

# Preface

Many real-world processes exhibit an inherent trend visible at a macroscopic level. A river flow has a direction and speed, various financial and economic indicators are noted to have near constant growth or decline rates when sufficiently long time series are taken. However, at a microscopic level, the macroscopic trends tend to be overshadowed by random fluctuations, be it a Brownian bombardment of a pollen grain by nearby water molecules, or intraday volatility in a stock market. How to make optimal decisions that highly depend on the macroscopic trend when one is only observing a development of the process at a microscopic scale is the main theme of this dissertation.

The thesis consists of two essays:

1. ‘Bayesian sequential testing of the drift of a Brownian motion’,
2. ‘Optimal liquidation of an asset under drift uncertainty’;

both written together with my PhD supervisor Prof. Erik Ekström.

In the first article, an optimal sequential statistical procedure for determining the sign of the drift of a Brownian motion is investigated in the Bayesian framework under a general prior distribution. The second paper studies a financial problem that concerns optimal timing for the sale of an indivisible asset under arbitrary prior beliefs about the drift of the price process.



## **Acknowledgements**

I want to express my wholehearted gratitude to my PhD supervisor Prof. Erik Ekström for his support and encouragement throughout this work, including the guidance received, invaluable comments on my arguments, not to mention innumerable expansive discussions that have shaped the ideas presented in this thesis.



# Bayesian sequential testing of the drift of a Brownian motion

Erik Ekström \*      Juozas Vaicenavicius

## Abstract

We study a classical Bayesian statistics problem of sequentially testing the sign of the drift of an arithmetic Brownian motion with the ‘0 – 1’ loss function and a constant cost of observation per unit of time for general prior distributions. The statistical problem is reformulated as an optimal stopping problem with the current conditional probability that the drift is non-negative as the underlying process. The volatility of this conditional probability process is shown to be non-increasing in time, which enables us to prove monotonicity and continuity of the optimal stopping boundaries as well as to characterize them completely in the finite-horizon case as the unique continuous solution to a pair of integral equations. In the infinite-horizon case, the boundaries are shown to solve another pair of integral equations and a convergent approximation scheme for the boundaries is provided. Also, we describe the dependence between the prior distribution and the long-term asymptotic behaviour of the boundaries.

## 1 Introduction

One of the classical questions in Sequential Analysis concerns the testing of two simple hypotheses about the sign of the drift of an arithmetic Brownian motion. More precisely, suppose that an observed process  $X_t$  is an arithmetic Brownian motion

$$X_t = Bt + W_t,$$

where the constant  $B$  is unknown and  $W$  is a standard driftless Brownian motion. Based on observations of the process  $X$ , one wants to test sequentially the hypotheses  $H_0 : B < 0$  and  $H_1 : B \geq 0$ . In the Bayesian formulation of this sequential testing problem, the drift  $B$  is a random variable with distribution  $\mu$ , corresponding to the hypothesis tester’s prior belief about the likeliness of the different values  $B$  may take. Moreover, it is assumed that  $B$  and  $W$  are independent. In this article, we consider a classical formulation of the problem in which the accuracy and urgency of a decision is governed by a ‘0 – 1’ loss function together with a constant cost  $c > 0$  of observation per unit time. The ‘0 – 1’ loss function means that the tester gains nothing for a right decision but pays a penalty of size 1 for being wrong. The overall goal is to find a decision rule minimising the expected total

---

\*Support from the Swedish Research Council (VR) is gratefully acknowledged.

cost (provided such a decision rule exists). If the decision is required to be made before a fixed predetermined time  $T > 0$ , the problem is said to have a finite horizon, and if there is no upper bound on the decision time, an infinite horizon.

In the classical literature [5] by Chernoff and [2] by Bather on Bayesian sequential testing procedures for the sign of a drift, the special case of a normal prior distribution is studied. While Bather considers the ‘0 – 1’ loss function described above as well as a few other alternatives, Chernoff deals with a different penalty function, which equals the magnitude  $|B|$  of the error. In these papers, it is argued that the sequential analysis problem reduces to a free-boundary problem for a function of time and the current value of the observation process, but, as in the case of most time-dependent free-boundary problems, the free-boundary problem lacks an explicit solution. Instead, the focus of these and many follow-up articles in the area, including [3], [4], [6], [7], and [14] to mention a few, is on asymptotic approximations for optimal stopping boundaries (for more references, see the survey article [15]). Only recently, [21] has characterised the optimal stopping boundaries for the original Chernoff’s problem in terms of an integral equation, which can be solved numerically.

In [20], the sequential testing problem is solved explicitly for a two-point prior distribution by utilising the connection with a time-homogeneous free-boundary problem. Notably, the natural spatial variable in this free-boundary problem is not the value of the observation process, but the conditional probability of the drift taking one of the two possible values. (Since there is a one-to-one correspondence between these two processes at each fixed time, the free-boundary problem could be transformed into one based on the observation process instead, but that formulation would introduce time-dependencies and thus make the explicit solution more difficult to find.)

The fact that the problem can be solved in a very special case of a two-point prior, raises a natural question – can the sequential testing problem be solved for a more general prior distribution? In this article, we investigate the sequential testing problem under a general prior distribution. Since this introduces time-dependencies in the problem, there is generally no hope for explicit solutions. Nevertheless, additional structure is found, which enables us to arrive at a fairly satisfactory answer.

To explain in some further detail, following standard arguments, the statistical problem is shown to admit an equivalent formulation as an optimal stopping problem, which we study to characterise optimal decision rules. The underlying process of the optimal stopping problem is chosen to be the current probability, conditional on observations of  $X$ , that the drift is non-negative, i.e.

$$\Pi_t := \mathbb{P}(B \geq 0 | \mathcal{F}_t^X).$$

The pay-off function of the associated optimal stopping problem is then concave in  $\Pi$ , so general results about preservation of concavity for optimal stopping problems may be employed to derive structural properties of the continuation region. Moreover, the volatility of the underlying process  $\Pi$  can be shown to be decreasing in time (except for the two-point distribution discussed above, in which case it is constant). These important facts allow us

to show that the optimal stopping boundaries are monotone, so techniques from the theory of free-boundary problems with monotone boundaries can be applied. In particular, the monotonicity of the boundaries enables us to prove the smooth-fit condition and the continuity of the boundaries, as well as to study the corresponding integral equations. In the finite-horizon case, we characterise the optimal boundaries as a unique continuous solution to a pair of integral equations. In the infinite-horizon case, the situation turns out to be more subtle. The boundaries are shown to solve another pair of integral equations, but whether the system admits a unique solution remains unanswered. Instead, we provide a converging approximation scheme for the optimal stopping boundaries, establishing that the optimal boundaries of the finite-horizon problem converge pointwise to the optimal boundaries of the infinite-horizon problem. Also, we determine the long-term asymptotes of the boundaries and describe their dependence on the prior distribution.

From a technical perspective, we tackle a number of issues stemming from the infinite-dimensionality of the parameter space of the underlying process  $\Pi$ , the particular form of the unbounded payoff function, as well as the presence of time-dependent infinite-horizon boundaries. Filtering and analytic techniques are used to understand the behaviour of the conditional probability  $\Pi$ , with a particular focus on the properties that are invariant under any prior distribution. Also, the generality of the prior makes the verification of the smooth-fit condition more involved than in standard situations. Moreover, the specific form of the payoff function with the additive unbounded time term requires some additional effort to prove optimality of the hitting time in the infinite-horizon case. Our approach to approximate the optimal infinite-horizon boundaries could possibly be utilised in other similar situations.

The paper is organised as follows. In Section 2, the sequential testing problem is formulated and reduced to an optimal stopping problem. In Section 3, filtering techniques are applied to find an expression for  $\Pi$  in terms of the observation process  $X$ , and its dynamics in terms of the innovation process are determined. We also study the volatility function of  $\Pi$ , and it is shown that this function is non-increasing in time. In Section 4, the optimal stopping problem is studied together with the corresponding free-boundary problem, and it is shown that the optimal stopping boundaries are continuous. In Section 5, integral equations for the boundaries are determined, and uniqueness of solutions is established in the finite-horizon case. The long-term asymptotic behaviour of the infinite-horizon boundaries is presented in Section 6. Finally, Section 7 is devoted to a special case of the normal prior distribution.

**Acknowledgement.** We are most grateful to Ioannis Karatzas for his suggestion to generalise the setting of an early version of the current paper, and for sharing his unpublished notes on a related problem with us.

## 2 Problem formulation and reduction to an optimal stopping problem

Let  $(\Omega, \mathbb{P}, \mathcal{F})$  be a complete probability space supporting a Brownian motion  $W$  and a random variable  $B$  with distribution  $\mu$  such that  $W$  and  $B$  are independent. Define

$$X_t = Bt + W_t.$$

Writing  $\mathbb{F}^X = \{\mathcal{F}_t^X\}_{t \geq 0}$  for the filtration generated by the process  $X$  and the null sets in  $\mathcal{F}$ , our goal is to find a pair  $(\tau, d)$  consisting of an  $\mathbb{F}^X$ -stopping time  $\tau$  and an  $\mathcal{F}_\tau^X$ -measurable decision rule  $d : \Omega \rightarrow \{0, 1\}$ , indicating which of the hypotheses  $H_0 : B < 0$  or  $H_1 : B \geq 0$  to accept, in order to minimise the Bayes risk

$$R(\tau, d) := \mathbb{E}[\mathbb{1}_{\{d=1, B < 0\}}] + \mathbb{E}[\mathbb{1}_{\{d=0, B \geq 0\}}] + c\mathbb{E}[\tau].$$

Since  $d$  is  $\mathcal{F}_\tau^X$ -measurable, we have

$$R(\tau, d) = \mathbb{E} \left[ \mathbb{E} [\mathbb{1}_{\{B < 0\}} | \mathcal{F}_\tau^X] \mathbb{1}_{\{d=1\}} + \mathbb{E} [\mathbb{1}_{\{B \geq 0\}} | \mathcal{F}_\tau^X] \mathbb{1}_{\{d=0\}} + c\tau \right], \quad (2.1)$$

which shows that, at a given stopping time  $\tau$ , the decision rule

$$d = \begin{cases} 1 & \text{if } \mathbb{P}(B \geq 0 | \mathcal{F}_\tau^X) \geq \mathbb{P}(B < 0 | \mathcal{F}_\tau^X), \\ 0 & \text{otherwise} \end{cases}$$

is optimal. Consequently, writing

$$\Pi_t := \mathbb{P}(B \geq 0 | \mathcal{F}_t^X),$$

the sequential testing problem (2.1) reduces to an optimal stopping problem

$$V = \inf_{\tau \in \mathcal{T}} \mathbb{E}[g(\Pi_\tau) + c\tau], \quad (2.2)$$

where  $g(\pi) = \pi \wedge (1 - \pi)$  and  $\mathcal{T}$  denotes the set of  $\mathbb{F}^X$ -stopping times. We also consider the same sequential testing problem but with a finite horizon  $T < \infty$ . The corresponding optimal stopping problem is then

$$V^T = \inf_{\tau \in \mathcal{T}^T} \mathbb{E}[g(\Pi_\tau) + c\tau],$$

where  $\mathcal{T}^T = \{\tau \in \mathcal{T} : \tau \leq T\}$ .

**Remark** By translation, our study readily extends to testing the hypotheses  $H_0 : B < \theta$  and  $H_1 : B \geq \theta$  for any given  $\theta \in \mathbb{R}$ . The methods also extend to the case when the two types of possible errors are associated with different costs, i.e. when

$$R(\tau, d) = a\mathbb{E}[\mathbb{1}_{\{d=1, B < 0\}}] + b\mathbb{E}[\mathbb{1}_{\{d=0, B \geq 0\}}] + c\mathbb{E}[\tau]$$

for constants  $a > 0$  and  $b > 0$  with  $a \neq b$ . For simplicity of the presentation, however, we assume throughout the article that  $\theta = 0$  and  $a = b = 1$ .



Note that in the cases when  $\mu((-\infty, 0)) = 0$  or  $\mu([0, \infty)) = 0$ , the sequential testing problem becomes trivial as we can make the correct statement about the sign of the drift at time zero. Hence, from now onwards, we always assume that

$$0 < \mu([0, \infty)) < 1. \quad (2.3)$$

### 3 Conditional probability of non-negative drift

In this section, we derive a filtering equation for the distribution of  $B$  conditional on the observations of  $X$ , which is then applied to prove some elementary results concerning the conditional distribution of the sign of  $B$ . We also show that there is an explicit one-to-one correspondence between  $\Pi$  and the observation process  $X$  at each fixed time, and we determine the dynamics of  $X$  and  $\Pi$  in terms of the innovation process.

#### 3.1 Filtering of the unknown drift

**Proposition 3.1.** *Assume that  $q : \mathbb{R} \rightarrow \mathbb{R}$  satisfies  $\int_{\mathbb{R}} |q(b)|\mu(db) < \infty$ . Then*

$$\mathbb{E} [q(B)|\mathcal{F}_t^X] = \frac{\int_{\mathbb{R}} q(b)e^{bX_t - b^2t/2}\mu(db)}{\int_{\mathbb{R}} e^{bX_t - b^2t/2}\mu(db)} \quad (3.1)$$

for any  $t > 0$ .

*Proof.* The proof is based on standard methods in filtering theory, see e.g. [1, Section 3.3], yet we include it for completeness. First define an enlarged filtration  $\mathbb{G} = \{\mathcal{G}_t\}_{0 \leq t < \infty}$  as the completion of  $\{\sigma(B, W_s : 0 \leq s \leq t)\}_{0 \leq t < \infty}$ . Clearly,  $\mathcal{F}_t^X \subseteq \mathcal{G}_t$  for any  $t \geq 0$ , so  $\mathbb{G}$  is an enlargement of  $\mathbb{F}^X$ . Observing that  $Z_t := e^{-BW_t - B^2t/2}$  is a  $\mathbb{G}$ -martingale, we define a new probability measure  $\mathbb{P}_*$  on the restriction  $(\Omega, \mathcal{G}_T)$  for some large enough  $T$  by

$$\frac{d\mathbb{P}_*}{d\mathbb{P}}|_{\mathcal{G}_T} := Z_T.$$

It can be shown that under  $\mathbb{P}_*$ ,  $X_t$  is a Brownian motion independent of  $\mathcal{G}_0$  and therefore also of  $B$  and that the law of  $B$  is  $\mu$  (see [1, Proposition 3.13]). Thus Bayes' rule (cf., for example, [16]) gives that

$$\mathbb{E} [q(B)|\mathcal{F}_t^X] = \frac{\mathbb{E}_* [q(B)/Z_t|\mathcal{F}_t^X]}{\mathbb{E}_* [1/Z_t|\mathcal{F}_t^X]} = \frac{\int_{\mathbb{R}} q(b)e^{bX_t - b^2t/2}\mu(db)}{\int_{\mathbb{R}} e^{bX_t - b^2t/2}\mu(db)}$$

for  $t > 0$  and for any function  $q : \mathbb{R} \rightarrow \mathbb{R}$  with  $\int_{\mathbb{R}} |q(b)|\mu(db) < \infty$ . □

#### 3.2 Conditional probability of non-negative drift

According to Proposition 3.1,

$$\Pi_t = \mathbb{E} [\mathbb{1}_{[0, \infty)}(B)|\mathcal{F}_t^X] = \pi(t, X_t),$$

where the function  $\pi(t, x) : (0, \infty) \times \mathbb{R} \rightarrow (0, 1)$  is given by

$$\pi(t, x) := \frac{\int_{[0, \infty)} e^{bx - b^2 t/2} \mu(db)}{\int_{\mathbb{R}} e^{bx - b^2 t/2} \mu(db)}. \quad (3.2)$$

Denoting by

$$\mu_{t,x}(db) := \frac{e^{bx - \frac{b^2}{2}t} \mu(db)}{\int_{\mathbb{R}} e^{bx - \frac{b^2}{2}t} \mu(db)} \quad (3.3)$$

the distribution of  $B$  at time  $t$  conditional on  $X_t = x$ , we thus have

$$\pi(t, x) = \int_{[0, \infty)} \mu_{t,x}(db).$$

**Proposition 3.2.** *Assume that  $q : \mathbb{R} \rightarrow \mathbb{R}$  is non-decreasing and satisfies  $\int_{\mathbb{R}} |q(b)| \mu(db) < \infty$ . Then the function*

$$u(t, x) := \frac{\int_{\mathbb{R}} q(b) e^{bx - b^2 t/2} \mu(db)}{\int_{\mathbb{R}} e^{bx - b^2 t/2} \mu(db)} = \int_{\mathbb{R}} q(b) \mu_{t,x}(db) \quad (3.4)$$

is non-decreasing in  $x$  for any fixed  $t > 0$ .

*Proof.* We will prove the claim by showing that  $u(t, \cdot)$  is differentiable with a non-negative derivative on  $\mathbb{R}$ . By a standard differentiation lemma (see, for example, [13, Theorem 6.28]), both the numerator and the denominator in (3.4) are differentiable with respect to  $x$  with their derivatives obtained by differentiating under the integral sign. Thus the derivative of  $u$  with respect to the second argument  $x$  is given by

$$\begin{aligned} \partial_2 u(t, x) &= \int_{\mathbb{R}} q(b) b \mu_{t,x}(db) - \int_{\mathbb{R}} q(b) \mu_{t,x}(db) \int_{\mathbb{R}} b \mu_{t,x}(db) \\ &= \mathbb{E}_{t,x} [q(B)B] - \mathbb{E}_{t,x} [q(B)] \mathbb{E}_{t,x} [B], \end{aligned} \quad (3.5)$$

where  $\mathbb{E}_{t,x}$  is the expectation operator under the probability measure  $\mathbb{P}_{t,x}(\cdot) := \mathbb{P}(\cdot | X_t = x)$ . Since

$$\begin{aligned} &\mathbb{E}_{t,x} [q(B)B] - \mathbb{E}_{t,x} [q(B)] \mathbb{E}_{t,x} [B] \\ &= \mathbb{E}_{t,x} [(B - \mathbb{E}_{t,x}[B])(q(B) - q(\mathbb{E}_{t,x}[B]))] \geq 0, \end{aligned} \quad (3.6)$$

this finishes the proof.  $\square$

**Corollary 3.3.** *Let  $a \in \mathbb{R}$  and  $t > 0$ . Then*

1.  $\mathbb{P}(B > a | X_t = x)$  is non-decreasing in  $x$ ,
2.  $\mathbb{P}(B < a | X_t = x)$  is non-increasing in  $x$ .

*Proof.* The first claim follows by applying Proposition 3.2 to the function  $q(b) = 1_{(a, \infty)}(b)$ . The second claim follows from  $\mathbb{P}(B < a | X_t = x) = 1 - \mathbb{P}(B \geq a | X_t = x)$  and by applying Proposition 3.2 to the function  $q(b) = 1_{[a, \infty)}(b)$ .  $\square$

**Proposition 3.4.** *For any given  $t > 0$ , the function  $\pi(t, \cdot) : \mathbb{R} \rightarrow (0, 1)$  defined in (3.2) is a strictly increasing continuous bijection.*

*Proof.* First note that

$$(B - \mathbb{E}_{t,x}[B])(\mathbb{1}_{[0, \infty)}(B) - \mathbb{1}_{[0, \infty)}(\mathbb{E}_{t,x}[B])) \geq 0$$

and that

$$\mathbb{P}_{t,x}((B - \mathbb{E}_{t,x}[B])(\mathbb{1}_{[0, \infty)}(B) - \mathbb{1}_{[0, \infty)}(\mathbb{E}_{t,x}[B])) > 0) > 0$$

since (2.3) implies  $\mu_{t,x}((-\infty, 0)) > 0$  and  $\mu_{t,x}([0, \infty)) > 0$ . Consequently, the inequality in (3.6) is strict, so  $x \mapsto \pi(t, x)$  is strictly increasing.

Next, note that

$$\pi(t, x) = \frac{1}{1 + A(t, x)},$$

where

$$A(t, x) = \frac{\int_{(-\infty, 0)} e^{bx - \frac{b^2}{2}t} \mu(db)}{\int_{[0, \infty)} e^{bx - \frac{b^2}{2}t} \mu(db)}. \quad (3.7)$$

By monotone convergence, we find that  $A(t, x) \rightarrow 0$  as  $x \rightarrow \infty$  and  $A(t, x) \rightarrow \infty$  as  $x \rightarrow -\infty$ . Consequently,  $\lim_{x \rightarrow \infty} \pi(t, x) = 1$  and  $\lim_{x \rightarrow -\infty} \pi(t, x) = 0$ , which finishes the proof.  $\square$

An immediate consequence of Proposition 3.4 is that for any fixed  $t > 0$ , the spatial inverse  $\pi(t, \cdot)^{-1} : (0, 1) \rightarrow \mathbb{R}$  exists. To facilitate intuition, we denote the inverse by  $x(t, \cdot)$ .

We end this subsection with a result that describes the long-term behaviour of the process  $\Pi$ .

**Proposition 3.5.**  *$\Pi_t \rightarrow \Pi_\infty$  a.s. as  $t \rightarrow \infty$ , where  $\Pi_\infty$  is a Bernoulli random variable with  $\mathbb{P}(\Pi_\infty = 0) = \mu((-\infty, 0))$  and  $\mathbb{P}(\Pi_\infty = 1) = \mu([0, \infty))$ .*

*Proof.* Firstly, since  $\Pi_t = \mathbb{E}[\mathbb{1}_{[0, \infty)}(B) | \mathcal{F}_t^X]$  is a bounded martingale, by the martingale convergence theorem, the pointwise limit  $\Pi_\infty := \lim_{t \rightarrow \infty} \Pi_t$  is a well-defined random variable closing the martingale  $\Pi$ . By the law of large numbers for Brownian motion and Proposition 6.1 below (the proof of which is independent of the current result), for any  $b < 0$  in the support of  $\mu$  we have

$$\mathbb{P}(\Pi_\infty = 0 | B = b) = 1,$$

so

$$\mathbb{P}(\Pi_\infty = 0) = \int_{(-\infty, 0)} \mathbb{P}(\Pi_\infty = 0 | B = b) \mu(db) = \mu((-\infty, 0)).$$

Hence as  $\Pi_\infty$  can only take values in  $[0, 1]$ , the fact that  $\mathbb{E}[\Pi_\infty] = \mathbb{E}[\Pi_0] = \mu([0, \infty))$  implies  $\mathbb{P}(\Pi_\infty = 1) = \mu([0, \infty))$ .  $\square$

### 3.3 SDE for the conditional probability of non-negative drift

Assuming that  $B$  has a first moment, the conditional expectation of  $B$  exists and is given by

$$\mathbb{E}[B|\mathcal{F}_t^X] = \frac{\int_{\mathbb{R}} be^{bX_t - b^2t/2} \mu(db)}{\int_{\mathbb{R}} e^{bX_t - b^2t/2} \mu(db)}, \quad (3.8)$$

compare (3.1). Moreover, the observation process  $X$  is represented in terms of the innovation process

$$\hat{W}_t := X_t - \int_0^t \mathbb{E}[B|\mathcal{F}_s^X] ds$$

as

$$dX_t = \mathbb{E}[B|\mathcal{F}_t^X] dt + d\hat{W}_t.$$

Here  $\hat{W}$  is a standard  $\mathbb{F}^X$ -Brownian motion (see [1, Proposition 2.30 on p. 33]). Moreover, 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 have  $\mathbb{F}^X = \mathbb{F}^{\hat{W}}$  (see the remark on p. 35 in [1]).

From now onwards, the following integrability condition on  $\mu$  will be imposed throughout the article.

**Assumption.**  $\int_{\mathbb{R}} e^{\epsilon b^2} \mu(db) < \infty$  for some  $\epsilon > 0$ . (3.9)

Note that this assumption is a minor restriction on our hypothesis testing problem since, given any probability distribution  $\mu$ , the distributions  $\mu_{t,x}$  all satisfy (3.9) for  $t > 0$ . In other words, no matter what prior distribution  $\mu$  one starts with, the condition (3.9) will be satisfied after any infinitesimal time of observation. Also, note that the assumption allows us to extend the definition of  $\mu_{t,x}$  in (3.3) to  $t = 0$ . Moreover, if we have a prior distribution  $\xi$  on  $\mathbb{R}$  given by

$$\xi(db) := \frac{e^{\epsilon b^2} \mu(db)}{\int_{\mathbb{R}} e^{\epsilon b^2} \mu(db)},$$

then

$$\mu_{0,x}(db) = \xi_{2\epsilon,x}(db) := \frac{e^{bx - b^2(2\epsilon)/2} \xi(db)}{\int_{\mathbb{R}} e^{bx - b^2(2\epsilon)/2} \xi(db)}. \quad (3.10)$$

Consequently, the distribution  $\mu_{0,x}$  can be identified with a conditional distribution at time 0 given that the prior distribution at time  $-2\epsilon$  was  $\xi$  and the current value of the observation process is  $x$ . This gives us a generalisation of the notion of the starting point of the observation process  $X$  to allow  $X_0 = x \neq 0$ , and we may regard time 0 as an interior point of the time interval.

A closer look at the condition (3.9) and the expression (3.2) assures that the standard differentiability lemma can be applied to differentiate  $\pi(t, x)$  with respect to both variables

multiple times inside  $(-2\epsilon, \infty) \times \mathbb{R}$ . Applying Ito's formula to  $\Pi_t = \pi(t, X_t)$ , we find that

$$\begin{aligned} d\Pi_t &= \left( \partial_1 \pi(t, X_t) + \mathbb{E}[B | \mathcal{F}_t^X] \partial_2 \pi(t, X_t) + \frac{1}{2} \partial_2^2 \pi(t, X_t) \right) dt + \partial_2 \pi(t, X_t) d\hat{W}_t \\ &= \partial_2 \pi(t, x(t, \Pi_t)) d\hat{W}_t, \end{aligned}$$

where the second equality is verified using the expression (3.8) and

$$\begin{aligned} \partial_1 \pi(t, x) &= - \int_{[0, \infty)} \frac{b^2}{2} \mu_{t,x}(db) + \int_{[0, \infty)} \mu_{t,x}(db) \int_{\mathbb{R}} \frac{b^2}{2} \mu_{t,x}(db), \\ \partial_2 \pi(t, x) &= \int_{[0, \infty)} b \mu_{t,x}(db) - \int_{\mathbb{R}} b \mu_{t,x}(db) \int_{[0, \infty)} \mu_{t,x}(db), \end{aligned} \quad (3.11)$$

$$\begin{aligned} \partial_2^2 \pi(t, x) &= \int_{[0, \infty)} b^2 \mu_{t,x}(db) - 2 \int_{\mathbb{R}} b \mu_{t,x}(db) \int_{[0, \infty)} b \mu_{t,x}(db) \\ &\quad - \int_{\mathbb{R}} b^2 \mu_{t,x}(db) \int_{[0, \infty)} \mu_{t,x}(db) + 2 \int_{[0, \infty)} \mu_{t,x}(db) \left( \int_{\mathbb{R}} b \mu_{t,x}(db) \right)^2. \end{aligned}$$

Thus the dynamics of  $\Pi_t$  are specified by a zero drift and the volatility

$$\sigma(t, \Pi_t) = \partial_2 \pi(t, x(t, \Pi_t)), \quad (3.12)$$

being a positive function of the current time and the current value of  $\Pi$ . Using (3.5), the volatility function can be expressed as

$$\begin{aligned} \sigma(t, \pi) &= \mathbb{E}_{t,x(t,\pi)}[B \mathbb{1}_{\{B \geq 0\}}] - \mathbb{P}_{t,x(t,\pi)}(B \geq 0) \mathbb{E}_{t,x(t,\pi)}[B] \\ &= (1 - \pi) \mathbb{E}_{t,x(t,\pi)}[B \mathbb{1}_{\{B \geq 0\}}] - \pi \mathbb{E}_{t,x(t,\pi)}[B \mathbb{1}_{\{B < 0\}}]. \end{aligned} \quad (3.13)$$

**Example (The two-point distribution).** Assume that  $\mathbb{P}(B = a_1) = 1 - p$  and  $\mathbb{P}(B = a_2) = p$  for some constants  $a_1 < 0 \leq a_2$  and  $p \in (0, 1)$ . Then

$$\mathbb{P}(B = a_2 | X_t = x) = \pi(t, x) = \frac{p e^{a_2 x - a_2^2 t / 2}}{(1 - p) e^{a_1 x - a_1^2 t / 2} + p e^{a_2 x - a_2^2 t / 2}}$$

and

$$\sigma(t, \pi) = (a_2 - a_1) \pi (1 - \pi).$$

This example with a two-point prior distribution is a special case of the Wonham filter.

**Example (The normal distribution).** Assume that  $\mu$  is the normal distribution with mean  $m$  and variance  $\gamma^2$ . Then the conditional distribution  $\mathbb{P}(\cdot | X_t = x) = \mu_{t,x}$  is also normal but with mean  $\frac{m + \gamma^2 x}{1 + t \gamma^2}$  and variance  $\frac{\gamma^2}{1 + t \gamma^2}$ . Consequently,

$$\pi(t, x) = \Phi \left( \frac{m + \gamma^2 x}{\gamma \sqrt{1 + t \gamma^2}} \right) \quad (3.14)$$

and

$$\sigma(t, \pi) = \varphi(\Phi^{-1}(\pi)) \frac{\gamma}{\sqrt{1 + t\gamma^2}},$$

where

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

and

$$\Phi(y) = \int_{-\infty}^y \varphi(z) dz$$

are the density and the cumulative distribution of the standard normal random variable, respectively. Note that this instance of a normal prior distribution is a special case of the well-known Kalman-Bucy filter, see for example [16, Chapter 6].

### 3.4 Volatility of the conditional probability process

In this section we study the volatility function  $\sigma$ . The main result, Corollary 3.8, states that the volatility is non-increasing as a function of time.

Let  $\pi \in (0, 1)$  be a fixed number and consider the map  $x(\cdot, \pi) : [0, \infty) \rightarrow \mathbb{R}$  sending  $t \mapsto x(t, \pi)$ . Note that the graph of this function is the trajectory that the process  $X_t$  has to follow in order for the conditional probability process  $\Pi_t$  to stay constant at the value  $\pi$ . Thus we call  $x(\cdot, \pi)$  *the  $\pi$ -level curve*. Some handy regularity of  $x(\cdot, \cdot)$  and  $\sigma(\cdot, \cdot)$  is brought to light in the following.

**Proposition 3.6.** *The functions  $x(\cdot, \cdot) : [0, \infty) \times (0, 1) \rightarrow \mathbb{R}$  and  $\sigma(\cdot, \cdot) : [0, \infty) \times (0, 1) \rightarrow \mathbb{R}$  are both  $C^1$ .*

*Proof.* Define  $F : (-\epsilon, \infty) \times \mathbb{R} \rightarrow (-\epsilon, \infty) \times (0, 1)$  by  $F(t, x) = (t, \pi(t, x))$ , where  $\epsilon$  is as in the assumption (3.9). The function  $F$  is  $C^1$ , which is evident by applying the standard differentiation lemma as in the proof of Proposition 3.2. The Jacobian matrix of  $F$  is

$$J_F(t, x) = \begin{pmatrix} 1 & 0 \\ \partial_1 \pi(t, x) & \partial_2 \pi(t, x) \end{pmatrix}.$$

Since  $F$  is invertible and  $\det(J_F(t, x)) = \partial_2 \pi(t, x) > 0$  for all  $t > -\epsilon$  and all  $x \in \mathbb{R}$ , the inverse function theorem tells us that the inverse of  $F$  is also  $C^1$ , with

$$J_{F^{-1}}(F(t, x)) = (J_F(t, x))^{-1}.$$

Consequently,  $x(\cdot, \cdot)$  is  $C^1$  on  $(-\epsilon, \infty) \times (0, 1)$  with the derivatives

$$\partial_1 x(t, \pi) = -\partial_1 \pi(t, x(t, \pi)) / \partial_2 \pi(t, x(t, \pi))$$

and

$$\partial_2 x(t, \pi) = 1 / \partial_2 \pi(t, x(t, \pi)).$$

Finally, since a product of continuous functions is continuous, by the chain rule,  $\sigma(\cdot, \cdot)$  is continuously differentiable on  $(-\epsilon, \infty) \times (0, 1)$  and so on  $[0, \infty) \times (0, 1)$ .  $\square$

Next, denoting the initial value  $\Pi_0$  by  $\pi_0 \in (0, 1)$ , we show that the tails of the conditional distribution  $\mu_{t,x}$  are decreasing along the level curve  $x(\cdot, \pi_0)$ .

**Proposition 3.7.**

1. If  $a \geq 0$ , then for any  $t > 0$ ,

$$\mathbb{P}(B > a | X_t = x(t, \pi_0)) - \mathbb{P}(B > a) \leq 0. \quad (3.15)$$

Supposing  $\mu((a, \infty)) > 0$ , the inequality above is strict if and only if  $\mu([0, a]) > 0$ .

2. Likewise, if  $a < 0$ , then for any  $t > 0$ ,

$$\mathbb{P}(B < a | X_t = x(t, \pi_0)) - \mathbb{P}(B < a) \leq 0. \quad (3.16)$$

Supposing  $\mu((-\infty, a)) > 0$ , the inequality above is strict if and only if  $\mu([a, 0]) > 0$ .

*Proof.* We prove only the first of the two claims as the proof of the second one follows the same argument with straightforward modifications.

In the case  $\mu((a, \infty)) = 0$ , the claim holds trivially with equality in (3.15). Thus we assume that  $\mu((a, \infty)) > 0$  in what follows. Writing  $x(t)$  instead of  $x(t, \pi_0)$  for brevity, we note that using  $\mu_{t,x(t)}([0, \infty)) = \mu([0, \infty))$ , the inequality (3.15) is easily seen to be equivalent to

$$\frac{\int_{(a, \infty)} e^{bx(t) - b^2 \frac{t}{2}} \mu(db)}{\int_{[0, \infty)} e^{bx(t) - b^2 \frac{t}{2}} \mu(db)} \leq \frac{\mu((a, \infty))}{\mu([0, \infty))}. \quad (3.17)$$

Now, we will split the proof into consideration of two separate cases.

Case 1:  $a \geq 2x(t)/t$ . Here

$$\begin{aligned} \frac{\int_{[0, a]} e^{bx(t) - b^2 \frac{t}{2}} \mu(db)}{\int_{(a, \infty)} e^{bx(t) - b^2 \frac{t}{2}} \mu(db)} &\geq \frac{e^{ax(t) - a^2 \frac{t}{2}} \mu([0, a])}{e^{ax(t) - a^2 \frac{t}{2}} \mu((a, \infty))} \\ &= \frac{\mu([0, a])}{\mu((a, \infty))}, \end{aligned}$$

which is equivalent to (3.17). Since  $\mu((a, \infty)) > 0$ , the inequality above is strict if and only if  $\mu([0, a]) > 0$ .

Case 2:  $x(t) > 0$  and  $0 < a < 2x(t)/t$ . Using that  $\mu_{t,x(t)}([0, \infty)) = \mu([0, \infty))$ , we get

$$\begin{aligned} \int_{[0, \infty)} e^{bx(t) - b^2 t/2} \mu(db) &= \mu([0, \infty)) \frac{\int_{(-\infty, 0)} e^{bx(t) - b^2 t/2} \mu(db)}{\mu((-\infty, 0))} \\ &< \mu([0, \infty)), \end{aligned}$$

where the inequality holds since  $x(t) > 0$ . Hence rewriting

$$\frac{\int_{(a, \infty)} e^{bx(t) - b^2 t/2} \mu(db)}{\mu((a, \infty))} = \frac{\int_{[0, \infty)} e^{bx(t) - b^2 t/2} \mu(db) - \int_{[0, a]} e^{bx(t) - b^2 t/2} \mu(db)}{\mu([0, \infty)) - \mu([0, a])}$$

and keeping in mind that  $0 < a < 2x(t)/t$ , one clearly sees that

$$\frac{\int_{(a,\infty)} e^{bx(t)-b^2t/2} \mu(\mathrm{d}b)}{\mu((a,\infty))} \leq \frac{\int_{[0,\infty)} e^{bx(t)-b^2t/2} \mu(\mathrm{d}b)}{\mu([0,\infty))} \quad (3.18)$$

with the strict inequality if and only if  $\mu([0, a]) > 0$ . As (3.18) is equivalent to (3.17), the proof is complete.  $\square$

**Corollary 3.8.** *For any  $\pi \in (0, 1)$  fixed, the volatility function  $\sigma(\cdot, \pi) : [0, \infty) \rightarrow \mathbb{R}$  defined in (3.12) is non-increasing in time. Moreover, it is strictly decreasing for any initial prior  $\mu$  except a two-point distribution in which case  $t \mapsto \sigma(t, \pi)$  is a constant function.*

*Proof.* A key to the proof is a realisation that it is sufficient to prove that  $\sigma(0, \pi_0) \geq \sigma(s, \pi_0)$  for any  $s > 0$ ; the rest will immediately follow by a ‘moving-frame’ argument. More precisely, by ‘moving-frame’ we mean that for any  $\pi \in (0, 1)$ ,  $t \geq 0$ , one can think of  $\mu_{t,x(t,\pi)}$  as the initial prior distribution at time zero and so immediately obtain that  $\sigma(t, \pi) \geq \sigma(t+s, \pi)$  for any  $s > 0$ .

Using a shorthand  $x(t)$  for  $x(t, \pi_0)$  as before, recall from (3.13) that

$$\sigma(t, \pi_0) = (1 - \pi_0) \mathbb{E}_{t,x(t)}[B \mathbb{1}_{\{B \geq 0\}}] - \pi_0 \mathbb{E}_{t,x(t)}[B \mathbb{1}_{\{B < 0\}}].$$

Consequently,

$$\begin{aligned} \sigma(0, \pi_0) - \sigma(t, \pi_0) &= (1 - \pi_0) (\mathbb{E}[B \mathbb{1}_{\{B \geq 0\}}] - \mathbb{E}_{t,x(t)}[B \mathbb{1}_{\{B \geq 0\}}]) \\ &\quad + \pi_0 (\mathbb{E}_{t,x(t)}[B \mathbb{1}_{\{B < 0\}}] - \mathbb{E}[B \mathbb{1}_{\{B < 0\}}]) \\ &\geq 0 \end{aligned}$$

by Proposition 3.7. Moreover, by the same proposition, the inequality reduces to an equality if and only if  $\mu$  is a two-point distribution.  $\square$

**Remark** It seems difficult to find an easy intuitive argument for the monotonicity of the volatility function. As an example, consider a symmetric prior distribution, and a strictly positive time-point  $t$  at which the observation process satisfies  $X_t = 0$ . Then the conditional distribution  $\mu_{t,0}$  is also symmetric, so  $\Pi_t = \Pi_0 = 1/2$ . One certainly expects that  $\mu_{t,0}$  is obtained from the prior distribution  $\mu$  by pushing mass towards zero (this is also verified in Proposition 3.7 above). One could expect that the  $\Pi$ -process of a distribution with a lot of mass close to zero is sensitive to small changes in the observation process since the mass easily may ‘spill over’ to the other side of zero, and thus such a distribution gives rise to a comparatively large volatility. On the other hand, a concentrated distribution makes it difficult to distinguish possible drifts from each other, and changes in the observation process would to a higher degree be attributed to the Brownian fluctuations. This implies a slow learning process, which indicates a small volatility. Corollary 3.8 shows that the latter effect outweighs the former one.



## 4 Analysis of the optimal stopping problem

In this section, we study the perpetual optimal stopping problem (2.2) and its finite-horizon counterpart under the integrability condition (3.9). Most of the time the emphasis is on the perpetual case, though the corresponding results also hold for the finite horizon case by the same arguments. If the analogy is straightforward, we only comment on it, otherwise, more details are provided.

### 4.1 The value function with arbitrary starting points

Recall that

$$d\Pi_t = \sigma(t, \Pi_t) d\hat{W}_t,$$

where

$$\sigma(t, \pi) = \partial_2 \pi(t, x(t, \pi)) > 0$$

for all  $(t, \pi) \in [0, \infty) \times (0, 1)$  (beware that  $\pi(\cdot, \cdot)$  is a function, while  $\pi$  is a real number). We embed the optimal stopping problem (2.2), in which the starting point of the process  $\Pi$  is given by  $\Pi_0 := \mathbb{P}(B \geq 0)$ , into the optimal stopping problem

$$v(t, \pi) = \inf_{\tau \in \mathcal{T}} \mathbb{E} \left[ g(\Pi_{t+\tau}^{t, \pi}) + c\tau \right], \quad (t, \pi) \in [0, \infty) \times (0, 1), \quad (4.1)$$

for the process  $\Pi^{t, \pi} = \Pi$  given by

$$\begin{cases} d\Pi_{t+s} = \sigma(t+s, \Pi_{t+s}) d\hat{W}_{t+s}, & (s > 0) \\ \Pi_t = \pi, \end{cases} \quad (4.2)$$

where  $\mathcal{T}$  denotes the set of stopping times with respect to the completed filtration of  $\{\Pi_{t+s}^{t, \pi}\}_{s \geq 0}$ . The SDE (4.2) possesses a unique solution since  $\sigma(\cdot, \cdot)$  is locally Lipschitz by Proposition 3.6. Furthermore, the embedding has a consistent interpretation also at time  $t = 0$ , which is given by (3.10) and the remark following it. Note that choosing  $\tau = 0$  gives  $v(t, \pi) \leq g(\pi)$ .

**Proposition 4.1.** *The value function  $v(t, \pi)$  is concave in  $\pi$  for any fixed  $t \geq 0$ .*

*Proof.* This follows by a standard approximation argument using optimal stopping problems where stopping is only allowed at a discrete set of time-points, compare [8].

To outline this, denote by  $\mathcal{T}_{t, n}$ , where  $t \leq n$ , the set of stopping times in  $\mathcal{T}$  taking values in  $\{k2^{-n}, k = 0, 1, \dots, n2^n\} \cap [0, n - t]$ ,  $n = 1, 2, \dots$ , and let

$$v_n(t, \pi) = \inf_{\tau \in \mathcal{T}_{t, n}} \mathbb{E} \left[ g(\Pi_{t+\tau}^{t, \pi}) + c\tau \right].$$

Then  $v_n(n, \pi) = g(\pi)$  is concave in  $\pi$ . By preservation of concavity for martingale diffusions, see [10] (the results of [10] extend to the current setting with both an upper and a

lower bound on the state space),  $\pi \mapsto v_n(t, \pi)$  is concave also for  $t \in (n - 2^{-n}, n)$ . Next, at time  $t = n - 2^{-n}$  the value is given by dynamic programming as

$$v_n(t, \pi) = \min \{g(\pi), \mathbb{E} [v_n(n, \Pi_n^{t, \pi}) + c2^{-n}]\},$$

which is concave (being the minimum of two concave functions). Proceeding recursively shows that  $v_n$  is concave in  $\pi$  at all times  $t \in [0, n]$ . Since  $v_n$  converges pointwise to  $v$  as  $n \rightarrow \infty$ , this implies that also  $v$  is concave in  $\pi$ .  $\square$

**Proposition 4.2.** *The value function  $v(t, \pi)$  is non-decreasing in  $t$  for every fixed  $\pi \in (0, 1)$ .*

*Proof.* This can be proven using approximation by Bermudan options as in the proof of Proposition 4.1 above. Indeed, for a fixed time  $t \geq 0$  one may approximate  $v(t, \pi)$  by the optimal value in the case when stopping times are restricted to take values in the set  $\{k2^{-n} : n \in \mathbb{N}, k \in \{0, 1, \dots, n2^n\}\}$ . Since the expected value of a concave function of a martingale diffusion is non-increasing in the volatility, see [10], the approximation is non-decreasing in  $t$  by Corollary 3.8. Letting  $n \rightarrow \infty$  finishes the proof.  $\square$

**Proposition 4.3.** *The value function  $v$  is continuous on  $[0, \infty) \times [0, 1]$ .*

*Proof.* By concavity of  $v$  in the second variable together with the bounds  $0 \leq v \leq g$ , we have that  $v$  is Lipschitz continuous in  $\pi$  for any fixed  $t$ , with Lipschitz coefficient 1. Thus it suffices to check that  $v$  is continuous in time. To do this, let  $t_2 > t_1 \geq 0$  and note that

$$v(t_1, \pi) \geq \mathbb{E}[v(t_2, \Pi_{t_2}^{t_1, \pi})] \geq v(t_2, \pi) - \mathbb{E}[|\Pi_{t_2}^{t_1, \pi} - \pi|],$$

where the first inequality holds since  $\mathbb{E}[v(t_2, \Pi_{t_2}^{t_1, \pi})]$  represents the value of a sequential testing problem, started at  $t_1$ , with the running cost of observation not started until time  $t_2$ , the second inequality holds by the concavity of  $v$  in the second variable and the bounds  $0 \leq v \leq g$ . Thus

$$0 \leq v(t_2, \pi) - v(t_1, \pi) \leq \mathbb{E}[|\Pi_{t_2}^{t_1, \pi} - \pi|].$$

Since the expected value of a convex function of a martingale diffusion is non-decreasing in the volatility (again by [10]) and  $\sigma(0, \cdot) \geq \sigma(\cdot, \cdot)$  on  $[0, \infty) \times (0, 1)$ , we deduce that  $\mathbb{E}[|\Pi_{t_2}^{t_1, \pi} - \pi|] \leq \mathbb{E}[|\Pi_{t_2-t_1}^{0, \pi} - \pi|] \rightarrow 0$  as  $t_2 - t_1 \searrow 0$ . This finishes the proof.  $\square$

**Lemma 4.4.** *We have  $v(t, 1/2) < g(1/2)$  for all times  $t \geq 0$ .*

*Proof.* Let  $t \geq 0$ , and define  $A_\epsilon := [t, t + \epsilon] \times [1/2 - (c + 1)\epsilon, 1/2 + (c + 1)\epsilon]$  for  $\epsilon$  small enough so that  $A_\epsilon \subseteq [t, \infty) \times (0, 1)$ . Let

$$\tau_\epsilon := \inf\{s \geq 0 : (t + s, \Pi_{t+s}^{t, 1/2}) \notin A_\epsilon\}$$

be the first exit time from  $A_\epsilon$ . By Proposition 3.6 and Corollary 3.8,  $\sigma(\cdot, \cdot)$  is continuous and strictly positive on  $[0, \infty) \times (0, 1)$ . Thus  $\sigma_\epsilon := \inf_{(s, \pi) \in A_\epsilon} \sigma(s, \pi)$  is strictly positive and non-increasing as a function of  $\epsilon$ , so  $\sigma_\epsilon$  is bounded away from 0 as  $\epsilon \rightarrow 0$ . Now,

$$\begin{aligned} g(1/2) - v(t, 1/2) &\geq 1/2 - \mathbb{E} \left[ g(\Pi_{t+\tau_\epsilon}^{t, 1/2}) + c\tau_\epsilon \right] \\ &\geq 1/2 - (1/2 - (c+1)\epsilon + c\epsilon)\mathbb{P}(\tau_\epsilon < \epsilon) - (1/2 + c\epsilon)\mathbb{P}(\tau_\epsilon = \epsilon) \\ &= \epsilon - (c+1)\epsilon\mathbb{P}(\tau_\epsilon = \epsilon). \end{aligned} \tag{4.3}$$

Here

$$\begin{aligned} \mathbb{P}(\tau_\epsilon = \epsilon) &= \mathbb{P} \left( \sup_{0 \leq s \leq \epsilon} \left| \int_0^s \sigma(t+u, \Pi_{t+u}^{t, 1/2}) d\hat{W}_{t+u} \right| \leq (c+1)\epsilon \right) \\ &\leq \mathbb{P} \left( \sup_{0 \leq s \leq \epsilon} \left| \sigma_\epsilon \hat{W}_s \right| \leq (c+1)\epsilon \right) \\ &\leq \mathbb{P} \left( \sup_{0 \leq s \leq \epsilon} \hat{W}_s \leq (c+1)\epsilon/\sigma_\epsilon \right) \rightarrow 0 \end{aligned} \tag{4.4}$$

as  $\epsilon \rightarrow 0$ , where the first inequality follows from [10, Lemma 10]. Consequently, (4.3) and (4.4) yield that  $g(1/2) - v(t, 1/2) > 0$ , which finishes the proof of the claim.  $\square$

## 4.2 The structure of an optimal strategy

Recalling that  $0 \leq v(t, \pi) \leq g(\pi)$ , we denote by

$$\mathcal{C} := \{(t, \pi) \in [0, \infty) \times (0, 1) : v(t, \pi) < g(\pi)\}$$

the continuation region, and by

$$\mathcal{D} := \{(t, \pi) \in [0, \infty) \times (0, 1) : v(t, \pi) = g(\pi)\}$$

the stopping region. Since  $v$  is continuous,  $\mathcal{C}$  is open and  $\mathcal{D}$  is closed. Resorting to intuition from optimal stopping theory, we expect that the stopping time

$$\tau^* := \inf\{s \geq 0 : (t+s, \Pi_{t+s}^{t, \pi}) \in \mathcal{D}\} \tag{4.5}$$

is an optimal stopping time in (4.1). (Note that standard optimal stopping theory does not apply since the pay-off process is not uniformly integrable.) The optimality of  $\tau^*$  is verified below, see Theorem 4.6.

**Proposition 4.5.** *There exist two functions  $b_1 : [0, \infty) \rightarrow [0, 1/2)$  and  $b_2 : [0, \infty) \rightarrow (1/2, 1]$  such that*

$$\mathcal{C} = \{(t, \pi) : b_1(t) < \pi < b_2(t)\}.$$

*The function  $b_1$  is non-decreasing and right-continuous with left limits. Similarly,  $b_2$  is non-increasing and right-continuous with left limits.*

*Proof.* The existence of  $b_1$  and  $b_2$  follows from the concavity of  $v$  and Lemma 4.4. The monotonicity properties are immediate consequences of Proposition 4.2. Moreover, by the continuity of  $v$ , the function  $b_1$  is upper semi-continuous and  $b_2$  is lower semi-continuous. Hence, they are right-continuous with left limits.  $\square$

Let us also consider the same optimal stopping problem with a finite horizon  $T > 0$ . It is written as

$$v^T(t, \pi) = \inf_{\tau \in \mathcal{T}_{T-t}} \mathbb{E} \left[ g(\Pi_{t+\tau}^{t, \pi}) + c\tau \right], \quad (4.6)$$

where  $\mathcal{T}_{T-t}$  denotes the set of stopping times less or equal to  $T - t$  with respect to the completed filtration of  $\{\Pi_{t+s}^{t, \pi}\}_{s \geq 0}$ . Note that all results for the perpetual problem (4.1) described above in this section also hold for the finite horizon problem (4.6), with the obvious modifications regarding the time horizon, by the same proofs. Moreover, the payoff process in (4.6) is continuous and bounded, so standard optimal stopping theory (see, for example, [19, Corollary 2.9 on p. 46]) yields that

$$\tau^T := \inf\{s \geq 0 : \Pi_{t+s}^{t, \pi} \notin (b_1^T(t+s), b_2^T(t+s))\}$$

is an optimal stopping time in (4.6), where  $b_1^T$  and  $b_2^T$  are the corresponding boundaries enclosing the finite-horizon continuation region

$$\mathcal{C}^T := \{(t, \pi) \in [0, T) \times (0, 1) : v^T(t, \pi) < g(\pi)\}.$$

The infinite-horizon problem can be approximated by finite-horizon problems in the following sense.

**Theorem 4.6.** *The functions  $v^T \searrow v$ ,  $b_1^T \searrow b_1$ , and  $b_2^T \nearrow b_2$  pointwise as  $T \nearrow \infty$ . The stopping times  $\tau^T \nearrow \tau^*$  a.s. as  $T \nearrow \infty$ , where  $\tau^*$  is defined in (4.5). Moreover,  $\tau^*$  is optimal in (4.1).*

*Proof.* Since  $v^T \geq v$ , we have that  $b_1 \leq b_1^T < b_2^T \leq b_2$  and  $\tau^T \leq \tau^*$ . By bounded and monotone convergence,  $v^T(t, \pi) \searrow v(t, \pi)$  pointwise as  $T \rightarrow \infty$ , so  $b_1^T \searrow b_1$  and  $b_2^T \nearrow b_2$  pointwise as  $T \rightarrow \infty$ . Consequently, by the monotonicity of  $b_i$  and  $b_i^T$ , it follows that  $\tau^T \nearrow \tau^*$  a.s. Thus

$$v^T(t, \pi) = \mathbb{E} \left[ g(\Pi_{t+\tau^T}^{t, \pi}) + c\tau^T \right] \rightarrow \mathbb{E} \left[ g(\Pi_{t+\tau^*}^{t, \pi}) + c\tau^* \right]$$

by bounded and monotone convergence. By uniqueness of limits,

$$v(t, \pi) = \mathbb{E} \left[ g(\Pi_{t+\tau^*}^{t, \pi}) + c\tau^* \right],$$

so  $\tau^*$  is optimal.  $\square$

### 4.3 Optimal stopping boundaries and the free-boundary problem

**Proposition 4.7.** *The boundaries  $b_1$  and  $b_2$  satisfy  $0 < b_1(t) < 1/2 < b_2(t) < 1$  for all times  $t \geq 0$ .*

*Proof.* The two middle inequalities are granted by Lemma 4.4. To see that  $b_1 > 0$  on  $[0, \infty)$ , without loss of generality, it is sufficient to show that  $b_1 > 0$  on  $(0, \infty)$ ; this is due to the possibility provided by (3.10) to start the process  $\Pi$  slightly earlier. Let us assume, to reach a contradiction, that  $b_1(t) = 0$  for some  $t > 0$ . Then, by monotonicity,  $b_1 \equiv 0$  on  $[0, t]$ . By the martingale inequality,

$$\mathbb{P}(\tau^* \leq t) \leq \mathbb{P}\left(\sup_{0 \leq s \leq t} \Pi_s^{0, \pi} \geq 1/2\right) \leq 2\pi.$$

Consequently,

$$\mathbb{E}[\tau^*] \geq t(1 - 2\pi),$$

so if  $\pi \leq ct/(2ct + 1)$ , then  $v(0, \pi) \geq c\mathbb{E}[\tau^*] \geq \pi$ . Thus  $b_1(0) \geq ct/(2ct + 1) > 0$ , which contradicts our assumption. Therefore  $b_1 > 0$  at all times. The proof that  $b_2 < 1$  is analogous.  $\square$

**Proposition 4.8.** *The triplet  $(v, b_1, b_2)$  satisfies the free boundary problem*

$$\begin{cases} \partial_1 v(t, \pi) + \frac{\sigma(t, \pi)^2}{2} \partial_2^2 v(t, \pi) + c = 0 & b_1(t) < \pi < b_2(t) \\ v(t, \pi) = \pi & \pi \leq b_1(t) \\ v(t, \pi) = 1 - \pi & \pi \geq b_2(t). \end{cases} \quad (4.7)$$

*Moreover, the smooth-fit condition holds in the sense that the function  $\pi \mapsto v(t, \pi)$  is  $C^1$  for all  $t \geq 0$ .*

*Proof.* The proof that the differential equation in (4.7) holds is based on the strong Markov property and the continuity of  $v$ . However, the procedure is standard and we therefore omit the argument, referring to the proof of [12, Theorem 7.7] for the details instead. The value of  $v$  for  $\pi \notin (b_1(t), b_2(t))$  follows from concavity and the definition of  $b_1$  and  $b_2$ .

For the smooth-fit condition, note that the value function  $\pi \mapsto v(t, \pi)$  is continuous on  $(0, 1)$  and  $C^1$  for  $\pi \in (b_1(t), b_2(t))$  as well as for  $\pi \in (0, b_1(t)) \cup (b_2(t), 1)$ . Thus it remains to check the  $C^1$  property at  $b_1(t)$  and  $b_2(t)$ . To prove the  $C^1$  property at  $b_1(t)$  (the  $C^1$  property at  $b_2(t)$  being completely analogous), note that since  $v$  is concave in  $\pi$ , it suffices to show that

$$\liminf_{\epsilon \downarrow 0} \frac{v(t, b(t) + \epsilon) - v(t, b(t))}{\epsilon} \geq 1. \quad (4.8)$$

Without loss of generality, we do this for  $t = 0$ , letting  $\pi = b_1(0)$ .

Let  $\epsilon \in (0, 1/2 - \pi)$  and denote by  $\tau^\epsilon$  the first hitting time of the stopping region for  $\Pi^{0, \pi + \epsilon}$ . Then

$$\begin{aligned} v(0, \pi + \epsilon) - v(0, \pi) &\geq \mathbb{E}\left[g(\Pi_{\tau^\epsilon}^{0, \pi + \epsilon}) - g(\Pi_{\tau^\epsilon}^{0, \pi})\right] \\ &\geq \epsilon - 2\mathbb{E}\left[(\Pi_{\tau^\epsilon}^{0, \pi + \epsilon} - \Pi_{\tau^\epsilon}^{0, \pi}) \mathbb{1}_{\{\Pi_{\tau^\epsilon}^{0, \pi + \epsilon} > 1/2\}}\right]. \end{aligned}$$

Thus, to prove (4.8) it suffices (by the Cauchy-Schwartz inequality) to show that

$$\mathbb{E} \left[ (\Pi_{\tau^\epsilon}^{0, \pi + \epsilon} - \Pi_{\tau^\epsilon}^{0, \pi})^2 \right] \mathbb{P} \left( \Pi_{\tau^\epsilon}^{0, \pi + \epsilon} > 1/2 \right) = o(\epsilon^2) \quad (4.9)$$

as  $\epsilon \rightarrow 0$ . To do this, first assume that  $\sigma$  is Lipschitz continuous in  $\pi$  on any compact time interval, and define

$$h(t) := \mathbb{E} \left[ (\Pi_{t \wedge \tau^\epsilon}^{0, \pi + \epsilon} - \Pi_{t \wedge \tau^\epsilon}^{0, \pi})^2 \right].$$

Fixing  $T > 0$ , for  $t \in [0, T]$  we have

$$\begin{aligned} h(t) &= \mathbb{E} \left[ \left( \epsilon + \int_0^{t \wedge \tau^\epsilon} \sigma(s, \Pi_s^{0, \pi + \epsilon}) - \sigma(s, \Pi_s^{0, \pi}) d\hat{W}_s \right)^2 \right] \\ &\leq \epsilon^2 + \int_0^t \mathbb{E} \left[ D(T)^2 (\Pi_s^{0, \pi + \epsilon} - \Pi_s^{0, \pi})^2 \mathbb{1}_{\{s \leq \tau^\epsilon\}} \right] ds \\ &\leq \epsilon^2 + D(T)^2 \int_0^t h(s) ds, \end{aligned}$$

where  $D(T)$  is a Lipschitz constant for  $\sigma$  on  $[0, T] \times (0, 1)$ . Consequently, Gronwall's inequality yields

$$h(T) \leq \epsilon^2 e^{D(T)^2 T}. \quad (4.10)$$

Next, denote by  $f(y) := \pi - \frac{\pi}{1-\pi}(y - \pi)$  the affine function satisfying  $f(\pi) = \pi$  and  $f(1) = 0$ , and note that  $f \leq g$  on  $[\pi, 1]$ . Therefore,

$$\begin{aligned} c\mathbb{E}[\tau^\epsilon] &= v(0, \pi + \epsilon) - \mathbb{E}[g(\Pi_{\tau^\epsilon}^{0, \pi + \epsilon})] \\ &\leq g(\pi + \epsilon) - \mathbb{E}[f(\Pi_{\tau^\epsilon}^{0, \pi + \epsilon})] \\ &= \pi + \epsilon - \left( \pi - \frac{\pi}{1-\pi} \mathbb{E}[\Pi_{\tau^\epsilon}^{0, \pi + \epsilon} - \pi] \right) \\ &= \epsilon / (1 - \pi), \end{aligned}$$

where the inequality follows from the monotonicity of  $b_1$  and the last equality by optional sampling. Thus, writing  $D = 1/(1 - \pi)$ , we have

$$\mathbb{P}(\tau_\epsilon > T) \leq D\epsilon / (cT). \quad (4.11)$$

Moreover, writing

$$\tau_{\pi, 1/2} := \inf\{s \geq 0 : \Pi_s^{0, \pi + \epsilon} \notin (\pi, 1/2)\},$$

we have

$$\mathbb{P} \left( \Pi_{\tau^\epsilon}^{0, \pi + \epsilon} > 1/2 \right) \leq \mathbb{P} \left( \Pi_{\tau_{\pi, 1/2}}^{0, \pi + \epsilon} = 1/2 \right) \leq \frac{\epsilon}{\frac{1}{2} - \pi} = C\epsilon \quad (4.12)$$

for  $C = 2/(1 - 2\pi)$ , where we used the martingality of  $\Pi$  to obtain the second inequality. Putting together (4.10), (4.11) and (4.12) yields

$$\begin{aligned} & \mathbb{E} \left[ (\Pi_{\tau_\epsilon}^{0, \pi+\epsilon} - \Pi_{\tau_\epsilon}^{0, \pi})^2 \right] \mathbb{P} \left( \Pi_{\tau_\epsilon}^{0, \pi+\epsilon} > 1/2 \right) \\ & \leq \left( \mathbb{E} \left[ (\Pi_{\tau_\epsilon}^{0, \pi+\epsilon} - \Pi_{\tau_\epsilon}^{0, \pi})^2 \mathbb{1}_{\{\tau_\epsilon \leq T\}} \right] + \mathbb{P}(\tau_\epsilon > T) \right) \mathbb{P} \left( \Pi_{\tau_\epsilon}^{0, \pi+\epsilon} > 1/2 \right) \\ & \leq \epsilon^2 C (\epsilon e^{D(T)^2 T} + D/(cT)). \end{aligned}$$

Given  $\delta > 0$ , it is possible to choose  $T$  large enough so that  $CD/(cT) \leq \delta/2$ , and then to choose  $\epsilon > 0$  small enough so that  $C\epsilon e^{D^2(T)^2 T} \leq \delta/2$ . This proves (4.9) and thus finishes the proof of the smooth-fit property if  $\sigma$  is Lipschitz in  $\pi$ , locally uniformly in  $t$ .

For a general  $\sigma$ , due to the  $C^1$  regularity of  $\sigma$  on  $[0, \infty) \times (0, 1)$ , one can find another volatility function  $\hat{\sigma}$  that is Lipschitz continuous in  $\pi$  on any given compact interval in time, and that satisfies  $0 \leq \hat{\sigma} \leq \sigma$  everywhere and  $\hat{\sigma} = \sigma$  on  $[0, \infty) \times [b_1(0), b_2(0)]$ . By monotonicity in the volatility, the corresponding value function  $\hat{v}$  satisfies  $\hat{v} \geq v$ . On the other hand, since  $\hat{\sigma} = \sigma$  on  $[0, \infty) \times [b_1(0), b_2(0)]$  and since  $\tau^*$  is optimal for the volatility  $\sigma$ , we also have  $\hat{v} \leq v$ , so  $\hat{v} = v$ . By the above argument,  $\hat{v}$  is  $C^1$ , which finishes the proof.  $\square$

**Theorem 4.9.** *The boundaries  $b_1$  and  $b_2$  are both continuous.*

*Proof.* Let us prove continuity of  $b_1$  (the proof for  $b_2$  is analogous). We know that  $b_1$  is right-continuous, so it suffices to assume for a contradiction that  $b_1$  is not continuous at some time  $t_0 > 0$ . By monotonicity,  $b_1(t_0) > b_1(t_0-)$ . In the continuation region,  $\partial_1 v \geq 0$ , so (4.7) yields

$$\frac{\sigma^2}{2} \partial_2^2 v \leq -c.$$

Since  $\sigma$  is locally bounded away from zero, this means that on each compact set we can find some constant  $d > 0$  such that  $\partial_2^2 v \leq -d$ . By Proposition 4.8, the map  $\pi \mapsto v(t, \pi)$  is  $C^1$  on  $[b_1(t), b_2(t)]$  for any  $t \geq 0$ , so for  $t < t_0$  and  $b_1(t) < \pi < b_1(t_0)$ , we have

$$\begin{aligned} v(t, \pi) - g(\pi) &= \int_{b_1(t)}^\pi \int_{b_1(t)}^w \partial_2^2 (v - g)(t, u) \, du \, dw \\ &\leq -d(\pi - b_1(t))^2/2. \end{aligned}$$

Choosing  $\pi = \frac{b_1(t_0-) + b_1(t_0)}{2}$  and letting  $t \rightarrow t_0$  gives

$$v(t_0, \frac{b_1(t_0-) + b_1(t_0)}{2}) - g(\frac{b_1(t_0-) + b_1(t_0)}{2}) \leq -d(b_1(t_0) - b_1(t_0-))^2/2 < 0.$$

This contradicts the assumption that  $(t_0, \frac{b_1(t_0-) + b_1(t_0)}{2})$  belongs to the stopping region, so  $b_1$  has to be continuous.  $\square$

**Remark** Even though, in this section, all the results are formulated for the perpetual problem (4.1), it is straightforward to check that the corresponding results for the finite-horizon problem (4.6) also hold. In that case, the boundaries  $b_1^T : [0, 1] \rightarrow (0, 1)$  and

$b_2^T : [0, 1] \rightarrow (0, 1)$  are continuous and monotone, with  $0 < b_1^T < 1/2 < b_2^T < 1$  on  $[0, T)$  and  $b_1^T(T) = b_2^T(T) = 1/2$ . Also, the assertions of Proposition 4.8 hold for  $(v^T, b_1^T, b_2^T)$  on the time interval  $[0, T)$  in place of  $(v, b_1, b_2)$ .

## 5 Integral equations for the boundaries

It is well-known that optimal stopping boundaries, under some conditions, can be characterized by certain integral equations, compare [9] and [18]. In this section, we study the integral equations for the optimal stopping boundaries arising in our sequential testing problem. For the problem (4.6) with finite horizon, a pair of integral equations is shown to completely characterise the optimal stopping boundaries within the class of continuous solutions. The situation in the perpetual case is more delicate, and uniqueness of solutions remains an open question.

### 5.1 A pair of integral equations for the finite-horizon boundaries

**Theorem 5.1.** *Assume that  $T < \infty$ . Then the pair  $(b_1^T, b_2^T)$  is the unique continuous solution of*

$$\begin{cases} c_1(t) = \mathbb{E} \left[ g(\Pi_T^{t, c_1(t)}) \right] + c \int_0^{T-t} \mathbb{P}(c_1(t+u) < \Pi_{t+u}^{t, c_1(t)} < c_2(t+u)) du \\ 1 - c_2(t) = \mathbb{E} \left[ g(\Pi_T^{t, c_2(t)}) \right] + c \int_0^{T-t} \mathbb{P}(c_1(t+u) < \Pi_{t+u}^{t, c_2(t)} < c_2(t+u)) du \end{cases} \quad (5.1)$$

such that  $0 < c_1(t) \leq 1/2 \leq c_2(t) < 1$  for all  $t \in [0, T]$ .

*Proof.* Applying Ito's formula (more precisely, an extension of Ito's formula, compare the results of [17], which can be applied thanks to the monotonicity of  $b_1^T$  and  $b_2^T$ ) and then taking expectations yields

$$\mathbb{E} \left[ g(\Pi_T^{t, \pi}) \right] = v^T(t, \pi) - c \int_0^{T-t} \mathbb{P}(b_1^T(t+u) < \Pi_{t+u}^{t, \pi} < b_2^T(t+u)) du.$$

Plugging in  $\pi = b_1^T(t)$  and  $\pi = b_2^T(t)$  shows that  $(b_1^T, b_2^T)$  solves (5.1).

For uniqueness, assume that  $(c_1, c_2)$  is another continuous solution to (5.1) with  $0 < c_1(t) \leq 1/2 \leq c_2(t) < 1$ , and define

$$V(t, \pi) := \mathbb{E} \left[ g(\Pi_T^{t, \pi}) \right] + c \int_0^{T-t} \mathbb{P}(c_1(t+u) < \Pi_{t+u}^{t, \pi} < c_2(t+u)) du.$$

Then  $V(t, c_1(t)) = c_1(t)$  and  $V(t, c_2(t)) = 1 - c_2(t)$  by (5.1), and  $V(T, \pi) = g(\pi)$ . Moreover, by the Markov property, the process

$$M_s := V(t+s, \Pi_{t+s}^{t, \pi}) + c \int_0^s \mathbb{1}_{(c_1(t+u), c_2(t+u))}(\Pi_{t+u}^{t, \pi}) du$$

is a martingale for any  $(t, \pi)$ . In particular, the process

$$\tilde{M}_s := v^T(t+s, \Pi_{t+s}^{t, \pi}) + c \int_0^s \mathbb{1}_{(b_1^T(t+u), b_2^T(t+u))}(\Pi_{t+u}^{t, \pi}) du$$



is also a martingale.

Claim 1:  $V(t, \pi) = g(\pi)$  for  $\pi \notin (c_1(t), c_2(t))$ .

Assume that  $\pi \leq c_1(t)$  (the case  $\pi \geq c_2(t)$  is similar), and let

$$\gamma_c := \inf\{s \geq 0 : \Pi_{t+s}^{t, \pi} \geq c_1(t+s)\} \wedge (T-t).$$

Then

$$V(t, \pi) = \mathbb{E} \left[ V(t + \gamma_c, \Pi_{t+\gamma_c}^{t, \pi}) \right] = \mathbb{E} \left[ \Pi_{t+\gamma_c}^{t, \pi} \right] = \pi = g(\pi),$$

whith the first equality being justified by optional sampling and the martingale property of  $M$ , the second by (5.1), and the third by optional sampling and the martingale property of  $\Pi$ .

Claim 2:  $V \geq v^T$ .

Take  $(t, \pi)$  such that  $c_1(t) < \pi < c_2(t)$ , and let

$$\tau_c := \inf\{s \geq 0 : \Pi_{t+s}^{t, \pi} \notin (c_1(t+s), c_2(t+s))\} \wedge (T-t).$$

Then

$$\begin{aligned} V(t, \pi) &= \mathbb{E} \left[ V(t + \tau_c, \Pi_{t+\tau_c}^{t, \pi}) \right] + c \mathbb{E} \left[ \int_0^{\tau_c} \mathbb{1}_{(c_1(t+u), c_2(t+u))}(\Pi_{t+u}^{t, \pi}) \, du \right] \\ &= \mathbb{E} \left[ g(\Pi_{t+\tau_c}^{t, \pi}) \right] + c \mathbb{E} [\tau_c] \geq v^T(t, \pi). \end{aligned}$$

From this and Claim 1, Claim 2 follows.

Claim 3:  $b_1^T \leq c_1$  and  $c_2 \leq b_2^T$ .

Assume that  $b_1^T(t) > c_1(t)$  for some  $t$ . Choose  $\pi = c_1(t)$ , and let

$$\gamma_b := \inf\{s \geq 0 : \Pi_{t+s}^{t, \pi} \geq b_1^T(t+s)\} \wedge (T-t).$$

Then, by right-continuity of  $b_1^T$  and  $c_1$ ,

$$\mathbb{E} \left[ \int_0^{\gamma_b} \mathbb{1}_{(c_1(t+u), c_2(t+u))}(\Pi_{t+u}^{t, \pi}) \, du \right] > 0. \quad (5.2)$$

On the other hand, by optional sampling and martingality of  $M$  and  $\tilde{M}$  we have

$$\begin{aligned} 0 &= V(t, \pi) - v^T(t, \pi) \\ &= \mathbb{E} \left[ V(t + \gamma_b, \Pi_{t+\gamma_b}^{t, \pi}) - v^T(t + \gamma_b, \Pi_{t+\gamma_b}^{t, \pi}) \right] \\ &\quad + c \mathbb{E} \left[ \int_0^{\gamma_b} \mathbb{1}_{(c_1(t+u), c_2(t+u))}(\Pi_{t+u}^{t, \pi}) \, du \right]. \end{aligned}$$

Since  $V \geq v$  by Claim 2, this contradicts (5.2) and thus  $b_1^T \leq c_1$ . The claim  $c_2 \leq b_2^T$  is proved similarly.

Claim 4:  $c_1 \leq b_1^T$  and  $b_2^T \leq c_2$ .

Assume that  $c_1(t) > b_1^T(t)$  for some  $t \geq 0$ , and pick  $\pi \in (b_1^T(t), c_1(t))$ . Let  $\tau_b := \inf\{s \geq 0 : \Pi_{t+s}^{t,\pi} \notin (b_1^T(t+s), b_2^T(t+s))\} \wedge (T-t)$ . Then

$$\mathbb{E} \left[ \int_0^{\tau_b} \mathbb{1}_{(b_1^T(t+u), c_1(t+u))}(\Pi_{t+u}^{t,\pi}) \right] > 0. \quad (5.3)$$

On the other hand,

$$\begin{aligned} \mathbb{E} \left[ g(\Pi_{t+\tau_b}^{t,\pi}) \right] &= \mathbb{E} \left[ V(t + \tau_b, \Pi_{t+\tau_b}^{t,\pi}) \right] \\ &= V(t, \pi) - c \mathbb{E} \left[ \int_0^{\tau_b} \mathbb{1}_{(c_1(t+u), c_2(t+u))}(\Pi_{t+u}^{t,\pi}) \, du \right] \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left[ g(\Pi_{t+\tau_b}^{t,\pi}) \right] &= \mathbb{E} \left[ v^T(t + \tau_b, \Pi_{t+\tau_b}^{t,\pi}) \right] \\ &= v^T(t, \pi) - c \mathbb{E} \left[ \int_0^{\tau_b} \mathbb{1}_{(b_1^T(t+u), b_2^T(t+u))}(\Pi_{t+u}^{t,\pi}) \, du \right]. \end{aligned}$$

Hence, using Claim 2 and Claim 3 above, we find that

$$\mathbb{E} \left[ \int_0^{\tau_b} \mathbb{1}_{(b_1^T(t+u), c_1(t+u))}(\Pi_{t+u}^{t,\pi}) \, du \right] \leq 0,$$

which contradicts (5.3). Consequently, such a point  $\pi$  does not exist, so  $b_1^T \geq c_1$ . Similarly,  $c_2 \geq b_2^T$ .

Finally, combining Claim 3 and Claim 4 finishes the proof of the theorem.  $\square$

## 5.2 A pair of integral equations for the infinite-horizon boundaries

**Theorem 5.2.** *The pair  $(b_1, b_2)$  is a solution of*

$$\begin{cases} b_1(t) = c \int_0^\infty \mathbb{P}(b_1(t+u) < \Pi_{t+u}^{t,b_1(t)} < b_2(t+u)) \, du \\ 1 - b_2(t) = c \int_0^\infty \mathbb{P}(b_1(t+u) < \Pi_{t+u}^{t,b_2(t)} < b_2(t+u)) \, du. \end{cases} \quad (5.4)$$

*Proof.* For fixed  $T > 0$ , an application of Itô's formula as in the preceding proof gives

$$\mathbb{E} \left[ v(T, \Pi_T^{t,\pi}) \right] = v(t, \pi) - c \int_0^{T-t} \mathbb{P}(b_1(t+u) < \Pi_{t+u}^{t,\pi} < b_2(t+u)) \, du.$$

Since  $v$  is bounded and  $\Pi_T^{t,\pi}$  converges to either 0 or 1 as  $T \rightarrow \infty$  by Proposition 3.5, we find that

$$v(t, \pi) = c \int_0^\infty \mathbb{P}(b_1(t+u) < \Pi_{t+u}^{t,\pi} < b_2(t+u)) \, du.$$

Plugging in  $\pi = b_1(t)$  and  $\pi = b_2(t)$  shows that  $(b_1, b_2)$  solves (5.4).  $\square$

**Remark** The main technical difficulty when trying to apply the uniqueness proof of Theorem 5.1 to the perpetual problem lies in the lack of a straightforward extension of the optional sampling theorem to unbounded, possibly infinite stopping times.

### 5.3 The case of a symmetric volatility function

Now assume that the volatility function is symmetric about  $\pi = 1/2$ , i.e.  $\sigma(t, \pi) = \sigma(t, 1 - \pi)$ . This is the case, for example, if the prior distribution  $\mu$  is symmetric about zero in the sense that  $\mu([0, a]) = \mu((-a, 0))$  for all  $a > 0$ . Then, by symmetry,  $b_1^T = 1 - b_2^T$ , and we set  $b^T := b_1^T$ . The following result is a straightforward consequence of Theorem 5.1.

**Theorem 5.3.** *Assume that  $\sigma$  is symmetric about  $\pi = 1/2$ . Then the boundary  $b^T$  is the unique continuous solution of*

$$c(t) = \mathbb{E} \left[ g(\Pi_T^{t, c(t)}) \right] + c \int_0^{T-t} \mathbb{P}(c(t+u) < \Pi_{t+u}^{t, c(t)} < 1 - c(t+u)) du \quad (5.5)$$

such that  $0 < c(t) \leq 1/2$  for all  $t \in [0, T]$ .

**Remark** Although not necessarily symmetric, all normal prior distributions as well as all two-point priors give rise to symmetric volatilities, compare Section 3.

## 6 Long-term asymptotics of the volatility and the boundaries

Since the boundaries  $b_1$  and  $b_2$  are monotone, the limits  $b_i(\infty) := \lim_{t \rightarrow \infty} b_i(t)$ ,  $i = 1, 2$ , exist with  $b_1(\infty) \leq 1/2$  and  $b_2(\infty) \geq 1/2$ . In this section we determine these limits. To do that, we first derive a few limiting properties of level curves as well as study the limit  $\sigma(\infty, \pi) := \lim_{t \rightarrow \infty} \sigma(t, \pi)$  of the volatility.

Let us define

$$r = \inf\{s \geq 0 : \mu([s, s + \epsilon]) > 0 \text{ for all } \epsilon > 0\} \quad (6.1)$$

and

$$l = \sup\{s < 0 : \mu((s - \epsilon, s]) > 0 \text{ for all } \epsilon > 0\}. \quad (6.2)$$

We write  $m = (l + r)/2$  for the midpoint between  $l$  and  $r$ .

The following proposition will serve as a useful device for understanding long-term volatility.

**Proposition 6.1.**

1. If  $\alpha > m$ , then  $\pi(t, \alpha t) \rightarrow 1$  as  $t \rightarrow \infty$ .
2. If  $\alpha < m$ , then  $\pi(t, \alpha t) \rightarrow 0$  as  $t \rightarrow \infty$ .

*Proof.* Given  $\alpha \in \mathbb{R}$ , define

$$h(t) := \frac{\int_{(-\infty, 0)} \exp\left(-\frac{(b - \alpha)^2 t}{2}\right) \mu(db)}{\int_{[0, \infty)} \exp\left(-\frac{(b - \alpha)^2 t}{2}\right) \mu(db)},$$

so that  $\pi(t, \alpha t) = 1/(1 + h(t))$ . We will prove the claims in two different cases separately.

(i) First case:  $l < r$ .

1. First note that, in view of Proposition 3.4, it suffices to treat the case  $\alpha \in (m, r)$ . For such  $\alpha$ , fix  $\gamma > r$  such that  $\gamma - \alpha < \alpha - l$ . Then

$$h(t) \leq \frac{\exp\left(-(\alpha - l)^2 \frac{t}{2}\right) \int_{(-\infty, 0)} \mu(db)}{\exp\left(-(\gamma - \alpha)^2 \frac{t}{2}\right) \int_{[0, \gamma]} \mu(db)} \rightarrow 0$$

as  $t \rightarrow \infty$ . Hence  $\pi(t, \alpha t) \rightarrow 1$  as  $t \rightarrow \infty$ .

2. For the second result, suppose that  $\alpha < m$ , and note that it suffices to treat the case  $\alpha \in (l, m)$ . Let  $\gamma < l$  be such that  $\alpha - \gamma < r - \alpha$ . Then

$$h(t) \geq \frac{\exp\left(-(\alpha - \gamma)^2 \frac{t}{2}\right) \int_{(\gamma, 0)} \mu(db)}{\exp\left(-(\alpha - r)^2 \frac{t}{2}\right) \int_{[0, \infty)} \mu(db)} \rightarrow \infty$$

as  $t \rightarrow \infty$ . Hence  $\pi(t, \alpha t) \rightarrow 0$  as  $t \rightarrow \infty$ .

(ii) Second case:  $l = r = 0$ .

1. Assume that  $\alpha > 0$ , and let  $\epsilon > 0$ . Then

$$\begin{aligned} h(t) &\leq \frac{\int_{(-\infty, -\epsilon)} \exp\left(-(\alpha + \epsilon)^2 \frac{t}{2}\right) \mu(db) + \int_{[-\epsilon, 0)} \exp\left(-\alpha^2 \frac{t}{2}\right) \mu(db)}{\int_{[0, \alpha]} \exp\left(-\alpha^2 \frac{t}{2}\right) \mu(db)} \\ &\rightarrow \frac{\mu([- \epsilon, 0))}{\mu([0, \alpha])} \end{aligned}$$

as  $t \rightarrow \infty$ . Thus, since  $\epsilon > 0$  is arbitrary and  $\mu([- \epsilon, 0)) \rightarrow 0$  as  $\epsilon \rightarrow 0$ , we conclude that  $h(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Consequently,  $\pi(t, \alpha t) \rightarrow 1$ .

2. Next, assume that  $\alpha < 0$ . Choosing  $\gamma \in (\alpha, 0)$  with  $\mu((\alpha, \gamma)) > 0$ , we find that

$$h(t) \geq \frac{\exp\left(-(\alpha - \gamma)^2 \frac{t}{2}\right) \int_{(\alpha, \gamma)} \mu(db)}{\exp\left(-\alpha^2 \frac{t}{2}\right) \int_{[0, \infty)} \mu(db)} \rightarrow \infty$$

as  $t \rightarrow \infty$ . Consequently,  $\pi(t, \alpha t) \rightarrow 0$  as  $t \rightarrow \infty$ , which finishes the proof. □

**Remark** Notice that Proposition 6.1 implies that for any fixed value  $\pi$ , the corresponding level curve  $x(\cdot, \pi)$  satisfies  $\lim_{t \rightarrow \infty} (ct - x(t, \pi)) = \infty$  if  $\alpha > m$ , and  $\lim_{t \rightarrow \infty} (\alpha t - x(t, \pi)) = -\infty$  if  $\alpha < m$ .

## 6.1 Long-term behaviour of the volatility

Now, we are in a position to determine the limit  $\sigma(\infty, \pi) := \lim_{t \rightarrow \infty} \sigma(t, \pi)$  of the volatility as time increases.

**Proposition 6.2.** *The long-term limit  $\sigma(\infty, \pi) = (r - l)\pi(1 - \pi)$ .*

**Remark** Note that if  $l = r = 0$ , then the volatility converges to zero as time tends to infinity. Also, note that if  $l < r$ , then the volatility tends to the volatility from the case of a two-point prior distribution.

*Proof of Proposition 6.2.* We first claim that

$$\mathbb{E}_{t,x(t,\pi)} [B\mathbb{1}_{[0,\infty)}(B)] \rightarrow \pi r$$

as  $t \rightarrow \infty$ . To see this, suppose that  $a > r$  and take  $\gamma \in (r, a)$  such that  $\gamma - r < a - \gamma$ . By Corollary 3.3 and Proposition 6.1, for all large enough  $t$ ,

$$\begin{aligned} \mathbb{E}_{t,x(t,\pi)} [B\mathbb{1}_{(a,\infty)}(B)] &\leq \mathbb{E}_{t,\gamma t} [B\mathbb{1}_{(a,\infty)}(B)] \\ &= \frac{\int_{(a,\infty)} b e^{-(b-\gamma)^2 \frac{t}{2}} \mu(db)}{\int_{\mathbb{R}} e^{-(b-\gamma)^2 \frac{t}{2}} \mu(db)} \\ &\leq \frac{\exp\left(-\frac{(a-\gamma)^2 t}{2}\right) \int_{(a,\infty)} b \mu(db)}{\exp\left(-\frac{(\gamma-r)^2 t}{2}\right) \mu([r, \gamma))}, \end{aligned}$$

which tends to 0 as  $t \rightarrow \infty$ . Now, the fact that  $\mathbb{P}_{t,x(t,\pi)}(B \in [0, r)) = 0$  for all  $t \geq 0$  finishes the claim.

Next, straightforward modifications of the arguments above show that

$$\mathbb{E}_{t,x(t,\pi)} [B\mathbb{1}_{(-\infty,0)}(B)] \rightarrow (1 - \pi)l$$

as  $t \rightarrow \infty$ . Since

$$\sigma(t, \pi) = (1 - \pi)\mathbb{E}_{t,x(t,\pi)} [B\mathbb{1}_{[0,\infty)}(B)] - \pi\mathbb{E}_{t,x(t,\pi)} [B\mathbb{1}_{(-\infty,0)}(B)],$$

this finishes the proof. □

**Remark** Similar arguments as in the proof above show that  $\mu_{t,x(t,\pi)} \Rightarrow (1 - \pi)\delta_l + \pi\delta_r$  as  $t \rightarrow \infty$ . Thus, along a level curve  $x(\cdot, \pi)$  the conditional distribution of  $B$  converges weakly to the two-point distribution with mass  $\pi$  at  $r$  and mass  $1 - \pi$  at  $l$ .

## 6.2 Long-term behaviour of the boundaries

**Theorem 6.3.**

- If  $l = r = 0$ , then  $b_1(\infty) = b_2(\infty) = 1/2$ .
- If  $l < r$ , then  $b_1(\infty) = b_1^{r-l}$  and  $b_2(\infty) = b_2^{r-l}$ , where  $b_1^{r-l} < 1/2 < b_2^{r-l}$  are the optimal boundaries for a two-point prior distribution with mass at points separated by 0 and at a distance  $r - l$  from each other.

*Proof.* Since the volatility  $\sigma(\cdot, \cdot)$  is non-increasing in time, Proposition 6.2 and Dini's theorem yield that  $\sigma(t, \cdot)$  converges to  $\sigma(\infty, \pi) = (r-l)\pi(1-\pi)$  uniformly on the compact interval  $[b_1(0), b_2(0)]$  as  $t \rightarrow \infty$ . Therefore, given  $\epsilon > 0$  we can find  $t_0$  large enough so that  $\sigma(t_0, \pi) \leq (\epsilon + r - l)\pi(1 - \pi)$  for  $\pi \in [b_1(0), b_2(0)]$ . Define

$$\hat{\sigma}(t, \pi) := \sigma(t, \pi) \mathbb{1}_{[b_1(0), b_2(0)]}(\pi),$$

and denote by  $\hat{v}$  the corresponding value function. Since the optimal stopping problem (4.1) is monotone in the volatility (compare e.g. [10, Lemma 10]), we have that  $\hat{v} \geq v$ . On the other hand, since  $\hat{\sigma} = \sigma$  on the continuation region  $\{(t, \pi) : b_1(t) < \pi < b_2(t)\}$ , we also have  $\hat{v} \leq v$ , so  $\hat{v} = v$ . Moreover, by monotonicity in the volatility,

$$v^{\epsilon+r-l} \leq \hat{v} = v \leq v^{r-l}$$

on  $[t_0, \infty) \times (0, 1)$ , where  $v^a$  denotes the value function corresponding to a volatility function  $a\pi(1 - \pi)$ . Since the value function  $v$  is squeezed in between the value functions  $v^{\epsilon+r-l}$  and  $v^{r-l}$  from time  $t_0$ , the optimal stopping boundaries  $b_1$  and  $b_2$  are squeezed in between the corresponding optimal stopping boundaries for  $v^{\epsilon+r-l}$  and  $v^{r-l}$ . By inspection of the explicit formulas in the two-point distribution case, see [19, Theorem 21.1], the gaps  $b_1^{r-l} - b_1^{\epsilon+r-l}$  and  $b_2^{\epsilon+r-l} - b_2^{r-l}$  between the boundaries vanish as  $\epsilon \rightarrow 0$ , which finishes the proof. □

**Remark** It is also of interest to determine  $b_i(0)$  for  $i = 1, 2$  in order to find the best bounds for the continuation region. It seems difficult to determine these quantities in general, but an upper bound for the continuation region initially (and thus at all times) can be established by solving the free-boundary problem for the time-homogeneous volatility  $\sigma(0, \pi)$ . However, we expect these bounds to be rather crude, and therefore do not provide any details.

## 7 The normal prior distribution

In this final section, we study the case of a normal prior distribution. In particular, we show that the kernel in the integral equations determined in Section 5 can be calculated explicitly for normal priors.

First, recall from Section 3 that a normal prior distribution with mean  $m$  and variance  $\gamma^2$  leads to a volatility surface  $\sigma(\cdot, \cdot)$  that is symmetric around the line  $\pi = 1/2$ . As a result, the stopping boundaries  $b_1$  and  $b_2$  are also symmetric around  $\pi = 1/2$  with  $b_2(t) = 1 - b_1(t)$ , so it suffices to solve a single integral equation to determine both boundaries. Next, recall that the conditional distribution  $\mu_{t,x}$  is normal with standard deviation  $\gamma(t) := \gamma/\sqrt{1+t\gamma^2}$ . Consequently, the  $x$ -value that gives  $\pi(t, x) = b(t)$  is such that the conditional drift equals

$$m(t) := \Phi^{-1}(b(t))\gamma/\sqrt{1+t\gamma^2}.$$

Now, given  $s > 0$ , let  $Y$  denote a  $N(m(t)s, s + s^2\gamma^2(t))$ -distributed random variable. Then using (3.14), we calculate

$$\begin{aligned} K(t, s, b(t), b(t+s)) &:= \mathbb{P}\left(b(t+s) < \Pi_{t+s}^{t, b(t)} < 1 - b(t+s)\right) \\ &= \mathbb{P}\left(b(t+s) < \Phi\left(\frac{m(t) + \gamma^2(t)Y}{\gamma(t)\sqrt{1 + s\gamma^2(t)}}\right) < 1 - b(t+s)\right) \\ &= \Phi(d_2) - \Phi(d_1), \end{aligned}$$

where

$$d_1 := \frac{\Phi^{-1}(b(t+s))\gamma(t)\sqrt{1 + s\gamma^2(t)} - m(t)(1 + s\gamma^2(t))}{\gamma^2(t)\sqrt{s + s^2\gamma^2(t)}}$$

and

$$d_2 := \frac{-\Phi^{-1}(b(t+s))\gamma(t)\sqrt{1 + s\gamma^2(t)} - m(t)(1 + s\gamma^2(t))}{\gamma^2(t)\sqrt{s + s^2\gamma^2(t)}}.$$

Thus the kernel  $K$  appearing in the integral equation (5.5) and in the corresponding equation for the infinite-horizon formulation is explicit.

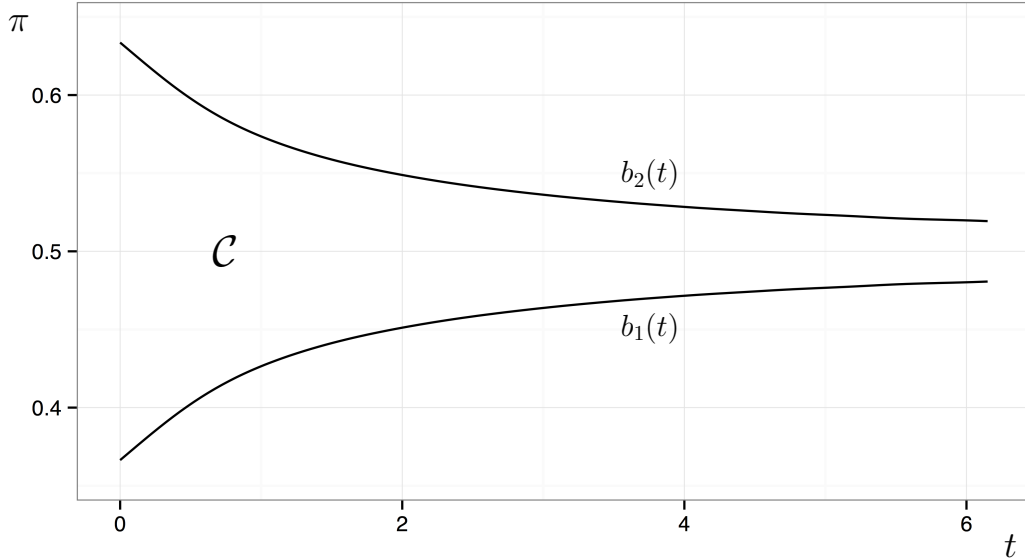


Figure 1: The boundaries  $b_1$  and  $b_2$  calculated numerically for the case of  $N(m, 1)$ -prior (note that the boundaries do not depend on  $m \in \mathbb{R}$ ) and the cost of observation  $c = 0.5$  per unit time.

## References

- [1] Bain, A. and Crisan, D. *Fundamentals of stochastic filtering*. Stochastic Modelling and Applied Probability, 60. Springer, New York, 2009.
- [2] Bather, J. A. Bayes procedures for deciding the sign of a normal mean. *Proc. Cambridge Philos. Soc.*, vol. 58, 1962, 599-620.
- [3] Bickel P. J. and Yahav, Y. A. On the Wiener process approximation to Bayesian sequential testing problems. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.*, vol. 1, 1972, 57-84.
- [4] Breakwell, J. and Chernoff, H. Sequential tests for the mean of a normal distribution. II. (Large t). *Ann. Math. Statist.*, vol. 35, 1964, 162-173.
- [5] Chernoff, H. Sequential tests for the mean of a normal distribution. *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, vol. 1, 1961, 79-91.
- [6] Chernoff, H. Sequential tests for the mean of a normal distribution III (small t). *Ann. Math. Statist.*, vol. 36, 1965, 28-54.
- [7] Chernoff, H. Sequential tests for the mean of a normal distribution IV (discrete case). *Ann. Math. Statist.*, vol. 36, 1965, 55-68.
- [8] Ekström, E. Properties of American option prices. *Stochastic Process. Appl.*, vol. 114, no. 2, 2004, 265-278.
- [9] Jacka, S. Optimal stopping and the American put. *Math. Finance*, vol. 1, 1991, 1-14.
- [10] Janson, S. and Tysk, J. Volatility time and properties of option prices. *Ann. Appl. Probab.*, vol. 13, no. 3, 2003, 890-913.
- [11] Karatzas, I. and Shreve, S. *Brownian motion and stochastic calculus. Second edition*. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.
- [12] Karatzas, I. and Shreve, S. *Methods of mathematical finance*. Applications of Mathematics, 39. Springer-Verlag, New York, 1998.
- [13] A. Klenke, *Probability Theory: A Comprehensive Course* Universitext. Springer-Verlag, London, 2008.
- [14] Lai, T. L. Nearly optimal sequential tests of composite hypotheses. *Ann. Statist.*, vol. 16, 1988, 856-886.
- [15] Lai, T. L. On optimal stopping problems in sequential hypothesis testing. *Statistica Sinica*, vol. 7, 1997, 33-51.



- [16] Øksendal, B. *Stochastic differential equations. An introduction with applications. Sixth edition.* Universitext. Springer-Verlag, Berlin, 2003.
- [17] Peskir, G. A change-of-variable formula with local time on curves. *J. Theoret. Probab.*, vol. 18, 2005, 499-535.
- [18] Peskir, G. On the American option problem. *Math. Finance*, vol. 15, no. 1, 2005, 169-181.
- [19] Peskir, G., Shiryaev, A. *Optimal stopping and free-boundary problems.* Lectures in Mathematics, ETH Zürich. Birkhäuser Verlag, Basel, 2006.
- [20] Shiryaev, A. N. Two problems of sequential analysis. *Cybernetics*, vol. 3, no. 2, 1967, 63-69.
- [21] Zhitlukhin, M., Muravlev, A. On Chernoff's Hypotheses Testing Problem for the Drift of a Brownian Motion, *Theory of Probability & Its Applications*, vol. 57, no. 4, 2013, 708-717.



# Optimal liquidation of an asset under drift uncertainty

Erik Ekström and Juozas Vaicenavicius

## Abstract

We study a problem of finding an optimal stopping strategy to liquidate an asset with unknown drift. Taking a Bayesian approach, we model any subjective initial beliefs of an individual about the drift parameter by allowing an arbitrary probability distribution to characterise the subjective uncertainty. 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 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 consider 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  $\tau^*$ , belonging to a set of admissible stopping times  $\mathcal{T}^T$ , such that

$$\mathbb{E}[S_{\tau^*}] = \sup_{\tau \in \mathcal{T}^T} \mathbb{E}[S_\tau]. \quad (1.1)$$

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  $S$ . Thus, we assume  $\mathcal{T}^T = \mathcal{T}_T^S$ .

In the context of the classical Black-Scholes model

$$dS_t = \alpha S_t dt + \sigma S_t dW_t, \quad (1.2)$$

where  $\alpha$ ,  $\sigma$  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  $\tau$  is optimal.

However, in applications, the known constant drift assumption is usually too strong. To obtain reasonable precision when estimating the drift one needs very long time-series, which are rarely available. 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. Furthermore, even in those few cases where enough past data is available, the benefit of accurate calibration of the historical drift is most likely to be outweighed by the model risk introduced. This is because most financial models, including Black-Scholes, are only plausible as short-term models; the simplistic constant parameter assumptions are non-viable over longer time periods. On the other hand, the assumption of known volatility parameter  $\sigma$  is justifiable as it can be estimated, at least in theory, from an arbitrarily short observation period.

Though the notoriety of the drift estimation pushed much of financial mathematics literature to focus on questions where the drift parameter can be avoided or at least does not play a crucial role (e.g. risk-neutral pricing and hedging), in the optimal liquidation problem, the drift can have a noticeable effect. Figure 1, containing the estimated model parameters of a few famous IPOs over the first year since going public, suggests that it is unlikely that the price change in all these cases was due to the volatility alone, leading us to believe in the significance of the drift contribution that 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

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

Admission that due to unattainable calibration, in many situations, modelling a price by a geometric Brownian motion with a known constant drift is ill-suited is not a reason to give up modelling, but a mere indication that the model should be improved to incorporate extra factors. As 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  $\alpha$  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  $\sigma$  is concerned, we stay with the known constant volatility assumption.

A conceivable practical application of this drift uncertainty modelling is in the optimal liquidation of an IPO share. A person possessing a share of an IPO has only beliefs about the drift of the price process as no past price data is available to calibrate the model. 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 to come up with such prior. One possible approach in the IPO example is to use the empirical

distribution of the returns of similar IPOs over the initial period of the same length as our investment horizon  $T$ . The similarity criteria could be the market sector, the country, the market share, 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 non-decreasing boundary that is the unique solution of a particular integral equation is shown to be optimal.

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 volatility 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 volatility 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 a unique non-positive right-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  $\sigma$  as well as the prior distribution. In addition, we 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 up to 10% for some feasible parameter regimes.

## 1.1 Literature review

Over the last three decades, investment problems with incomplete information about the drift has received much attention from both financial mathematicians and financial economists. Some of the most distinct works on portfolio optimisation include [5], regarded as the first incomplete information problem studied in financial literature, and the general

portfolio problems studied in [15, 16]; see also the recent article [3] 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 [17] 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 [22]).

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 the very restrictive case of a two-point prior. The optimal liquidation of an asset with unknown drift has been studied in [7] and of an asset with unknown jump intensity in [14]. For option valuation problems under incomplete information, see [4] and [10]. 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 various different formulations of the optimal selling problem in the case of complete information about the parameters have been studied in [6], [9], and [11].

## 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

$$dS_t = XS_t dt + \sigma S_t dW_t, \quad (2.1)$$

where the volatility  $\sigma > 0$  is a constant. We write  $\mathbb{F}^S = \{\mathcal{F}_t^S\}_{t \geq 0}$  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  $\mu$ , 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

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

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  $\mu$  is contained in  $[0, \infty)$ , then an optimal strategy is to stop at the terminal time  $T$ . Similarly, if the support of  $\mu$  is contained in  $(-\infty, 0]$ , then

an optimal strategy is to stop immediately. To exclude these trivial cases, we from now on impose the assumption that  $\mu((-\infty, 0)) \neq 0$  and  $\mu((0, \infty)) \neq 0$ .

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

$$d\tilde{S}_t = (X - r)\tilde{S}_t dt + \sigma\tilde{S}_t dW_t, \quad (2.3)$$

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  $\mu(\cdot + r)$ .

## 2.1 Equivalent reformulation under a measure change

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

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

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  $\{\sigma(\hat{W}_s : 0 \leq s \leq t)\}_{t \geq 0}$ , it is worth remarking 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

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

and writing

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

we have

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

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

$$V = \sup_{\tau \in \mathcal{T}_T^S} \mathbb{E}^{\mathbb{Q}} [e^{\int_0^\tau \hat{X}_s ds}]. \quad (2.4)$$

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  $S_t = S_0 e^{Y_t - \frac{\sigma^2}{2}t}$ , implying that 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  $q : \mathbb{R} \rightarrow \mathbb{R}$  satisfy  $\int_{\mathbb{R}} |q(u)| \mu(du) < \infty$ . Then*

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

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

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

and

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

for any  $t > 0$ , where

$$f(t, y) = \int_{\mathbb{R}} u \mu_{t,y}(du) = \frac{\int_{\mathbb{R}} u e^{\frac{2uY_t - u^2t}{2\sigma^2}} \mu(du)}{\int_{\mathbb{R}} e^{\frac{2uY_t - u^2t}{2\sigma^2}} \mu(du)}$$

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  $\mu$  is imposed.

**Assumption 3.2.** *The prior distribution  $\mu$  satisfies*

$$\int_{\mathbb{R}} e^{au^2} \mu(du) < \infty \quad (3.3)$$

for some  $a > 0$ .

This assumption is an insignificant restriction on our optimal liquidation problem, since, given any probability distribution  $\mu$ , the distributions  $\mu_{t,y}$  in (3.1) satisfy (3.3) 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  $\mu$  satisfies (3.3) with  $a = \epsilon/(2\sigma^2)$  for some  $\epsilon > 0$ . Defining a probability distribution  $\xi$  on  $\mathbb{R}$  by

$$\xi(\mathrm{d}u) := \frac{e^{\frac{\epsilon u^2}{2\sigma^2}} \mu(\mathrm{d}u)}{\int_{\mathbb{R}} e^{\frac{\epsilon u^2}{2\sigma^2}} \mu(\mathrm{d}u)}, \quad (3.4)$$

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  $\xi$  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  $I_\mu$  denote the interior of the smallest closed interval containing the support of  $\mu$ , i.e.  $I_\mu = (\inf(\text{supp}(\mu)), \sup(\text{supp}(\mu)))$ ,

**Lemma 3.3.** *For any given  $t \geq 0$ , the function  $f(t, \cdot) : \mathbb{R} \rightarrow I_\mu$  defined above 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  $f(0, y) \rightarrow \sup I_\mu$  as  $y \rightarrow \infty$  and  $f(0, y) \rightarrow \inf I_\mu$  as  $y \rightarrow -\infty$ . 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

$$\int_{\mathbb{R}} u e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) - \theta \int_{\mathbb{R}} e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) = \int_{\mathbb{R}} (u - \theta) e^{\frac{uy}{\sigma^2}} \mu(\mathrm{d}u) = e^{\frac{\theta y}{\sigma^2}} \int_{\mathbb{R}} w e^{\frac{wy}{\sigma^2}} \mu(\theta + \mathrm{d}w), \quad (3.5)$$

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,

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

as  $y \rightarrow \infty$  by monotone convergence. Consequently, from (3.5) follows that  $f(0, y) \geq \theta$  for all large enough  $y$ . Since  $\theta \in I_\mu$  was arbitrary, we conclude that  $f(0, y) \rightarrow \sup I_\mu$  as  $y \rightarrow \infty$ , which finishes the proof.  $\square$

Writing  $\mathbb{F}^{\hat{X}} = \{\mathcal{F}_t^{\hat{X}}\}_{t \geq 0}$  for the completion of the filtration generated by  $\hat{X}$  and  $\mathcal{T}_T^{\hat{X}}$  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

$$V = \sup_{\tau \in \mathcal{T}_T^{\hat{X}}} \mathbb{E}^{\mathbb{Q}}[e^{\int_0^\tau \hat{X}_s ds}]. \quad (3.6)$$

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

$$d\hat{X}_t = \sigma \partial_2 f(t, Y_t) d\hat{W}_t. \quad (3.7)$$

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

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

and, from (3.7) above, obtain a stochastic differential equation

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

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

$$d\hat{X}_t = \sigma \psi(t, \hat{X}_t) dt + \psi(t, \hat{X}_t) dZ_t. \quad (3.8)$$

The volatility  $\psi$  can be expressed more explicitly (by differentiating  $f$  under the integral sign) as

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

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  $\delta_l, \delta_h$  denote the Dirac measures at  $l, h \in \mathbb{R}$  respectively. Then  $\psi(t, x) = \frac{1}{\sigma} (h - x)(x - l)$ .

**Example (The normal prior)** Suppose  $\mu$  is the normal distribution with mean  $m$  and variance  $\gamma^2$ . 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 Volatility of the conditional mean

The following inequality will be the key to understanding the volatility function  $\psi$ .

**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.  $\square$

**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 volatility function  $\psi$ ).

1. *For any  $x \in I_\mu$ , the function  $t \mapsto \psi(t, x)$  is non-increasing. It is strictly decreasing unless  $\mu$  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  $\mu$  is a two-point distribution, in which case we have equality.*
3. *If  $\mu$  is compactly supported, then  $\psi$  is bounded.*

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

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

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

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

which substituted into the last expression above yields

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

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  $\psi$ , 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

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

by implicit differentiation. Hence

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

It remains to establish the inequality

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

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

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

Further differentiation yields that

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

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

### 3. The well-known identity

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

ensures that  $\psi$  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 volatility  $\psi$  is unbounded. For this, think of a discrete probability measure supported on an infinite number of points  $x_1 < x_2 < \dots < x_n < \dots$  such that  $x_n - x_{n-1} \rightarrow \infty$  as  $n \rightarrow \infty$ . Using the notation  $\bar{x}_n := (x_{n-1} + x_n)/2$  for the mean between neighbouring points, the value  $\psi(t, \bar{x}_n) = \text{Var}_{t, \bar{x}_n}(X)/\sigma \rightarrow \infty$  as  $n \rightarrow \infty$  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  $\psi$ . For instance, we know that  $\psi$  is bounded in the case of a normal prior as seen in the example on page 8. Though a rigorous investigation into precise technical conditions on the prior for the boundedness of  $\psi$  appears to be involved enough to be omitted in this article, we conjecture, based on numerical investigations, that  $\psi$  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  $\psi$  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  $\mu$  is such that  $\text{Var}_{0,y}(X) < \infty$  for all  $y \in \mathbb{R}$ .*

### 3.3 The Markovian value function and the optimal strategy

Using the dynamics (3.8) of  $\hat{X}$ , we are able to embed the optimal stopping problem (3.6) into a Markovian framework. To do that, define

$$v(t, x) = \sup_{\tau \in \mathcal{T}_{T-t}} \mathbb{E}^{\mathbb{Q}} \left[ e^{\int_0^\tau \hat{X}_{t+s}^{t,x} ds} \right], \quad (3.9)$$

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

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

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

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

$$\tau_{\mathcal{D}} := \inf\{s \geq 0 : (t+s, \hat{X}_{t+s}^{t,x}) \in \mathcal{D}\} \quad (3.10)$$

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 (3.10) 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 [13], to prove the claims it suffices to show that

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

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

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

If  $m > 0$  is a constant dominating  $\psi$ , then

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

Thus

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

where the equality comes from the reflection principle.  $\square$

Since  $\psi$  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 the slightly stronger assumption of Lipschitz continuity on the whole of  $[0, T] \times I_\mu$ .

**Assumption 3.9.** *The function  $\psi$  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 function  $h : [0, T] \rightarrow (-\infty, 0]$  that is non-decreasing, right-continuous with left limits, and satisfying  $\mathcal{C} = \{(t, x) \in [0, T] \times I_\mu : x > h(t)\}$ .
5. The value function  $(t, x) \mapsto v(t, x)$  solves the free-boundary problem

$$\begin{cases} \partial_1 v + \sigma \psi(t, x) \partial_2 v + \frac{1}{2} \psi(t, x)^2 \partial_2^2 v + xv = 0, & x \in \mathcal{C}, \\ v = 1, & x \in \mathcal{D}. \end{cases} \quad (3.11)$$

Furthermore, the smooth-fit property holds in that the function  $x \rightarrow v(t, x)$  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 of the value function

$$v(t, x) = \sup_{\tau \in \mathcal{T}_{T-t}} \mathbb{E}^{\mathbb{Q}} \left[ e^{\int_0^\tau \hat{X}_{t+s}^{t,x} ds} \right] \quad (3.12)$$

together with a comparison theorem, see [21, Theorem IX.3.7].

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

$$d\hat{R}_t = -\sigma \psi(t, -R_t) dt - \psi(t, -\hat{R}_t) dZ_t.$$

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 [8, 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 volatility  $\psi$  is decreasing in  $t$ , so the claim follows by the Bermudan approximation argument for the value function and Theorem 6.1 in [8].
3. First, let  $l > 0$  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  $(-\infty, l]$ . 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  $\{t_n\}_{n \geq 0} \subset [0, 1]$  such that the sequence of left-derivatives  $\partial_2^- v(t_n, l)$  diverges to  $\infty$ . However, this would imply that  $v(t_n, l+1) \rightarrow \infty$ , contradicting the fact that  $v(t_n, l+1) \leq v(0, l+1) < \infty$ .
- 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

$$\inf_{(t,x) \in \mathcal{R}} v(t, x) > v(t_0, x_0 + \delta). \quad (3.13)$$

Thus  $\mathcal{R} \subseteq \mathcal{C}$ . Let  $t \in (t_0 - \delta, t_0)$  and  $\tau_{\mathcal{R}} := \inf\{u \geq t : \hat{X}^{t,x_0} \notin \mathcal{R}\}$ . Then

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

as  $t \rightarrow t_0$ , which contradicts (3.13).

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

$$\sup_{(t,x) \in \mathcal{R}} v(t, x) < v(t_0, x_0). \quad (3.14)$$

In particular,  $\mathcal{R} \subseteq \mathcal{C}$  and we have

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

as  $\epsilon \searrow 0$ , which contradicts (3.14).

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

$$\inf_{x \in (x_0 - \delta, x_0)} v(t_0, x) > v(t_0+, x_0) = 1. \quad (3.15)$$

Thus  $(t_0, T] \times (x_0 - \delta, x_0) \subseteq \mathcal{D}$  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 (3.15).

4. Existence of a non-decreasing boundary  $h : [0, T] \rightarrow [-\infty, \infty]$  satisfying  $\mathcal{C} = \{(t, x) \in [0, T] \times I_\mu : x > h(t)\}$  is a direct consequence of the first two parts above. Non-positivity of  $h$  is clear from the expression (3.9), 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  $\xi$  as in (3.4), 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  $\xi$ . Writing  $v_\xi$  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

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

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  $t \in [0, T]$ .

Finally, note that continuity of the value function implies that  $h$  is right-continuous with left limits.

5. The proof of (3.11) is along standard lines (e.g. see [13, Theorem 7.7 in Chapter 2]), so we do not include it here.

Let us next establish the smooth-fit property. 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  $\tau_\epsilon$ , we have

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

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  $\tau_\epsilon \rightarrow 0$  a.s. as  $\epsilon \searrow 0$ . Hence, by the dominated convergence theorem,

$$\mathbb{E}^{\mathbb{Q}}[\tau_\epsilon e^{2 \int_0^{\tau_\epsilon} \hat{X}_u^{0,x+\epsilon} du}] \rightarrow 0 \quad \text{as } \epsilon \searrow 0$$

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

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

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

To this end,

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

where  $c$  is a constant dependent on  $T$ ,  $\sigma$ , and the Lipschitz constant of  $\psi$ . In the above, the first inequality comes from Jensen's inequality, 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.  $\square$

**Remark** Although we expect the stopping boundary  $h$  to be continuous, to push through a standard proof (e.g. see p. 382 of [19]) we need that for every  $t_0 \in (0, T)$  there is a constant  $d > 0$  such that  $\partial_2^2 v \geq d$  on  $\mathcal{C} \cap [0, t_0) \times (-\infty, 0]$ , a condition that we have not been able to verify. However, just a slightly weaker condition  $\partial_2^2 v \geq 0$  holds due to convexity of the value function in the second variable.

## 4 An integral equation for the boundary

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

$$\mathbb{E}^{\mathbb{Q}} \left[ e^{\int_t^T \hat{X}_{t+u}^{t,h(t)} du} \right] = 1 + \int_t^T \mathbb{E}^{\mathbb{Q}} \left[ e^{\int_t^s \hat{X}_u du} \hat{X}_s^{t,h(t)} 1_{\{\hat{X}_s^{t,h(t)} \leq h(s)\}} \right] ds \quad (4.1)$$

*in the class of non-positive right-continuous functions.*

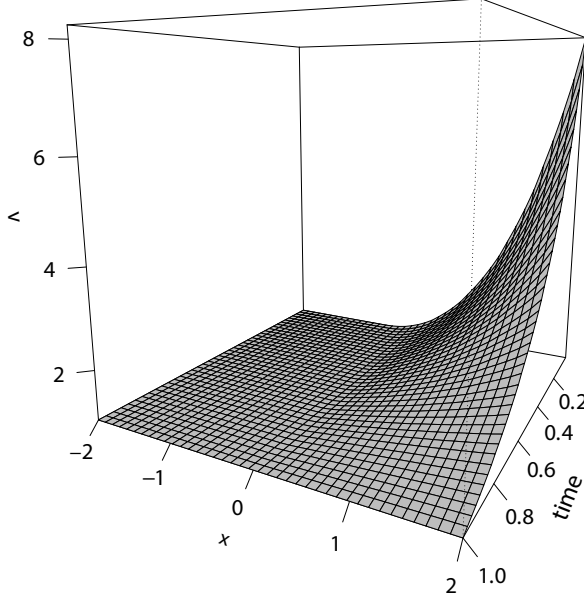


Figure 2: The value function  $v(t, x)$  in the case of a normal prior with standard deviation  $\gamma = 0.5$ , the market volatility  $\sigma = 0.2$ .

*Proof.* An application of Itô's formula (more precisely, its extension proved in [18], which can be applied thanks to the monotonicity of  $h$ ) to  $v(s, \hat{X}_s^{t,x})e^{\int_t^s \hat{X}_u^{t,x} du}$  yields

$$\begin{aligned} v(s, \hat{X}_s^{t,x})e^{\int_t^s \hat{X}_u^{t,x} du} &= v(t, \hat{X}_t^{t,x}) + \int_t^s e^{\int_t^r \hat{X}_u^{t,x} du} \left( \mathcal{L}_{\hat{X}^{t,x}} v(r, \hat{X}_r^{t,x}) + \hat{X}_r^{t,x} v(r, \hat{X}_r^{t,x}) \right) dr \\ &\quad + \int_t^s e^{\int_t^r \hat{X}_u^{t,x} du} \psi(r, \hat{X}_r^{t,x}) \partial_2 v(r, \hat{X}_r^{t,x}) dZ_r. \end{aligned} \quad (4.2)$$

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

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

from (4.2) we get

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

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

$$\mathbb{E}^{\mathbb{Q}} [v(T, \hat{X}_T^{t,x}) e^{\int_t^T \hat{X}_u^{t,x} du}] = v(t, x) + \int_t^T \mathbb{E}^{\mathbb{Q}} \left[ e^{\int_t^r \hat{X}_u^{t,x} du} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \leq h(r)\}} \right] dr.$$

Here, the left-hand side is obtained by dominated convergence as  $v(T \wedge \tau_n, \hat{X}_{T \wedge \tau_n}^{t,x}) e^{\int_t^{T \wedge \tau_n} \hat{X}_u^{t,x} du}$  is dominated by  $e^{2T(\sup_{t \leq u \leq T} \hat{X}_u^{t,x} \vee 0)}$ , which is integrable; the right-hand side comes from

monotone convergence. Substitution  $x = h(t)$  in (4.3) gives

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

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

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

$$\tilde{v}(t, x) := \mathbb{E}^{\mathbb{Q}}[e^{\int_t^T \hat{X}_u^{t,x} du}] - \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{\int_t^r \hat{X}_u^{t,x} du} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \leq k(r)\}} dr \right].$$

The two processes defined for  $s \in [t, T]$  as

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

and

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

are easily verified to be  $\mathbb{Q}$ -martingales using the Markov property.

**Claim 1:**  $\tilde{v}(t, x) = 1$  for  $x \leq k(t)$ .

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

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

where the first equality holds by the definitions of  $\tilde{v}$  and  $M^{\tilde{v}}$ , the second by the martingality of  $M^{\tilde{v}}$ , the third by the optional sampling theorem, and the last one by (4.1).

**Claim 2:**  $\tilde{v} \leq v$ .

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

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

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.

**Claim 3:**  $h \leq k$ .

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

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

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 right-continuity and the non-positivity of  $k$  and  $h$ . Hence we have obtained a contradiction. Claim 4:  $k \leq h$ .

Assume for a contradiction that  $k(t) > h(t)$ . Let  $x \in (h(t), k(t))$  and define  $\tau_h := \inf\{s \geq 0 : \hat{X}_{t+s}^{t,x} \leq h(t+s)\} \wedge (T-t)$ . Recall that

$$\mathbb{E}^{\mathbb{Q}} \left[ e^{\int_t^{t+\tau_h} \hat{X}_u^{t,x} du} \right] = v(t, x),$$

though, alternatively,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[ e^{\int_t^{t+\tau_h} \hat{X}_u^{t,x} du} \right] &= \mathbb{E}^{\mathbb{Q}} \left[ \tilde{v}(t + \tau_h, \hat{X}_{t+\tau_h}^{t,x}) e^{\int_t^{t+\tau_h} \hat{X}_u^{t,x} du} \right] \\ &= \tilde{v}(t, x) + \mathbb{E}^{\mathbb{Q}} \left[ \int_t^{t+\tau_h} e^{\int_t^r \hat{X}_u^{t,x} du} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \leq k(r)\}} dr \right]. \end{aligned}$$

Equating these different expressions, we get that

$$\mathbb{E}^{\mathbb{Q}} \left[ \int_t^{t+\tau_h} e^{\int_t^r \hat{X}_u^{t,x} du} \hat{X}_r^{t,x} \mathbb{1}_{\{\hat{X}_r^{t,x} \leq k(r)\}} dr \right] = v(t, x) - \tilde{v}(t, x).$$

In this equality, the left hand side is strictly negative, while the right hand side is non-negative by Claim 2 above; a contradiction.

To conclude, the last two claims imply  $k = h$  giving us the uniqueness result we set out to prove. □

## 5 Parameter dependence

### 5.1 Dependence of the value function on the market volatility

A large volatility  $\sigma$  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  $\sigma$ . While a full proof of this intuitive remark appears to be challenging, we have the following sufficient condition which guarantees monotonicity in  $\sigma$ .

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

*Proof.* If  $\sigma\psi(t, x)$  is non-increasing in  $\sigma$ , then both the drift term and the diffusion term of  $\hat{X}$  are non-increasing in  $\sigma$ . Therefore, Theorem 6.1 from [8] can be applied to prove that the value function  $v$  is decreasing in  $\sigma$ . □

**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  $\sigma$ .

**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  $\sigma$ . 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  $\sigma$  lies in the fact that it is not true in general that the Markovian value function  $v$  in (3.9) is decreasing in  $\sigma$ . We can see this in the case of a normal prior in Figure 3. The picture depicts the difference between two Markovian value functions for the same normal prior with standard deviation  $\gamma = 0.5$ , but different volatilities  $\sigma$ . Nevertheless, the same picture shows that at time  $t = 0$ , the difference is positive, so conforming with our intuitive conjecture.

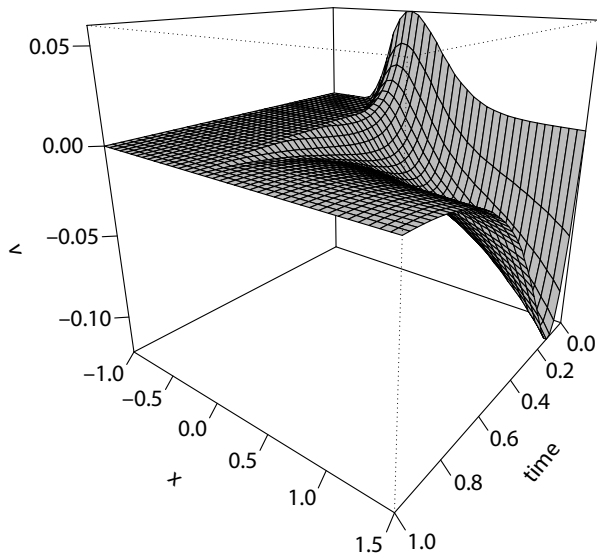


Figure 3: The difference  $v_{0.2} - v_{0.5}$  between two value functions;  $v_{0.2}$  and  $v_{0.5}$  denote the value functions in the cases of the market volatility  $\sigma$  being equal to 0.2 and 0.5, respectively.

As far as the optimal stopping boundaries are concerned, the lack of monotonicity of the Markovian value function in the volatility  $\sigma$  manifests in that the stopping boundaries for different values of  $\sigma$  may intersect. An example of this appears in Figure 4. The same graph also provides intuition about how the shape of the boundary changes as one varies the parameter  $\sigma$ . In particular, we get an impression what boundary to expect as  $\sigma$  approaches zero or grows to infinity.

## 5.2 Dependence of the value function on the initial prior

**Theorem 5.2.** *Assume that  $\mu_1$  and  $\mu_2$  are two prior distributions such that the corresponding volatilities  $\psi_1$  and  $\psi_2$  satisfy  $\psi_1(t, x) \leq \psi_2(t, x)$  for all  $(t, x) \in [0, T] \times \mathbb{R}$ . Then the corresponding value functions  $v_1$  and  $v_2$  satisfy  $v_1 \leq v_2$ .*

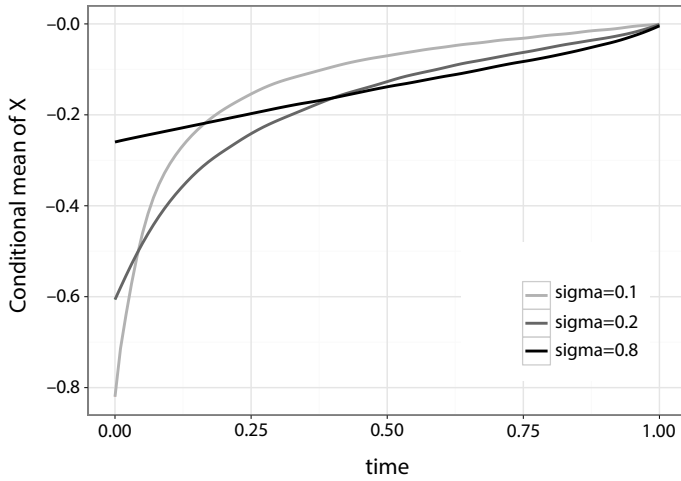


Figure 4: Optimal stopping boundaries for different values of market volatility  $\sigma$  in the case a normal prior with standard deviation  $\gamma = 0.5$ .

*Proof.* Again, Theorem 6.1 from [8] can be applied to prove that the value function  $v$  is increasing in  $\psi$ .  $\square$

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  $\gamma$  of the prior. Hence Theorem 5.2 applies and the Markovian value function  $v$  increases in  $\gamma$ . A consequence of this is that optimal stopping boundaries are ordered by the size of  $\gamma$  as shown in Figure 5.

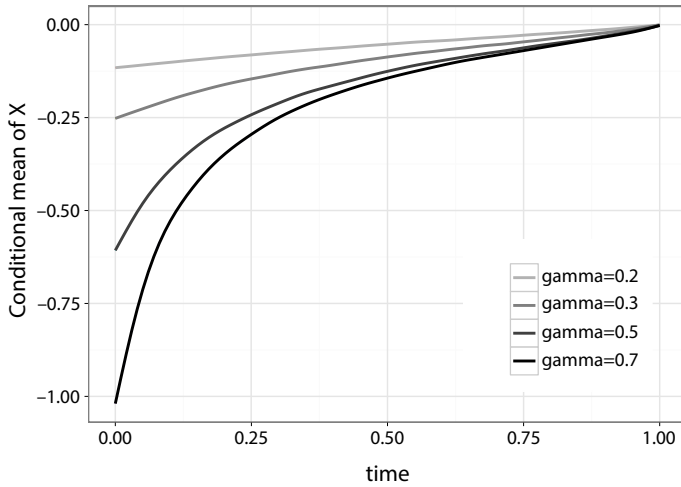


Figure 5: Optimal stopping boundaries for different values of standard deviation  $\gamma$  in the case of a normal prior when the market volatility  $\sigma = 0.2$ .

For compactly supported distributions, Theorem 5.2 offers a way to construct a lower bound of the Markovian value function  $v$ . Suppose the prior  $\mu$  is a compactly supported distribution. Since  $\psi_\mu$  is bounded, by the two-point prior example on page 8, we can find a two-point distribution  $\eta := (1 - \pi)\delta_a + \pi\delta_b$  with  $\int_{\mathbb{R}} u\mu(du) = \int_{\mathbb{R}} u\eta(du)$  such that  $\psi_\mu \leq \psi_\eta$ . Then Theorem 5.2 yields that  $v_\mu \leq v_\eta$  and so the stopping boundaries satisfy



$h_\mu \leq h_\eta$ . As a result,  $V_\mu \geq \mathbb{E}[S_{\tau_{h_\eta}}]$ , where  $V_\mu$  denotes the initial value for the prior  $\mu$ . Thus the described procedure provides a lower bound for the value function and a strategy to achieve it.

### 5.3 The value of filtering: numerical investigation

Having introduced and solved the optimal liquidation problem for an arbitrary prior, a pragmatic question arises - how much is there to be gained from the elaborate sequential liquidation strategy with real-time filtering in comparison to a naive optimal selling strategy without filtering? In this part, we address the question within our model from a numerical point of view in the normal prior case.

Let us consider an agent who wants to liquidate an asset evolving according to (2.1) before time  $T$ ; the agent's prior  $\mu$  for the drift  $X$  is the normal distribution  $N(m, \gamma)$ . If the agent does not know about the possibility of real-time filtering, he does not utilise any valuable information from the asset price observations and so, at time 0, will make a decision whether to sell immediately, i.e. at time 0, or liquidate at time  $T$ . The expected value from an optimal liquidation strategy with selling allowed only at times 0 and  $T$  is

$$V_{\{0,T\}} := \mathbb{E}[e^{X_T}] \vee 1 = e^{mT + \frac{1}{2}(\gamma T)^2} \vee 1.$$

However, if the agent is aware of the optimal sequential liquidation procedure involving filtering, the expected value from an optimal selling is

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

In Figure 6, we see the two values  $V_{\{0,T\}}$  and  $V$  calculated for two different priors at a range of different market volatilities. In addition, Figure 7 depicts the percentage improvement  $(V - V_{\{0,T\}})/V_{\{0,T\}}$  that the sequential procedure with filtering brings over the simple strategy. Dependence of the two values on the standard deviation  $\gamma$  of the normal prior is illustrated by Figure 8.

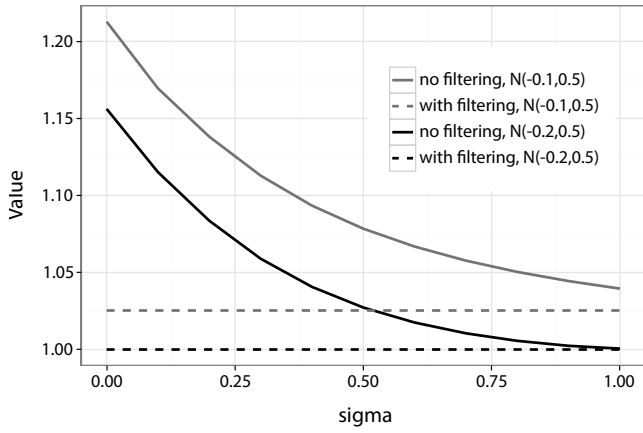


Figure 6: **Initial value as a function of market volatility.** The solid gray curve corresponds to the optimal liquidation value and the dashed gray line to the value without filtering - both for the normal prior  $N(-0.1, 0.5)$ . Similarly, the solid black curve corresponds to the optimal liquidation value and the dashed black line to the value without filtering for the normal prior  $N(-0.2, 0.5)$ .

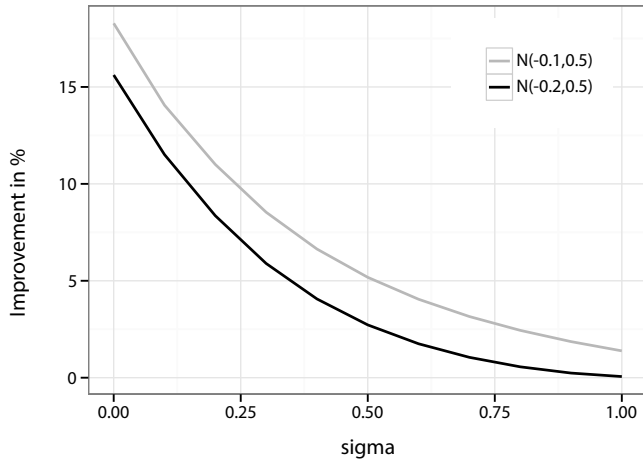


Figure 7: **Improvement due to filtering.** The percentage improvement  $(V - V_{\{0,T\}})/V_{\{0,T\}}$  over the strategy without filtering.

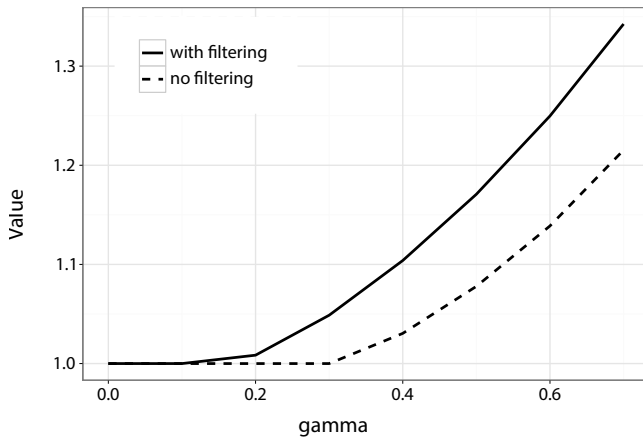


Figure 8: **Initial value as a function of standard deviation of the normal prior.** The solid curve corresponds to the optimal liquidation value while the dashed line to the value without filtering - both for the normal prior with mean  $-0.05$ ; the market volatility  $\sigma = 0.2$ .

## 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] 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,
- [4] Décamps, J.-P., Mariotti, T., Villeneuve, S. Investment timing under incomplete information. Math. Oper. Res., vol. 30, no. 2, 2005, 472-500.
- [5] 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.
- [6] du Toit, J., Peskir, G. Selling a stock at the ultimate maximum. Annals of Applied Probability, vol. 19, no. 3, 2009, 983-1014.
- [7] Ekström, E., Lu, B. Optimal Selling of an Asset under Incomplete Information. International Journal of Stochastic Analysis, vol. 2011, 2011, ID 543590.
- [8] Ekström, E., Tysk, J. Convexity theory for the term structure equation. Finance and Stochastics. vol. 12, no. 1, 2008, 117-147.
- [9] Evans, J., Henderson, V., Hobson, D. Optimal timing for an indivisible asset sale. Mathematical Finance, vol. 18, no. 4, 2008, 545-567.
- [10] 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.
- [11] Henderson, V., and 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.
- [12] Karatzas, I., Shreve, S. *Brownian motion and stochastic calculus. Second edition*. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.
- [13] Karatzas, I., Shreve, S. *Methods of Mathematical Finance*. Applications of Mathematics, 39. Springer-Verlag, New York, 1998.
- [14] Lu, B. Optimal Selling of an Asset with Jumps Under Incomplete Information. Applied Mathematical Finance, vol. 20, no. 6, 2013, 599610.
- [15] Lakner, P. Utility maximization with partial information. Stochastic Processes and their Applications, vol. 56, no. 2, 1995, 247-249.

- [16] 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.
- [17] 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.
- [18] Peskir, G. A change-of-variable formula with local time on curves. *Journal of Theoretical Probability*, vol. 18, no. 3, 2005, 499-535.
- [19] Peskir, G., Shiryaev, A. *Optimal stopping and free-boundary problems*. Lectures in Mathematics, ETH Zürich. Birkhäuser Verlag, Basel, 2006.
- [20] Protter, P. E. *Stochastic Integration and Differential Equations. Version 2.1*. Stochastic Modelling and Applied Probability, 21. Springer, 2005.
- [21] Revuz, D., Yor, M. *Continuous martingales and Brownian motion. Third edition*. Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, 1999.
- [22] Ziegler, A. *Incomplete Information and Heterogeneous Beliefs in Continuous- Time Finance*. Springer, 2003.