Optimal liquidation of an asset under drift uncertainty

Erik Ekström and Juozas Vaicenavicius

January 22, 2016

Abstract

We study a problem of finding an optimal stopping strategy to liquidate an asset with unknown drift. Taking a Bayesian approach, we model the initial beliefs of an individual about the drift by allowing an arbitrary probability distribution to characterise the uncertainty about the drift parameter. Filtering theory is used to describe the evolution of the posterior beliefs about the drift once the price process is being observed. An optimal stopping time is determined as the first passage time of the posterior mean below a monotone boundary, which can be characterised as the unique solution to a non-linear integral equation. We also study monotonicity properties with respect to the prior distribution and the asset volatility.

1 Introduction

It is an inevitable feature of human economic activity that prices of goods vary in time. Thus, naturally, a person participating in trade cares much about the best time to perform a transaction. Let us think about an individual who possesses an indivisible asset with price evolution $\{S_t\}_{t\geq 0}$ and wants to sell it before time $T \geq 0$. Assuming a liquid market, how should the seller choose a selling time to maximise his/her profit from the sale? Mathematically, the question is about finding a stopping time τ^*, belonging to a set of admissible stopping times $\mathcal{T}_T$, such that

$$\mathbb{E}[S_{\tau^*}] = \sup_{\tau \in \mathcal{T}_T} \mathbb{E}[S_\tau]. \tag{1.1}$$

A natural set of admissible stopping times $\mathcal{T}_T$ to consider is the set $\mathcal{T}_T^S$ of stopping times with respect to the price process S, i.e. at any point in time, the decision whether to sell the asset or not must be based solely on the price history of S. Throughout this article we assume $\mathcal{T}_T = \mathcal{T}_T^S$.

This selling problem (1.1) corresponds to the situation when the seller does not want or cannot easily buy the asset back and then sell it again. One example when repurchasing the asset is not typically feasible is when a significant stake in a firm is sold in a single transaction to another company, otherwise known as an acquisition. Another example of an asset that cannot be bought back is the employee stock option (ESO), which is an American call option given to employees as part of the remuneration. The employees can only exercise such options; also, they are often barred from trading in the company's shares to prevent insider trading.

In the context of the classical Black-Scholes model

$$\mathrm{d}S_t = \alpha S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t \,, \tag{1.2}$$

where α, σ are known constant parameters, the answer to the optimal selling question (1.1) is straightforward: if $\alpha > 0$, then the optimal strategy is to sell at the terminal time T; if $\alpha < 0$, then the optimal strategy is to sell immediately, i.e. at time 0; if $\alpha = 0$, then any stopping time τ is optimal.

However, in applications, the known constant drift assumption of the Black-Scholes model is usually too strong. To obtain reasonable precision when estimating the drift one needs extraordinarily long time-series spanning decades or even centuries, which are hardly ever available (e.g. see [28, Section 4.2 on p. 144] for a discussion). An extreme example of the lack of data is a stock of an initial public offering (IPO) for which the price history simply does not exist. The estimated Black-Scholes model parameters of a few famous IPOs over the first year since going public (see Table 1) suggest that it is unlikely that the price change in all those cases was due to the volatility alone, leading us to believe in the significance of the drift contribution, which needs to be addressed.

IPO	$\log(S_1/S_0)$	$\hat{\alpha}$	$\hat{\sigma}$
Amazon (1997)	1.34	1.68	0.83
Google (2004)	1.03	1.11	0.41
Facebook (2012)	-0.42	-0.27	0.55
Vonage (2006)	-1.53	-1.29	0.70

Table 1: The estimates $\hat{\alpha}$ and $\hat{\sigma}$ of the drift α and the volatility σ, respectively, calculated over the first year of an IPO using the daily closing prices. Data source: Google Finance.

Admitting that in practice the exact value of the drift parameter is unknown, we choose to model the inherent uncertainty about the drift by a probability distribution. More precisely, we extend the geometric Brownian motion model (1.2) by replacing the constant drift α by a random variable whose distribution (called 'a prior' in Bayesian statistics) encapsulates all the knowledge available to us concerning the uncertainty about the drift. As far as the volatility σ is concerned, we stay with the known constant volatility assumption as it can be estimated much more easily – in theory, from an arbitrarily short price observation period.

Though in this article, we view the prior distribution as subjective beliefs whose origin we do not question, one can also think of transparent constructive approaches for choosing the prior. For example, if one is interested in selling a stake in a newly established company whose market valuation can be continuously observed, one possibility for the prior is to use the empirical distribution of the returns of similar startups over the initial period of the same length as our investment horizon T. The similarity criteria could be the market sector, the region, etc.

In this article, we solve the optimal liquidation problem (1.1) within the proposed model under an arbitrary prior distribution for the drift. The first time the posterior mean of the drift passes below a specific non-decreasing curve is shown to be optimal; the stopping boundary is characterised as the unique solution of a particular integral equation.

To include more details, our investigation of the optimal strategy can be briefly described in the following. The original problem with incomplete information about the drift is reformulated as a complete information problem by projecting the price evolution onto the observable filtration using filtering theory. The mean of the posterior distribution becomes the underlying process of a new equivalent optimal stopping problem with a stochastic killing/creation rate and a constant payoff function. This conditional mean is shown to satisfy a stochastic differential equation driven by the innovation process. The dispersion coefficient of the SDE is proved to be decaying in time as well as satisfy a special condition on the second spatial derivative. Embedding the value function into a Markovian framework and making a suitable connection with the term-structure equation, the established dispersion function properties enable us to employ the available convexity results to prove convexity of the Markovian value function in the spatial variable. Moreover, the value function is shown to be continuous and decreasing in time. These significant facts allow us to show that the first passage time below a monotone boundary is an optimal stopping time, so techniques from the theory of free-boundary problems with monotone boundaries can be applied. Specifically, the monotonicity of the boundary enables us to prove the smooth-fit property and to investigate the corresponding integral equation. The optimal stopping boundary is characterised as the unique non-positive and continuous solution to a non-linear integral equation.

Besides the examination of the optimal strategy, we investigate monotonicity properties of the expected optimal liquidation value with respect to the asset volatility and the prior distribution. Notwithstanding that all-inclusive theorems about parameter dependence appear currently to be beyond reach, we derive some sufficient conditions for monotonicity in the volatility σ as well as the prior distribution. In addition, we discuss the effects of parameter uncertainty and filtering on the expected optimal selling payoff and conduct numerical experiments in the case of the normal prior. Some results reinforce standard intuition, others illustrate inherent subtleties. In particular, additional value that an optimal strategy involving filtering brings over an optimal strategy without filtering is calculated, exhibiting an improvement of close to 10% for some feasible parameter regimes.

As far as extensions of this work are concerned, solving the optimal liquidation problem for more general diffusions, with possibly time- and level-dependent coefficients, is more problematic. Such extensions typically lead to the loss of the useful one-dimensional Markovian structure present in the classical geometric Brownian motion setting; an optimal decision then depends on the whole price trajectory rather than the current spot price alone. Clearly, a complete treatment of the resulting optimal stopping problem is much more complicated.

1.1 Literature review

Over the last three decades, investment problems with incomplete information about the drift have received much attention from both financial mathematicians and financial economists. Some of the most distinct works on portfolio optimisation include [7], regarded as the first incomplete information problem studied in financial literature, and the general portfolio problems studied in [18, 19]; see also the recent article [4] proposing a general framework for most of the earlier works as well as containing an excellent survey with references. Hedging in an incomplete market under partial information about a constant drift was addressed in [20] in the case of the Kalman-Bucy filter. In addition, incomplete information models have been investigated in the financial economics literature (see the survey paper [2] as well as the monograph [29]).

In contrast, there have been surprisingly few attempts to tackle financial optimal stopping problems under incomplete information such as the optimal liquidation problem above, with the existing works focusing mainly on a very restrictive case, namely, the two-point prior. The optimal liquidation of an asset with unknown drift has been studied in [3], [9], and of an asset with unknown jump intensity in [17]. For option valuation problems under incomplete information, see [6] and [12]. The financial optimal stopping articles above typically assume a two-point prior distribution; having in mind that the prior represents the beliefs about all the different values the parameter could possibly take, the two-point prior strikes as a simplistic and unrealistic assumption. Overcoming this assumption is one of the main contributions of the present article. It is also worth mentioning that variations of the optimal selling problem in the case of complete information about the parameters have been studied in [8], [11], [13], and Example 10.2.2 in [22]. In the case with additional insider information available, the optimal selling problem is covered in [21] as optimal exercise of an employee stock option with zero strike is an equivalent mathematical problem.

2 The model and problem formulation

We consider a financial market living on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ satisfies the usual conditions and the measure $\mathbb{P}$ denotes the physical probability measure. The basis supports a Brownian motion W and a random variable X such that W and X are independent. We assume that the observed price process S evolves according to

$$\mathrm{d}S_t = XS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t, \tag{2.1}$$

where the volatility $\sigma > 0$ is a known constant. We write $\mathbb{F}^S = \left\{\mathcal{F}^S_t\right\}_{t\geq 0}$ for the filtration generated by the price process S and augmented by the null sets of $\mathcal{F}$. In this article, $\mathbb{F}^S$ corresponds to the only available source of information, i.e. an agent can only observe the price process S, but not the random driver W or the drift X. The distribution of X, which we denote by μ, represents the subjective beliefs of the individual about the likeliness of the different values the mean return rate X may take.

The optimal selling problem that we are interested in is

$$V = \sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}[S_\tau], \tag{2.2}$$

where $\mathcal{T}_T^S$ denotes the set of $\mathbb{F}^S$-stopping times that are less or equal to a specified time horizon $T > 0$.

Note that if the support of μ is contained in $[0, \infty)$, then an optimal strategy is to stop at the terminal time T. Similarly, if the support of μ is contained in $(-\infty, 0]$, then an optimal strategy is to stop immediately. To exclude these trivial cases, we from now on impose the following assumption:

Assumption 2.1. *$\mu((-\infty, 0)) \neq 0$ and $\mu((0, \infty)) \neq 0$.*

Before proceeding, we make the following remarks on the optimal stopping problem (2.2) and possible extensions thereof.

- The formulation (2.2) of the liquidation problem as an optimal stopping problem stems from the assumption that the asset is indivisible, i.e. that it cannot be sold in pieces. However, this assumption is only for notational convenience since it would be suboptimal for an agent with risk neutral preferences to divide the sale of an asset into several trades.

- The inclusion of a constant discount rate $r > 0$ is straightforward. Indeed, the discounted price $\tilde{S}_t := e^{-rt} S_t$ satisfies

  $$\mathrm{d}\tilde{S}_t = (X - r)\tilde{S}_t \,\mathrm{d}t + \sigma \tilde{S}_t \,\mathrm{d}W_t, \tag{2.3}$$

  and so the optimal stopping problem

  $$\sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}[e^{-r\tau} S_\tau]$$

  reduces to (2.2) but with the prior distribution replaced by $\mu(\cdot + r)$.

- Including a power utility function to account for the position's riskiness is also straightforward. In fact, if $\gamma \in (0, 1)$, then the optimal stopping problem

  $$\sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}[S_\tau^\gamma]$$

  is the same as

  $$\sup_{\tau \in \mathcal{T}_T^{\hat{S}}} \mathbb{E}[\hat{S}_\tau],$$

  where $\hat{S} := S^\gamma$ is a geometric Brownian motion with drift $\gamma(X + (\gamma - 1)\sigma^2/2)$ and volatility $\gamma\sigma$.

2.1 Equivalent reformulation under a measure change

Assuming that μ has a first moment, $\hat{X}_t := \mathbb{E}[X \mid \mathcal{F}_t^S]$ exists, and the process

$$\hat{W}_t := \frac{1}{\sigma}\int_0^t (X - \hat{X}_s)\,\mathrm{d}s + W_t,$$

known as the innovation process, is an $\mathbb{F}^S$-Brownian motion (see [1, Proposition 2.30 on p. 33]). Writing $\mathbb{F}^{\hat{W}} = \{\mathcal{F}_t^{\hat{W}}\}_{t\geq 0}$ for the completion of the filtration $\{\sigma(\hat{W}_s : 0 \leq s \leq t)\}_{t\geq 0}$, we note that $\mathbb{F}^S = \mathbb{F}^{\hat{W}}$ (see the remark on p. 35 in [1]).

Defining a change of measure by the random variable

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = e^{\sigma \hat{W}_T - \frac{\sigma^2}{2}T},$$

and writing

$$\begin{aligned} S_t &= S_0 e^{Xt + \sigma W_t - \frac{\sigma^2}{2}t} \\ &= S_0 e^{\int_0^t \hat{X}_s\,\mathrm{d}s + \sigma \hat{W}_t - \frac{\sigma^2}{2}t}, \end{aligned}$$

we have

$$\mathbb{E}\left[S_\tau\right] = \mathbb{E}^{\mathbb{Q}}\left[S_0 e^{\int_0^\tau \hat{X}_s\,\mathrm{d}s}\right] = S_0 \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^\tau \hat{X}_s\,\mathrm{d}s}\right],$$

where $\tau \in \mathcal{T}_T^S$; the measures $\mathbb{P}$ and $\mathbb{Q}$ are clearly equivalent. Without loss of generality, we assume $S_0 = 1$ throughout the article; the optimal stopping problem (2.2) then becomes

$$V = \sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}^{\mathbb{Q}}[e^{\int_0^\tau \hat{X}_s\,\mathrm{d}s}]. \tag{2.4}$$

We also note that, by Girsanov's theorem, the process $Z_t := -\sigma t + \hat{W}_t$ is a $\mathbb{Q}$-Brownian motion on $[0, T]$.

3 Analysis of the optimal stopping problem

3.1 Projecting onto the observable filtration

Let us introduce $Y_t := Xt + \sigma W_t$ so that $S_t = S_0 e^{Y_t - \frac{\sigma^2}{2}t}$. Clearly, the processes Y and S generate the same filtrations. The following proposition describes the conditional distribution of X given observations of the stock price in terms of the current value of the process Y. For its proof, see Proposition 3.16 in [1].

Proposition 3.1. *Let $q : \mathbb{R} \to \mathbb{R}$ satisfy $\int_{\mathbb{R}} |q(u)|\mu(\mathrm{d}u) < \infty$. Then*

$$\mathbb{E}\left[q(X)|\mathcal{F}_t^S\right] = \mathbb{E}\left[q(X)|Y_t\right] = \frac{\int_{\mathbb{R}} q(u) e^{\frac{2uY_t - u^2 t}{2\sigma^2}}\mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{2uY_t - u^2 t}{2\sigma^2}}\mu(\mathrm{d}u)}$$

for any $t \geq 0$.

By Proposition 3.1, the distribution $\mu_{t,y}$ of X at time t conditional on $Y_t = y$ is given by

$$\mu_{t,y}(\mathrm{d}u) := \frac{e^{\frac{2uy-u^2t}{2\sigma^2}}\mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{2uy-u^2t}{2\sigma^2}}\mu(\mathrm{d}u)}, \tag{3.1}$$

and

$$\hat{X}_t = \mathbb{E}[X \mid \mathcal{F}_t^S] = \mathbb{E}\left[X|Y_t\right] = f(t, Y_t) \tag{3.2}$$

for any $t > 0$, where

$$f(t,y) = \int_{\mathbb{R}} u\mu_{t,y}(\mathrm{d}u) = \frac{\int_{\mathbb{R}} ue^{\frac{2uy-u^2t}{2\sigma^2}}\mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{2uy-u^2t}{2\sigma^2}}\mu(\mathrm{d}u)} \tag{3.3}$$

As a shorthand, we denote by $\mathbb{E}_{t,y}$ the expectation operator under the probability measure $\mathbb{P}_{t,y}(\cdot) := \mathbb{P}(\cdot|Y_t = y)$.

From now onwards, the following integrability condition on μ is imposed.

Assumption 3.2. *The prior distribution μ satisfies*

$$\int_{\mathbb{R}} e^{au^2}\mu(\mathrm{d}u) < \infty \tag{3.4}$$

for some $a > 0$.

This assumption is an insignificant restriction on our optimal liquidation problem, since, given any probability distribution μ, the distributions $\mu_{t,y}$ in (3.1) satisfy (3.4) for any $t > 0$. The main benefit of Assumption 3.2 is that it allows us to extend the definition of $\mu_{t,y}$ in (3.1) to $t = 0$. Indeed, suppose that μ satisfies (3.4) with $a = \epsilon/(2\sigma^2)$ for some $\epsilon > 0$. Defining a probability distribution ξ on $\mathbb{R}$ by

$$\xi(\mathrm{d}u) := \frac{e^{\frac{\epsilon u^2}{2\sigma^2}}\mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{\epsilon u^2}{2\sigma^2}}\mu(\mathrm{d}u)}, \tag{3.5}$$

the measure

$$\mu_{0,y}(\mathrm{d}u) = \frac{e^{\frac{uy}{\sigma^2}}\mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{uy}{\sigma^2}}\mu(\mathrm{d}u)}$$

coincides with

$$\xi_{\epsilon,y}(\mathrm{d}u) := \frac{e^{\frac{2uy-u^2\epsilon}{2\sigma^2}}\xi(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{2uy-u^2\epsilon}{2\sigma^2}}\xi(\mathrm{d}u)}.$$

Consequently, the distribution $\mu_{0,y}$ can be identified with a conditional distribution at time 0 given that the prior distribution at time $-\epsilon$ was ξ and the current value of the observation process is y. This gives us a generalisation of the notion of the starting point of the observation process Y to allow $Y_0 = y \neq 0$, so we may regard time 0 as an interior point of the time interval.

Next, we establish a bijective correspondence between S_t and $\hat{X}_t$. For it, let I_μ denote the interior of the smallest closed interval containing the support of μ, i.e. $I_\mu = (\inf(\mathrm{supp}(\mu)), \sup(\mathrm{supp}(\mu)))$.

Lemma 3.3. *For any given $t \geq 0$, the function $f(t, \cdot) : \mathbb{R} \to I_\mu$ defined in* (3.3) *is a strictly increasing continuous bijection.*

Proof. Thanks to Assumption 3.2, it suffices to prove the claim only for $t = 0$. Differentiation of f under the integral sign yields $\partial_2 f(0, y) = \frac{1}{\sigma^2}(\mathbb{E}_{0,y}[X^2] - \mathbb{E}_{0,y}[X]^2)$, which is strictly positive and finite for all $y \in \mathbb{R}$. As a result, $y \mapsto f(0, y)$ is strictly increasing. For surjectivity, we need that $f(0, y) \to \sup I_\mu$ as $y \to \infty$ and $f(0, y) \to \inf I_\mu$ as $y \to -\infty$. We only prove the first claim as the second one then follows immediately by symmetry.

Let $\theta \in I_\mu \cap (0, \infty)$, $y > 0$, and consider

$$\int_{\mathbb{R}} u e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) - \theta \int_{\mathbb{R}} e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) = \int_{\mathbb{R}} (u - \theta) e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) = e^{\frac{\theta y}{\sigma^2}} \int_{\mathbb{R}} w e^{\frac{wy}{\sigma^2}} \mu(\theta + \mathrm{d}w), \quad (3.6)$$

where $w := u - \theta$. As the minimum of $w \mapsto w e^{\frac{wy}{\sigma^2}}$ is attained at $w = -\sigma^2/y$, we have

$$\int_{(-\infty,0]} w e^{\frac{wy}{\sigma^2}} \mu(\theta + \mathrm{d}w) \geq -\frac{\sigma^2 e^{-1}}{y} \int_{(-\infty,0]} \mu(\theta + \mathrm{d}w) \geq -\frac{\sigma^2}{y}.$$

Furthermore,

$$\int_{(0,\infty)} w e^{\frac{wy}{\sigma^2}} \mu(\theta + \mathrm{d}w) \to \infty$$

as $y \to \infty$ by monotone convergence. Consequently, from (3.6) follows that $f(0, y) \geq \theta$ for all large enough y. Since $\theta \in I_\mu$ was arbitrary, we conclude that $f(0, y) \to \sup I_\mu$ as $y \to \infty$, which finishes the proof. □

Writing $\mathbb{F}^{\hat{X}} = \{\mathcal{F}_t^{\hat{X}}\}_{t \geq 0}$ for the completion of the filtration generated by $\hat{X}$ and writing $\mathcal{T}_T^{\hat{X}}$ for the set of $\mathbb{F}^{\hat{X}}$-stopping times not exceeding T, we formulate the following immediate corollary.

Corollary 3.4. $\mathbb{F}^S = \mathbb{F}^{\hat{X}}$ *and* $\mathcal{T}_T^S = \mathcal{T}_T^{\hat{X}}$.

A consequence of this corollary is that the optimal stopping problem (2.4) can be rewritten as

$$V = \sup_{\tau \in \mathcal{T}_T^{\hat{X}}} \mathbb{E}^{\mathbb{Q}}[e^{\int_0^\tau \hat{X}_s \, \mathrm{d}s}]. \quad (3.7)$$

Looking for a more tractable characterisation of $\hat{X}$, we find an SDE representation of $\hat{X}$ with respect to the observations filtration $\mathbb{F}^S$. An application of Itô's formula to $\hat{X}_t = f(t, Y_t)$ yields

$$\mathrm{d}\hat{X}_t = \sigma \partial_2 f(t, Y_t) \, \mathrm{d}\hat{W}_t. \quad (3.8)$$

Introducing the notation $f_t := f(t, \cdot)$, we define

$$\psi(t,x) := \sigma \partial_2 f(t, f_t^{-1}(x))$$

and, from (3.8) above, obtain a stochastic differential equation

$$\mathrm{d}\hat{X}_t = \psi(t, \hat{X}_t)\,\mathrm{d}\hat{W}_t$$

for the conditional mean $\hat{X}_t$. Rewriting of the equation in terms of the $\mathbb{Q}$-Brownian motion $Z_t = -\sigma t + \hat{W}_t$ results in

$$\mathrm{d}\hat{X}_t = \sigma\psi(t, \hat{X}_t)\,\mathrm{d}t + \psi(t, \hat{X}_t)\,\mathrm{d}Z_t. \tag{3.9}$$

The dispersion ψ can be expressed more explicitly (by differentiating f under the integral sign) as

$$\psi(t,x) = \frac{1}{\sigma}\left(\mathbb{E}_{t,y_x(t)}[X^2] - \mathbb{E}_{t,y_x(t)}[X]^2\right) = \frac{1}{\sigma}\,\mathrm{Var}_{t,y_x(t)}(X),$$

where the notation $y_x(t) := f_t^{-1}(x)$ is used (note that $y_x(t)$ is the unique value of the observation process Y_t that yields $\hat{X}_t = x$).

Example (The two-point prior) Suppose $\mu = \pi\delta_h + (1-\pi)\delta_l$, where $\pi \in (0,1)$, the symbols δ_l, δ_h denote the Dirac measures at $l < 0$ and $h > 0$, respectively. Then $\psi(t,x) = \frac{1}{\sigma}(h-x)(x-l)$.

Example (The normal prior) Suppose μ is the normal distribution with mean m and variance γ^2. Then the conditional distribution $\mathbb{P}(\cdot|Y_t = y) = \mu_{t,y}$ is also normal but with mean $\frac{\sigma^2 m + \gamma^2 y}{\sigma^2 + t\gamma^2}$ and variance $\frac{\sigma^2\gamma^2}{\sigma^2+t\gamma^2}$. Consequently, $\psi(t,x) = \frac{\sigma\gamma^2}{\sigma^2+t\gamma^2}$.

3.2 Dispersion of the conditional mean

The following inequality will be the key to understanding the dispersion function ψ.

Proposition 3.5. *Let X be a random variable with $\mathbb{E}[X^4] < \infty$. Then*

$$\mathbb{E}[X^4]\mathbb{E}[X^2] + 2\mathbb{E}[X^3]\mathbb{E}[X^2]\mathbb{E}[X] - \mathbb{E}[X^4]\mathbb{E}[X]^2 - \mathbb{E}[X^3]^2 - \mathbb{E}[X^2]^3 \geq 0$$

with the equality if and only if X has a one-point or a two-point distribution.

Proof. Let X, Y, Z be independent and identically distributed random variables with $\mathbb{E}[X^4] < \infty$. Observe that

$$\begin{aligned}
&\mathbb{E}[(X-Y)^2(Y-Z)^2(Z-X)^2] \\
&= \mathbb{E}[X^4(Y^2+Z^2) + Y^4(Z^4+X^4) + Z^4(X^4+Y^4)] \\
&\quad -2\mathbb{E}[X^4YZ + Y^4ZX + Z^4XY] \\
&\quad +2\mathbb{E}[X^3(Y^2Z+Z^2Y) + Y^3(Z^2X+X^2Z) + Z^3(X^2Y+Y^2X)] \\
&\quad -2\mathbb{E}[X^3Z^3 + Y^3X^3 + Z^3Y^3] - 6\mathbb{E}[X^2Y^2Z^2] \\
&= 6(\mathbb{E}[X^4]\mathbb{E}[X^2] - \mathbb{E}[X^4]\mathbb{E}[X]^2 + 2\mathbb{E}[X^3]\mathbb{E}[X^2]\mathbb{E}[X] - \mathbb{E}[X^3]^2 - \mathbb{E}[X^2]^3),
\end{aligned}$$

where the last equality holds because X, Y, Z are i.i.d. It is clear that

$$\mathbb{E}[(X-Y)^2(Y-Z)^2(Z-X)^2] \geq 0$$

with the equality if and only if X has a one-point or a two-point distribution. This finishes the proof of the claim. □

Remark We are grateful to Johan Tysk for providing an alternative proof of the above proposition based on the Pythagorean theorem in L^2 spaces.

Proposition 3.6 (Properties of the dispersion function ψ)**.**

1. *For any $x \in I_\mu$, the function $t \mapsto \psi(t,x)$ is non-increasing. It is strictly decreasing unless μ is a two-point distribution, in which case $t \mapsto \psi(t,x)$ is a constant.*
2. *$\partial_2^2 \psi \geq -\frac{2}{\sigma}$ with a strict inequality unless μ is a two-point distribution, in which case we have equality.*
3. *If μ is compactly supported, then ψ is bounded.*

Proof. 1. Recall the notation $y_x(t) = f_t^{-1}(x)$, and consider

$$\begin{aligned}
\partial_1 \psi(t,x) &= \frac{\partial}{\partial t}\left(\frac{1}{\sigma}\left(\mathbb{E}_{t,y_x(t)}[X^2] - \mathbb{E}_{t,y_x(t)}[X]^2\right)\right) \\
&= \frac{\partial}{\partial t}\left(\frac{1}{\sigma}\left(\mathbb{E}_{t,y_x(t)}[X^2] - x^2\right)\right) \\
&= \frac{1}{\sigma}\Bigg(\mathbb{E}_{t,y_x(t)}\left[X^2\left(\frac{y_x'(t)}{\sigma^2}X - \frac{1}{2\sigma^2}X^2\right)\right] \\
&\quad - \mathbb{E}_{t,y_x(t)}\left[X^2\right]\mathbb{E}_{t,y_x(t)}\left[\frac{y_x'(t)}{\sigma^2}X - \frac{1}{2\sigma^2}X^2\right]\Bigg) \\
&= \frac{1}{\sigma^3}\Big(y_x'(t)\left(\mathbb{E}_{t,y_x(t)}[X^3] - \mathbb{E}_{t,y_x(t)}[X^2]\mathbb{E}_{t,y_x(t)}[X]\right) \\
&\quad - \frac{1}{2}\left(\mathbb{E}_{t,y_x(t)}[X^4] - \mathbb{E}_{t,y_x(t)}[X^2]^2\right)\Big)
\end{aligned}$$

Implicit differentiation using the identity $x = f(t, y_x(t))$ gives that

$$y_x'(t) = \frac{1}{2}\frac{\mathbb{E}_{t,y_x(t)}[X^3] - \mathbb{E}_{t,y_x(t)}[X^2]\mathbb{E}_{t,y_x(t)}[X]}{\mathrm{Var}_{t,y_x(t)}(X)},$$

which substituted into the last expression above yields

$$\begin{aligned}
\partial_1 \psi(t,x) &= \frac{\left(\mathbb{E}_{t,y_x(t)}[X^3] - \mathbb{E}_{t,y_x(t)}[X^2]\mathbb{E}_{t,y_x(t)}[X]\right)^2 - \mathrm{Var}_{t,y_x(t)}(X^2)\,\mathrm{Var}_{t,y_x(t)}(X)}{2\sigma^3\,\mathrm{Var}_{t,y_x(t)}(X)} \\
&= \frac{-1}{2\sigma^3\,\mathrm{Var}_{t,y_x(t)}(X)}\Big(\mathbb{E}_{t,y_x(t)}[X^4]\mathbb{E}_{t,y_x(t)}[X^2] \\
&\quad + 2\mathbb{E}_{t,y_x(t)}[X^3]\mathbb{E}_{t,y_x(t)}[X^2]\mathbb{E}_{t,y_x(t)}[X] \\
&\quad - \mathbb{E}_{t,y_x(t)}[X^4]\mathbb{E}_{t,y_x(t)}[X]^2 - \mathbb{E}_{t,y_x(t)}[X^3]^2 - \mathbb{E}_{t,y_x(t)}[X^2]^3\Big).
\end{aligned}$$

Now, the claim follows from Proposition 3.5 applied to the term between the parentheses.

2. By the chain rule applied to the definition of ψ, we have

$$\partial_2\psi(t,x) = \sigma\partial_2^2 f(t, y_x(t))\partial_2 y(t,x),$$

where $y(t,x) := y_x(t)$. Here

$$\begin{aligned}\partial_2^2 f(t,y) &= \frac{1}{\sigma^2}\frac{\partial}{\partial y}\left(\mathbb{E}_{t,y}[X^2] - \mathbb{E}_{t,y}[X]^2\right)\\ &= \frac{1}{\sigma^4}\left(\mathbb{E}_{t,y}[X^3] - 3\mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X] + 2\mathbb{E}_{t,y}[X]^3\right)\end{aligned}$$

by straightforward differentiation under the integral sign, and

$$\partial_2 y(t,x) = \frac{1}{\partial_2 f(t, y(t,x))} = \frac{\sigma^2}{\mathrm{Var}_{t,y_x(t)}(X)}.$$

by implicit differentiation. Hence

$$\partial_2\psi(t,x) = \frac{1}{\sigma}\left(\frac{\mathbb{E}_{t,y_x(t)}[X^3] - \mathbb{E}_{t,y_x(t)}[X^2]\mathbb{E}_{t,y_x(t)}[X]}{\mathrm{Var}_{t,y_x(t)}(X)} - 2x\right).$$

It remains to establish the inequality

$$\partial_2^2\psi(t,x) + \frac{2}{\sigma} = \frac{1}{\sigma}\frac{\partial}{\partial y}\left(\frac{\mathbb{E}_{t,y}[X^3] - \mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X]}{\mathrm{Var}_{t,y}(X)}\right)\Bigg|_{y=y_x(t)}\partial_2 y(t,x) \geq 0.$$

As $\partial_2 y > 0$, equivalently, it suffices to prove the non-negativity of

$$\begin{aligned}q(t,y) &:= \mathrm{Var}_{t,y}(X)^2\frac{\partial}{\partial y}\left(\frac{\mathbb{E}_{t,y}[X^3] - \mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X]}{\mathrm{Var}_{t,y}(X)}\right)\\ &= \frac{\partial}{\partial y}\left(\mathbb{E}_{t,y}[X^3]\right)\mathrm{Var}_{t,y}(X) - \mathbb{E}_{t,y}[X^3]\frac{\partial}{\partial y}\left(\mathrm{Var}_{t,y} X\right)\\ &\quad -\frac{\partial}{\partial y}\left(\mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X]\right)\mathrm{Var}_{t,y}(X) + \mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X]\frac{\partial}{\partial y}\mathrm{Var}_{t,y}(X).\end{aligned}$$

Further differentiation yields that

$$\begin{aligned}q(t,y) = \frac{1}{\sigma^2}\Big(&\mathbb{E}_{t,y}[X^4]\mathbb{E}_{t,y}[X^2] + 2\mathbb{E}_{t,y}[X^3]\mathbb{E}_{t,y}[X^2]\mathbb{E}_{t,y}[X]\\ &-\mathbb{E}_{t,y}[X^4]\mathbb{E}_{t,y}[X]^2 - \mathbb{E}_{t,y}[X^3]^2 - \mathbb{E}_{t,y}[X^2]^3\Big).\end{aligned}$$

Thus, by Proposition 3.5, $q \geq 0$; moreover, $q > 0$ for all priors μ except the two-point distribution in which case $q = 0$.

3. The well-known identity

$$\mathbb{E}_{t,y_x(t)}[|X|^2] = 2\int_{[0,\infty)} u\mathbb{P}_{t,y_x(t)}(|X| > u)\,\mathrm{d}u$$

ensures that ψ is bounded for compactly supported distributions.

□

Remark

1. It is possible to come up with a contrived example of a prior distribution for which the dispersion ψ is unbounded. For this, think of a discrete probability measure supported on an infinite number of points $x_1 < x_2 < \ldots < x_n < \ldots$ such that $x_n - x_{n-1} \to \infty$ as $n \to \infty$. Using the notation $\bar{x}_n := (x_{n-1} + x_n)/2$ for the mean between neighbouring points, the value $\psi(t, \bar{x}_n) = Var_{t,\bar{x}_n}(X)/\sigma \to \infty$ as $n \to \infty$ by comparison with a two-point distribution concentrated at the points x_{n-1} and x_n.

2. Let us stress that compact support of the prior is by no means a necessary condition for the boundedness of ψ. For instance, we know that ψ is bounded in the case of a normal prior as seen in the example on page 9. Though a rigorous investigation into precise technical conditions on the prior for the boundedness of ψ appears to be involved enough to be omitted in this article, we conjecture, based on numerical investigations, that ψ is bounded for any prior admitting a density that monotonically approaches zero outside a large enough finite-length interval around the origin.

As the boundedness of ψ appears to be satisfied by any conceivable prior of interest in practical applications, we make it an assumption in the rest of the article.

Assumption 3.7. *The prior distribution* μ *is such that* $\mathrm{Var}_{0,y}(X) < \infty$ *for all* $y \in \mathbb{R}$.

3.3 The Markovian value function and the optimal strategy

Using the dynamics (3.9) of $\hat{X}$, we are able to embed the optimal stopping problem (3.7) into a Markovian framework. To do that, define

$$v(t, x) = \sup_{\tau \in \mathcal{T}_{T-t}} \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^\tau \hat{X}^{t,x}_{t+s}\,\mathrm{d}s}\right], \quad (t, x) \in [0, T] \times I_\mu, \tag{3.10}$$

where the process $\hat{X} = \hat{X}^{t,x}$ is given by

$$\begin{cases} \mathrm{d}\hat{X}_{t+s} = \sigma\psi(t+s, \hat{X}_{t+s})\,\mathrm{d}s + \psi(t+s, \hat{X}_{t+s})\,\mathrm{d}Z_{t+s} \qquad (s > 0), \\ \hat{X}_t = x, \end{cases}$$

and $\mathcal{T}_{T-t}$ denotes the set of stopping times less or equal to $T - t$ with respect to the completed filtration of $\{\hat{X}^{t,x}_{t+s}\}_{s\geq 0}$.

Let us define the sets

$$\mathcal{C} = \{(t, x) \in [0, T] \times I_\mu : v(t, x) > 1\}$$

and

$$\mathcal{D} = \{(t, x) \in [0, T] \times I_\mu : v(t, x) = 1\},$$

which we will soon show to correspond respectively to continuation and stopping sets of an optimal strategy. Note that

$$v \geq 1 \tag{3.11}$$

everywhere and that $v(T, x) = 1$, so $\mathcal{C} \cup \mathcal{D} = [0, T] \times I_\mu$ and the random time

$$\tau_{\mathcal{D}} := \inf\{s \geq 0 : (t + s, \hat{X}^{t,x}_{t+s}) \in \mathcal{D}\} \tag{3.12}$$

satisfies $\tau_{\mathcal{D}} \leq T - t$.

Proposition 3.8 (Optimal stopping time)**.** *The value function v is finite, and the time $\tau_{\mathcal{D}}$ defined in* (3.12) *is an optimal stopping time.*

Proof. Without loss of generality, assume that $t = 0$, and let $x \in \mathbb{R}$. By Theorem D.12 in [16], to prove the claims it suffices to show that

$$\mathbb{E}^{\mathbb{Q}}\left[\sup_{0 \leq t \leq T} \exp\left(\int_0^t \hat{X}^{0,x}_s \, \mathrm{d}s\right)\right] < \infty.$$

By the Dambis-Dubins-Schwartz theorem, there exists (possibly on a larger probability space) a Brownian motion B such that

$$\int_0^t \psi(s, \hat{X}^{0,x}_s) \, \mathrm{d}Z_s = B_{\int_0^t \psi(t, \hat{X}^{0,x}_s)^2 \, \mathrm{d}s}.$$

If $m > 0$ is a constant dominating ψ, then

$$\begin{aligned}
\sup_{0 \leq t \leq T} \exp\left(\int_0^t \hat{X}^{0,x}_s \, \mathrm{d}s\right) &\leq \exp\left(T \sup_{0 \leq t \leq T} \hat{X}^{0,x}_t\right) \\
&\leq \exp\left(T\left(x + \sigma m T + \sup_{0 \leq t \leq T} B_{\int_0^t \psi(s, \hat{X}^{0,x}_s)^2 \, \mathrm{d}s}\right)\right) \\
&\leq \exp\left(T\left(x + \sigma m T + \sup_{0 \leq t \leq m^2 T} B_t\right)\right).
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\sup_{0 \leq t \leq T} \exp\left(\int_0^T \hat{X}^{0,x}_s \, \mathrm{d}s\right)\right] &\leq \mathbb{E}^{\mathbb{Q}}\left[\exp\left(T\left(x + \sigma m T + \sup_{0 \leq t \leq m^2 T} B_t\right)\right)\right] \\
&= \exp\left(T\left(x + \sigma m T\right)\right) \mathbb{E}^{\mathbb{Q}}\left[\exp\left(T |B_{m^2 T}|\right)\right] \\
&< \infty,
\end{aligned}$$

where the equality comes from the reflection principle. □

Since ψ is continuously differentiable, it is Lipschitz continuous on any compact subset of $[0,T] \times I_\mu$. To avoid additional technical complications, from now on we impose a slightly stronger assumption of Lipschitz continuity on the whole of $[0,T] \times I_\mu$

Assumption 3.9. *The function ψ is Lipschitz continuous in the second variable, i.e. there exists $K > 0$ such that $|\psi(t,x) - \psi(t,y)| \leq K|x-y|$ for all $t \in [0,T]$ and all $x, y \in I_\mu$.*

We remark that our canonical examples of the normal and the two-point prior both fulfill Assumption 3.9.

Theorem 3.10 (Properties of the value function)**.**

1. *The function $x \mapsto v(t,x)$ is non-decreasing and convex for any fixed $t \in [0,T]$.*
2. *The function $t \mapsto v(t,x)$ is non-increasing for any fixed $x \in I_\mu$.*
3. *The value function v is continuous on $[0,T] \times I_\mu$.*
4. *There exists a non-decreasing and continuous function $h : [0,T] \to (-\infty, 0]$ with $h(T) = 0$ such that $\mathcal{C} = \{(t,x) \in [0,T) \times I_\mu : x > h(t)\}$.*
5. *The value function $(t,x) \mapsto v(t,x)$ solves the boundary value problem*

$$\begin{cases} \partial_1 v + \sigma\psi(t,x)\partial_2 v + \frac{1}{2}\psi(t,x)^2\partial_2^2 v + xv = 0, & x \in \mathcal{C}, \\ v = 1, & x \in \mathcal{D}. \end{cases} \tag{3.13}$$

   Furthermore, the smooth-fit property holds in that the function $x \mapsto v(t,x)$ is C^1 for all $t \in [0,T]$.

Proof.

1. (i) The monotonicity of $x \mapsto v(t,x)$ is clear from the representation

$$v(t,x) = \sup_{\tau \in \mathcal{T}_{T-t}} \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^\tau \hat{X}_{t+s}^{t,x}\,\mathrm{d}s}\right] \tag{3.14}$$

   of the value function together with a comparison theorem, see [27, Theorem IX.3.7].

   (ii) Let us define $v_E(t,x) := \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{T-t} \hat{X}_{t+s}^{t,x}\,\mathrm{d}s}\right]$ and $u_E(t,r) := \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T-t} \hat{R}_{t+s}^{t,r}\,\mathrm{d}s}\right]$, where $\hat{R} = -\hat{X}$ and so

$$\mathrm{d}\hat{R}_t = -\sigma\psi(t, -\hat{R}_t)\,\mathrm{d}t - \psi(t, -\hat{R}_t)\,\mathrm{d}Z_t.$$

   Then $v_E(t,x) = u_E(t,-x)$. Now, the convexity result follows by approximating the value function, starting with v_E as the first approximation, by Bermudan options, which preserve convexity by [10, Theorem 5.1] with the needed condition $\partial_2^2\psi \geq -\frac{2}{\sigma}$ for the theorem to hold ensured by Proposition 3.6.

2. From Proposition 3.6, the dispersion ψ is non-increasing in t, so the claim follows by the Bermudan approximation argument for the value function and Theorem 6.1 in [10].

3. First, let $l \in I_\mu$ and we will show that there exists a constant $K > 0$ such that, for every $t \in [0, T]$, the map $x \mapsto v(t, x)$ is K-Lipschitz continuous on $(-\infty, l] \cap I_\mu$. Assume for a contradiction that there is no such K. Recall that convexity of a single-variable function implies continuity and existence of one-sided derivatives. Hence using a characterisation of convexity saying that a real-valued function f defined on an interval is convex if and only if the function $(x_1, x_2) \mapsto (f(x_2) - f(x_1))/(x_2 - x_1)$ is increasing in both x_1 and x_2, we obtain that there is a sequence $\{t_n\}_{n\geq 0} \subset [0, T]$ such that the sequence of left-derivatives $\partial_2^- v(t_n, l)$ diverges to ∞. However, taking $\epsilon \in (0, \sup I_\mu - l)$, this would imply that $v(t_n, l + \epsilon) \to \infty$, contradicting the fact that $v(t_n, l + \epsilon) \leq v(0, l + \epsilon) < \infty$ for all $n \in \mathbb{N}$.

   To finish the proof of the continuity of v, it suffices to show that $v(t, x)$ is continuous in t. To reach a contradiction, assume that $t \mapsto v(t, x_0)$ is not continuous at $t = t_0$ for some x_0. By time-decay, this means that v has a negative jump.

   First consider the case when $v(t_0-, x_0) > v(t_0, x_0)$. By Lipschitz continuity in the second variable, there exists a rectangle $\mathcal{R} = (t_0 - \delta, t_0) \times (x_0 - \delta, x_0 + \delta)$ with $\delta > 0$ such that

$$\inf_{(t,x)\in\mathcal{R}} v(t, x) > v(t_0, x_0 + \delta). \tag{3.15}$$

   Thus $\mathcal{R} \subseteq \mathcal{C}$. Let $t \in (t_0 - \delta, t_0)$ and $\tau_{\mathcal{R}} := \inf\{u \geq 0 : (t + u, \hat{X}^{t,x_0}_{t+u}) \notin \mathcal{R}\}$. Then, by martingality in the continuation region (see [16, Appendix D]),

$$\begin{aligned} v(t, x_0) &= \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{\tau_{\mathcal{R}}} \hat{X}^{t,x_0}_{t+u}\,\mathrm{d}u} v(t + \tau_{\mathcal{R}}, \hat{X}^{t,x_0}_{t+\tau_{\mathcal{R}}})\right] \\ &\leq \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{t_0-t} \hat{X}^{t,x_0}_{t+u}\vee 0\,\mathrm{d}u} v(t, x_0 + \delta)\mathbb{1}_{\{t+\tau_{\mathcal{R}}<t_0\}}\right] \\ &\quad + \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{t_0-t} \hat{X}^{t,x_0}_{t+u}\vee 0\,\mathrm{d}u} v(t_0, x_0 + \delta)\mathbb{1}_{\{t+\tau_{\mathcal{R}}=t_0\}}\right] \\ &\leq e^{(t_0-t)(x_0+\delta)^+} v(t, x_0 + \delta)\mathbb{Q}(t + \tau_{\mathcal{R}} < t_0) + e^{(t_0-t)(x_0+\delta)^+} v(t_0, x_0 + \delta) \\ &\to v(t_0, x_0 + \delta) \end{aligned}$$

   as $t \to t_0$, which contradicts (3.15).

   Next, consider the case when $v(t_0, x_0) > v(t_0+, x_0)$. We begin by investigating the situation $v(t_0, x_0) > v(t_0+, x_0) > 1$. By Lipschitz continuity of v in the second variable and its decay in time, there exists $\mathcal{R} = (t_0, t_0 + \epsilon] \times [x_0 - \delta, x_0 + \delta]$ with $\epsilon > 0$ and $\delta > 0$ such that

$$v(t_0, x_0) > \sup_{(t,x)\in\mathcal{R}} v(t, x) \geq \inf_{(t,x)\in\mathcal{R}} v(t, x) > 1. \tag{3.16}$$

In particular, $\mathcal{R} \subseteq \mathcal{C}$ and writing $\tau_{\mathcal{R}} := \inf\{u \geq 0 : (t_0 + u, \hat{X}^{t_0,x_0}_{t_0+u}) \notin \mathcal{R}\}$ we have

$$\begin{aligned}
v(t_0, x_0) &= \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{\tau_{\mathcal{R}}} \hat{X}^{t_0,x_0}_{t_0+u}\,\mathrm{d}u} v(t_0 + \tau_{\mathcal{R}}, \hat{X}^{t_0,x_0}_{t_0+\tau_{\mathcal{R}}})\right] \\
&\leq \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{\epsilon} \hat{X}^{t_0,x_0}_{t_0+u} \vee 0\,\mathrm{d}u} v(t_0, x_0 + \delta)\mathbb{1}_{\{\tau_{\mathcal{R}} < \epsilon\}}\right] \\
&\quad + \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{\epsilon} \hat{X}^{t_0,x_0}_{t_0+u} \vee 0\,\mathrm{d}u} v(t_0 + \epsilon, x_0 + \delta)\mathbb{1}_{\{\tau_{\mathcal{R}} = \epsilon\}}\right] \\
&\leq e^{\epsilon(x_0+\delta)^+} v(t_0, x_0 + \delta)\mathbb{Q}(\tau_{\mathcal{R}} < \epsilon) + e^{\epsilon(x_0+\delta)^+} v(t_0 + \epsilon, x_0 + \delta) \\
&\to v(t_0+, x_0 + \delta)
\end{aligned}$$

as $\epsilon \searrow 0$, which contradicts (3.16).

Alternatively, suppose that $v(t_0, x_0) > v(t_0+, x_0) = 1$. By Lipschitz continuity in the second variable, there exists $\delta > 0$ such that

$$\inf_{x \in (x_0 - \delta, x_0)} v(t_0, x) > v(t_0+, x_0) = 1. \tag{3.17}$$

Then $(t_0, T] \times (x_0 - \delta, x_0) \subseteq \mathcal{D}$ and so the process $\hat{X}^{t_0, x_0 - \delta/2}$ hits the stopping region immediately, implying that $(t_0, x_0 - \delta/2) \in \mathcal{D}$; this contradicts (3.17).

4. Existence of a non-decreasing boundary $h : [0, T) \to [-\infty, \infty]$ satisfying $\mathcal{C} = \{(t, x) \in [0, T) \times I_\mu : x > h(t)\}$ is a direct consequence of the first two parts above, and we can then define $h(T) = \lim_{t \nearrow T} h(t)$. Non-positivity of h is clear from the expression (3.10), since, for any starting point $(t, x) \in [0, T) \times (0, \infty)$, the strategy of stopping at the first time $\hat{X}^{t,x}$ hits 0 gives a value strictly greater than 1.

   To show that h is bounded from below, assume for a contradiction that $\{0\} \times (-\infty, \infty) \subseteq \mathcal{C}$. Hence, defining ξ as in (3.5), we know that $(-\epsilon, 0] \times \mathbb{R} \subseteq \mathcal{C}_\xi$, where $\mathcal{C}_\xi$ denotes the continuation region for the optimal selling problem started at time $-\epsilon < 0$ with the prior ξ. Writing v_ξ to denote the Markovian value function for the selling problem from time $-\epsilon$, let $-t' \in (-\epsilon, 0)$ and let $a < 0$ be such that $v_\xi(-t', a) < e^{-at'}$. Now, let $x \in (-\infty, a)$, and observe that

$$\begin{aligned}
v_\xi(-t', x) &\leq e^{at'} v_\xi(-t', a)\mathbb{P}\left(\sup_{0 \leq u \leq t'} \hat{X}^{-t',x}_{-t'+u} < a\right) + v_\xi(-t', a)\mathbb{P}\left(\sup_{0 \leq u \leq t'} \hat{X}^{-t',x}_{-t'+u} \geq a\right) \\
&\to e^{at'} v_\xi(-t', a) < 1
\end{aligned}$$

   as $x \searrow -\infty$. This gives a contradiction since $v_\xi \geq 1$ by definition. As a result, we can conclude that $h(t) \in (-\infty, 0]$ for all $t \in [0, T]$.

   For the continuity of h, note that continuity together with time-decay of v imply that h is right-continuous with left limits. Assume for a contradiction that $h(t_0-) < h(t_0)$ for some $t_0 \in (0, T)$. Take points x_1, x_2 with $h(t_0-) < x_1 < x_2 < h(t_0)$, let

$x = (x_1 + x_2)/2$, and consider the rectangle $\mathcal{R} = (t_0 - \delta, t_0) \times (x_1, x_2) \subseteq \mathcal{C}$ for some $\delta > 0$. For $t \in (t_0 - \delta, t_0)$,

$$\begin{aligned} v(t,x) &= \mathbb{E}^{\mathbb{Q}}\left[e^{\int_0^{\tau_{\mathcal{R}}} \hat{X}^{t,x}_{t+u}\,\mathrm{d}u} v(t+\tau_{\mathcal{R}}, \hat{X}^{t,x}_{t+\tau_{\mathcal{R}}})\right] \\ &\leq \mathbb{Q}\left(\tau_{\mathcal{R}} < t_0 - t\right) v(t_0 - \delta, x_2) + e^{x_2(t_0-t)}. \end{aligned} \tag{3.18}$$

Now, estimating $\tau_{\mathcal{R}}$ above with the leaving time of $\mathcal{R}$ for a Brownian motion with drift (compare the proof of Proposition 3.8), it is straightforward to check that $\mathbb{Q}\left(\tau_{\mathcal{R}} < t_0 - t\right) = o(t_0 - t)$. Since $x_2 < 0$, (3.18) thus implies that $v(t,x) < 1$ for t close to t_0, which contradicts (3.11).

The above argument also works to show that $h(T-) = 0$.

5. The proof of (3.13) is along standard lines (e.g. see [16, Theorem 7.7 in Chapter 2]), so we do not include it here.

Let us next establish the smooth-fit property. Since $x \mapsto v(t,x)$ is non-decreasing, it suffices to show that

$$\lim_{\epsilon \downarrow 0} \frac{v(t, h(t)+\epsilon) - v(t, h(t))}{\epsilon} \leq 0.$$

Without loss of generality, let $t = 0$. Writing $x = h(0)$, it is enough to show that

$$v(t, x+\epsilon) - v(t,x) = o(\epsilon) \quad \text{as } \epsilon \searrow 0.$$

Denoting the optimal stopping time when starting at the point $(0, x+\epsilon)$ by τ_ϵ, we have

$$\begin{aligned} v(t,x+\epsilon) - v(t,x) &\leq \mathbb{E}^{\mathbb{Q}}\big[e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u}\big] - \mathbb{E}^{\mathbb{Q}}\big[e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x}_u\,\mathrm{d}u}\big] \\ &= \mathbb{E}^{\mathbb{Q}}\big[e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u}(1 - e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x}_u - \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u})\big] \\ &\leq \mathbb{E}^{\mathbb{Q}}\big[e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u} \int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u - \hat{X}^{0,x}_u\,\mathrm{d}u\big] \\ &\leq \mathbb{E}^{\mathbb{Q}}\big[e^{\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u}\big(\tau_\epsilon \int_0^{\tau_\epsilon} (\hat{X}^{0,x+\epsilon}_u - \hat{X}^{0,x}_u)^2\,\mathrm{d}u\big)^{1/2}\big] \\ &\leq \mathbb{E}^{\mathbb{Q}}\big[\tau_\epsilon e^{2\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u}\big]^{1/2} \mathbb{E}^{\mathbb{Q}}\big[\int_0^{\tau_\epsilon} (\hat{X}^{0,x+\epsilon}_u - \hat{X}^{0,x}_u)^2\,\mathrm{d}u\big]^{1/2}, \end{aligned}$$

where the penultimate inequality follows from Jensen's inequality and the last one from Cauchy-Schwartz. Since the boundary h is non-decreasing, with the help of Lévy's modulus of continuity theorem as well as the law of the iterated logarithm, we see that $\tau_\epsilon \to 0$ a.s. as $\epsilon \searrow 0$. Hence, by the dominated convergence theorem,

$$\mathbb{E}^{\mathbb{Q}}\big[\tau_\epsilon e^{2\int_0^{\tau_\epsilon} \hat{X}^{0,x+\epsilon}_u\,\mathrm{d}u}\big] \to 0 \quad \text{as } \epsilon \searrow 0$$

with the dominating function as in the proof of Proposition 3.8.

To complete the proof of smooth-fit, it suffices to show that

$$\mathbb{E}^{\mathbb{Q}}\left[\left(\int_0^{\tau_\epsilon} \hat{X}_u^{0,x+\epsilon} - \hat{X}_u^{0,x}\,\mathrm{d}u\right)^2\right] = O(\epsilon^2) \quad \text{as } \epsilon \to 0.$$

To this end,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\left(\int_0^{\tau_\epsilon} \hat{X}_u^{0,x+\epsilon} - \hat{X}_u^{0,x}\,\mathrm{d}u\right)^2\right] &\leq T\mathbb{E}^{\mathbb{Q}}\left[\int_0^T (\hat{X}_u^{0,x+\epsilon} - \hat{X}_u^{0,x})^2\,\mathrm{d}u\right] \\ &\leq T^2\mathbb{E}^{\mathbb{Q}}\left[\sup_{0\leq u\leq T}(\hat{X}_u^{0,x+\epsilon} - \hat{X}_u^{0,x})^2\right] \\ &\leq c\epsilon^2,\end{aligned}$$

where c is a constant dependent on T, σ, and the Lipschitz constant of ψ. In the above, the first inequality comes from Jensen's inequality, and the last inequality is a standard estimate coming from an application of Gronwall's inequality combined with Doob's L^2 inequality. This finishes the proof of the claim.

□

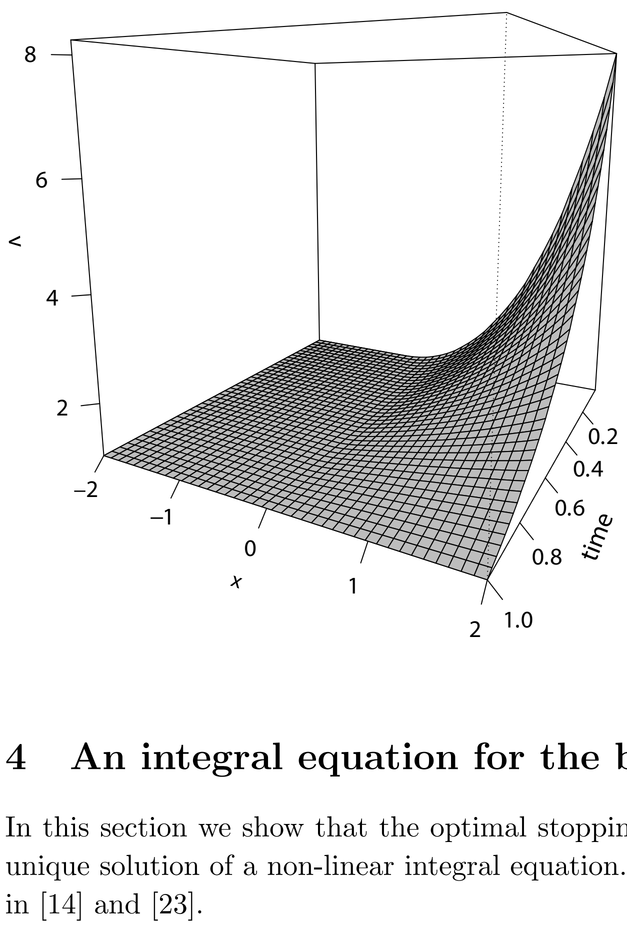

Figure 1: The value function $v(t,x)$ in the case of a normal prior with standard deviation $\gamma = 0.5$, the market volatility $\sigma = 0.2$.

4 An integral equation for the boundary

In this section we show that the optimal stopping boundary can be characterised as the unique solution of a non-linear integral equation. The proof follows along similar lines as in [14] and [23].

Theorem 4.1 (Optimal stopping boundary)**.** *The stopping boundary h is the unique solution to the integral equation*

$$\mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^T \hat{X}_u^{t,h(t)}\,\mathrm{d}u}\right] = 1 + \int_t^T \mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^s \hat{X}_u^{t,h(t)}\,\mathrm{d}u}\hat{X}_s^{t,h(t)}1_{\{\hat{X}_s^{t,h(t)}\le h(s)\}}\right]\mathrm{d}s \tag{4.1}$$

in the class of non-positive continuous functions.

Proof. An application of Itô's formula (more precisely, its extension proved in [24], which can be applied thanks to the monotonicity of h) to $v(s,\hat{X}_s^{t,x})e^{\int_t^s \hat{X}_u^{t,x}\,\mathrm{d}u}$ yields

$$\begin{aligned} v(s,\hat{X}_s^{t,x})e^{\int_t^s \hat{X}_u^{t,x}\,\mathrm{d}u} &= v(t,\hat{X}_t^{t,x}) + \int_t^s e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\Big(\mathcal{L}_{\hat{X}^{t,x}}v(r,\hat{X}_r^{t,x}) + \hat{X}_r^{t,x}v(r,\hat{X}_r^{t,x})\Big)\,\mathrm{d}r \\ &\quad + \int_t^s e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\psi(r,\hat{X}_r^{t,x})\partial_2 v(r,\hat{X}_r^{t,x})\,\mathrm{d}Z_r\,. \end{aligned} \tag{4.2}$$

Let us introduce a localising sequence $\tau_n := \inf\{r\ge t\,:\,\hat{X}_r^{t,x}\ge n\}\wedge T$; it satisfies $\tau_n \nearrow T$ a.s. as $n\to\infty$. Since, for all $n\in\mathbb{N}$,

$$\mathbb{E}^{\mathbb{Q}}\left[\int_t^{s\wedge\tau_n} e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\psi(r,\hat{X}_r^{t,x})\partial_2 v(r,\hat{X}_r^{t,x})\,\mathrm{d}Z_r\right] = 0,$$

from (4.2) we get

$$\mathbb{E}^{\mathbb{Q}}[v(T\wedge\tau_n,\hat{X}_{T\wedge\tau_n}^{t,x})e^{\int_t^{T\wedge\tau_n}\hat{X}_u^{t,x}\,\mathrm{d}u}] = v(t,x) + \mathbb{E}^{\mathbb{Q}}\left[\int_t^{T\wedge\tau_n} e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\hat{X}_r^{t,x}\mathbb{1}_{\{\hat{X}_r^{t,x}\le h(r)\}}\,\mathrm{d}r\right].$$

Letting $n\to\infty$, the equation becomes

$$\mathbb{E}^{\mathbb{Q}}[v(T,\hat{X}_T^{t,x})e^{\int_t^T \hat{X}_u^{t,x}\,\mathrm{d}u}] = v(t,x) + \int_t^T \mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\hat{X}_r^{t,x}\mathbb{1}_{\{\hat{X}_r^{t,x}\le h(r)\}}\right]\mathrm{d}r\,. \tag{4.3}$$

Here, the left-hand side is obtained by dominated convergence as $v(T\wedge\tau_n,\hat{X}_{T\wedge\tau_n}^{t,x})e^{\int_t^{T\wedge\tau_n}\hat{X}_u^{t,x}\,\mathrm{d}u}$ is dominated by $e^{2T(\sup_{t\le u\le T}\hat{X}_u^{t,x}\vee 0)}$, which is integrable; the right-hand side comes from monotone convergence. Substitution $x=h(t)$ in (4.3) gives

$$\mathbb{E}^{\mathbb{Q}}[e^{\int_t^T \hat{X}_u^{t,h(t)}\,\mathrm{d}u}] = 1 + \int_t^T \mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^r \hat{X}_u^{t,h(t)}\,\mathrm{d}u}\hat{X}_r^{t,h(t)}\mathbb{1}_{\{\hat{X}_r^{t,h(t)}\le h(r)\}}\right]\mathrm{d}r\,,$$

which shows that h solves the integral equation (4.1).

For uniqueness, assume that $t\mapsto k(t)$ is another non-positive continuous solution to (4.1) and define

$$\tilde{v}(t,x) := \mathbb{E}^{\mathbb{Q}}[e^{\int_t^T \hat{X}_u^{t,x}\,\mathrm{d}u}] - \mathbb{E}^{\mathbb{Q}}\left[\int_t^T e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\hat{X}_r^{t,x}\mathbb{1}_{\{\hat{X}_r^{t,x}\le k(r)\}}\,\mathrm{d}r\right]. \tag{4.4}$$

Using (4.3), (4.4) and the Markov property, the two processes defined for $s\in[t,T]$ as

$$M_s^{\tilde{v}} := \tilde{v}(s,\hat{X}_s^{t,x})e^{\int_t^s \hat{X}_u^{t,x}\,\mathrm{d}u} - \int_t^s e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u}\hat{X}_r^{t,x}\mathbb{1}_{\{\hat{X}_r^{t,x}\le k(r)\}}\,\mathrm{d}r$$

and

$$M_s^v := v(s, \hat{X}_s^{t,x}) e^{\int_t^s \hat{X}_u^{t,x}\,\mathrm{d}u} - \int_t^s e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \le h(r)\}}\,\mathrm{d}r$$

are easily verified to be $\mathbb{Q}$-martingales.
<u>Claim 1</u>: $\tilde{v}(t,x) = 1$ for $x \le k(t)$.

Let $x \le k(t)$ and define $\gamma_k := \inf\{s \ge 0 : \hat{X}_{t+s}^{t,x} \ge k(t+s)\} \wedge (T-t)$. Then

$$\begin{aligned} M_{t+\gamma_k}^{\tilde{v}} &= \tilde{v}(t+\gamma_k, \hat{X}_{t+\gamma_k}^{t,x}) e^{\int_t^{t+\gamma_k} \hat{X}_u^{t,x}\,\mathrm{d}u} - \int_t^{t+\gamma_k} e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \le k(r)\}}\,\mathrm{d}r \\ &= e^{\int_t^{t+\gamma_k} \hat{X}_u^{t,x}\,\mathrm{d}u} - \int_t^{t+\gamma_k} e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u} \hat{X}_r^{t,x}\,\mathrm{d}r \\ &= 1, \end{aligned}$$

where the second equality follows from (4.1). By optional sampling,

$$\tilde{v}(t,x) = M_t^{\tilde{v}} = \mathbb{E}^{\mathbb{Q}}[M_{t+\gamma_k}^{\tilde{v}}] = 1,$$

which finishes the proof of Claim 1.
<u>Claim 2</u>: $\tilde{v} \le v$.

Suppose $x > k(t)$ and define $\tau_k := \inf\{s \ge 0 : \hat{X}_{t+s}^{t,x} \le k(t+s)\} \wedge (T-t)$. Then

$$\begin{aligned} \tilde{v}(t,x) &= \mathbb{E}^{\mathbb{Q}}\left[\tilde{v}(t+\tau_k, \hat{X}_{t+\tau_k}^{t,x}) e^{\int_t^{t+\tau_k} \hat{X}_u^{t,x}\,\mathrm{d}u}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^{t+\tau_k} \hat{X}_u^{t,x}\,\mathrm{d}u}\right] \\ &\le v(t,x), \end{aligned}$$

with the first equality following from the martingality of $M^{\tilde{v}}$ and the optional sampling theorem, the second by the definition of $\tilde{v}$ and (4.1). Combining this with Claim 1, the result is obtained.
<u>Claim 3</u>: $h \le k$.

Assume for a contradiction that $h(t) > k(t)$ for some t. Let $x = k(t)$ and define $\gamma_h := \inf\{s \ge 0 : \hat{X}_{t+s}^{t,x} \ge h(t+s)\} \wedge (T-t)$. Then

$$\begin{aligned} 0 &= v(t,x) - \tilde{v}(t,x) \\ &= \mathbb{E}^{\mathbb{Q}}[M_{t+\gamma_h}^v - M_{t+\gamma_h}^{\tilde{v}}] \\ &= \mathbb{E}^{\mathbb{Q}}\left[e^{\int_t^{t+\gamma_h} \hat{X}_u^{t,x}\,\mathrm{d}u}(v(t+\gamma_h, \hat{X}_{t+\gamma_h}^{t,x}) - \tilde{v}(t+\gamma_h, \hat{X}_{t+\gamma_h}^{t,x}))\right] \\ &\quad - \mathbb{E}^{\mathbb{Q}}\left[\int_t^{t+\gamma_h} e^{\int_t^r \hat{X}_u^{t,x}\,\mathrm{d}u} \hat{X}_r^{t,x} \mathbb{1}_{\{X_r^{t,x} \in (k(r),h(r)]\}}\,\mathrm{d}r\right]. \end{aligned}$$

In the first equality above, $v(t,x) = 1$ by the assumption $h(t) > k(t)$, and $\tilde{v}(t,x) = 1$ by the definition of $\tilde{v}$ and (4.1). The second equality comes from optional sampling. In the final expression, the first term is non-negative by Claim 2, the second term (including the minus sign in front) is strictly positive by the assumption $h(t) > k(t)$ together with the continuity and the non-positivity of k and h. Hence we have obtained a contradiction, which proves the claim.

It follows from (4.3), (4.4), Claim 2, and Claim 3 that $v = \tilde{v}$. Since $v(t,x) > 1$ for $x > h(t)$, Claim 1 yields that $h \geq k$. In view of Claim 3, this finishes the proof. □

5 Parameter dependencies

5.1 Dependence of the value function on the market volatility

A large volatility σ makes the observation process noisy, slowing down the speed of learning about the drift. Since the fluctuations are trend-free, the intuition is that the agent should benefit from a smaller market volatility σ. While a full proof of this intuitive remark appears to be challenging, we have the following sufficient condition which guarantees monotonicity in σ.

Proposition 5.1. *Assume that the dispersion function ψ is such that $\sigma\psi(t,x)$ is non-increasing in σ. Then the value V in* (2.2) *is non-increasing in σ.*

Proof. If $\sigma\psi(t,x)$ is non-increasing in σ, then both the drift term and the diffusion term of $\hat{X}$ are non-increasing in σ. Therefore, Theorem 6.1 from [10] can be applied to prove that the value function v is decreasing in σ. □

Example Suppose X has a two-point prior distribution $\mu = (1-\pi)\delta_l + \pi\delta_h$, where $l < h$. Then $\sigma\psi(t,x) = (h-x)(x-l)$, so V is decreasing in σ.

Example Suppose the prior distribution of X is $N(m,\gamma^2)$. Then $\sigma\psi(t,x) = \frac{\sigma^2\gamma^2}{\sigma^2+\gamma^2 t}$, which is *increasing* in σ. Thus Theorem 5.1 does not apply.

The difficulty in proving the intuitive conjecture that the initial value V in (2.2) should be decreasing in the volatility σ lies in the fact that it is not true in general that the Markovian value function v in (3.10) is decreasing in σ. We can see this in the case of a normal prior in Figure 2. The picture depicts the difference between two Markovian value functions for the same normal prior with standard deviation $\gamma = 0.5$, but different volatilities σ. Nevertheless, the same picture shows that at time $t = 0$, the difference is positive, so conforming with our intuitive conjecture.

As far as the optimal stopping boundaries are concerned, the lack of monotonicity of the Markovian value function in the volatility σ manifests in that the stopping boundaries for different values of σ may intersect. An example of this appears in Figure 3. The same graph also provides intuition about how the shape of the boundary changes as one varies the parameter σ. In particular, we get an impression what boundary to expect as σ approaches zero or grows to infinity.

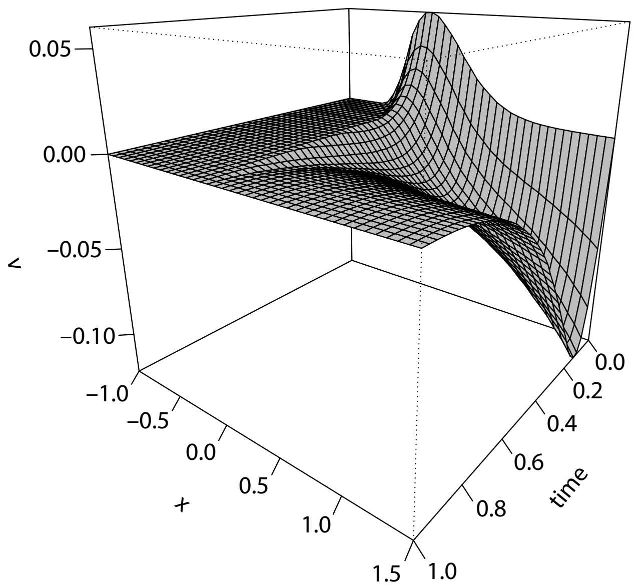

Figure 2: The difference $v_{0.2} - v_{0.5}$ between two value functions; $v_{0.2}$ and $v_{0.5}$ denote the value functions in the cases of the market volatility σ being equal to 0.2 and 0.5, respectively, and $T = 1$.

5.2 Dependence of the value function on the initial prior

Proposition 5.2. *Assume that μ_1 and μ_2 are two prior distributions such that the corresponding volatilities ψ_1 and ψ_2 satisfy $\psi_1(t,x) \leq \psi_2(t,x)$ for all $(t,x) \in [0,T] \times \mathbb{R}$. Then the corresponding Markovian value functions v_1 and v_2 satisfy $v_1 \leq v_2$.*

Proof. Again, Theorem 6.1 from [10] can be applied to prove that the value function v is increasing in ψ. □

In the case of the normal prior, the function $\psi(t,x) = \frac{\sigma\gamma^2}{\sigma^2 + t\gamma^2}$ is monotonically increasing in the standard deviation γ of the prior. Hence Theorem 5.2 applies and the Markovian value function v increases in γ. A consequence of this is that optimal stopping boundaries are ordered by the size of γ as shown in Figure 4.

For compactly supported distributions, Theorem 5.2 offers a way to construct an upper bound for the Markovian value function v. Suppose the prior μ is a compactly supported distribution. Since ψ_μ is bounded, by the two-point prior example on page 9, we can find a two-point distribution $\eta := (1-\pi)\delta_a + \pi\delta_b$ with $\int_{\mathbb{R}} u\mu(\,\mathrm{d}u) = \int_{\mathbb{R}} u\eta(\,\mathrm{d}u)$ such that $\psi_\mu \leq \psi_\eta$. Then Theorem 5.2 yields that $v_\mu \leq v_\eta$ and so the stopping boundaries satisfy $h_\eta \leq h_\mu$. As a result, $V_\mu \leq \mathbb{E}_\eta[S_{\tau_{h_\eta}}]$, where V_μ denotes the initial value for the prior μ, $\mathbb{E}_\eta$ denotes the expectation operator under which the prior is η instead of μ.

5.3 Effects of drift uncertainty and filtering

Having introduced and solved the optimal liquidation problem for an arbitrary prior, two pragmatic questions arise. First, how much does the drift uncertainty influence the optimal liquidation value? Second, how much is there to be gained from the elaborate liquidation

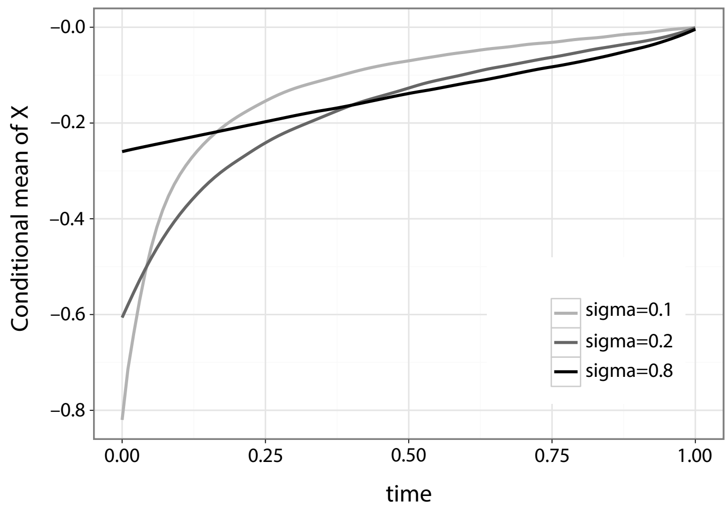

Figure 3: Optimal stopping boundaries for different values of market volatility σ in the case a normal prior with standard deviation $\gamma = 0.5$ and $T = 1$.

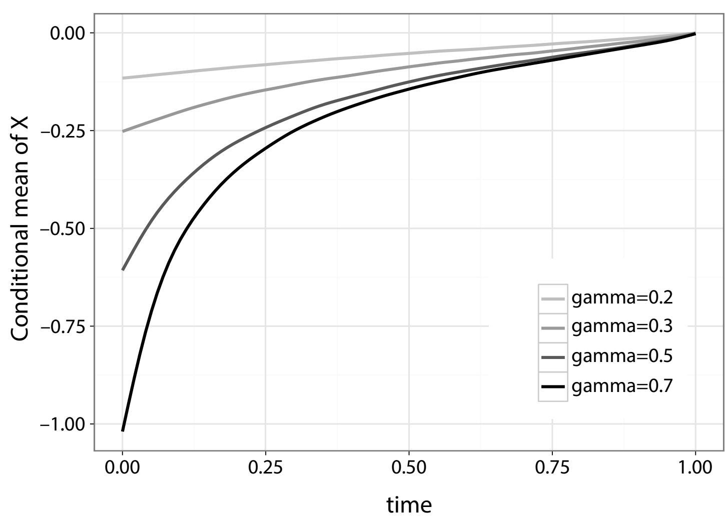

Figure 4: Optimal stopping boundaries for different values of standard deviation γ in the case of a normal prior when the market volatility $\sigma = 0.2$ and $T = 1$.

strategy with real-time filtering in comparison to a naive optimal selling strategy without filtering? In this part, we look at these questions within our model, providing numerical illustrations in the normal prior case.

We suppose the asset price evolves according to (2.1) for some random variable X and consider three different scenarios under which an agent wants to liquidate the asset before time T. In the first scenario, the asset has a known drift m (i.e. $X = m$), and the expected gain from optimal liquidation is then

$$V_m := e^{mT} \vee 1 = e^{m^+ T}.$$

In the second scenario, the agent recognises an uncertainty about the drift parameter X and uses a probability distribution μ with mean m to quantify the uncertainty. Moreover, he is unaware of the possibility to apply real-time filtering. Hence the agent does not utilise any information from the asset price observations and so makes a decision to sell immediately (i.e. at time 0) or at time T. The expected income from an optimal liquidation strategy in this scenario is

$$V_{\{0,T\}} := \mathbb{E}[e^{XT}] \vee 1\,.$$

In the third scenario, the agent knows the optimal liquidation procedure with filtering, hence the expected gain of the asset liquidation is then

$$V = \sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}[S_\tau]$$

(this is the case considered thus far in the article). Notice that

$$V_m \leq V_{\{0,T\}} \leq V \tag{5.1}$$

holds for any distribution μ, where the first inequality comes from Jensen's inequality. Verbally, (5.1) says that the uncertainty about the drift always increases the optimal liquidation value, which is highest if filtering is also employed.

For a concrete example, suppose that the prior μ is the normal distribution $N(m, \gamma^2)$ with $m \in \mathbb{R}$, $\gamma > 0$. Figure 5 depicts the values $V_{\{0,T\}}$ and V calculated for two different values of $m < 0$ at a range of different market volatilities; also, note that $V_m \equiv 1$ in both cases. We see that even if the mean drift is negative, the uncertainty about the drift and filtering each additionally increase the expected income from optimal selling. Moreover, Figure 6 depicts the percentage improvement $(V - V_{\{0,T\}})/V_{\{0,T\}}$ that the liquidation strategy with filtering brings over a more naive selling strategy. Lastly, the monotone dependence of the two values on the standard deviation γ quantifying the drift uncertainty is illustrated in Figure 7.

Remark According to Figure 6, the lower the volatility σ, the greater the gain from using the filtering procedure. Thus for an agent interested in optimally liquidating only the idiosyncratic (i.e. asset-specific) component of a stock price, the filtering procedure is particularly beneficial due to smaller volatility. To see this, a simple structural Black-Scholes model including an idiosyncratic and a market factor (e.g. see [5]) suggests that the idiosyncratic price component

$$S_t/I_t \propto e^{(X' - \sigma'^2/2)t + \sigma' W'_t}.$$

Here I_t is a large-basket index representing the market factor, X', σ', W' denote the idiosyncratic drift, the idiosyncratic volatility, and the idiosyncratic Brownian driver, respectively; the idiosyncratic volatility satisfies $\sigma' < \sigma$.

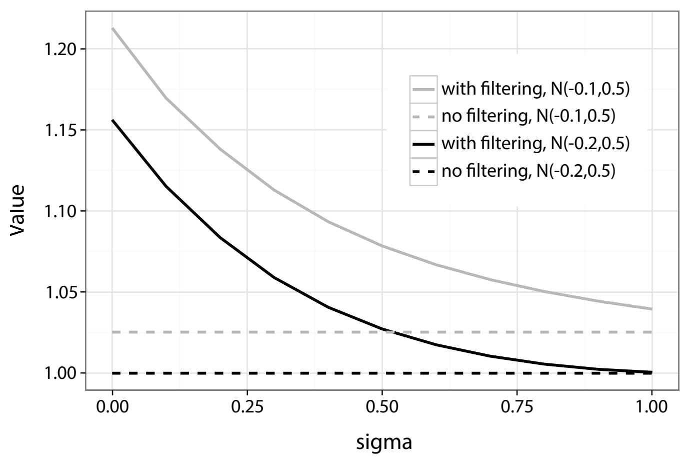

Figure 5: **Initial value as a function of market volatility.** The solid gray curve corresponds to the optimal liquidation value and the dashed gray line to the value without filtering - both for the normal prior $N(-0.1, 0.5)$ and $T = 1$. Similarly, the solid black curve corresponds to the optimal liquidation value and the dashed black line to the value without filtering for the normal prior $N(-0.2, 0.5)$.

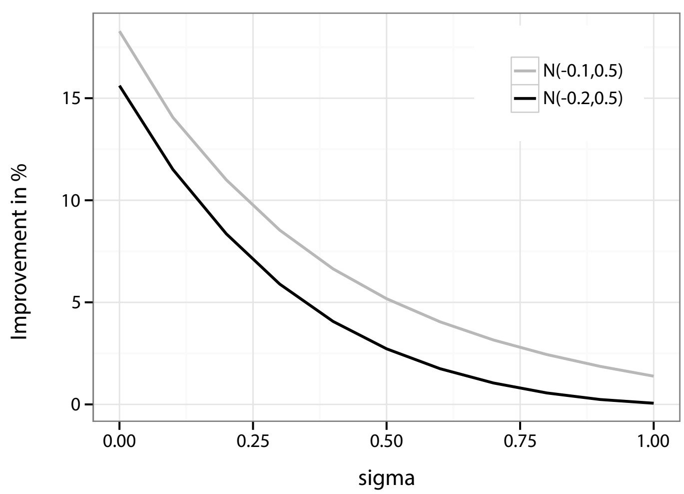

Figure 6: **Improvement due to filtering**. The percentage improvement $(V - V_{\{0,T\}})/V_{\{0,T\}}$ over the strategy without filtering.

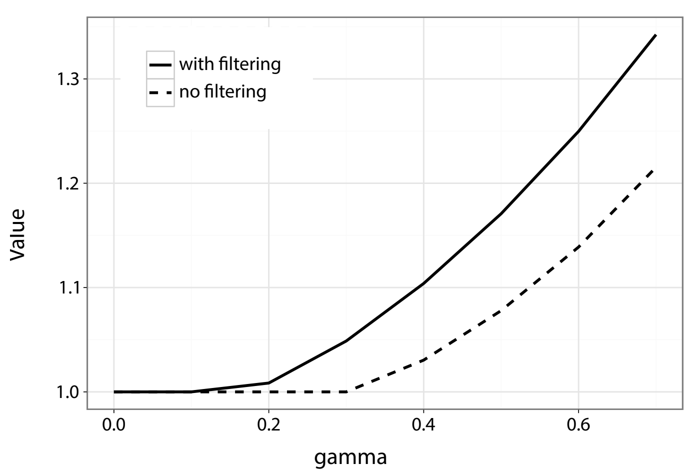

Figure 7: **Initial value as a function of standard deviation of the normal prior.** The solid curve corresponds to the optimal liquidation value while the dashed line to the value without filtering. Here the prior is normal with mean -0.05, the market volatility $\sigma = 0.2$.

References

[1] Bain, A., Crisan, D. *Fundamentals of stochastic filtering.* Stochastic Modelling and Applied Probability, 60. Springer, New York, 2009.

[2] Back, K. Incomplete and asymmetric information in asset pricing theory. Stochastic Methods in Finance, Lecture Notes in Mathematics, vol. 1856, 2004, 1-25.

[3] Beibel, M., Lerche, H.R. A New Look at Optimal Stopping Problems related to Mathematical Finance. Statistica Sinica, vol. 7, no. 1, 1997, 93-108.

[4] Björk, T., Davis, M. H. A., Landén, C. Optimal investment under partial information. Mathematical Methods of Operations Research, vol. 71, no. 2, 2010, 371-399.

[5] Bush, N., Hambly, B.M., Haworth, H., Jin, L., Reisinger, C. Stochastic evolution equations in portfolio credit modelling. SIAM Journal on Financial Mathematics, vol. 2, no. 1, 2011, 627-664.

[6] Décamps, J.-P., Mariotti, T., Villeneuve, S. Investment timing under incomplete information. Math. Oper. Res., vol. 30, no. 2, 2005, 472-500.

[7] Dothan, M., Feldman, D. Equilibrium interest rates and multiperiod bonds in a partially observable economy. Journal of Finance, vol. 41, no. 2, 1986, 369-382.

[8] du Toit, J., Peskir, G. Selling a stock at the ultimate maximum. Annals of Applied Probability, vol. 19, no. 3, 2009, 983-1014.

[9] Ekström, E., Lu, B. Optimal selling of an asset under incomplete information. International Journal of Stochastic Analysis, vol. 2011, 2011, ID 543590.

[10] Ekström, E., Tysk, J. Convexity theory for the term structure equation. Finance and Stochastics. vol. 12, no. 1, 2008, 117-147.

[11] Evans, J., Henderson, V., Hobson, D. Optimal timing for an indivisible asset sale. Mathematical Finance, vol. 18, no. 4, 2008, 545-567.

[12] Gapeev, P. Pricing of perpetual American options in a model with partial information. International Journal of Theoretical and Applied Finance, vol. 15, no. 1, 2012, ID 1250010.

[13] Henderson, V., Hobson, D. An explicit solution for an optimal stopping/optimal control problem which models an asset sale. Annals of Applied Probability, vol. 18, no. 5, 2008, 1681-1705.

[14] Jacka, S. Optimal stopping and the American put. Mathematical Finance 1 (1991), no. 2, 1-14.

[15] Karatzas, I., Shreve, S. *Brownian motion and stochastic calculus. Second edition.* Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.

[16] Karatzas, I., Shreve, S. *Methods of Mathematical Finance.* Applications of Mathematics, 39. Springer-Verlag, New York, 1998.

[17] Lu, B. Optimal selling of an asset with jumps under incomplete information. Applied Mathematical Finance, vol. 20, no. 6, 2013, 599610.

[18] Lakner, P. Utility maximization with partial information. Stochastic Processes and their Applications, vol. 56, no. 2, 1995, 247-249.

[19] Lakner, P. Optimal trading strategy for an investor: the case of partial information. Stochastic Processes and their Applications, vol. 76, no. 1, 1998, 77-97.

[20] Monoyios, M. Optimal investment and hedging under partial and inside information. Advanced Financial Modelling, Radon Series on Computational and Applied Mathematics, vol. 8, 2009, 371-410.

[21] Monoyios, M., Ng, A. Optimal exercise of an executive stock option by an insider. International Journal of Theoretical and Applied Finance, vol. 14, no. 1, 2011, 83-106.

[22] Øksendal, B. *Stochastic Differential Equations: an introduction with applications. Sixth edition.* Springer, 2007.

[23] Peskir, G. On the American option problem. Mathematical Finance 15 (2005), no. 1, 169-181.

[24] Peskir, G. A change-of-variable formula with local time on curves. Journal of Theoretical Probability, vol. 18, no. 3, 2005, 499-535.

[25] Peskir, G., Shiryaev, A. *Optimal stopping and free-boundary problems.* Lectures in Mathematics, ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[26] Protter, P. E. *Stochastic Integration and Differential Equations. Version 2.1.* Stochastic Modelling and Applied Probability, 21. Springer, 2005.

[27] Revuz, D., Yor, M. *Continuous martingales and Brownian motion. Third edition.* Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, 1999.

[28] Rogers, L.C.G. *Optimal Investment.* Springer Briefs in Quantitative Finance. Springer, 2013.

[29] Ziegler, A. *Incomplete Information and Heterogeneous Beliefs in Continuous- Time Finance.* Springer, 2003.