P_T - K)^+],$$

*we have*

$$\mathbb{Q}(P_T > K) = e^{rT} \left| \frac{\partial}{\partial K} \mathcal{C}(K, T) \right|, \tag{3}$$

*and, provided $\mathcal{C}$ is twice-differentiable in $K$*

$$\mathbb{Q}(P_T \in dK) = e^{rT} \frac{\partial^2}{\partial K^2} \mathcal{C}(K, T) \tag{4}$$

If the law of $P_T$ under $\mathbb{Q}$ has atoms, then $\mathbb{Q}(P_T > K)$ is given by the right derivative in (3), and $\mathbb{Q}(P_T \geq K)$ by the left derivative. In this case (4) must be understood in a distributional sense. However, for the present we assume that there are no atoms and $\mathcal{C}$ is smooth.

It follows from the Lemma that call prices, as represented by the set of decreasing convex functions are in one-to-one correspondence with the marginal laws of $P_T$ under the pricing measure. The above result applies for individual maturities. If we also have information about the dependence of option prices on maturity we can recover a price process.

**Theorem 2.3 (Krylov [39], Gyöngy [28], Dupire [23])** *Suppose call prices are known for every $K \in (0, \infty)$ and every $T \in (0, \overline{T})$. Assuming that $\mathcal{C}(T, K)$ is sufficiently differentiable, there exists a unique diffusion of the form*

$$dP_t = P_t \sigma(P_t, t) dB_t + rP_t dt$$

*such that*

$$\mathbb{E}[e^{-rT}(P_T - K)^+] = \mathcal{C}(K, T)$$

*In particular, $\sigma(p, t)$ solves*

$$\frac{1}{2} K^2 \sigma(K, T)^2 \mathcal{C}_{KK} - rK\mathcal{C}_K - \mathcal{C}_T = 0.$$

The idea in Krylov [39] is to take any process $(Y_t)_{t \geq 0}$ and to show that there exists a diffusion process $X_t$ with the same marginal distributions. Gyöngy refined the analysis and derived the relationship $\sigma(k,t)^2 = \mathbb{E}[(dY)_t^2|Y_t = k]$. Dupire has a slightly different starting point in that rather than assuming the existence of a process $Y$ he begins with the call price surface $\mathcal{C}(K,T)$.

The Dupire construction gives a unique Markovian martingale which agrees with the doubly infinite family of call option prices. In principle, it is then possible to give the prices of any other derivatives. However, the diffusion assumption is a strong assumption (and is inconsistent with stochastic volatility, for example), and although the Dupire diffusion calibrates perfectly to the initial structure of call prices, there is no guarantee that it will match derivative prices at any later time. (We consider in Section 6.3, the existence of other martingales which also match the marginals implicit in the call price function.)

Both the Breeden and Litzenberger [8] and Dupire [23] results give us information about how the marginal distributions of the asset (under the pricing measure) can be determined from call prices. More generally, given any set of option prices we can ask:

- is there a model consistent with those prices? (i.e. can we find $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ such that $\mathbb{E}^{\mathbb{Q}}[e^{-rT}F_T] = f$ for each claim $F$ with associated traded price $f$)

- if such a model exists is it unique?

- if there is no such model, is there an arbitrage?

## 2.4  Notation

It is very convenient to remove the impact of interest rates by switching to a notation in which prices are expressed in discounted terms. Define $S_t = e^{-rt}P_t$, then $(S_t)_{t \geq 0}$ is the forward price, and is a martingale under a pricing measure. Furthermore, we write $C(k)$ for the time-zero price of an option on the forward $S$ with strike $k$ (and maturity $T$):

$$C(k) = \mathbb{E}^{\mathbb{Q}}[(S_T - k)^+] = \mathbb{E}^{\mathbb{Q}}[e^{-rT}(P_T - ke^{rT})^+] = \mathcal{C}(ke^{rT})$$

where $\mathcal{C}$ denotes the price of a call option on $P_T$. Then $\mathbb{Q}(S_T > k) = |C'(k+)|$.

When written without a superscript $C$ denotes the traded set of options prices. However, under a martingale model for which $S_T \sim \mu$, we also have that call prices are given by

$$C_\mu(k) = \mathbb{E}[(X - k)]^+, \qquad X \sim \mu;$$

and we use $C$ with a subscript to denote call prices which correspond to $S_T$ having a given law. Later we use $P_\mu$ to denote put prices under law $\mu$:

$$P_\mu(k) = \mathbb{E}[(k - X)]^+, \qquad X \sim \mu.$$

In the financial context, $\mu$ will be the law of a positive random variable, often with mean 1. However, the quantities $C_\mu$ and $P_\mu$ are well defined for any $\mu$ with first moment, and we will frequently consider the case where $\mu$ is the law of a centred random variable.

## 2.5 The pricing problem

The goal of the next two sections is to develop a formal framework within which we can attempt to determine model-independent bounds and hedging strategies. Our set-up is an extension of the framework developed by Föllmer and Schied [27] and especially Cox and Obłój [18, 19].

Let $\Omega^S$ describe the space of possible paths for the asset price process. A natural candidate is to let $\Omega^S$ be the space of continuous non-negative trajectories with a given initial value $S_0$. The advantage of using this choice is that we can then define the natural filtration associated with the coordinate process, and then first hitting times will be stopping times with respect to this filtration. Note that we do not want to specify the probability triple (and especially the probability measure) exogenously, but rather to construct models (and measures) which are consistent with observed prices. See Cox and Obłój [19] for further discussion on this issue.

Suppose we are given a family $\Theta$ of admissible trading strategies $(\theta_t)_{0 \le t \le T}$ with associated terminal gains from trade processes $G^\theta = \int_0^T \theta_t dS_t$. The space $G^\Theta = \{G^\theta; \theta \in \Theta\}$ represents a space of costless, forward transactions. Implicitly, we are assuming zero transactions costs for trades on the forward market, or at least that such costs are negligible on strategies that we class as admissible.

Suppose we are given a family $\{F_\alpha, \alpha \in A\}$ of traded payoffs (random variables), and suppose that $V$ is a vector space with basis $F_\alpha, \alpha \in A$.

Suppose we are also given a map (pricing functional), $\mathcal{P} : V \oplus G^\Theta \to \mathbb{R}$ such that $\mathcal{P}$ is both linear, and monotone:

$$
\begin{aligned}
\mathcal{P}(v + G^\theta) &= \mathcal{P}(v); \\
\mathcal{P}(\lambda_1 F_1 + \lambda_2 F_2) &= \lambda_1 \mathcal{P}(F_1) + \lambda_2 \mathcal{P}(F_2); \\
F_1 \ge F_2 &\Rightarrow \mathcal{P}(F_1) \ge \mathcal{P}(F_2).
\end{aligned}
$$

Note that the monotonicity plays the role of a no-arbitrage condition. The idea is that $\mathcal{P}$ represents the prices at which the traded payoffs can be bought in the market. In the analysis we will assume that these payoffs can be bought and sold at the same price, but once we attempt to value a hedging portfolio of traded payoffs we can use the market bid or ask prices as appropriate.

Now suppose we add a family of payoffs $\{F_{\tilde\alpha}; \tilde\alpha \in \tilde A\}$, and let $\tilde V$ be the vector space with basis $\{F_\alpha; \alpha \in A\} \cup \{F_{\tilde\alpha}; \tilde\alpha \in \tilde A\}$. The problem is to decide when $\mathcal{P}$ can be extended to a linear function from $\tilde V \oplus G^\theta$ to $\mathbb{R}$.

**Definition 2.4** *The family $\{F_\alpha; \alpha \in A\}$ is the set of vanilla securities.*
*The family $\{F_{\tilde\alpha}; \tilde\alpha \in \tilde A\}$ is the set of exotic securities.*

Equivalently the set $\{F_\alpha; \alpha \in A\}$ is the set of payoffs whose prices are precisely determined by the market, and the set $\{F_{\tilde\alpha}; \tilde\alpha \in \tilde A\}$ is the set of payoffs for which we wish to determine a price, relative to the prices of the vanilla securities.

We have not yet been specific about the set of admissible trading strategies $\Theta$, and the definition of $\Theta$ may be chosen according to the context. We certainly want to exclude strategies which allow for arbitrage. The main example is when $\Theta$ is the linear space generated by $\theta^\tau = (\theta_t^\tau)_{0 \le t \le T}$ for stopping times $\tau$, where

$\theta_t^\tau = I_{\{t \in (\tau, T]\}}$. These are the gains from trade obtained from simple combinations of forward contracts. However, in Section 5.1 we will want to allow for more general strategies.

## 2.6   The hedging problem

**Definition 2.5** *We say $\Pi^V$ is a static vanilla portfolio if*

$$\Pi^V = \sum_{i=1}^{n} \pi_{\alpha_i} F_{\alpha_i} \qquad n < \infty, \alpha_1, \ldots \alpha_n \in A$$

*We say $\Pi$ is a semi-static portfolio if $\Pi$ is the sum of a static vanilla portfolio and an admissible trading portfolio.*

**Definition 2.6** *$\Pi$ is a semi-static superhedging portfolio for $\hat{F}$ if $\Pi$ is a semi-static portfolio and $\Pi \geq \hat{F}$ almost surely.*

**Lemma 2.7** *If $\mathcal{P}(F_\alpha) = f_\alpha$ for $\alpha \in A$ and if $\Pi = \sum_{i=1}^{n} \pi_{\alpha_i} F_{\alpha_i} + G^\theta$ is a semi-static superhedging portfolio for $\hat{F}$, then for any extension $\mathcal{P}$ from $V$ to $\tilde{V}$ we must have*

$$\mathcal{P}(\hat{F}) \leq \sum_{i=1}^{n} \pi_{\alpha_i} f_{\alpha_i}$$

**Proof.**
If $\mathcal{P}$ can be extended to the linear space spanned by $\{F_\alpha : \alpha \in A\} \cup \hat{F}$ then

$$\mathcal{P}(\hat{F}) \leq \mathcal{P}(\Pi) = \sum_{i=1}^{n} \pi_{\alpha_i} f_{\alpha_i}$$

by monotonicity and linearity. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

**Remark 2.8** In the set-up of Section 2.3 an implicit assumption is that option prices are both bid and ask prices and that there are zero transaction costs. This is a necessary assumption in order to identify a precise correspondence between prices and probabilities under the risk-neutral measure for use in the pricing problem. However, once we have identified a candidate model-independent super-replicating strategy involving vanilla options, in the hedging problem we can use market bid or ask prices (depending on whether the hedge involves a short or long position) as appropriate. In this way the theory extends to situations in which there are bid-ask spreads, even on liquidly traded securities. Note, moreover, that the position in vanilla securities is static over time, so these transaction costs are incurred once only, and unlike classical model-dependent $\delta$-hedging, the semi-static hedging strategy remains feasible in a model with transaction costs.

By restricting the class of admissible trading strategies in the forward, we can also allow for transaction costs on the forward. For example, in the next section we could restrict the class of admissible trading portfolios to those which involve at most a single forward transaction, and therefore incur transaction costs at most once.

## 2.7 An example: a digital option

We return to the example first discussed in the opening section; namely given the prices of call options what are the possible prices of a one-touch digital option, a security which makes a unit payment at $T$ if ever $S_t \geq B$ for some $t \leq T$.

We assume that forward contracts with maturity $T$ are traded at zero cost, and that

$$\Theta = \left\{ \theta : \theta_t = \sum_{i=1}^{p} \gamma_i I_{\{\tau_i < t \leq T\}} \right\} \qquad p \in \mathbb{N}, \gamma_i \in \mathbb{R}$$

where $(\tau_i)_{i \leq p}$ are stopping times. In fact the only property we need for $\Theta$ is that $\Theta \supseteq \{c I_{\{H_B < t \leq T\}}; c \in \mathbb{R}\}$.

We assume that the vanilla options are precisely the set of call options with maturity $T$, parameterised by the strike $K$, and that a continuum of such calls are traded (one for each possible strike), together with the bond which pays a unit value at time $T$. In the notation of this section, $\{F_\alpha; \alpha \in \mathcal{A}\} = \{F_K; K \in \mathbb{R}_+\} = \{(S_T - K)^+; K \in \mathbb{R}_+\}$, and for $K \in (0, \infty)$ the payoff $F_K = (S_T - K)^+$ has price $f_K = C(K)$. Note that to preclude arbitrage in the vanilla options we must have that $C$ is a decreasing convex function, and that by our assumption that $S$ is a forward price $C(0) = S_0$, $C'(S_0) \geq -1$ and $C(K) \geq (S_0 - K)^+$. We also make the very natural assumption that $\lim_{K \uparrow \infty} C(K) = 0$. (As observed by Davis and Hobson [20] this does not follow from no arbitrage alone, but rather from a martingale assumption, and can be violated if the asset price allows a bubble, see e.g. Cox and Hobson [16]. Davis and Hobson [20] introduce the notion of a *weak arbitrage* and Cox and Obłój [19] introduce the parallel notion of a *weak free lunch with vanishing risk*, which they use to describe some of the consequences of the failure of this assumption.)

Now we consider the contract $\tilde{F} = I_{\{H_B \leq T\}}$ with $B \geq S_0$. For any $K < B$ we have (recall (1))

$$\tilde{F} \leq \frac{1}{(B - K)} F_K + G^\theta$$

where $F_K$ is the payoff of a vanilla option $(S_T - K)^+$, and $\theta_t = -I_{\{H_B < t \leq T\}}/(B - K)$ so that $G^\theta = (S_{H_B} - S_T)/(B - K)$. There are two cases to be checked, namely when the left-hand-side is zero or one. If the barrier is not reached, then the left-hand-side is null and there is only one non-zero term on the right-hand-side, which is non-negative. If the barrier is reached, then the second term on the right-hand-side of (1) is present and moreover $S_{H_B} \geq B$ (with equality if $S$ is continuous). Taking the value $(S_T - K)$ rather than the positive part $(S_T - K)^+$ again only makes the right-hand-side smaller, and leads to equality in (1) on $(H_B \leq T)$.

It follows immediately that $\mathcal{P}(\tilde{F}) \leq C(K)/(B - K)$, and since $K$ is arbitrary we conclude $\mathcal{P}(\tilde{F}) \leq \inf_{K < B} C(K)/(B - K) = C(K^*)/(B - K^*) =: \overline{D}$, where $K^* = \arg\inf_{K < B}(C(K)/(B - K))$.

Now we want to argue that this bound is best possible in the given framework. To do this we exhibit a model under which the fair prices (ie expected values) of the vanilla securities are equal to the bound $\overline{D}$. Since we are interested in pricing issues, it makes sense to work directly in a martingale measure, so that there is no distinction between $\mathbb{P}$ and $\mathbb{Q}$.
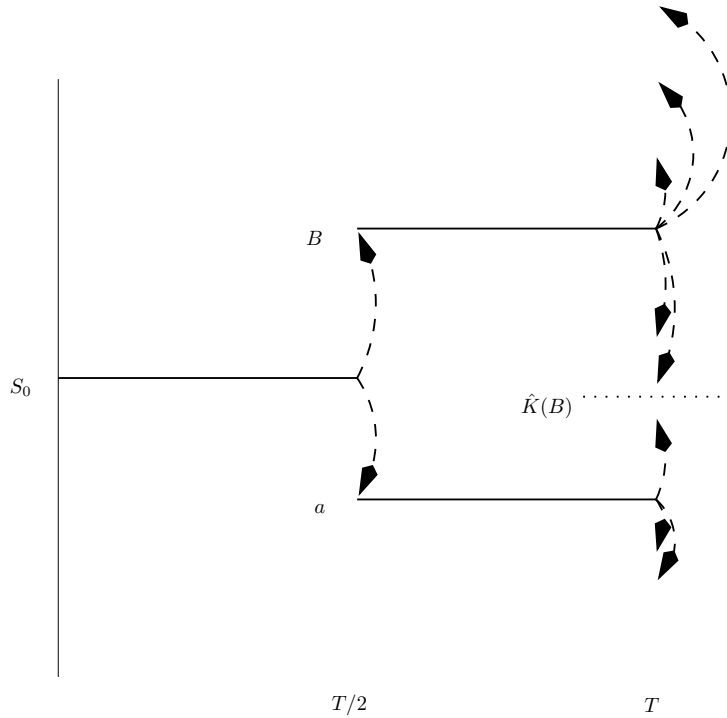
Figure 2: A jump model which matches options prices, and for which the maximum value of the digital option is attained. The price process is constant except for jumps at $T/2$ and $T$. Those paths which are at $B$ (respectively $a$) at $T/2$ end up above (below) $\hat{K}(B)$ after the jump at time $T$.

Write $\mu$ for the law of $S_T$. Let $b(K) = \mathbb{E}[S_T | S_T \geq K] = \int_{y \geq k} y \mu(dy) / \int_{y \geq k} \mu(dy)$; $b$ is called the barycentre function. By definition it is left-continuous, increasing and $b(K) \geq K \vee \mathbb{E}[S_T]$, at least for $K$ smaller than the upper bound on the support of $\mu$, where it is defined. Furthermore, if $\mu$ has a density then $b(K)$ is continuous.

Given $B$ there exists $\hat{K} = \hat{K}(B)$ such that $b(\hat{K}+) \geq B \geq b(\hat{K})$ and then we can find a set $A$ with

$$(S_T > \hat{K}) \subseteq A \subseteq (S_T \geq \hat{K})$$

for which $B = \int_A y \mu(dy) / \int_A \mu(dy)$, or in other words

$$0 = \int_A (y - B)\mu(dy) = \int_A (y - \hat{K})\mu(dy) + (\hat{K} - B)\mu(A).$$

It follows that $\mu(A) = C(\hat{K})/(B - \hat{K})$ so that

$$\mathbb{P}(S_T > \hat{K}) \leq \frac{C(\hat{K})}{B - \hat{K}} \leq \mathbb{P}(S_T \geq \hat{K}).$$

Let $a = \int_{A^c} y \mu(dy) / \int_{A^c} \mu(dy)$, and let $S_{T/2}$ take the values $B$ or $a$ with probabilities $(x - a)/(B - a)$ and $(B - x)/(B - a)$ respectively.

On the set $S_{T/2} = B$, let $S_T$ have the law $\mu$ restricted to $A$. Similarly, on $S_{T/2} = a$, let $S_T$ have the law $\mu$ restricted to $A^c$.

Suppose $S_t$ is constant except for a pair of jumps at $T/2$ and $T$. See Figure 2. Then, by construction, $\mathbb{E}[S_T; A] = B\mathbb{P}(A)$ and $\mathbb{E}[S_T; A^c] = a\mathbb{P}(A^c)$ so that $S_t$ is a
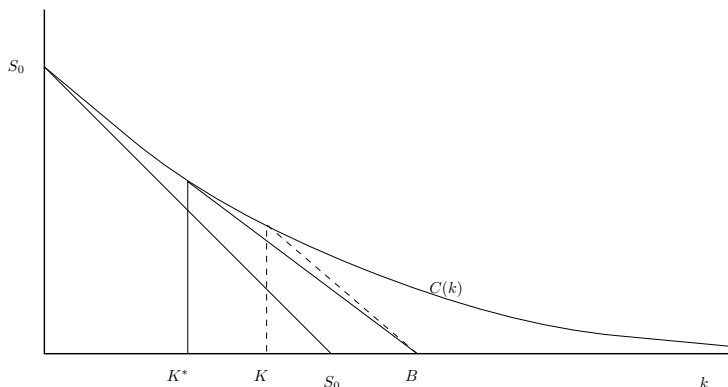
11

Figure 3: The minimiser $K^* = K^*(B)$ is chosen such that the tangent to $C$ at $K^*$ crosses the $x$-axis at $B$.

martingale. Moreover, if prices $\mathcal{P}$ are given by expectation with respect to $\mathbb{P}$ then

$$\mathcal{P}(\tilde{\mathcal{F}}) = \mathbb{P}(H_B \leq T) = \mathbb{P}(S_{T/2} = B) = \mu(A) = \frac{C(\hat{K})}{B - \hat{K}}.$$

It remains to show that $\hat{K}$ is a minimiser of $C(K)/(B-K)$. This can be deduced from the fact that the bound $\mathcal{P}(\tilde{F}) \leq C(K)/(B-K)$ holds for any martingale model with the correct marginals. We have

$$\mathbb{P}(H_B \leq T) = \mu(A) = \frac{C(\hat{K})}{(B - \hat{K})} \geq \inf_{K<B} \frac{C(K)}{(B - K)} \geq \sup \tilde{\mathbb{P}}(H_B \leq T) \geq \mathbb{P}(H_B \leq T)$$

where the supremum is taken over martingale models, denoted by $\tilde{\mathbb{P}}$, with the correct marginal laws at time $T$. Hence there is equality throughout and $\hat{K}$ is a minimiser.

The above analysis exhibits a pure-jump model for which $S_T \sim \mu$ (and the price of the barrier option is equal to the robust upper bound $C(K^*)/(B - K^*)$). Although potentially the simplest model, it is not unique. Indeed, we can construct a continuous price process model with the same properties.

Let $(W_t^0)_{t\geq 0}$ be Brownian motion, started at zero, and for $\alpha < 0 < \beta$, let $H_{\alpha,\beta}^0 = \inf\{u : W_u^0 \notin (\alpha, \beta)\}$. Then $W_{H_{\alpha,\beta}^0}^0 \in \{\alpha, \beta\}$ and $\mathbb{P}(W_{H_{\alpha,\beta}^0}^0 = \alpha) = \beta/(\beta - \alpha) =: p$, and we have embedded a two-point distribution in Brownian motion. Now set $X_u = W_{u/(1-u)\wedge H_{\alpha,\beta}^0}^0$; then $X$ is a martingale with $X_1 \sim p\delta_\alpha + (1-p)\delta_\beta$.

With this in mind, define $H_{a,B}^s = \inf\{u : W_u^s \notin (a, B)\}$ where $W_t^s = S_0 + W_t^0$, and $S_0 = s$. By construction $S_0 + W_{H_{a,B}^s}^0 \in \{a, B\}$. If we set

$$S_t = S_0 + W_{2t/(T-2t)\wedge H_{a-s,B-s}^0}^0 = W_{2t/(T-2t)\wedge H_{a,B}^s}^s$$

then $(S_t)_{0\leq t\leq T/2}$ is a continuous martingale which takes values in $\{a, B\}$ at time $T/2$. To complete the construction (i.e. to construct the process $S_t$ over $(T/2, T]$) we need to discuss how to embed any centred distribution in Brownian motion (and not just a two point distribution). This is the Skorokhod embedding problem.

12

# 3 The Skorokhod Embedding Problem

## 3.1 The classical Skorokhod embedding problem

Suppose we are given a stochastic process $X = (X_t)_{t \geq 0}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t \geq 0}, \mathbb{P}, (X_t)_{t \geq 0})$. Suppose $X$ has state space $I$.

The Skorokhod embedding problem (SEP) for $(X, \mu)$ is, given $\mu$ on $I$, to find (where possible) a stopping time $\tau$ such that the law of the stopped process is $\mu$. We write $\tau \in \bar{\mathcal{T}}(X, \mu)$. The classical version of the problem takes $X \equiv W$, a Brownian motion on $\mathbb{R}$, started at 0, and $\mu$ a centred probability measure, in which case we suppress $W$ from the notation and write $\tau \in \bar{\mathcal{T}}(\mu)$.

Obłój [43] provides a comprehensive survey of the literature of solutions to the Skorokhod embedding problem. He lists more than twenty solutions, some of which we will discuss below.

In the sequel the key properties of Brownian motion that we will use are the fact that Brownian motion leaves any interval in finite time almost surely, (so $H_a := \inf\{u : W_u = a\} < \infty$ a.s.); the martingale property; the consequent fact that for $a < x < b$, $\mathbb{P}^x(H_a < H_b) = (x - a)/(b - a)$; and the fact that Brownian motion has no intervals of constancy. Only occasionally will we use the speed of Brownian motion, but then we have that $\mathbb{E}^x[H_a \wedge H_b] = (b - x)(x - a)$.

We shall often decompose Brownian motion into its constituent excursions away from zero, which are parameterised by the local time $L_t$. See Rogers [49] for an introduction to Brownian excursion theory. If we write $J_t$ for the maximum process $J_t = \sup_{0 \leq s \leq t}\{W_s\}$ (we reserve the notation $M_t$ for martingales, and $S_t$ for forward prices) then by Lévy's Theorem $(L_t, |W_t|) \sim (J_t, J_t - W_t)$, and there is an equivalent decomposition of a Brownian path into the excursions below the maximum. The local time is parameterised in such a way that the rate of excursions with maximum (in modulus) greater than $a$ is $a^{-1}$, for any positive $a$.

## 3.2 Doob's solution

Our first solution is attributed by Rogers and Williams [50] to Doob, and is not intended as a genuine solution to the SEP, but rather to show that without some refinements on the class of solutions to the SEP, the problem is trivial.

Given $\mu$, define $F_\mu$ to the the distribution function $F_\mu(x) = \mu((-\infty, x]) = \mathbb{P}(X \leq x | X \sim \mu)$. Let $\Phi$ be the distribution function of a standard normal distribution, and let $Z = F_\mu^{-1}(\Phi(W_1))$. Then $Z$ has law $\mu$:

$$\mathbb{P}(Z \leq x) = \mathbb{P}(F_\mu^{-1}(\Phi(W_1)) \leq x) = \mathbb{P}(W_1 \leq \Phi^{-1}(F_\mu(x))) = F_\mu(x).$$

Now let $\tau = \inf\{u \geq 1 : W_u = F_\mu^{-1}(\Phi(W_1))\}$. Then $W_\tau \sim F_\mu^{-1}(\Phi(W_1)) \sim \mu$. See Figure 4.

Note that in general $F_\mu^{-1}(\Phi(x)) \neq x$ and thus, conditional on $W_1 = x$, the expected value of $\tau$ is infinite. Then, unless $\mu$ has a standard normal distribution, $\mathbb{E}[\tau] = \infty$. In particular, $(W_{t \wedge \tau})_{t \geq 0}$ is not uniformly integrable.
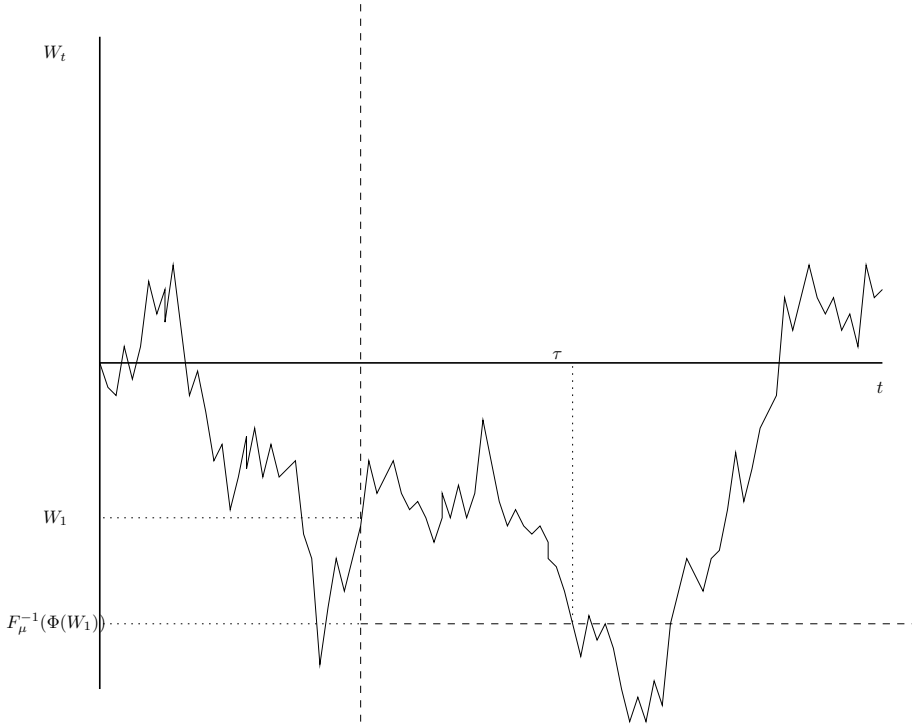
Figure 4: The Doob construction. Run the Brownian motion until time 1. Conditional on the value at time 1, construct a new target level and run the Brownian motion until it first reaches this level. For a carefully chosen rule for selecting the target level, this construction embeds $\mu$.

## 3.3 Hall's solution

Let $c = \int_0^\infty x\mu(dx)$. By the centring property we must also have $c = \int_{-\infty}^0 |x|\mu(dx)$. Choose a pair of random variables $U \in (-\infty, 0)$, $V \in [0, \infty)$ with joint law $\rho$ where

$$\rho(du, dv) = \frac{(|u| + v)}{c}\mu(du)\mu(dv)$$

(For example, if $\mu$ is a uniform distribution on $(-1, 1)$ then $c = 1/4$ and $\rho(du, dv) = (|u| + v)$.)

Then, conditional on $(U = u, V = v)$, let $\tau_{u,v} = \inf\{t \geq 0 : W_t \notin (u, v)\}$. We claim that $\tau = \tau_{U,V}$ embeds $\mu$. See Figure 5.

We have, for $u < 0$,

$$
\begin{aligned}
\mathbb{P}(W_\tau \in du) &= \int_{v \in [0,\infty)} \mathbb{P}(U \in du, V \in dv)\mathbb{P}(W_{\tau_{u,v}} = u | U \in du, V \in dv) \\
&= \int_{v=0}^\infty \rho(du, dv)\frac{v}{|u| + v} = \mu(du)\left(\int_{v=0}^\infty \mu(dv)\frac{v}{c}\right) = \mu(du).
\end{aligned}
$$

A similar calculation applies for $v > 0$, and the Hall embedding is a solution of the SEP for $(W, \mu)$.

For this construction we can also calculate $\mathbb{E}[\tau]$. Again conditioning on the exogenous random variables $U$ and $V$,

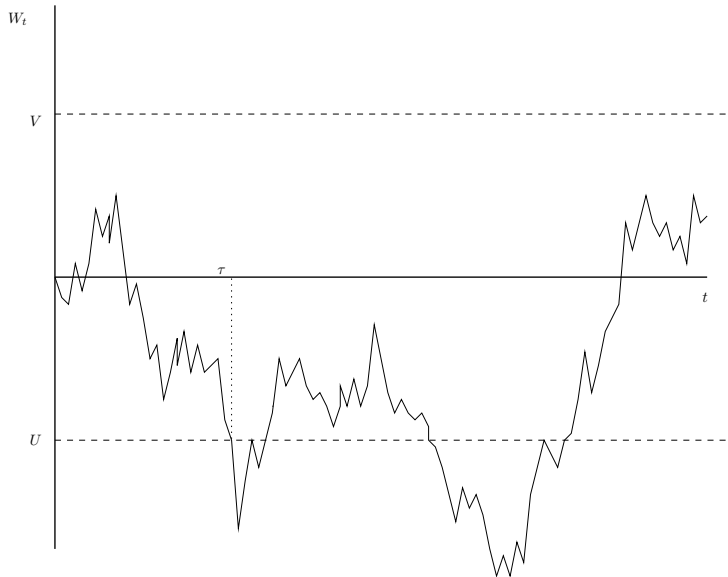$$\mathbb{E}[\tau] = \mathbb{E}[\mathbb{E}[\tau_{U,V} | U, V]]$$

14

Figure 5: The Hall construction. Given random variables $U < 0 < V$, we stop the Brownian motion the first time it leaves the interval $[U, V]$. For the appropriate joint distribution of $(U, V)$, this leads to an embedding of $\mu$.

$$
\begin{aligned}
&= \int_{-\infty}^{0} \int_{0}^{\infty} dv |u| v \rho(du, dv) \\
&= \int_{-\infty}^{0} \mu(du) \int_{0}^{\infty} \mu(dv) \frac{|u|^2 v + |u| v^2}{c} \\
&= \int_{-\infty}^{0} u^2 \mu(du) \int_{0}^{\infty} \mu(dv) \frac{v}{c} + \int_{-\infty}^{0} \frac{|u|}{c} \mu(du) \int_{0}^{\infty} \mu(dv) v^2 \\
&= \int_{-\infty}^{\infty} u^2 \mu(du),
\end{aligned}
$$

so that at least in the case where $\mu$ has finite variance, $\mathbb{E}[\tau]$ is equal to that variance. As we shall see below this is best possible. In this sense, the Hall solution is optimal; however, it has the undesirable feature of requiring independent randomisation.

### 3.4   Properties of good solutions

Briefly we return to the general setting of a stochastic process $X$ on $I$, and let $\tau$ and $\sigma$ be stopping times.

**Definition 3.1** *The stopping time $\tau$ is* minimal *if $\sigma \leq \tau$ and $X_\sigma \sim X_\tau$ implies $\sigma = \tau$ almost surely.*

Minimality is a natural concept for good solutions of the SEP in a general context, but now we return to the Brownian setting.

**Theorem 3.2 (Monroe [42])**   *1. Minimal stopping times exist; given $\tau$ there exists a minimal stopping time $\sigma$ with $W_\sigma \sim W_\tau$.*

15

2. *Hitting times are minimal; all stopping times with finite expectation are minimal.*

3. *Suppose $\tau$ is such that $\mathbb{E}[W_\tau] = 0$. Then $\tau$ is minimal if and only if $W_{t \wedge \tau}$ is uniformly integrable.*

The equivalence between uniformly integrable stopping times and minimal stopping times in the centred case allows us to translate an analytic condition which is useful for concluding that $\mathbb{E}[W_{\sigma \wedge \tau}] = 0$ into a path-wise condition describing good stopping times, where the adjective 'good' might better be expressed as 'not obviously inefficient'.

We prove some of Theorem 3.2.

**Proof that stopping times with finite expectation are minimal.**

If $\mathbb{E}[\tau] < \infty$ then $\mathbb{E}[W_\tau] = 0$ and $\mathbb{E}[W_\tau^2] = \mathbb{E}[\tau]$. If $\sigma \leq \tau$ and $W_\sigma \sim W_\tau$, then

$$\mathbb{E}[\sigma] = \mathbb{E}[W_\sigma^2] = \mathbb{E}[W_\tau^2] = \mathbb{E}[\tau]$$

and $\tau = \sigma$ almost surely.

**Corollary 3.3** *If $\mu$ has a second moment and $\tau \in \bar{\mathcal{T}}(\mu)$ then either $\mathbb{E}[\tau] = \int x^2 \mu(dx)$ or $\mathbb{E}[\tau] = \infty$.*

**Proof that if $W_{t \wedge \tau}$ is uniformly integrable, then $\tau$ is minimal.**

Suppose that $\sigma \leq \tau$ and $W_\sigma \sim W_\tau$. Then for all $a$,

$$\mathbb{E}[W_\tau - a; W_\tau \geq a] = \mathbb{E}[W_\sigma - a; W_\sigma \geq a] = \mathbb{E}[W_\tau - a; W_\sigma \geq a]$$

where we use the identity in law, and the martingale and uniformly integrability properties. Now $\mathbb{E}[W_\tau - a; A]$ is maximised over sets $A \subset \Omega$ by choices $A$ for which, modulo null sets $(W_\tau > a) \subseteq A \subseteq (W_\tau \geq a)$. But $\mathbb{E}[W_\tau - a; W_\sigma \geq a]$ attains this maximum and hence modulo null sets, $(W_\tau > a) \subseteq (W_\sigma \geq a) \subseteq (W_\tau \geq a)$. Hence $W_\tau = W_\sigma$ almost surely.

Now let $\eta$ be any other stopping time with $\sigma \leq \eta \leq \tau$. Then

$$W_\eta = \mathbb{E}[W_\tau | \mathcal{F}_\eta] = \mathbb{E}[W_\sigma | \mathcal{F}_\eta] = W_\sigma = W_\tau, \qquad \text{almost surely.}$$

Hence $W$ is constant on $(\sigma, \tau)$, and since $W$ has no intervals on constancy we must have $\sigma = \tau$ as required. $\qquad \square$

In the centred Brownian case we write $\tau \in \mathcal{T}(\mu)$ if $\tau \in \bar{\mathcal{T}}(\mu)$ and $\tau$ is minimal.

**Corollary 3.4** *If $\mu$ is centred and has support contained in an interval $I$ and if $\tau \in \mathcal{T}(\mu)$ then $\tau$ is less than or equal to the first exit time from $I$.*

## 3.5  The Azéma-Yor solution

The Azéma-Yor [3] solution is based on a stopping rule which is the first entrance of the joint process $(W_t, J_t)$ into a domain $\mathcal{D}_\mu^J$. Let $\beta$ be an increasing (but not necessarily continuous, nor strictly increasing) process such that $\beta(j) \leq j$. (If $j_0$
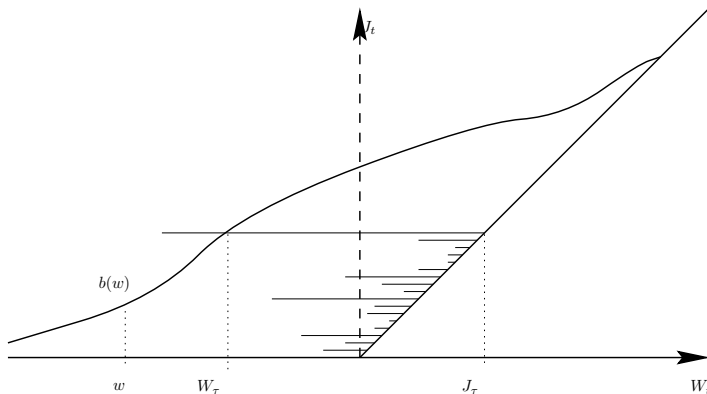
Figure 6: The Azéma-Yor embedding. The figure shows a plot of the excursions of $W_t$ down from the maximum, represented by the horizontal lines in $(W_t, J_t)$ space. The Azéma-Yor stopping time is the first time that the Brownian path crosses the line given by the barycentre $b$.

is such that $\beta(j_0) = j_0$ then we insist that $\beta(j) = j$ for all $j \geq j_0$.) Let $b$ be the left-continuous inverse to $\beta$ so that $b = \beta^{-1}$, and let $\tau_\beta = \inf\{t > 0 : W_t \leq \beta(J_t)\}$. See Figure 6.

Our approach is based on Rogers [49]. By the representation of the Brownian path in terms of excursions down from the maximum we have that the probability that $J_\tau$ is greater than $j$ is equal to the probability that no excursion down from the maximum gets below $\beta(J.)$ before $J$ reaches $j$. Let $n$ denote Itô excursion measure and let $A_z$ be the set where the excursion down from the maximum when the maximum equals $z$ gets below $\beta(z)$: $A_z = \{\max_{0<u<\xi_z}(e_z(u)) > z - \beta(z)\}$. Here $\xi_z$ is the lifetime of the excursion down from the maximum which occurs when the maximum is at $z$, and $(e_z(u))_{0<u<\xi_z}$ is an excursion away from 0 of Brownian motion. Then

$$\mathbb{P}(J_\tau \geq j) = \mathbb{P}(n\left(\cup_{z \leq j} A_z\right) = 0) = \exp\left(-\int_0^j \frac{dz}{z - \beta(z)}\right)$$

where we use the fact that the rate of excursions whose maximum modulus exceeds $a$ is $a^{-1}$, and the fact that for a Poisson random variable $Z$ of mean $\lambda$, $\mathbb{P}(Z = 0) = e^{-\lambda}$.

Having described the construction it remains to describe how to choose $\beta$ so that $\tau$ is a solution of the Skorokhod embedding problem for $(W, \mu)$. Note that the law of $J_\tau$ has a continuous density, except perhaps for an atom at $j_0$. Moreover, for $j$ such that $\beta(j) < j$,

$$(W_\tau > \beta(j)) \subseteq (J_\tau \geq j) \subseteq (W_\tau \geq \beta(j)), \tag{5}$$

and, at least where $b$ is continuous, we have

$$\mathbb{P}(W_\tau \geq w) = \exp\left(-\int_0^{b(w)} \frac{dz}{z - \beta(z)}\right), \tag{6}$$

and, if $b$ is differentiable,

$$-\frac{\mathbb{P}(W_\tau \in dw)}{\mathbb{P}(W_\tau \geq w)} = -\frac{b'(w)}{b(w) - w}.$$

17

If $W_\tau \sim \mu$ then $\mathbb{P}(W_\tau \geq w) = \mu([w, \infty)) =: \bar{\mu}(w)$, and if $w$ is a continuity point then $(b(w) - w)\mu(dw) = b'(w)\bar{\mu}(w)$. It follows that

$$\frac{d}{dw}\left(\int_w^\infty z\mu(dz)\right) = -w\mu(dw) = b'(w)\bar{\mu}(w) - b(w)\mu(dw) = \frac{d}{dw}\left(b(w)\int_w^\infty \mu(dz)\right).$$

Integrating, and using the fact that at $w$ equal to the lower limit on the support of $\mu$ we have $\int_w^\infty z\mu(dz) = 0$ we find that $b$ is equal to the barycentre function:

$$b(w) = \frac{\int_w^\infty z\mu(dz)}{\int_w^\infty \mu(dz)} = \mathbb{E}^{X\sim\mu}[X|X \geq w]. \tag{7}$$

**Example 3.5** If $\mu \sim U[-1, 1]$ then $b(w) = (w + 1)/2$ (for $-1 \leq w \leq 1$) and $\beta(j) = 2j - 1$ for $0 \leq j \leq j_0 = 1$.

If $E_1$ is an exponential random variable rate 1 and $\mu$ is the centred distribution given by $\mu \sim (E_1 - 1)$ then $b(w) = w + 1$ (for $w > -1$) and $\beta(j) = j - 1$ for $j < j_0 = \infty$.

If $\mu = (p\delta_1 + p\delta_{-1} + (1 - 2p)\delta_0)$ with $p < 1/2$ then for $w \leq -1$, $b(w) = 0$; for $-1 < w \leq 0$, $b(w) = p/(1-p)$ and for $0 < w < 1$, $b(w) = 1$. Then $\beta$ is also piecewise constant (for $j \leq 1$): for $0 \leq j < p/(1-p)$ we have $\beta(j) = -1$; for $p/(1-p) \leq j < 1$ we have $\beta(j) = 0$, and $\beta(j) = j$ for $j \geq 1$.

We have that the Azéma-Yor construction gives a solution of the SEP for $(W, \mu)$. The question arises, is it minimal? By Monroe's result (Theorem 3.2), it is equivalent to ask if $W_{t\wedge\tau}$ is uniformly integrable, and for the Azéma-Yor stopping rule it is possible to verify by calculation that this is the case.

By a Theorem of Azéma, Gundy and Yor [2] a (necessary and) sufficient condition for uniform integrability is that $\lim_{x\uparrow\infty}\mathbb{P}(\sup_{t\leq\tau}|W_t| \geq x) \to 0$. For the Azéma-Yor stopping time,

$$\mathbb{P}(\sup_{t\leq\tau}|W_t| \geq x) = \mathbb{P}(H_{-x} < H_{b(-x)})$$

$$+ \mathbb{P}(H_{b(-x)} < H_{-x})\mathbb{P}\left(\sup_{t\leq\tau} W_t \geq x \,\middle|\, H_{b(-x)} < H_{-x}\right)$$

$$= \frac{b(-x)}{x + b(-x)} + \frac{x}{x + b(-x)}\frac{\mathbb{P}(W_\tau \geq \beta(x))}{\mathbb{P}(W_\tau \geq -x)}.$$

Now $b(-x) \to 0$ and $\bar{\mu}(-x) \to 1$ so that

$$\lim_{x\uparrow\infty} x\mathbb{P}(\sup_{t\leq\tau}|W_t| \geq x) = \lim_{x\uparrow\infty} x\bar{\mu}(\beta(x)) = \lim_{y\uparrow\infty} b(y)\bar{\mu}(y)$$

$$= \lim_{y\uparrow\infty} \mathbb{E}[Y|Y \geq y]\mathbb{P}(Y \geq y) = \lim_{y\uparrow\infty} \mathbb{E}[Y; Y \geq y] = 0.$$

Thus, if we are given $\mu$ and we define the barycentre function and its inverse, and the Azéma-Yor stopping time relative to this inverse, then $W_{t\wedge\tau}$ is uniformly integrable. However, if we consider a general increasing function $\beta$ (for example, $\beta(j) = 1 - 1/j$), then there is no reason for the corresponding stopping rule to be associated with a uniformly integrable stopped process $W_{t\wedge\tau}$.

Note that from (7) we have

$$(b(w) - w)\mathbb{P}(W_\tau \geq w) = \int_w^\infty (z - w)\mu(dz) = \mathbb{E}[(W_\tau - w)^+].$$

Furthermore, by Doob's submartingale inequality, for any uniformly integrable stopping time, and any $j$,

$$
\begin{aligned}
0 = \mathbb{E}[W_\tau - j; J_\tau \geq j] &= \mathbb{E}[W_\tau - w; J_\tau \geq j] + (w - j)\mathbb{P}(J_\tau \geq j) \\
&\leq \mathbb{E}[W_\tau - w, W_\tau \geq w] + (w - j)\mathbb{P}(J_\tau \geq j). \quad (8)
\end{aligned}
$$

Then

$$
\mathbb{P}(J_\tau \geq j) \leq \inf_{w < j} \frac{\mathbb{E}[(W_\tau - w)^+]}{j - w} = \inf_{w < j} \frac{C_\mu(w)}{j - w}. \quad (9)
$$

This is suggestive of the model independent bound $\inf_{K < B} C(K)/(B - K)$ from (2).

There is equality in (8) if and only if

$$
(W_\tau > w) \subseteq (J_\tau \geq j) \subseteq (W_\tau \geq w). \quad (10)
$$

Comparing (10) with (5) we see that there is equality in (8) for the Azéma-Yor construction provided $w = \beta(j)$. We have the following result:

**Proposition 3.6** *Amongst all minimal solutions of the SEP for $(W, \mu)$ the Azéma-Yor stopping rule maximises $\mathbb{P}(J_\tau \geq j)$ simultaneously for all $j$.*

It is crucial in the proposition that we restrict attention to uniformly integrable stopping times. If we consider all embeddings then the problem is degenerate, and for each $j > 0$ there is an embedding for which $\mathbb{P}(J_\tau \geq j) = 1$. To see this run the Brownian motion until the first return to zero after the first hit on $j$, and then use a favourite embedding thereafter.

A corollary of the proposition is that there exists a model for which the forward price is continuous and the price of a one-touch digital option is equal to the upper bound $C(K^*)/(B - K^*)$. Moreover, the bound is attained by a process in which the sets where the maximum is large correspond exactly to the sets where the terminal value is large.

## 3.6 Solutions of the SEP and candidate price processes

Let $(M_t)_{t \geq 0}$ be a continuous martingale, null at 0, such that $M_T \sim \mu$. Then by the Dambis-Dubins-Schwarz Theorem (see e.g. Revuz and Yor [48, Theorem V.1.6]), $M_t = W_{\langle M \rangle_t}$ so that $\langle M \rangle_T$ is a solution of the SEP for $(W, \mu)$.

Conversely, if $\tau \in \mathcal{T}(\mu)$ then

$$
M_t = W_{t/(T-t) \wedge \tau}
$$

is a martingale with $M_T \sim \mu$.

Further, if $\mu$ is an integrable but non-centred measure (with mean $m_\mu$), then we can construct a centred measure $\mu^0$ by a parallel shift $\mu^0((-\infty, x]) = \mu((-\infty, x + m_\mu])$. If $M_T^0$ and $W_\tau^0$ embed $\mu^0$, then $m_\mu + M_T^0$ and $m_\mu + W_\tau^0$ embed $\mu$.

Putting these two ideas together, there is a one-to-one correspondence between minimal solutions of the Skorokhod Embedding Problem, and candidate martingale price processes whose marginal distribution at $T$ is consistent with call prices. If $\mu$ is supported on $\mathbb{R}^+$, then by Corollary 3.4, $(M_t)_{0 \leq t \leq T}$ is non-negative.

## 3.7 Azéma martingales, the Azéma-Yor embedding and super-replication

Suppose we are given an increasing (differentiable) function $F(j)$, and consider the problem of maximising $\mathbb{E}[F(J_\tau)]$ over minimal embeddings of $\mu$. For ease of exposition suppose that $\mu$ is such that the barycentre $b$ and its inverse $\beta$ are continuous.

Given a positive function $h$ define $H(j,w) = \int_0^j h(r)(r-w)dr$. The Azéma martingale is $H(J_t, W_t) = \int_0^{J_t} h(r)(r - W_t)dr$. It is easy to see from Itô's formula that

$$dH(J_t, W_t) = -\left(\int_0^{J_t} h(r)dr\right) dW_t + h(J_t)(J_t - W_t)dJ_t.$$

The final term in this expression is zero since $J_t = W_t$ whenever the finite variation process $J_t$ is not constant, and hence $H(J_t, W_t)$ is a martingale.

We aim to design $H$ such that

$$\arg\max_j\{F(j) - H(j,w)\} = b(w).$$

This requires the choice $h(j) = F'(j)/(j - \beta(j))$ and then

$$H(j,w) = \int_0^j F'(r)\frac{(r-w)}{r - \beta(r)}dr = F(j) - F(0) + \int_0^j F'(r)\frac{(\beta(r) - w)}{r - \beta(r)}dr,$$

so that

$$F(j) - H(j,w) = F(0) + \int_0^j F'(r)\frac{(w - \beta(r))}{r - \beta(r)}dr.$$

We want to maximise this expression. Note that the integrand is positive for $\beta(r) < w$ (i.e. $r < b(w)$) and negative for $\beta(r) > w$ (i.e. $r > b(w)$). Hence the expression is maximised at $j = b(w)$ and

$$F(j) - H(j,w) \leq F(0) + \int_0^{b(w)} F'(r)\frac{(w - \beta(r))}{r - \beta(r)}dr \equiv G(w)$$

where the above is taken as the definition of $G$. In particular, $F(j) \leq G(w) + H(j,w)$ (uniformly in $j > 0, w \in \mathbb{R}$) and applying this result at $(J_\tau, W_\tau)$ we obtain

$$F(J_\tau) \leq G(W_\tau) + H(J_\tau, W_\tau),$$

on a path-by-path basis, with equality for the Azéma-Yor embedding.

Taking expectations, then provided $\tau$ is such that $H(J_{t\wedge\tau}, W_{t\wedge\tau})$ is uniformly integrable, it follows that $\mathbb{E}[F(J_\tau)] \leq \mathbb{E}[G(W_\tau)]$. If we further restrict attention to minimal stopping times which are solutions of the SEP for $(W, \mu)$ then

$$\sup_{\tau \in \mathcal{T}(\mu)} \mathbb{E}[F(J_\tau)] \leq \int_{\mathbb{R}} G(x)\mu(dx).$$

Although we have derived this result for differentiable $F$ and regular $\mu$ it is clear that it can be extended to arbitrary (increasing) functions. For example, for the increasing function $F(j) = I_{\{j \geq B\}}$ we have

$$\sup_{\tau \in \mathcal{T}(\mu)} \mathbb{P}[J_\tau \geq B] \leq \mathbb{E}\left[\frac{W_\tau - \beta(B)}{B - \beta(B)} I_{\{W > \beta(B)\}}\right].$$

This yields another proof of Proposition 3.6.

## 3.8 The Perkins solution

The Azéma-Yor stopping time is based on a stopping rule which compares the value of the Brownian motion to a function of the running maximum. The stopping time has the property that it maximises the law of the maximum amongst the class of uniformly integrable embeddings.

The Perkins [47] embedding has the property that it minimises the law of the maximum of the stopped Brownian motion, amongst all embeddings. (There is no need to impose a minimality criterion here; if $\tau$ is not minimal then there exists $\sigma \leq \tau$ with $W_\sigma = W_\tau$ and then $J_\sigma \leq J_\tau$). In fact the Perkins embedding has the property that it simultaneously minimises the law of the maximum $J_\tau$ and maximises the law of the minimum $I_\tau = \inf\{W_s : 0 \leq s \leq \tau\}$.

Our interpretation of the Perkins construction is based on the following pathwise inequality, which plays an analogous role to (1):

$$I_{\{H_B \leq \tau\}} \geq I_{\{W_\tau \geq B\}} + \frac{(W_\tau - B)^+}{(B - K)} - \frac{(K - W_\tau)^+}{(B - K)} + \frac{W_{H_B} - W_\tau}{(B - K)} I_{\{H_B \leq \tau\}} \quad (11)$$

There is equality in (11) if on $H_B > \tau$ we have $K \leq W_\tau < B$ and on $H_B \leq \tau$ we have either $W_\tau \geq B$ or $W_\tau \leq K$.

Under the assumption that $\tau$ is minimal, taking expectations in (11) we find

$$\mathbb{P}(J_\tau \geq B) \geq \mathbb{P}(W_\tau \geq B) + \frac{1}{B - K}\left(C_\mu(B) - P_\mu(K)\right) \quad (12)$$

where $P_\mu(K) = \mathbb{E}^{X \sim \mu}[(K - X_T)^+]$. Since (12) holds for any $K$ we can find a best bound of this type by taking the supremum over $K < B$, and then we find that $K$ is chosen to maximise

$$\frac{C_\mu(B) - P_\mu(K)}{B - K}. \quad (13)$$

The final bound is

$$\mathbb{P}(J_\tau \geq B) \geq |C'_\mu(B-)| + \sup_{K < B} \frac{1}{B - K}\left(C_\mu(B) - P_\mu(K)\right). \quad (14)$$

If we can find a model such that the bound is attained, then it will follow that (14) is a best bound.

Fix $B$ and suppose $X \sim \mu$. Then there exists $\hat{K}$ and $A$ with $(X < \hat{K}) \cup (X \geq B) \subseteq A \subseteq (X \leq \hat{K}) \cup (X \geq B)$ such that

$$0 = \int_A (k - B)\mu(dk) = \int_{k \geq B}(k - B)\mu(dk) + \int_{A \setminus (X \geq B)}(k - B)\mu(dk). \quad (15)$$

Suppose $\tau_1$ is such that $W_{\tau_1} \in \{\hat{K}, B\}$. On $W_{\tau_1} = B$ we embed the mass $\mu$ restricted to $A$; on $W_{\tau_1} = \hat{K}$ we embed the mass $\mu$ restricted to $A^c$. Let $\tau_P$ be this second embedding, then $\tau_P$ embeds $\mu$. By construction, $(J_{\tau_P} \geq B) \equiv (W_{\tau_1} = B)$, and from (15), $C_\mu(B) = (B - \hat{K})(\mu(A) - \mu([B, \infty))) + P_\mu(\hat{K})$. Then,

$$\begin{aligned}
\mathbb{P}(J_{\tau_P} \geq B) = \mu(A) &= \mu([B, \infty)) + \frac{C_\mu(B) - P_\mu(\hat{K})}{(B - \hat{K})} \\
&\leq \mu([B, \infty)) + \sup_{K < B} \frac{C_\mu(B) - P_\mu(K)}{(B - K)} \\
&\leq \inf_{\tau \in \mathcal{T}(\mu)} \mathbb{P}(J_\tau \geq B) \\
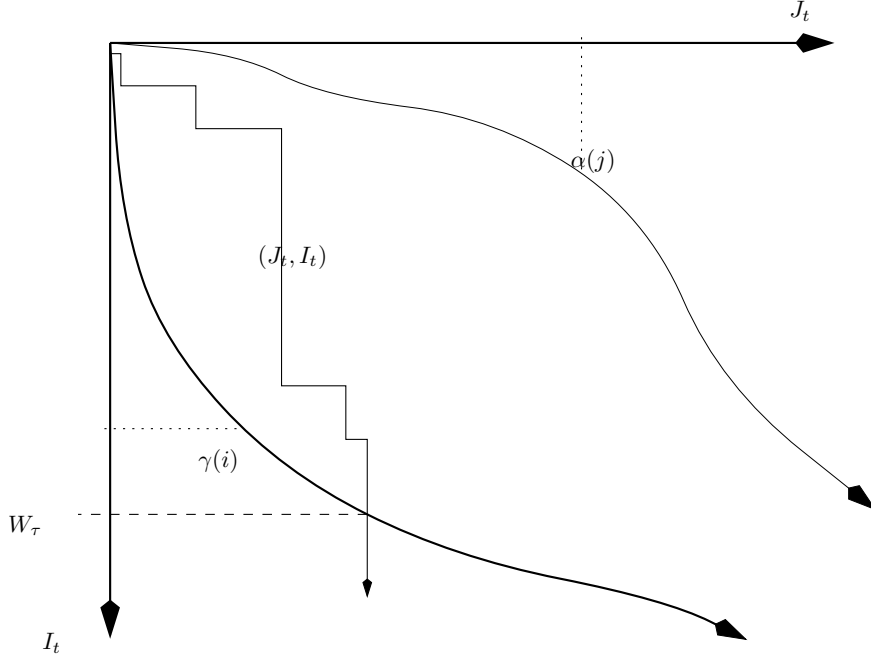&\leq \mathbb{P}(J_{\tau_P} \geq B)
\end{aligned}$$

Figure 7: The Perkins embedding. Given monotonic $\alpha(j)$ and $\gamma(i)$, the stopping rule $\tau = \inf\{u : W_u \notin \{\alpha(J_u), \gamma(I_u)\}\}$ is such that when $W$ is stopped it is either at a new maximum, or at a new minimum.

Hence, there is equality throughout, and $\hat{K}$ maximises (13).

The above construction gives a solution of the Skorokhod embedding which minimises $\mathbb{P}(J_\tau \geq B)$ for a fixed $B$. The Perkins embedding has the property that it minimises $\mathbb{P}(J_\tau \geq B)$ over all embeddings simultaneously for all $B > 0$. (It also minimises $\mathbb{P}(I_\tau \leq C)$ for all $C < 0$, where $I_t = \inf\{W_u; u \leq t\}$.) Define

$$\alpha(j) = \arg\max_{\alpha < 0} \frac{C_\mu(j) - P_\mu(\alpha)}{j - \alpha} \qquad \gamma(i) = \arg\min_{\gamma > 0} \frac{P_\mu(i) - C_\mu(\gamma)}{\gamma - i}$$

Let $\tau = \inf\{U : W_u \notin (\alpha(J_u), \gamma(I_u))\}$ Then $W_\tau \sim \mu$ and $\mathbb{P}(J_\tau \geq B)$ attains the lower bound from (14) uniformly in B. See Figure 7.

## 3.9  Financial interpretation of the Perkins construction

Suppose that $S_t$ is a forward price, and that calls are traded with maturity $T$ on $S$. Suppose that such calls are traded with a continuum of strikes $K$ with prices $C(K)$ which are consistent with no-arbitrage. This is equivalent to assuming we know the marginal distribution of $S_T$.

One candidate model for which option prices are given by $C$ is when the price process $(S_t)_{0 \leq t < T}$ is constant and then there is a jump at $T$, taken such that $\mathbb{E}[(S_T - K)^+] = C(K)$. For this price process, for $B > S_0$, $\mathbb{P}(J_T^S \geq B) = \mathbb{P}(S_T \geq B)$, where $J_t^S = \sup_{0 \leq u \leq t} S_u$. This is the lowest possible value for $\mathbb{P}(J_T^S \geq B)$ given the law of $S_T$.

Now suppose that the price process $S_t$ is assumed to be continuous. The effect of this assumption is to ensure that if we buy $S$ the first time that the price gets to

level $B$ or above, then the price paid is $B$.

Consider the one-touch digital option with barrier $B > S_0$ and maturity $T$. In particular, consider the payoff $\tilde{F} = I_{\{\sup_{0 \le t \le T} S_t \ge B\}}$, and consider the problem of finding a highest, model-independent, lower bound on the price of this option.

In the setting of the general theory we have $F_K = (S_T - K)^+$, with $f_K = C(K)$, and $\Theta \supseteq \{cI_{(H_B, T]}\}$. Then, modulo translation to non-centred random variables, the analysis of the previous section applies and

$$\mathcal{P}(\tilde{F}) \ge |C'(B-)| + \sup_{K < B} \frac{C(B) - P(K)}{(B - K)},$$

with equality for any model for which $(S_T < \hat{K}(B)) \subseteq (H_B \le T) \setminus (S_T \ge B) \subseteq (S_T \le \hat{K}(B))$ where $\hat{K}(B) = \arg\max_{K < B}(C(B) - P(K))/(B - K)$.

## 3.10  Barrier options

As another example of a model-independent bound for an exotic option consider a knock-in barrier call option on the forward price $S_t$ with maturity $T$, strike $K$ and barrier $B > \max\{S_0, K\}$. The option has payoff $(S_T - K)^+ I_{\{H_B \le T\}}$, paid at $T$.

If we place ourselves in the setting of Section 2.7 then we have the fundamental inequality, for $k \in [K, B)$;

$$(S_T - K)^+ I_{\{H_B \le T\}} \le \frac{B - K}{B - k}(S_T - k)^+ + \frac{k - K}{B - k}(S_{H_B} - S_T)I_{\{H_B \le T\}}.$$

The first term is a semi-static hedging strategy in calls, and the second is the result of a forward investment in the underlying. This yields the price bound on the barrier option

$$\mathcal{P}((S_T - K)^+ I_{\{H_B \le T\}}) \le (B - K) \inf_{k \in [K, B)} \frac{C(k)}{B - k} \tag{16}$$

Recall the definition of the barycentre function $b$ in (7) and its inverse $\beta$. To emphasize the dependence of these quantities on the law $\mu$ representing the marginal distribution of $S_T$ we use a subscript and write $b_\mu$ and $\beta_\mu$. Then, for $K \ge \beta_\mu(B)$ we find that the infimum in (16) is attained at $k = K$ and the upper bound on the knock-in barrier option is equal to the price of the call without the knock-in feature. However, for $K < \beta_\mu(B)$ the upper bound is strictly less than the call, and is equal to $(B - K)C(\beta_\mu(B))/(B - \beta_\mu(B))$. It is possible to construct a model based on the Azéma-Yor embedding and a time-change to show that this price can be attained.

For the lower bound, if there are no assumptions on the underlying then we have the trivial inequality

$$(S_T - K)^+ I_{\{H_B \le T\}} \ge (S_T - K)^+ I_{\{S_T \ge B\}}$$

with associated price bound

$$\mathcal{P}((S_T - K)^+ I_{\{H_B \le T\}}) \ge C(B) + (B - K)|C'(B-)|.$$

This price bound is attained for a jump model with a single jump at time $T$.

If we are prepared to assume that the underlying price process is continuous then we can refine the lower bound. (This assumption has the effect that $S_{H_B} \equiv B$, so that if we go short the forward the first time that the underlying equals or exceeds the barrier, then the gains from trade are $(B - S_T)$). For $k \in [K, B)$ we have the inequality

$$
\begin{aligned}
(S_T - K)^+ I_{\{H_B \leq T\}} \quad \geq \quad & (S_T - K)^+ - \frac{B - K}{B - k} \left\{ (S_T - k)^+ - (S_T - B)^+ \right\} \\
& + (B - K) I_{\{S_T \geq B\}} + \frac{k - K}{B - k} (S_{T \wedge H_B} - B)
\end{aligned}
$$

(To see this, split into two cases, $(H_B \leq T)$ and $(H_B > T)$, and observe that the right-hand-side is piecewise linear, so that it is sufficient to check that the inequality holds where there are changes in value or kinks, i.e. at $K, k$ and $B$.) Then

$$
\begin{aligned}
\mathcal{P}((S_T - K)^+ I_{\{H_B \leq T\}}) \quad \geq \quad & C(K) + (B - K)|C'(B-)| \\
& - \inf_{k \in [K,B)} \left\{ \frac{B - K}{B - k} (C(k) - C(B)) + \frac{(k - K)(B - S_0)}{(B - k)} \right\}
\end{aligned}
$$

This bound can be attained by a time-change of the Perkins solution to the Skorokhod embedding problem.

See Brown et al [9] for a further discussion of barrier options, including knock-out barriers, and knock-in and knock-out puts, and Cox and Oblój [18, 19] for a discussion of two-sided barriers.

## 3.11  Potential theory and the Chacon-Walsh embedding

To date we have supposed that the initial law of $W$ is trivial (and then by a parallel shift we may as well assume that $W_0 \sim \delta_0$). Now we suppose that $W_0 \sim \mu_0$. Given centred probability measures $\mu_0$ and $\mu_1$, we ask when does there exist a stopping time $\tau$ such that

$$
W_0 \sim \mu_0, W_\tau \sim \mu_1 \text{ and } W_{t \wedge \tau} \text{ is uniformly integrable?} \tag{17}
$$

Clearly, if we can decide when it is possible to solve (17), and if we can find constructions where they exist, then we can solve the iterated problem: given $(\mu_i)_{0 \leq i \leq n}$ find an increasing sequence of stopping times $0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_n$ (where possible) such that

$$
W_{\tau_i} \sim \mu_i, \qquad W_{t \wedge \tau_n} \text{ is uniformly integrable.} \tag{18}
$$

**Theorem 3.7 (Röst [52])** *A necessary and sufficient condition for the existence of a solution to (17) is $C_{\mu_0}(K) \leq C_{\mu_1}(K)$ for all $K$.*

**Proof.** Necessity follows from an application of the conditional Jensen's inequality. Since $(w - K)^+$ is convex as a function of $w$,

$$
C_{\mu_1}(K) = \mathbb{E}[(W_\tau - K)^+] = \mathbb{E}[\mathbb{E}[(W_\tau - K)^+|\mathcal{F}_0]] \geq \mathbb{E}[(W_0 - K)^+] = C_{\mu_0}(K).
$$

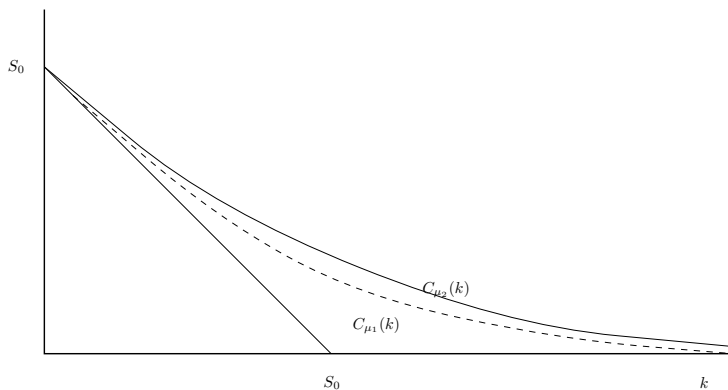Sufficiency follows from one of the constructions given below. $\qquad \square$

Figure 8: Given a pair of call price functions satisfying the no-arbitrage conditions, there exists a model for the forward price which is consistent with those prices provided the call prices are increasing in maturity.
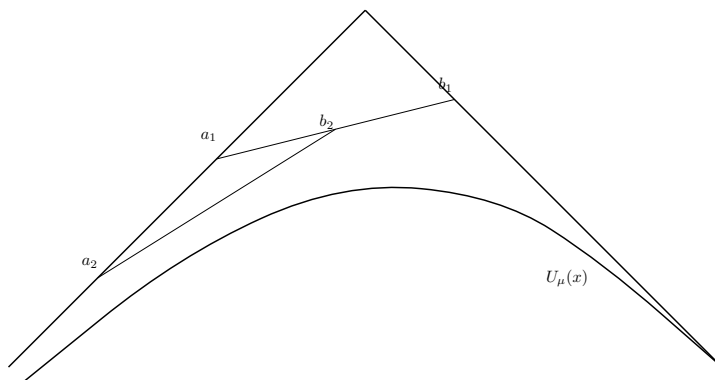


Figure 9: Potentials in the Chacon-Walsh picture.

Put-call parity is a consequence of the identity $(y-k)^+ - (y-k)^- = y-k$. Given also $(y-k)^+ + (y-k)^- = |y-k|$ we have $2(y-k)^+ = |y-k| + (y-k)$. Hence $C_{\mu_0}(K) \leq C_{\mu_1}(K)$ is equivalent to $U_{\mu_0}(K) \geq U_{\mu_1}(K)$ where $U_\eta(K) = -\mathbb{E}^{X\sim\eta}|X - K|$ is the potential (or minus the price of a straddle). Then the fact that $C$ is a convex function, with $C_\eta(k) \geq k^-$ implies that $U_\eta$ is concave and $U_\eta(x) \leq -|x|$. Kinks in the derivative $U_\eta'$ correspond to atoms of $\eta$; conversely straight segments of $U_\eta$ correspond to intervals where $\eta$ places no mass.

The Chacon-Walsh solution of the SEP is based on potentials. We return to the case where $W_0 = 0$ and $\mu$ is a centred probability measure, but it is immediate from the inductive nature of the construction that provided the potentials are decreasing, then the same construction will work as an embedding for any initial law.

Let $U_0(x) = U_{\delta_0}(x) = -|x|$, and $\tau_0 = 0$.

Fix $a_1 < 0 < b_1$. Set $\tau_1 = \inf\{t > \tau_0 : W_t \notin (a_1, b_1)\}$, $\mu_1 \sim \mathcal{L}(W_{\tau_1})$ and $U_1 \equiv U_{\mu_1}$. Then $U_1(x) = -|x|$ for $x \leq a$ and $x \geq b$, and $U_1(x) = a_1 - (b_1 + a_1)(x - a_1)/(b_1 - a_1)$ for $x \in (a_1, b_1)$. See Figure 9.

We now proceed inductively. Fix $a_n < b_n$. Set $\tau_n = \inf\{t > \tau_{n-1} : W_t \notin (a_n, b_n)\}$, $\mu_n \sim \mathcal{L}(W_{\tau_n})$, $U_n \equiv U_{\mu_n}$. Then $U_n$ is a continuous concave function with $U_n(x) = U_{n-1}(x)$ outside $(a_n, b_n)$ and $U_n$ linear on $(a_n, b_n)$.
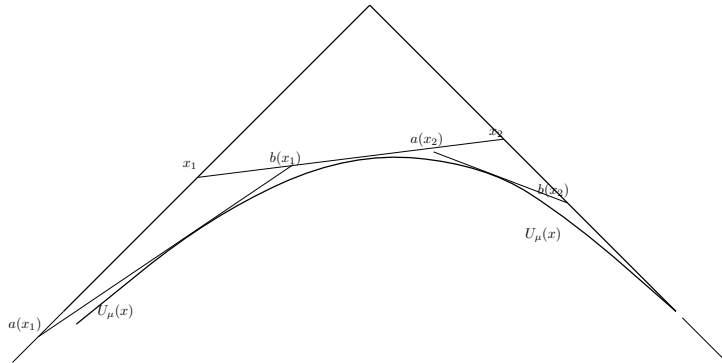
Figure 10: The Dubins construction in the Chacon-Walsh picture. Shown are the potentials at the first and second stages of the iteration, together with the starting and terminal potentials.

The idea is to choose the points $(a_n, b_n)$ appropriately in such a way that $U_n \downarrow U_\mu$. Then the increasing sequence of stopping times $\tau_n$ converges (to $\tau$ say) and $W_\tau \sim \mu$.

There is a large amount of freedom in the choice of the sequence $(a_n, b_n)$. A natural choice is given by the Dubins [22] solution to the SEP. In the Dubins construction the sequence $(a_n, b_n)$ is defined as follows.

Let $\mathcal{A}_0 = \{0\}$. Given $U_\mu$ let $(a(0), b(0))$ be the $x$-coordinates of the points where the tangent to $U_\mu$ at $x$ crosses $U_{\delta_0}(x) = -|x|$. (If there is an atom of $\mu$ at zero, then $U_\mu$ has a kink there, and there are many candidate tangents to $U_\mu$. We are free to choose any of these; for definiteness we can choose the tangent with slope equal to the right derivative of $U_\mu$ at $x_0$.) Set $\mathcal{A}_1 = \{a(0), b(0)\}$ and let $\mu_1$ be the law with corresponding potential $U_1(x) = \min\{U_0(x), U_\mu(0) + xU'_\mu(0+)\}$. See Figure 10.

Given $U_\mu$, $\mathcal{A}_n$ and potential $U_n$, for each $x \in \mathcal{A}_n$ let $(a(x), b(x))$ be the $x$-coordinates of the points where the tangent to $U_\mu$ (with slope $U'_\mu(x+)$) intersects $U_n$. Note that if $x_i < x_j \in \mathcal{A}_n$ then $a(x_i) \le b(x_i) \le a(x_j) \le b(x_j)$. Then let $\mathcal{A}_{n+1} = \{a(x_i), b(x_i); x_i \in \mathcal{A}_n\}$, and $\mu_{n+1}$ the law with corresponding potential $U_{n+1} = \min_{x_i \in \mathcal{A}_n}\{U_n(x), U_n(x_i) + (x - x_i)U'_n(x_i+)\}$.

The Azéma-Yor embedding [3] can also be phrased in terms of a Chacon-Walsh potential picture. Rather than considering a sequence of intervals, consider a continuum of intervals, parameterised by $j$, such that $b(j) = j$ and $a(j)$ is the $x$-coordinate of the point where the tangent to $U_\mu$ passes through the point $(j, -j)$. (Note that $a(j)$ is not necessarily unique, but it is increasing, so we can make it unique by making the function $a$ right continuous.) We have

$$\tau_j = \inf\{t \ge \tau_{j-}; W_t \notin (a(j), j)\}$$

A few lines of calculus show that $a(j)$ is the inverse barycentre function.

## 3.12 The Skorokhod solution in the potential picture

Skorokhod's solution [54] of the Skorokhod embedding problem can also be represented in the Chacon-Walsh picture. The construction is similar to the Hall solution,
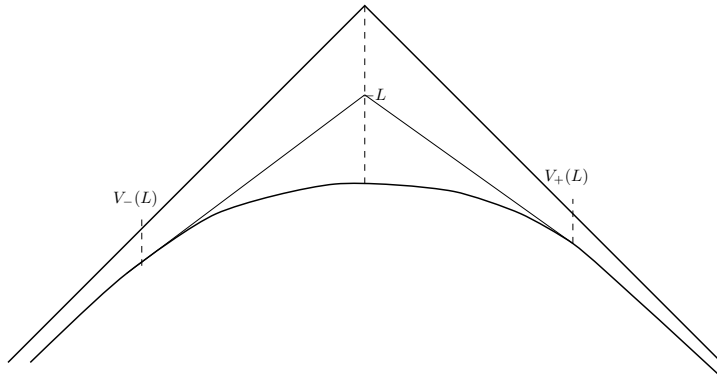
Figure 11: The Skorokhod embedding in the potential picture. The points $V_\pm(l)$ are determined by the (unique) tangents to $U_\mu$ which pass through $(0, -l)$.

and the idea is to choose a single exogenous random variable $L$, to use that random variable to define an interval, and to set $\tau$ to be the first time that Brownian motion leaves that interval.

Let $c = \int_{(0,\infty)} x\mu(dx)$, so that $U_\mu(0) = -2c$. For $l \in (0, 2c)$, let $V_+(l)$ be the $x$-coordinate of the point in $\mathbb{R}^+$ such that the tangent to $U_\mu$ at $V_+(l)$ crosses the $y$-axis at $y = -l$. (If $U_\mu$ has linear sections then there may be an interval of such $V_+(l)$, in which case we take the largest. With this definition the increasing function $V_+(l)$ is right-continuous.) Similarly, define the decreasing function $V_-(l)$ to be the $x$-coordinate of the point in $\mathbb{R}^-$ such that the tangent to $U_\mu$ at $V_+(l)$ crosses the $y$-axis at $y = -l$. See Figure 11.

If $\mu$ has a continuous distribution function then we take $L$ to have distribution given by $\mathbb{P}(L \geq l) = \mu((V_-(l), V_+(l)))$, for $0 \leq l \leq 2c$. More generally, to allow for atoms we take $\mathbb{P}(L \geq l)$ to equal one half the difference in slopes between the tangents joining $(V_-(l), U_\mu(V_-(l)))$ with $(0, -l)$ and $(0, -l)$ with $(V_+(l), U_\mu(V_+(l)))$. Finally we set $\tau_l = \inf\{u : W_u \notin (V_-(l), V_+(l))\}$, and then $W_{\tau_L} \sim \mu$.

Rather than prove this result, we give an example. Suppose $\mu$ is the law of a uniform random variable on $(-1, 1)$. Then $U_\mu(x) = -(x^2 + 1)/2$, $c = 1/4$, and $V_\pm(l) = \pm\sqrt{1 - 2l}$. Then, by symmetry, for $w \in (0, 1)$,

$$\mathbb{P}(W_{\tau_L} \leq w) - \mathbb{P}(W_{\tau_L} \geq 0) = \frac{1}{2}\mathbb{P}(V_+(L) \leq w) = \frac{1}{2}\mathbb{P}(L \geq (1 - w^2)/2) = \frac{w}{2}$$

as required.

**Remark 3.8** The original Skorokhod construction assumes the existence of a non-negative exogenous random variable $Z$, and then sets $\tau$ to be the first time the Brownian motion leaves the interval $(\rho(Z), Z)$ for a non-positive decreasing function $\rho$. From our representation of the Skorokhod solution in the potential picture we see $Z \sim V_+(L)$ and $\rho(z) = V_-(V_+^{-1}(z))$.

## 3.13 The Bass solution

Whereas our philosophy has been to construct solutions of the Skorokhod embedding problem for Brownian motion with the aim of generating models for the martingale

price process via a time-change, Bass [5] constructs a (time-inhomogeneous) martingale diffusion such that $X_1 \sim \mu$, and then uses this to construct a solution of the SEP for $(W, \mu)$. As such the Bass construction is interesting in the finance context because it gives directly (a family) of diffusion processes with marginal law $\mu$ at time 1.

Suppose that on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq 1}, \mathbb{P})$, the diffusion process $Z$ solves the stochastic differential equation $dZ_t = \eta(Z_t, t)dB_t + \gamma(Z_t, t)dt$ with $Z_0 = z$, and is such that the law of $Z_1$ is known and given by the atom-free centred distribution function $F_Z$. Suppose further that we can calculate the time-1 law of $Z_t$ from any starting point $(Z_t, t)_{0 \leq t < 1}$. (The simplest case is when $Z$ is itself a Brownian motion, but other constructions could be based on a Brownian bridge, a Bessel process or an Ornstein-Uhlenbeck process.) Given $\mu$, (and associated distribution function $F_\mu$), fix $g(z) = F_\mu^{-1}(F_Z(z))$. Then $g$ is increasing and $g(Z_1) \sim \mu$.

Now set $X_t = \mathbb{E}[g(Z_1)|\mathcal{F}_t]$. Then $X_t = h(Z_t, t)$, for a function $h$ which is increasing in its first argument, and has inverse $Z_t = H(X_t, t)$ say. Then $X_1 \sim \mu$ and

$$dX_t = h'(Z_t, t)\eta(Z_t, t)dB_t = a(X_t, t)dB_t$$

where $a(x, t) = h'(H(x, t), t)\eta(H(x, t), t)$. Note that $X_0 = \mathbb{E}[g(Z_1)] = 0$.

Given a Brownian motion $W$ and the function $a$, define the increasing additive functional $\Gamma$ and its inverse $A$ via

$$\Gamma_0 = 0, \qquad \frac{d\Gamma_u}{du} = \frac{1}{a(W_u, \Gamma_u)^2}, \qquad A_t = \inf\{u : \Gamma_u > t\}.$$

Set $X_t = W_{A_t}$. Then, with $u = A_t$,

$$\frac{dA_t}{dt} = \frac{1}{d\Gamma_u/du} = a(W_u, \Gamma_u)^2 = a(X_t, t)^2,$$

and for some Brownian motion $\hat{W}_t$,

$$dX_t = \sqrt{dA_t/dt} \ d\hat{W}_t = a(X_t, t)d\hat{W}_t.$$

Finally, since $X_1 \sim \mu$, we have $W_{A_1} \sim \mu$ and hence $\tau = A_1$ is a solution of the Skorokhod embedding problem for $(W, \mu)$.

## 3.14 A general schematic for solutions

Solutions of the Skorokhod embedding problem can be classified, very approximately, into three main types.

The first type (such as Dubins [22] and also Azéma-Yor [3]) fall under the general heading of potential theoretic constructions, and have a pictorial representation in the Chacon-Walsh [13] picture. It is also possible to place the original embedding due to Skorokhod in this category.

For the second type of embedding the aim is to solve the SEP subject to an additional optimality criterion. In particular, given some functional $H(W_t; 0 \leq t \leq \tau)$, the aim is to solve the SEP for $(W, \mu)$ such that $H$ is maximised, in the sense of first order stochastic dominance, or sometimes in the weaker sense of maximised

in expectation. As we have seen, the Azéma-Yor [3] and Perkins [47] embeddings fall into this category, but so do the Root/Röst [51, 53] constructions (based on the additive functional $t$), the Vallois [55, 56] embedding (based on the local time) and the Obłój-Yor [44] solutions (based on the age of the excursion), amongst others. We will discuss the Root and Röst solutions in detail below, because of the direct financial interpretations.

Finally, there are miscellaneous embeddings which fall into neither of the above categories. These include the Hall [29] and Bass [5] solutions described above, and also the Bertoin-Le Jan [7] solution.

# 4 Other applications of the Skorokhod Embedding Problem

## 4.1 The central limit theorem

These notes are primarily concerned with the connection between Skorokhod embeddings and derivative pricing and hedging. However, there is one other important application which deserves comment, namely the use of Skorokhod embeddings to prove the Central Limit Theorem.

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables, with law $\mu$ such that $\mu$ has mean zero and variance 1. Let $S_n = \sum_{k \leq n} X_k$.

Let $B^{(i)}$ be Brownian motions, and let $\tau^{(i)}$ be any of the minimal embeddings of $\mu$ in $B^{(i)}$ described to date. The times $\tau^{(i)}$ are independent identically distributed random variables with mean 1, and $B^{(i)}_{\tau^{(i)}}$ are independent. The sequence $(X_i)_{i \geq 1} = (B^{(i)}_{\tau^{(i)}})_{i \geq 1}$ has the properties listed in the previous paragraph.

Let $T_0 = 0$ and $T_j = \sum_{i \leq j} \tau^{(i)}$. Now define a Brownian motion $(W_t)_{t \geq 0}$ by

$$W_t = \sum_{i \leq j} B^{(i)}_{\tau^{(i)}} + B^{(j+1)}_t \qquad T_j \leq t < T_{j+1} \tag{19}$$

Then $W_{T_n} = S_n$, and by Brownian scaling, $S_n/\sqrt{n} = W_{T_n}/\sqrt{n} \sim W_{T_n/n}$.

The idea is that, since $T_n/n \to 1$ and since at fixed times Brownian motion has Gaussian marginals, then in the limit $S_n/\sqrt{n}$ will inherit a standard normal distribution.

**Theorem 4.1** *With the notation and hypotheses of this section $S_n/\sqrt{n}$ converges in law to a standard normal distribution.*

**Proof.** Let $W$ be the Brownian motion defined in (19). Let $Z_t^{(n)} = W_{nt}/\sqrt{n}$. Then, for each $n$, $Z^{(n)}$ is a Brownian motion. By the above construction we have

$$\frac{S_n}{\sqrt{n}} = \frac{W_{T_n}}{\sqrt{n}} = Z_{T_n/n}^{(n)} \sim Z_{T_n/n},$$

for a Brownian motion $Z$. Fix $\epsilon > 0$ and choose $\delta > 0$ so that

$$\mathbb{P}\left( \sup_{t:|1-t|<\delta} |Z_t - Z_1| > \epsilon \right) < \frac{\epsilon}{2}$$

By the strong law of large numbers $T_n/n \to 1$ almost surely, so that we can choose $N$ large, such that for all $n \geq N$,

$$\mathbb{P}(|T_n/n - 1| > \delta) < \frac{\epsilon}{2}$$

Combining these last two displayed equations we conclude that

$$\mathbb{P}(|W_{T_n}/\sqrt{n} - W_n/\sqrt{n}| > \epsilon) = \mathbb{P}(|Z_{T_n/n} - Z_1| > \epsilon) < \epsilon$$

Then $\mathbb{P}(W_n/\sqrt{n} < x - \epsilon) - \epsilon \leq \mathbb{P}(S_n/\sqrt{n} < x) \leq \mathbb{P}(W_n/\sqrt{n} < x + \epsilon) + \epsilon$ and the result follows.

## 4.2 Non-centred target laws

There is a very easy way to embed integrable, but non-centred, target laws in Brownian motion started at zero — simply run Brownian motion until it hits the mean, and thereafter use a favourite embedding for a centred distribution. Such an embedding is minimal, see Cox and Hobson [17]. Note however that whereas in the centred case if $\mu$ has support in an interval $I = [a, b]$ then for a minimal embedding $\tau \leq H_a \wedge H_b$, in the non-centred case this is no longer true. If the mean of the target distribution is positive then it is no longer possible to have an embedding for which $\tau \leq H_z$ almost surely, for any $z < 0$. (If it were, then $(W_{t \wedge \tau})_{t \geq 0}$ would be bounded below and Fatou's lemma would imply that the law of $W_\tau$ had negative mean.)

## 4.3 Embeddings in time-homogeneous diffusions

Suppose that $X$ is a time-homogeneous diffusion with state space $I \subseteq \mathbb{R}$ with absorbing or inaccessible endpoints, and consider the SEP for $(X, \mu)$. The two fundamental questions are: when is it possible to embed $\mu$, and if it is possible, how can it be done?

In fact the problem can easily be reduced to the case of Brownian motion. The relevant insight is due to Azéma and Yor [4], see also Pedersen and Peskir [46] and Cox and Hobson [15].

Suppose that $X$ solves $dX = a(X)dB + b(X)dt$. If $s$ solves $a(x)^2 s''(x)/2 + b(x)s'(x) = 0$, then $s$ is a scale function for $X$ and $Y = s(X)$ is a local martingale with state space $s(I)$. We have $dY = s'(s^{-1}(Y))a(s^{-1}(Y))dB$. We can choose $s$ such that $s(x_0) = 0$ and then by the Dambis-Dubins-Schwarz Theorem $Y_t = W_{\Gamma_t}$ for some Brownian motion $W$, where $\Gamma_t = \langle Y \rangle_t$.

Let $\nu$ be the image of $\mu$ under $s$. If we can solve the problem for $(W, \nu)$ then we can construct a solution $\tau$ for $(Y, \nu)$ via time-change, and then $\tau$ is a solution of the SEP for $(X, \mu)$. Thus, if $W_\tau \sim \mu$ and if $A$ is the inverse to $\Gamma$, then $Y_{A_\tau} = W_\tau \sim \nu$, and $X_{A_\tau} \sim s(Y_{A_\tau}) \sim \mu$.

Note that there is no reason to expect that $\nu$ is centred as a measure on $\mathbb{R}$. By the remarks in the previous section we can construct embeddings for non-centred target distributions provided we do not impose any restrictions that $\tau$ is less than the first exit time from some interval. However, for the purposes of embedding in

non-martingale diffusions we must have that $\tau$ is smaller than the first exit time from $s(I)$. If we set $m_\nu = \int_{\mathbb{R}} x\nu(dx)$, then we find that the answer to the question about whether $\nu$ can be embedded in $Y$ depends on the joint properties of $s(I)$ and $m_\nu$:

if $s(I) = \mathbb{R}$, then there is a solution for the SEP for any $\nu$,

if $s(I) = (-a, \infty)$ then there is a solution for the SEP for $\nu$ if and only if $m_\nu \le 0$,

if $s(I) = (-\infty, b)$ then there is a solution for the SEP for $\nu$ if and only if $m_\nu \ge 0$,

if $s(I) = (-a, b)$ then there is a solution for the SEP for $\nu$ if and only if $m_\nu = 0$.

# 5 The Root and Röst solutions and options on volatility

## 5.1 The Root and Röst embeddings

Return to the classical setting of minimal solutions of the SEP for $(W, \mu)$ where $\mu$ is a centred probability measure. Suppose $\mu$ has finite variance. Then there are embeddings $\tau$ such that $\mathbb{E}[\tau] = \int_{\mathbb{R}} x^2 \mu(dx)$, (and as described in Corollary 3.3 above, for any embedding either $\mathbb{E}[\tau] = \int_{\mathbb{R}} x^2 \mu(dx)$, or $\tau$ has infinite expectation). In particular, it does not make sense to attempt to find the solution of the SEP for which $\mathbb{E}[\tau]$ is minimised; any minimal stopping time will satisfy this criterion. The task of maximising $\mathbb{E}[\tau]$ over minimal stopping times is equally trivial.

However, the problem of searching for the embedding which minimises (over all stopping times) the variance of $\tau$, or alternatively the embedding which maximises (over uniformly integrable stopping times) the variance of $\tau$ leads to a unique solution.

Since we are looking for an embedding which optimises a functional of $t$ evaluated at the stopping time, it is plausible that the stopping rule should depend on the joint distribution of $W_t$ and the additive functional $t$.

Given $b : \mathbb{R} \to \mathbb{R}_+$, define $\mathcal{R} \equiv \mathcal{R}_b = \{(t, x); t \ge b(x)\}$. Let $\tau_b = \inf\{u : (u, W_u) \in \mathcal{R}_b\}$. We will describe the function $b$ (and sometimes the region $\mathcal{R}_b$) as a barrier.

**Theorem 5.1 (Root [51], Röst [53])**     1. *Given a centred probability measure $\mu$ there exists $b$ such that $\tau_b$ is a solution of the SEP for $(W, \mu)$; moreover $\tau_b$ is minimal.*

   2. *$\tau_b$ minimises the variance of $\tau$ amongst all embeddings of the law of $W_{\tau_b}$. More generally $\tau_b$ minimises $\mathbb{E}\left[\int_{t \wedge \tau}^{\tau} h(W_s) ds\right]$ for all non-negative functions $h$ and times $t$.*

The embedding property of the solution is due to Root [51], and henceforth we will label this solution as the Root solution. The optimality properties were conjectured by Kiefer [38] and proved by Röst [53]. Taking $h = 1$ we find that the Root solution minimises the *residual expectation* $\mathbb{E}[(\tau - t)^+] \equiv \mathbb{E}[(\tau - \tau \wedge t)]$ for all $t$. Since any convex function can be decomposed into a positive combination of such functions (and since $\mathbb{E}[\tau]$ is invariant across minimal embeddings) it follows that the Root barrier solution minimises $\mathbb{E}[f(\tau)]$ for any convex $f$.
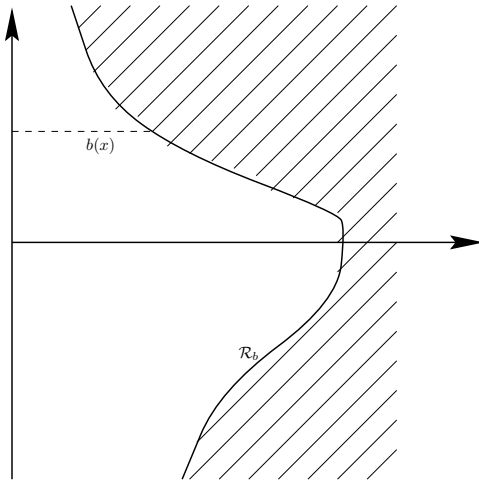
Figure 12: The Root barrier solution. Given $b : \mathbb{R} \to \mathbb{R}_+$, or equivalently $\mathcal{R} \equiv \mathcal{R}_b = \{(t,x); t \geq b(x)\}$, let $\tau_b = \inf\{u : (u, W_u) \in \mathcal{R}_b\}$ be the time of first entry by $(u, W_u)$ into the shaded region. The idea is to choose $b$ such that $W_\tau \sim \mu$.

**Example 5.2**     1. Suppose $\mu \sim N(0,1)$. Then the barrier is given by $b(x) = 1$, $\mathcal{R}_b = \{(t,x); t \geq 1\}$ and $\tau_b = 1$.

2. Suppose $\mu \sim (\delta_{-1} + \delta_{+1})/2$. Then $\mathcal{R}_b = \{(x,t) : x \leq -1, x \geq 1\}$, and $\tau_b = H_{\pm 1}$.

3. Suppose $\mu \sim p\delta_{-1} + p\delta_{+1} + (1 - 2p)\delta_0$ for $0 \leq p \leq 1/2$. Then $\mathcal{R}_b = \{(x,t) : x \leq -1, x \geq 1\} \cup \{(0,t); t \geq t_0(p)\}$. The quantity $t_0(p)$ does not have a simple expression, but it is clear that $t_0(0) = 0$, $t_0(1/2) = \infty$ and that $t_0(p)$ is strictly increasing in $p$. Hence the existence of a barrier which embeds $\mu$ is guaranteed.

There is also a converse to the Root embedding, which is based on a reverse barrier. The embedding is attributed to Röst by Meilijson [41], see also Oblój [43].

We assume that $\mu$ has no atom at zero. The construction can be extended to include an atom at 0 by independent randomisation.

**Theorem 5.3 (Röst)** *Given $b : \mathbb{R} \to \mathbb{R}_+$, define $\bar{\mathcal{R}} \equiv \bar{\mathcal{R}}_b = \{(t,x); t \leq b(x)\}$. Let $\bar{\tau}_b = \inf\{u : (u, W_u) \in \bar{\mathcal{R}}_b\}$. Then,*

1. *given a probability measure $\mu$ with no atom at 0, there exists $b$ such that $\bar{\tau}_b$ is a solution of the SEP for $(W, \mu)$; moreover $\bar{\tau}_b$ is minimal;*

2. *$\bar{\tau}_b$ maximises the variance of $\tau$ amongst all uniformly integrable embeddings of the law of $W_{\bar{\tau}_b}$. More generally $\bar{\tau}_b$ maximises $\mathbb{E}[\int_{t \wedge \tau}^{\tau} h(W_s)ds]$ for all non-negative functions $h$ and times $t$.*

**Remark 5.4** In fact both the Root and Röst embeddings can be extended to any continuous Markov process $(X_t)_{t \geq 0}$, and the optimality properties still hold. This will be clear from the fact that the proofs we sketch below do not use any properties of Brownian motion beyond the Markov property and continuity of paths.
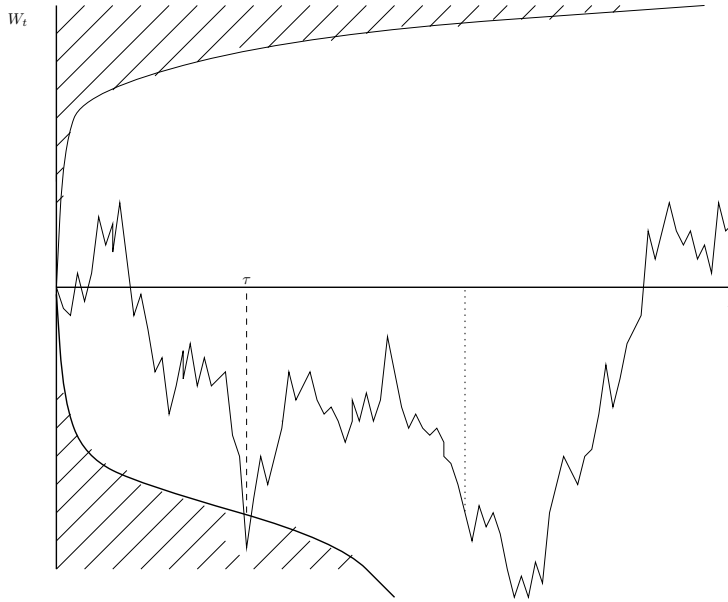
Figure 13: The Röst reversed barrier. The embedding $\tau$ is the first time Brownian motion $W$ enters the shaded region.

**Sketch of proof of Theorem 5.1.** Suppose $\mu$ consists of a finite collection of atoms at locations $\mathcal{X} = \{x_0 < x_1 < \ldots < x_N < x_{N+1}\}$, with associated probabilities $\{p_0, p_1, \ldots, p_N, p_{N+1}\}$. Suppose also that $\mu$ is centred, so that $\sum_{i=0}^{N+1} x_i p_i = 0$.

We search for an embedding for $(W, \mu)$ of barrier form. Since the embedding is minimal, by Corollary 3.4 we must have that the process is stopped before it leaves the interval $[x_0, x_{N+1}]$. Furthermore, the process should only stop at the points $x_i \in \mathcal{X}$, hence the barrier $\mathcal{R}$ must be of the form $\mathcal{R} = \{(x, b); x_i \in \mathcal{X}, b \geq b_i\}$. Moreover, we must have $b_0 = 0 = b_{N+1}$.

Let $\gamma = \{\gamma_1, \ldots, \gamma_N\}$ be a vector in $\mathbb{R}_+^N$, and augment it with the values $\gamma_0 = 0 = \gamma_{N+1}$. Let $\tau_\gamma = \inf\{u : B_u = x_i, \quad u \geq \gamma_i; 0 \leq i \leq N + 1\}$. Set $\Gamma_\mu = \{\gamma : \mathbb{P}(B_{\tau_\gamma} = x_i) \leq p_i; 1 \leq i \leq N\}$. Note that if $\gamma \in \Gamma_\mu$ we must have $\mathbb{P}(B_{\tau_\gamma} = x_0) + \mathbb{P}(B_{\tau_\gamma} = x_{N+1}) \geq p_0 + p_{N+1}$ to compensate. Then $\gamma \in \Gamma_\mu$ if the associated stopping rule never embeds more than allowed under $\mu$ at any point in the interior of the support of $\mu$. The excess mass is embedded at the endpoints.

We claim that if $\hat{\gamma}$ and $\tilde{\gamma}$ are elements of $\Gamma_\nu$ then so is $\underline{\gamma}$ where $\underline{\gamma}_i = \hat{\gamma}_i \wedge \tilde{\gamma}_i$. To verify the claim, fix $i \in \{1, \ldots, N\}$. Without loss of generality we may assume that $\hat{\gamma}_i \leq \tilde{\gamma}_i$. Then $\tau_{\underline{\gamma}} = \tau_{\hat{\gamma}}$ on $B_{\tau_{\underline{\gamma}}} = x_i$, and $\tau_{\underline{\gamma}} \leq \tau_{\hat{\gamma}}$ otherwise. See Figure 14. Thus $(\omega : B_{\tau_{\underline{\gamma}}} = x_i) \subseteq (\omega : B_{\tau_{\hat{\gamma}}} = x_i)$ and $\mathbb{P}(B_{\tau_{\underline{\gamma}}} = x_i) \leq p_i$. Since $i$ was arbitrary, the claim is proved.

It follows from the claim that $\Gamma_\mu$ has a minimal element. Moreover, this minimal element must embed $\mu$; if not then for some $i$ we must have $\mathbb{P}(B_{\tau_\gamma} = x_i) < p_i$, and for this $i$ we can reduce the value of $\gamma$ slightly without violating the condition $\mathbb{P}(B_{\tau_\gamma} = x_i) \leq p_i$ and only making the values of $(\mathbb{P}(B_{\tau_\gamma} = x_j))_{j \neq i}$ smaller.

It remains to extend from atomic distributions $\mu_n$ to general $\mu$. Root [51] achieves this by mapping $[0, \infty] \times [-\infty, \infty]$ to $[0, 1] \times [-1, 1]$ and putting a dis-
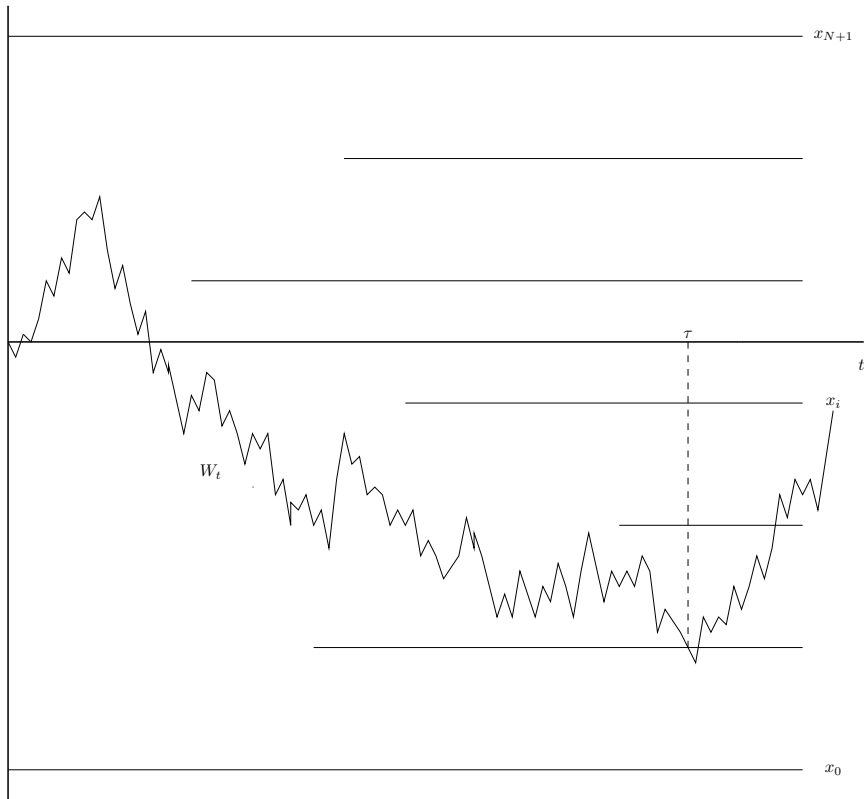
Figure 14: The embedding property of the Root barrier for atomic measures $\mu$ on a finite number of points. The stopping time $\tau_\gamma$ is the first time the Brownian path is at one of the $x_i$ at a time later than $\gamma_i$. If $\gamma$ is such the probability of stopping at $x_i$ is less than $p_i$, then making $\gamma_i$ smaller, will increase the probability that $B_{\tau_\gamma} = x_i$, but can only decrease the probability of stopping at any other point.

tance metric on barriers expressed as sets in the new space. Given the images of the barriers associated with $\mu_n$, there must be a convergent subsequence in the compact space $[0, 1] \times [-1, 1]$. This must take the form of a barrier, and mapping back to the original space we obtain the solution of the SEP for $\mu$.

$\square$

**Sketch of proof of optimality of the Root construction.**

Fix $x$ and $t$ and suppose that $\tau \in \mathcal{T}(\mu)$. We show that the Root barrier stopping time maximises $\mathbb{E}[L_{t\wedge\tau}^x]$ (simultaneously for all $x$ and $t$) amongst all minimal solutions of the Skorokhod embedding problem for $\mu$.

If so, then the Root stopping time maximises $\mathbb{E}[\int_0^{t\wedge\tau} h(W_s)ds]$ for all positive $h$, and since $\mathbb{E}[\int_0^\tau h(W_s)ds]$ is constant across minimal embeddings of $\mu$, it also minimises $\mathbb{E}[\int_{t\wedge\tau}^\tau h(W_s)ds]$.

Let $\tau$ be any embedding of $\mu$ and suppose that there are paths with $(W_\tau \in dx, \tau < t)$ and $(W_\sigma \in dx, t \leq \sigma < \tau)$. Then we can take the terminal mass of $W_\tau$ from those paths which pass through $(\sigma, x)$ and embed that mass by extending the paths from $(\tau, x)$ which are currently stopped at $(W_\tau \in dx, \tau < t)$. See Figure 15.
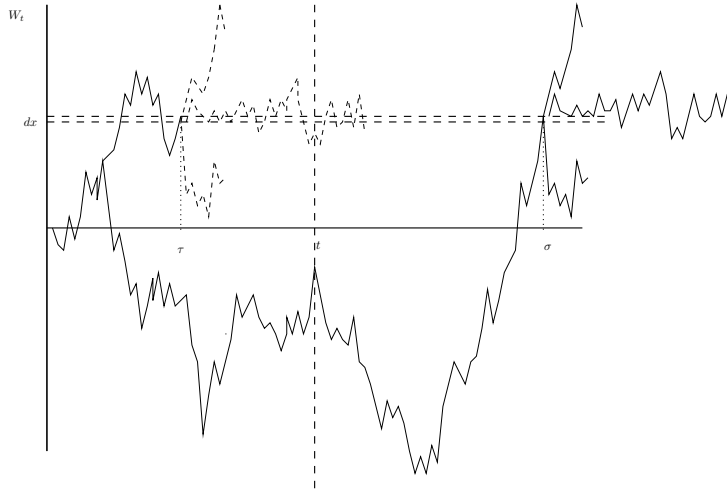
34

Figure 15: Optimality of the Root construction. If for some $(x, t)$ there are paths which stop at $(x, \tau)$ with $\tau < t$, and there are other paths which continue from $(x, \sigma)$ with $t < \sigma < \tau$, then by shifting the continuations of paths from $(x, \sigma)$ so that they are continuations from $(x, t)$ with $t < \tau$, we can only increase the time spent at $x$ before time $t$. We stop some paths which pass through $(x, \sigma)$ early, and extend other paths at $(x, \tau)$ to compensate.

This modification of the stopping rule must increase the value of the local time $L^x_{t \wedge \tau}$.

Hence, if $\tau$ is optimal, in the sense that it maximises $\mathbb{E}[L^x_{t \wedge \tau}]$ amongst minimal solutions of the Skorokhod embedding problem for $\mu$, then for any pair $(x, t)$ we have that either the stopping rule never involves stopping paths at $x$ before $t$, or, no paths cross $x$ after $t$.

Let $\mathcal{R}_t = \{x : \text{no paths cross } x \text{ after } t\}$. Then $\mathcal{R}_t$ is increasing in $t$ and $\mathcal{R} = \cup_{t \geq 0} \mathcal{R}_t$ defines a barrier.

$\square$

The justification for the embedding property and optimality of the Röst reverse barrier is similar. Suppose $\tau$ is such that there are paths with $(W_\sigma \in dx, \sigma < \tau < t)$ and $(W_\tau \in dx, \tau > t)$. Then we can take (some of) the terminal mass of $W_\tau$ from those paths which aren't stopped at $\sigma$ (but do stop by time $t$) and embed that mass by continuing paths from $(x, \tau)$ which are currently stopped at $(W_\tau \in dx, \tau < t)$. This modification of the stopping rule must decrease the value of the local time $L^x_{t \wedge \tau}$.

**Remark 5.5** Note that the proofs of the embedding property and of the optimality of the Root and Röst constructions do not rely on any properties of Brownian motion, beyond the Markov property and the continuity of paths. We will exploit this fact when we discuss applications to the pricing of volatility derivatives below.
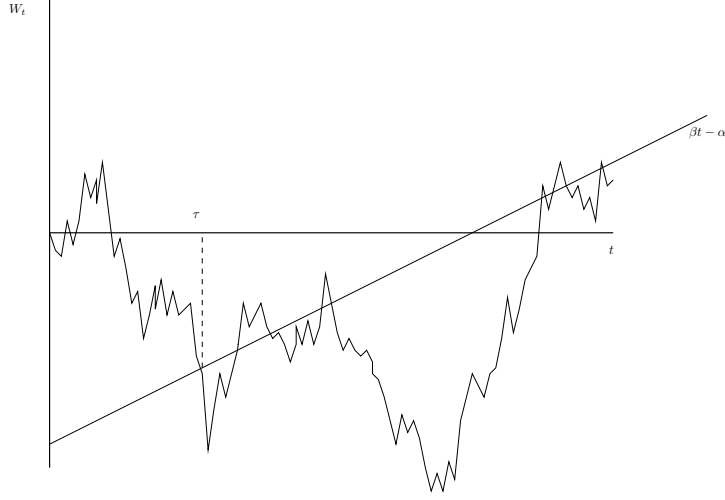
Figure 16: The law $\mu$ is such that Brownian motion stopped the first time it hits the straight line $x = \beta t - \alpha$ has law $\mu$.

**Example 5.6** Suppose, that for $\alpha, \beta > 0$

$$\mu(dx) = \frac{\alpha\sqrt{\beta}e^{-\beta x^2/(2(x+\alpha))}}{\sqrt{2\pi}(x+\alpha)^{3/2}} \, dx; \qquad x > -\alpha$$

The associated barrier function is $b(x) = (x + \alpha)/\beta$, which corresponds to the straight line $x = \beta t - \alpha$. See Figure 16.

Suppose $h$ is the convex function $h(t) = e^{\eta t}$ (with $\eta < \beta^2/2$) and consider the problem of minimising $\mathbb{E}[h(\tau)]$ over stopping times $\tau$ which are solutions of the SEP for $(W, \mu)$.

Define the constants $\phi = \beta - \sqrt{\beta^2 - 2\eta}$ (note that $\phi > 0$) and $\lambda = 2\eta e^{\phi\alpha}/\phi^2$, and set $M(t, x) = \lambda - \lambda e^{\phi x - \phi^2 t/2}$, so that $M(0,0) = 0$ and $M = (M(t, W_t))_{t \geq 0}$ is a martingale.

Define $\Gamma(x) = \inf_{t>0}[e^{\eta t} - M(t, x)]$. The infimum is attained at $t$ solving

$$\eta e^{\eta t} = \frac{\lambda \phi^2}{2} e^{\phi x - \phi^2 t/2}$$

(at least for $x > -\alpha$) which reduces to

$$t = \frac{2\phi(x+\alpha)}{(\phi^2 + 2\eta)} = \frac{x+\alpha}{\beta} = b(x),$$

where we use $(\phi - \beta)^2 = \beta^2 - 2\eta$ so that $\phi^2 + 2\eta = 2\phi\beta$. It also follows that

$$\Gamma(x) = \left(1 + \frac{2\eta}{\phi^2}\right) e^{\eta(x+\alpha)/\beta} - \frac{2\eta}{\phi^2} e^{\phi\alpha}.$$

Then, by construction,

$$h(t) \geq M(t, x) + \Gamma(x), \qquad \forall t, x,$$

with equality at $t = (x + \alpha)/\beta$. In particular, at $t = \tau$ we have $h(\tau) \geq M(\tau, W_\tau) + \Gamma(W_\tau)$ so that for any $\tau$ with $M(t \wedge \tau, W_{t\wedge\tau})$ uniformly integrable,

$$\mathbb{E}[h(\tau)] \geq \mathbb{E}[\Gamma(W_\tau)].$$

36

Further, for all $\tau$ which are solutions to the Skorokhod embedding problem for $\mu$ we have

$$\mathbb{E}[e^{\eta\tau}] \geq \int_{\mathbb{R}} \Gamma(x)\mu(dx) = e^{\alpha\phi}$$

with equality for the Root embedding $\tau = \inf\{u : W_u = \beta u - \alpha\}$.

## 5.2 Pricing options on volatility

Let $(P_t)_{t\geq 0}$ be the price process of an asset, (not necessarily a forward price) and suppose $P_t$ is continuous. Denote the quadratic variation by $\langle \ln P \rangle_t$. The problem is to derive robust, model independent bounds on the prices of call options on volatility, and, following Dupire [24] and Carr and Lee [12] the idea is to use to Root and Röst stopping times, and their optimality properties.

Suppose interest rates are deterministic. (For ease of exposition we will assume that they are constant, and equal to $r$.) Then $(S_t)_{t\geq 0}$ given by $S_t = e^{-rt}P_t$ is a forward price process, which we may assume to be a martingale under a pricing measure. Note that $\ln S_t = \ln P_t - rt$ so that $\langle \ln P \rangle_t = \langle \ln S \rangle_t$.

Write $X_t$ as shorthand for $\ln S_t$. Then

$$\langle X \rangle_T = \int_0^T \left(\frac{dS_u}{S_u}\right)^2 = \int_0^T \left(\frac{dP_u}{P_u}\right)^2.$$

On the other hand,

$$X_T - X_0 = \int_0^T \frac{dS_u}{S_u} - \frac{1}{2}\int_0^T \left(\frac{dS_u}{S_u}\right)^2,$$

so that

$$\langle X \rangle_T = -2\ln S_T + 2\ln S_0 + \int_0^T \frac{2}{S_u}dS_u.$$

Suppose that call options are traded for all strikes so that $\ln S_T$ is effectively a traded asset. Suppose further that the strategy $\theta_t = 2/S_t$ is an admissible dynamic hedging strategy. Then the integrated squared volatility $\langle X \rangle_T$ has a model free price equal to $2(\ln S_0 - \ln S_T)$.

Now consider call options on $\langle X \rangle_T = \langle \ln P \rangle_T$. In particular, consider the security with payoff $(\langle X \rangle_T - Q)^+$.

Without loss of generality we may assume that $P_0 = S_0 = 1$, so that $X_0 = 0$. The goal is to derive model-independent bounds on the price of $(\langle X \rangle_T - Q)^+$, assuming that the law of $S_T$ is given (by $\mu$ say).

Define $M_t = \int_0^t dS_t/S_t$. Then $X_t = M_t - \langle M \rangle_T/2 = M_t - \langle X \rangle_T/2$. We know that the continuous local martingale $M$ can be written as a time-change of Brownian motion: $M_t = W_{\langle X \rangle_t}$ for some Brownian motion $W$. We also have

$$S_t = e^{X_t} = e^{W_{\langle X \rangle_t} - \langle X \rangle_t/2} = Z_{\langle X \rangle_t}$$

where $(Z_u)_{u\geq 0}$ is the exponential Brownian motion $Z_u = e^{W_u - u/2}$.

Since $S_T \sim \mu$ we have $Z_{\langle \ln S \rangle_T} \sim \mu$, so that $\langle \ln S \rangle_T$ is a solution of the Skorokhod embedding problem for $(Z, \mu)$. Conversely, if $\sigma$ is a solution of the SEP for $(Z, \mu)$, and if $S_t = Z_{\sigma \wedge (t/(T-t))}$ then

$$S_T = Z_\sigma \sim \mu \qquad \text{and} \qquad \langle \ln S \rangle_T = \langle X \rangle_T = \sigma.$$

If $\sigma$ is the Root solution of the SEP for $(Z, \mu)$ then we say that $(S_t)_{0 \leq t \leq T}$ given by $S_t = Z_{\sigma \wedge t/(T-t)}$ is the Root model. This model is consistent with observed call prices as represented by $\mu$. Then, combining the time-change with Theorem 5.1 and Remark 5.4 we have the following theorem.

**Theorem 5.7 (Dupire [24], Carr-Lee [12])** *Suppose call prices satisfy $C(K) = C_\mu(K)$ for some distribution $\mu$ with mean $S_0$. Let $\sigma_R$ be the Root barrier solution of the Skorokhod embedding problem for $(Z, \mu)$.*

*Let $((S_t)_{0 \leq t \leq T}, \mathbb{P}^R)$ be a model for the forward price for which $S_t = Z_{\sigma_R \wedge (t/(T-t))}$. Let $((S_t)_{0 \leq t \leq T}, \tilde{\mathbb{P}})$ be any alternative model for which $S_t$ is a martingale and $S_T \sim \mu$. Then, for any $Q \geq 0$,*

$$\mathbb{E}^R[(\langle \ln S \rangle_T - Q)^+] \leq \tilde{\mathbb{E}}[(\langle \ln S \rangle_T - Q)^+]$$

*and the price of a call option on quadratic variation for a given set of vanilla call prices is minimised under the Root model.*

*The Röst (reverse barrier) construction gives an upper bound.*

## 5.3 A path-wise hedging strategy for volatility options

Theorem 5.7 gives bounds on the prices of call options on volatility, but it does not give an associated super-replicating or sub-replicating strategy. In particular, it does not give a path-wise inequality which could be used to enforce the bounds, in the way that the inequality (1) enforces no-arbitrage bounds on the prices of one-touch digitals. We show below that, subject to solving a variant on the heat equation with a rather strange 'boundary condition', a sub-replicating strategy exists.

Let call prices (for all strikes and maturity $T$) be such that $S_T \sim \mu$. (By scaling we may assume that $S_0 = 1$ and $\mu$ has support $\mathbb{R}^+$ and mean 1.) Suppose that the Root solution to the Skorokhod embedding problem for $(Z, \mu)$ is associated with the barrier $b_\mu$. Consider the problem of pricing an option with payoff $h(\langle \ln S \rangle_T)$ with $h(t)$ convex.

**Theorem 5.8** *Suppose there exists $m(t, z)$ such that $m(0, 1) = 0$, $\frac{z^2}{2} m'' + \dot{m} = 0$ and such that*

$$\arg \inf_{t \geq 0}[h(t) - m(t, z)] = b_\mu(z).$$

*Suppose further that $\theta$ given by*

$$\theta_t = m'(\langle \ln S \rangle_t, S_t)$$

*is admissible, in the sense that $G^\theta = \int_0^T \theta_t \, dS_t \in G^\Theta$.*

*If we define $\Gamma(z) = \inf_{t > 0}\{h(t) - m(t, z)\} = h(b_\mu(z)) - m(t, b_\mu(z))$ then*

$$h(\langle \ln S \rangle_T) \geq \Gamma(S_T) + G^\theta,$$

*path-wise, and thus*

$$\mathcal{P}(h\langle \ln S \rangle_T) \geq \int_{\mathbb{R}^+} \Gamma(s) \mu(ds),$$

*with equality for the Root model.*

**Proof.** Since $h(t) - m(t, z) \geq \Gamma(z)$ uniformly in $t$ and $z$ we have

$$h(\langle \ln S \rangle_T) \geq \Gamma(S_T) + m(\langle \ln S \rangle_T, S_T).$$

But, by Itô's formula, and using $d\langle \ln S \rangle_t = (dS_t)^2/S_t^2$,

$$
\begin{aligned}
m(\langle \ln S \rangle_T, S_T) &= m(0, 1) + \int_0^T \left[ \dot{m} \, d\langle \ln S \rangle_t + m' dS_t + \frac{S_t^2}{2} m'' \frac{(dS_t)^2}{S_t^2} \right] \\
&= \int_0^T m'(\langle \ln S \rangle_t, S_t) dS_t \\
&= G^\theta.
\end{aligned}
$$

$\square$

**Example 5.9** Suppose $h(t) = t$ and $m(t, z) = 2 \ln z + t$. Then $m(0, 1) = 0$ and $\dot{m} + z^2 m''/2 = 0$. Furthermore, $h(t) - m(t, z)$ is independent of $t$ so that, for any $\mu$ with Root barrier $b_\mu(z)$,

$$\Gamma(z) = \inf_t \{h(t) - m(t, z)\} = h(b_\mu(z)) - m(b_\mu(z), z) = -2 \ln z.$$

Then, provided $\theta$ given by $\theta_t = (2/S_t)$ is admissible we have

$$\mathcal{P}(\langle \ln S \rangle_T) = \mathcal{P}(-2 \ln S_T) = \int_{\mathbb{R}^+} (-2 \ln s) \mu(ds)$$

and the fair price of a security paying $\langle \ln S \rangle_T$ is model independent.

**Example 5.10** In this second example the payoff is non-linear and Theorem 5.8 gives a model-independent bound, rather than the unique no-arbitrage price. We show how to construct a sub-replicating strategy which enforces the lower bound. The example is based on Example 5.6.

Suppose that $h(t) = e^{\eta t}$. Suppose further that $S_T \sim \mu$ where

$$\mu(dz) = \frac{\alpha \sqrt{\gamma}}{\sqrt{2\pi}(\alpha + \ln z)^{3/2}} \exp\left(\frac{(\alpha - (2\gamma - 1)\ln z)^2}{8\gamma(\alpha + \ln z)}\right) dz, \qquad z > e^{-\alpha}.$$

Then $b_\mu(z) = (\alpha + \ln z)/\gamma$.

Fix $v = e^{-\alpha}$ and let $\psi$ be the smallest root of $\psi^2 - (1+2\gamma)\psi + 2\eta = 0$. (We assume that $\eta$ is small enough so that this equation has a solution.) Define $\phi = \psi(\psi - 1)/2$ and $\lambda = \eta \phi^{-1} v^{-\psi}$.

Now set $m(t, z) = \lambda - \lambda z^\psi e^{-\phi t}$. If $b(z) = \arg\inf_{t>0}[h(t) - m(t, z)]$ then $b(z)$ solves

$$\eta e^{\eta b(z)} = \phi \lambda z^\psi e^{-\phi b(z)}.$$

The choices of constants are such that $b(z) = b_\mu(z)$. Provided $\theta$ given by $\theta_t = -\lambda \psi S_t^{\psi-1} e^{-\phi \langle X \rangle_t}$ is an admissible strategy then, in the notation of Theorem 5.8, $\Gamma(S_T) + G^\theta$ is a sub-replicating strategy for the volatility option with payoff $h(\langle X \rangle_T)$.

The model-independent lower bound on the price of the option is $\int_{\mathbb{R}^+} \Gamma(s) \mu(ds)$, where $\Gamma(s) = h(b(s)) - m(b(s), s)$.

# 6 Multiple stopping times; option price data for several maturities.

To date we have assumed that there are a continuum of option prices available, for vanilla European options of a single maturity. But what if options trade with two (or more) maturities? Are the option prices consistent with no arbitrage? Can the call prices from an earlier maturity be used to refine the price bounds? Is it possible to derive bounds for forward starting options?

In terms of the Skorokhod embedding problem, the equivalent problems relate to the existence and properties of solutions when the initial law of the Brownian motion is non-trivial. Recall from Theorem 3.7 that there exist solutions of the Skorokhod embedding problem for centred initial and terminal laws if and only if $C_\mu(x) \leq C_\nu(x)$. This has an immediate Corollary in the finance setting; a necessary and sufficient condition for the absence of arbitrage between a pair of (decreasing, convex) call price functions is that for each fixed strike (in the bond numeraire), call prices are increasing in maturity. (There are extensions of this result to the case where only a finite number of strikes are traded, in which case the answer depends on the ability to extrapolate between the traded strikes to generate convex (in strike) call price functions whose prices are increasing in maturity. See Bühler [11], Davis and Hobson [20] or Cousot [14].)

So, suppose $\mu$ and $\nu$ are such $C_\mu(x) \leq C_\nu(x)$, or equivalently $U_\mu(x) \geq U_\nu(x)$. The aim is to find stopping rules $\tau$ which maximise a functional $F(W_t, 0 \leq \sigma \leq t \leq \tau)$ (or in the forward-starting version of the problem $F(W_t, \sigma \leq t \leq \tau)$) where $W_\sigma \sim \mu$ and $W_\tau \sim \nu$. the question arises: what are the appropriate generalisations of the Azéma-Yor (or Perkins, Root, Röst ...) embeddings to non-zero initial law? In the financial context, new derivatives become relevant, such as the forward-starting straddle option with payoff $|S_{T_2} - S_{T_1}|$. An analysis of this can be considered as a first step towards pricing the discrete time Asian option with payoff $(\sum_{i \leq n} S_{T_i} - K)^+$, and thence its continuous time analogue.

## 6.1 Maximising the law of the maximum, with intermediate constraints

Consider the problem: find
$$\sup_\tau \mathbb{P}(J_\tau \geq j)$$
where the supremum is taking over uniformly integrable stopping times $\tau$ such that there exists $\sigma \leq \tau$ for which

$$W_0 \sim \delta_0, \qquad W_\sigma \sim \mu, \qquad W_\tau \sim \nu. \tag{20}$$

This is directly linked to finding candidate models for which $S_{T_1} \sim \mu$ and $S_{T_2} \sim \nu$. It turns out that there are two cases, an easy case, and a hard case.

In the easy case, the barycentres $b_\mu$ and $b_\nu$ are ordered so that $b_\mu(x) \leq b_\nu(x)$ for all $x$. See Figure 17. Then the inverse barycentres $\beta_\mu$ and $\beta_\nu$ satisfy $\beta_\nu(j) \leq \beta_\mu(j)$. In particular, if we define $\sigma = \tau_{b_\mu} = \inf\{u : W_u \leq \beta_\mu(J_u)\}$ and $\tau = \tau_{b_\nu} = \inf\{u : W_u \leq \beta_\nu(J_u)\}$ then $\sigma \leq \tau$ embed $\mu$ and $\nu$ respectively.
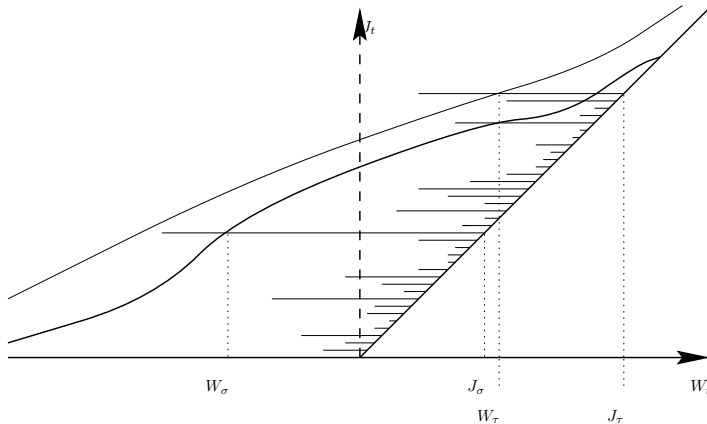
Figure 17: A pair of barycentres. In this case the barycentres are ordered and it is possible to find $\sigma$ and $\tau$ with $\sigma \leq \tau$ where both separately represent the Azéma-Yor construction of the associated measure.

Moreover, since $\tau$ is the Azéma-Yor embedding for $\nu$, $\tau$ maximises $\mathbb{P}(J_\tau \geq j)$ over all minimal solutions of the Skorokhod embedding problem for $(W, \nu)$, and hence must maximise $\mathbb{P}(J_\tau \geq j)$ over the smaller set of embeddings which are subject to the intermediate constraint.

Unfortunately, although $U_\mu(x) \geq U_\nu(x)$ is a necessary condition for $b_\mu(x) \leq b_\nu(x)$, $\forall x$, it is not sufficient. Suppose for example $\mu \sim (\delta_{-1} + \delta_1)/2$ and $\nu \sim p\delta_{-2} + (1 - 2p)\delta_0 + p\delta_2$. For $1/4 \leq p \leq 1/2$ we have $-1 = U_\mu(0) \geq U_\nu(0) = -4p$ and then $U_\mu(x) \geq U_\nu(x)$ for all $x$. However, for $-1 < x < 0$ we have $b_\mu(x) = 1$ and $b_\nu(x) = 2p/(1 - p)$ so that in order to have $b_\mu(x) \leq b_\nu(x)$ we must have $p \geq 1/3$. For $p < 1/3$ the barycentres are not ordered.

It follows that the intermediate constraint $W_\sigma \sim \mu$ has an impact on the possible values of $\mathbb{P}(J_\tau \geq j)$. For more on this situation, and the implications for model-independent bounds, see Brown et al [9].

A similar situation arises if we try to extend the Perkins [47], Root [51], Röst [41] or Vallois [55, 56] constructions to multiple time points. All these constructions utilise a stopping rule based on the first hitting time of the joint process $(W_u, A_u)$ of a domain, where $A_t$ is an additive functional of the Brownian path. If these domains posses a natural ordering, then the construction of optimal stopping times $\sigma, \tau$ satisfying (20) is straightforward. Otherwise the construction of the optimal stopping rule (i.e. the one which maximises $A_\tau$) is much more challenging. See Figure 18 for an illustration in the Root barrier case.

## 6.2 Maximising the law of the maximum, with non-trivial initial law

Now consider the problem:
$$\sup_\tau \mathbb{P}(J_\tau \geq j)$$
where the supremum is taking over stopping times $\tau$ such that

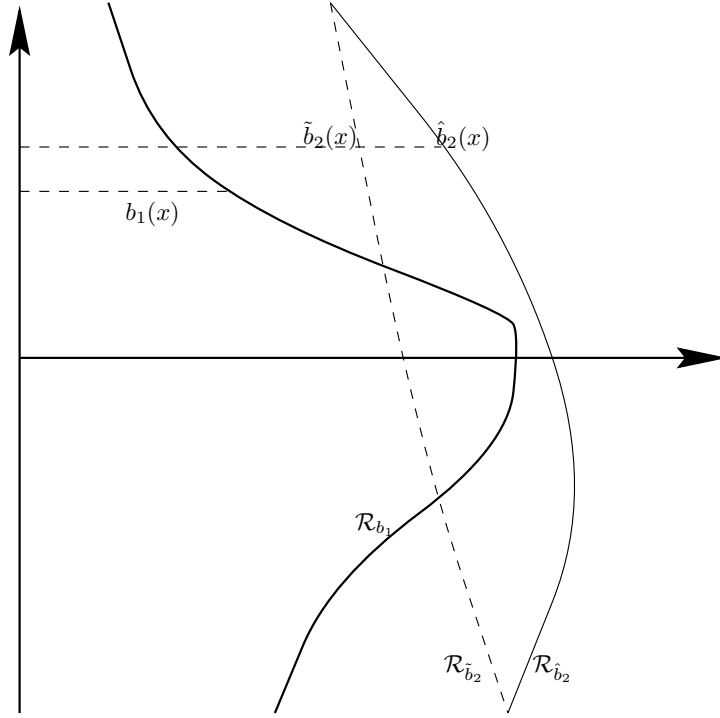$$W_0 \sim \mu, \qquad W_\tau \sim \nu \qquad (21)$$

41

Figure 18: A pair of Root barriers. Suppose we try to solve the problem (20) of constructing an embedding $\tau$ of $\nu$ subject to the existence of $\sigma \leq \tau$ for which $W_\sigma \sim \mu$. Let $b_1$ with associated stopping region $\mathcal{R}_{b_1}$ embed $\mu$; for (different) $\hat\nu$ and $\tilde\nu$ let $\hat b_2$ and $\tilde b_2$ be the associated Root barriers. If the barriers are ordered, e.g. $\hat b_2 \geq b_1$, then we can define $\sigma = \inf\{u : (u, W_u) \in \mathcal{R}_{b_1}\}$ and $\hat\tau = \inf\{u : (u, W_u) \in \mathcal{R}_{\hat b_2}$ and the problem is solved. But, if the barriers are not ordered (e.g. $b_1$ and $\tilde b_2$), then a more complicated construction is needed.
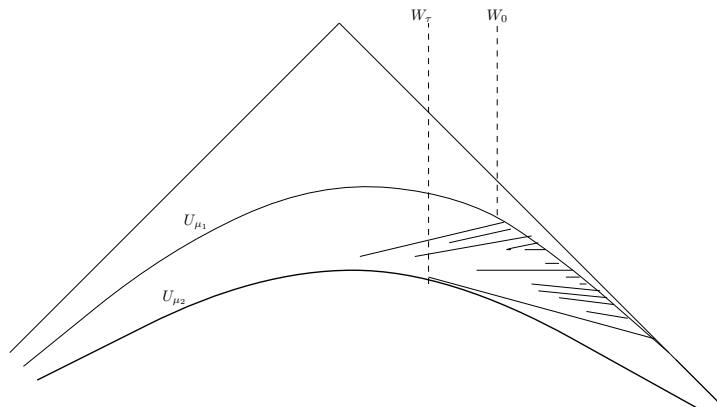


Figure 19: Potential theory picture of Azéma-Yor story with non-trivial starting law. The idea is to run excursions down from the current maximum along tangents to the potential $U_{\mu_2}$ which intersect $U_{\mu_1}$ at the current value of the maximum. See Hobson [32].

where $\mu$ and $\nu$ are centred probability measures with $U_\mu(x) \geq U_\nu(x)$.

An alternative but equivalent problem is to search for pairs of stopping times $\sigma \leq \tau$ for which (20) holds, but to only calculate the maximum over the interval $(\sigma, \tau)$. This corresponds to searching for model-independent derivative prices for forward-starting options, e.g. for barrier options where the option is knocked-in only if it is crossed in the period $(T_1, T_2)$.

Several generalisations of solutions of embeddings to the situation with non-zero initial laws are known. Hobson [32] shows how to adapt the potential picture to extend the Azéma-Yor embedding to this case (see Figure 19) and Hobson and Pedersen [36] consider the extension of the Perkins embedding, and give an application to bounds on the prices of barrier options. Indeed, in the case of a Root [51] barrier solution, the same form of construction still holds, and $\tau = \inf\{u : (u, W_u) \in \mathcal{R}_b\}$ is an embedding for a suitably chosen barrier $b$ depending on both $\mu$ and $\nu$.

## 6.3 Maximising the law of the maximum, with a continuum of marginals

Suppose we are given the marginal distributions of a martingale $X$ for every $t$. For there to exist a process with those marginals we must have that the potentials are decreasing in $t$, (or equivalently the call option prices are increasing in maturity). By the results of Krylov [39], Gyöngy [28] and Dupire [23], if the marginals arose from a continuous process then we can identify a unique diffusion with the same marginals. Now we ask, what other processes might have lead to the same marginals? The Dupire construction gives a canonical process with a given set of marginals, but it is not unique.

In particular, suppose $X_t \sim N(0, t)$, and that $X_t$ is a martingale. Clearly one process with these marginals is Brownian motion, and Brownian motion is the unique continuous Markovian martingale with $N(0, t)$ marginals.

Following Oleszkiewicz [45] we can define a *fake Brownian motion* to be a martingale with marginals $X_t \sim N(0, t)$, which is not Brownian motion. Given Dupire's result, in order to find such a process we need to relax either the assumption of continuity of paths, or the Markov assumption.

It is apparent from the discussion in Section 6.2 that several of the construction methods for solutions of the Skorokhod embedding problem for multiple time points extend to the setting of a family of marginals with parameter $t$, at least under some restrictions on that family. A first example is the Azéma-Yor construction.

Let $\mu_t \sim N(0, t)$, and let $b_t := b_{\mu_t}$. By scaling we have that $b_t(x) = b_1(x\sqrt{t})$, so that the family $b_t$ is increasing in $t$ for each fixed $x$. In particular, if $\tau_t = \inf\{u : W_u \leq \beta_t(J_u)\}$ then $\tau_t$ is an increasing family of stopping times such that $\tau_t$ embeds $\mu_t$. This is the Madan and Yor [40] fake Brownian motion. The resulting process is a Markovian martingale, for which paths have strictly decreasing continuous parts interleaved with positive jumps.

By extending the methods of Section 6.1 from two to a continuum of time-points we can also use the Perkins [47] or Vallois [55, 56] construction to generate other fake Brownian motions. (Note however that the generalisation of the Root [51]

construction does not give a fake Brownian motion, since the barriers are simply vertical lines and the resulting stopping times are $\tau_t \equiv t$.)

Interest in the problem of finding fake Brownian motions began with a paper by Hamza and Klebaner [30]. They use a remarkably clever decomposition and recombination of the normal distribution to produce a pure-jump martingale with Gaussian marginals. They also asked the question if there existed continuous (non-Brownian) martingales with Gaussian marginals. This was answered in the affirmative by Albin [1]. Finally, Oleszkiewicz [45] gave an elegant and simple construction of a continuous fake Brownian motion. In essence this construction makes use of the fact that $R_t \sin \Theta_{A_t}$ is a Brownian motion where $(R_t)_{t \geq 0}$ is a two-dimensional Bessel process, $(\Theta_u)_{u \geq 0}$ is a Brownian motion on the unit circle, and $(A_t)_{t \geq 0}$ is the increasing additive functional such that $A_t = \int_0^t R_s^{-2} ds$. The idea is to write $X_t = \sqrt{t} R \sin(U + W_{\ln t})$ where $R$ has the same distribution as $R_1$, $U$ is uniform distribution on $[0, 2\pi)$ and $W$ is Brownian motion.

The existence of fake Brownian motions implies that although the Dupire [23] construction gives a model which is consistent with the continuum of traded option prices (assuming such a family exists) — and perhaps gives the *canonical* model consistent with these prices — it is not the unique model with this property. Hence, although the Dupire diffusion might be used to give a guide price for exotic derivatives, these prices are not the unique prices consistent with no-arbitrage.

## 6.4 Model independent bounds on basket options

Underlying the discussion in these notes is a philosophy whereby the prices of exotic options are related to the prices of vanilla (traded) options by the construction of super-hedges involving those vanilla options. The (primal) pricing problem is related to a (dual) hedging problem.

The same philosophy can be applied in other contexts, one such being the pricing of basket (exotic) options, given the prices of (vanilla) call options on the individual constituents of the basket. It turns out that the model independent upper bound on the price of a basket option is associated with a model when the assets are co-monotonic, see Dhaene et al [21] or Hobson et al [34].

# 7 Closing remarks

In a mature market, the prices of liquidly traded vanilla options are not the expectations of the payoff under a stochastic model, but rather they are fixed by supply and demand, or market sentiment. From a derivative pricing perspective the objects of interest become the less liquid exotic derivatives which must be priced relative to the vanilla options.

In the standard case where the vanilla options are precisely the puts and calls, knowledge of vanilla prices is equivalent to knowledge of the marginal distribution of the underlying asset. Then, finding candidate models which fit option prices is equivalent to finding solutions of the Skorokhod embedding problem, and finding

the range of no-arbitrage prices for an exotic option is equivalent to finding extremal embeddings which maximise functionals on the Brownian paths.

The associated no-arbitrage bounds may be quite wide. (Conversely, see Brown et al [10], and Example 5.10 above, in some special cases the upper and lower bounds coincide, and there is a unique model-independent no-arbitrage price for the exotic option.) However, each bound is associated with a model which achieves the bound, and by investigating features of this model it is possible to determine the characteristics of models which lead to high option payouts. (For example, for the model which attains the upper bound on the price of a lookback option, the overall maximum is an increasing function of the time-$T$ price. Whilst in exponential Brownian motion model the correlation between the final value and the maximum is not perfect, the correlation is quite strong, and therefore for realistic parameter values the model price can be expected to be quite close to the theoretical upper bound.)

The second major advantage of the Skorokhod-embedding-model-independent-bound approach is that since no assumptions are made on the model the resulting strategies must be very simple. Even if an agent believes in a particular stochastic model there may be advantages (especially in the form of lower transaction costs) in following a semi-static hedging strategy involving calls, when compared with a delta-hedging strategy with infinite trading. Cox and Oblój [18] compare the performance of a model independent hedge with the Black Scholes hedge for no-touch options, and conclude that the robust hedge frequently outperforms the classical delta-hedge, even when transaction costs are small.

The optimal strategy involves purchasing an initial portfolio of traded options, but this portfolio is held constant over time, and there is no assumption that elements of the portfolio can be sold at intermediate times. Hence the success of any super-replicating strategy is not contingent upon any model for movements of option prices over time.

# References

[1] Albin, J. M. P.; A continuous non-Brownian motion martingale with Brownian motion martingale distributions, *Statistics and Probability Letters*, **78**, 6, 682–686. 2008.

[2] Azéma, J., Gundy, R. F. and Yor, M.; Sur l'intégrabilité uniforme des martingales continues. *Séminaire de Probabilités*, **XIV**, *Lecture Notes in Math.*, **784**, 53–61, Springer, Berlin, 1980.

[3] Azéma, J. and Yor, M.; Une solution simple au problème de Skorokhod, *Séminaire de Probabilités*, **XIII**, *Lecture Notes in Math.*, **721**, 90–115, Springer, Berlin, 1979.

[4] Azema Azéma, J. and Yor, M.; Une solution simple au problème de Skorokhod, *Séminaire de Probabilités*, **XIII**, *Lecture Notes in Math.*, **721**, 625–633, Springer, Berlin, 1979.

[5] Bass, R. F.; Skorokhod imbedding via stochastic integrals, *Séminaire de Probabilités*, **XVII**, *Lecture Notes in Math.*, 986, 221–224, Springer, Berlin, 1983.

[6] Bergman, Y.Z., Grundy, B.D., and Wiener, Z.; General properties of option prices, *J. Finance*, **51**, 1573–1610, 1996.

[7] Bertoin, J. and Le Jan, Y.; Representation of measures by balayage from a regular recurrent point, Ann. Probab., **20**, 1, 538–548, 1992.

[8] Breeden, D.T. and Litzenberger, R.H.; Prices of state-contingent claims implicit in options prices, *J. Business*, **51**, 621–651, 1978.

[9] Brown, H., Hobson, D. and Rogers, L. C. G.; The maximum maximum of a martingale constrained by an intermediate law, *Probab. Theory Related Fields*, **119**, 558–578, 2001.

[10] Brown, H., and Hobson, D. and Rogers, L. C. G.; Robust hedging of barrier options, *Math. Finance*, **11**, 3, 285–314, 2001.

[11] Bühler, H.; Expensive martingales, *Quantitative Finance*, **6**, 207–218, 2006.

[12] Carr, P. and Lee, R.; Hedging variance options on semi-martingales. *Preprint*, 2008.

[13] Chacon, R. V. and Walsh, J. B.; One-dimensional potential embedding, *Séminaire de Probabilités*, **X**. *Lecture Notes in Math.*, **511** 19–23 Springer, Berlin, 1976.

[14] Cousot, L.; Conditions on options prices for the absence of arbitrage and exact calibration, *J. Banking and Finance*, **31**(11), 3377–3397, 2007.

[15] Cox, A.M.G. and Hobson, D.G.; An optimal embedding for diffusions, *Stochastic Process. Appl.*, **111**, 1, 17–39, 2004.

[16] Cox, A.M.G. and Hobson, D.G., Local martingales, bubbles and options prices, *Finance and Stochastics*, **9**, 477–492, 2005.

[17] Cox, A.M.G. and Hobson, D.G., Skorokhod embeddings, minimality and non-centred target distributions, *Prob. Th. Rel. Fields*, **135**, 395–414, 2006.

[18] Cox, A.M.G, and Oblój, J.; Robust hedging of double touch barrier options, *Preprint*, 2008. arXiv:0808.4012

[19] Cox, A.M.G, and Oblój, J.; Robust pricing and hedging of double no-touch options, *Preprint*, 2009. arXiv:0901.0674

[20] Davis, M.H.A and Hobson, D.G.; The range of traded options prices, *Math. Finance*, **17**, 1–14, 2007.

[21] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D.; The concept of comonotonicity in actuarial science and finance: applications, *Insurance: Mathematics and Economics*, **31**, 133–161, 2002.

[22] Dubins, L. E.; On a theorem of Skorohod, *Ann. Math. Statist.*, **39**, 2094–2097, 1968.

[23] Dupire, B.; Pricing with a smile, *Risk Magazine*, 18–20, 1994.

[24] Dupire, B.; Arbitrage bounds for volatility derivatives, Presentation at *PDE and Mathematical Finance* KTH, Stockholm, 2005.

[25] El Karoui, N., Jeanblanc, M. and Shreve, S.E.; Robustness of the Black-Scholes formula, *Mathematical Finance*, **8**, 93–126, 1998.

[26] Figlewski, S.; Assessing the incremental value of option pricing theory relative to an informationally passive benchmark. *J. Derivatives*, Fall, 80–96, 2002.

[27] Föllmer, H. and Schied, A.; *Stochastic Finance: An Introduction in Discrete Time*, 2nd Edition, de Gruyter, Berlin, 2002.

[28] Gyöngy, I.; Mimicking the one-dimensional marginal distributions of processes having an Itô differential, *Prob. Th. Rel. Fields*, **71**, 501–516, 1986.

[29] Hall, W.; On the Skorokhod embedding theorem, *Technical Report* 33, Stanford University, Dept. of Statistics, 1968.

[30] Hamza, K. and Klebaner, F.C.; A family of non-Gaussian martingales with Gaussian marginals. *J. Appl. Math. and Stochastic Analysis*, Article Id 92723, 2007.

[31] Hobson, D.G.; Volatility misspecification, option pricing and super-replication via coupling, *Annals Appl. Prob.*, **8**, 193–205, 1998.

[32] Hobson, D.G.; The maximum maximum of a martingale, *Séminaire de Probabilités*, **XXXII**, *Lecture Notes in Math.*, **1686**, 250–263, Springer, Berlin, 1998.

[33] Hobson, D.G.; Robust Hedging of the Lookback Option, *Finance Stoch.*, **2**, 4, 329–347, 1998.

[34] Hobson, D.G., Lawrance, P. and Wang, T-H.; Static-arbitrage upper bounds for the prices of basket options, *Quantitative Finance*, **5**, 329–342, 2005.

[35] Hobson, D.G. and Neuberger, A.; Robust bounds for Forward Start options, *Preprint*, 2008.

[36] Hobson, D.G. and Pedersen, J. L.; The minimum maximum of a continuous martingale with given initial and terminal laws, Ann. Probab., **30**, 2, 978–999, 2002.

[37] Janson, S. and Tysk, J.; Volatility time and properties of options prices, *Annals Appl. Prob.*, **13**, 890–913, 2003.

[38] Kiefer, J.; Skorokhod imbedding of multivariate RV's and the sample DF. *Z. Wahrscheinlichkeitstheorie*, **24**, 1–35, 1972.

[39] Krylov, N.V.; On the relation between differential operators of second order and the solutions of differential equations. *Steklov Seminar 1984* Ed N.V. Krylov et al, Inc. Publications Division, New York, 214–229, 1985.

[40] Madan, D. B. and Yor, M.; Making Markov martingales meet marginals: with explicit constructions, *Bernoulli*, **8**, 4, 509–536, 2002.

[41] Meilijson, I.; Skorokhod's problem — embeddings in Brownian motion. *Notes from the 3rd Winter School of Probability, Chile*, 1983.

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

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

[44] Obłój, J. and Yor, M.; An explicit Skorokhod embedding for the age of Brownian excursions and Azéma martingale, *Stochastic Process. Appl.*, **110**, 1, 83–110, 2004.

[45] Oleszkiewicz, K.; On fake Brownian motions. *Statistics and Probability Letters*, **78**, 1251–1254, 2008.

[46] Pedersen, J. L. and Peskir, G.; The Azéma-Yor embedding in non-singular diffusions, *Stochastic Process. Appl.*, **96**, 2, 305–312, 2001.

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

[48] Revuz, D. and Yor, M.; *Continuous martingales and Brownian motion*, 3rd Edition, Springer, 1999.

[49] Rogers, L. C. G.; A guided tour through excursions, *Bull. London Math. Soc.*, **21**, 4, 305–341, 1989.

[50] Rogers, L.C.G. and Williams, D.; *Diffusions, Markov Processes and Martingales, Volume 2, Itô Calculus*, CUP, Cambridge, 2000.

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

[52] Rost, H.; The stopping distributions of a Markov Process, *Invent. Math.*, **14**, 1–16, 1971.

[53] Rost, H.; Skorokhod stopping times of minimal variance, *Séminaire de Probabilités*, **X**, 194–208. *Lecture Notes in Math.*, **511**, Springer, Berlin, 1976.

[54] Skorokhod, A.V.; *Studies in the theory of random processes.* Addison-Wesley, Reading, Mass., 1965.

[55] Vallois, P.; Le problème de Skorokhod sur ℝ: une approche avec le temps local, *Seminar on probability*, **XVII**, *Lecture Notes in Math.*, 986, 227–239, Springer, Berlin, 1983.

[56] Vallois, P.; Quelques inégalités avec le temps local en zero du mouvement Brownien, *Stochastic Process. Appl.*, **41**, 1, 117–155, 1992.