



Finite-horizon optimal investment with transaction costs: construction of the optimal strategies

Christoph Belak¹ · Jörn Sass²

Received: 17 March 2016 / Accepted: 14 June 2019 / Published online: 5 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract We revisit the problem of maximising expected utility of terminal wealth in a Black–Scholes market with proportional transaction costs. While it is known that the value function of this problem is the unique viscosity solution of the HJB equation and that the HJB equation admits a classical solution on a reduced state space, it has been an open problem to verify that these two coincide. We establish this result by devising a verification procedure based on superharmonic functions. In the process, we construct optimal strategies and provide a detailed analysis of the regularity of the value function.

Keywords Utility maximisation · Transaction costs · Reflected diffusions · Superharmonic functions

Mathematics Subject Classification (2010) 93E20 · 35R35 · 91G80

JEL Classification G11 · C61

1 Introduction

The aim of this paper is to solve the problem of maximising expected utility of terminal wealth for an investor facing proportional transaction costs in a Black–Scholes

✉ C. Belak
belak@math.tu-berlin.de

J. Sass
sass@mathematik.uni-kl.de

¹ Faculty II – Mathematics and Natural Sciences, Institute of Mathematics, Technische Universität Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany

² Department of Mathematics, University of Kaiserslautern, Erwin-Schrödinger Strasse, 67653 Kaiserslautern, Germany

market. While this problem has been studied extensively in the literature, it is still an open problem in the finite-horizon case to construct optimal strategies and verify their optimality. The aim of this paper is to close this gap.

Starting with the seminal contribution of Magill and Constantinides [15], the continuous-time optimal investment problem with proportional transaction costs has been studied extensively over the last decades, and three different approaches to tackle this type of problem have emerged: (1) The primal approach based on stochastic control and viscosity solution theory, in which one studies the Hamilton–Jacobi–Bellman (HJB) equation of the problem; (2) the dual approach based on shadow prices, in which one determines an auxiliary frictionless market with unfavourable price processes yielding the same optimal strategy as the original problem; and (3) asymptotics for vanishing costs.

For the problem of optimal consumption over an infinite horizon, the primal approach was utilised, among others, by Davis and Norman [23], Shreve and Soner [37], Akian et al. [1], Kabanov and Klüppelberg [32] and de Vallière et al. [25], whereas Kallsen and Muhle-Karbe [34], Choi et al. [13] and Herczegh and Prokaj [29] used the dual approach to solve this problem. Asymptotic optimality results were obtained by Janeček and Shreve [31] and Gerhold et al. [27]. Moreover, Akian et al. [3] use the primal approach and Gerhold et al. [28] and Gerhold et al. [26] the dual approach to determine closed-form solutions for the problem of maximising the asymptotic growth rate under (small) transaction costs.

In the present paper, we focus on the finite-horizon optimal terminal wealth problem without intermediate consumption in a Black–Scholes market, which was introduced in Akian et al. [2]. The HJB equation of this problem has also been studied in Davis et al. [24] in the context of utility indifference pricing. An adaptation of the results of Davis et al. [24] implies that the value function is a viscosity solution of the HJB equation, and uniqueness holds in the case of bounded utility functions. Belak et al. [9] extend the uniqueness result to more general utility functions including log and power utility. Dai and Yi [21] show that the HJB equation admits a classical solution in the case of log and power utility if the state space is reduced to positive stock holdings, and Dai et al. [20] and Chen et al. [12] extend this result to the problem with intermediate consumption and CARA utility, respectively. Moreover, Czichowsky et al. [19] and Czichowsky and Schachermayer [16, 17, 18] use the shadow price approach to establish existence of optimal strategies for general price processes extending beyond semimartingales. Numerical schemes in the Black–Scholes setting can be found in Kunisch and Sass [36], Dai and Zhong [22] and Herzog et al. [30]. Finally, Bichuch [11] in the Black–Scholes setting and Kallsen and Muhle-Karbe [35] and Kallsen and Li [33] for more general price processes solve the finite-horizon problem asymptotically for small transaction costs. Summing up the results on the finite-horizon problem, it is known that

- 1) the value function \mathcal{V} is the unique viscosity solution of the HJB equation;
- 2) there exists a classical solution V of the HJB equation if the state space is reduced to positive stock holdings; and
- 3) there exists a frictionless market in which the optimal strategy coincides with the optimal strategy in the transaction costs market.

It is, however, not immediate that the classical solution V coincides with the value function \mathcal{V} , nor if the optimal strategy obtained from the dual approach is a reflected diffusion in the no-trading region implied by the HJB equation. More precisely:

1) V is a classical solution on the reduced state space, whereas the value function \mathcal{V} is a viscosity solution on the entire state space. While it is not particularly difficult to verify that the classical solution V is also a viscosity solution on the reduced state space, this does not imply that $V = \mathcal{V}$ on the reduced state space since the existing literature only provides uniqueness results for viscosity solutions on the entire state space. Thus, in order to rigorously conclude that $V = \mathcal{V}$, one either needs to (a) prove uniqueness of viscosity solutions on the reduced state space, or (b) show that V extends to a viscosity solution on the entire state space. In both cases, a careful inspection of the behaviour of V and \mathcal{V} at the boundary of the reduced state space is necessary.

2) No link between the auxiliary frictionless market and the HJB equation is known. Hence, while existence of an optimal strategy is guaranteed by the dual approach, it is an open question whether optimal strategies are determined by the trading regions implied by the HJB equation. In particular, without establishing this link, it is not clear if the trading regions obtained from solving the HJB equation numerically determine an optimal strategy.

3) The classical approach of verifying optimality of candidate strategies via the primal approach requires a sufficiently smooth value function \mathcal{V} to justify the application of Itô's formula along *all* controlled state processes (and hence on the *entire* state space). We shall see, however, that the value function is not of class $C^{1,2}$ everywhere (see Theorem 4.14 for the precise statement), and thus there is a need for a verification argument requiring less regularity on \mathcal{V} .

To show that V and \mathcal{V} coincide and that optimal strategies are determined by the HJB equation, we devise a novel verification procedure which only requires to evaluate the candidate value function along the uncontrolled state process. Since the uncontrolled state process naturally avoids states in which the HJB equation degenerates (i.e., regularity fails), this puts Itô's formula at our disposal to establish the claims. For this, we characterise the value function as the smallest continuous function which is superharmonic with respect to the uncontrolled state process and nondecreasing in the direction of transactions. More precisely, we proceed as follows:

1) We define trading regions in terms of the classical solution V of the HJB equation and show that for every initial state, there exists a trading strategy which turns the corresponding state process into a diffusion reflected at the boundaries between the trading regions.

2) We define a function h_0 which maps the initial state to the expected utility obtained by following the trading strategy constructed in Step 1) and show that h_0 is superharmonic and nondecreasing in the direction of transactions. A simple argument shows that h_0 coincides with the classical solution V on the reduced state space, and by construction, h_0 is dominated by the value function \mathcal{V} .

3) We argue that every superharmonic function which is nondecreasing in the direction of transactions is a viscosity supersolution of the HJB equation. By the previous step and the comparison principle in Belak et al. [9], this implies that h_0 dominates the value function. By Step 2), this shows that h_0 and \mathcal{V} coincide, yielding that

\mathcal{V} and V are equal on the reduced state space and the trading strategies constructed in Step 1) are optimal.

In particular, as we characterise the value function as the smallest superharmonic function, our approach may be seen as an alternative duality theory for singular control problems. Moreover, since our approach naturally avoids points of singularity of the infinitesimal generator of the underlying stochastic process (which are typically the points at which it is difficult to verify regularity of the value function), it is conceivable that the verification argument can be applied to other singular control problems as well.

Our verification argument is inspired by recent results of Christensen [14] and Belak et al. [8] in the context of stochastic impulse control. However, since optimal trading regions for singular control problems are given in terms of first-order derivatives of the value function, as opposed to the value function itself for impulse control problems, the mathematical analysis in the present paper differs significantly from the corresponding impulse control results.

Moreover, our superharmonic function approach can be seen as a version of the stochastic Perron method; see Bayraktar and Sirbu [4, 5, 6] for early developments and Bayraktar and Zhang [7] for the case of a singular control problem with transaction costs. In contrast to the stochastic Perron method, we require the superharmonicity property along the uncontrolled state process together with the monotonicity in the direction of transactions, whereas for the stochastic Perron method, one would typically ask for superharmonicity along every controlled state process (or at least a subset of state processes containing a maximising sequence). As a consequence, it is more involved in our setting to argue that the superharmonic functions dominate the value function (we rely on viscosity arguments for this, whereas it is immediate in the setting of [7]). On the other hand, our definition makes verification significantly easier since we only need to verify superharmonicity for the uncontrolled state process. We note, however, that our setup implies superharmonicity along any state processes obtained from a piecewise constant strategy (see Remark 4.2 below) and hence the two concepts coincide as soon as there exists a maximising sequence consisting of piecewise constant strategies.

The remainder of this article is structured as follows. In Sect. 2, we set up the market model, recall existing results from the literature, and discuss implications of our results. In Sect. 3, we construct the candidate optimal strategies as well as the corresponding reflected diffusions. Our main results can be found in Sect. 4, where we present the verification theorem to show that these candidate optimal strategies are indeed optimal, analyse the regularity of the value function in detail, and prove that the classical solution of Dai and Yi [20] coincides with the value function on the reduced state space.

2 Market model and problem formulation

2.1 The market model

We let $W = (W(t))_{t \geq 0}$ be a standard Brownian motion defined on the canonical Wiener space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $t \geq 0$, we denote the augmented filtration gen-

erated by $(W(u) - W(t))_{u \geq t}$ by $\mathbb{F}^t = (\mathcal{F}^t(u))_{u \geq t}$ and set $\mathbb{F} := \mathbb{F}^0$. Moreover, we fix some terminal time $T > 0$ as well as some initial time $t \in [0, T)$.

We consider a Black–Scholes market $(P^0, P^1) = (P^0(u), P^1(u))_{u \in [t, T]}$ with

$$\begin{aligned} dP^0(u) &= 0, & u \in [t, T], & P^0(t) = 1, \\ dP^1(u) &= \alpha P^1(u) du + \sigma P^1(u) dW(u), & u \in [t, T], & P^1(t) = 1. \end{aligned}$$

Here, $\alpha > 0$ and $\sigma > 0$ denote the excess return and volatility of the stock, respectively. With this, we assume that the investor buys shares of the stock at the ask price $(1 + \lambda)P^1$, where $\lambda > 0$, and sells shares of the stock at the bid price $(1 - \mu)P^1$, where $\mu \in (0, 1)$.

Next, to model trading strategies, we take \mathbb{F}^t -adapted, nondecreasing, càdlàg processes $L = (L(u))_{u \in [t, T]}$ and $M = (M(u))_{u \in [t, T]}$ with $L(t-) = M(t-) = 0$. Here, L and M represent the cumulative units of money used for purchases and sales of the stock, respectively. With this, we denote by $B = B_{t,b}^{L,M} = (B_{t,b}^{L,M}(u))_{u \in [t, T]}$ and $S = S_{t,s}^{L,M} = (S_{t,s}^{L,M}(u))_{u \in [t, T]}$ the investor’s wealth invested in the bond and the stock, respectively. Assuming that the strategy (L, M) is self-financing, the evolution of B and S can be written as

$$\begin{aligned} dB(u) &= -(1 + \lambda) dL(u) + (1 - \mu) dM(u), & u \in [t, T], \\ dS(u) &= \alpha S(u) du + \sigma S(u) dW(u) + dL(u) - dM(u), & u \in [t, T], \end{aligned}$$

where the initial values are given by $B(t-) = b$ and $S(t-) = s$, respectively. The net wealth $X = X_{t,b,s}^{L,M} = (X_{t,b,s}^{L,M}(u))_{u \in [t, T]}$ of the investor after liquidation of the stock position is then given by

$$X(u) := B(u) + (1 - \mu)S(u)\mathbb{1}_{\{S(u) > 0\}} + (1 + \lambda)S(u)\mathbb{1}_{\{S(u) \leq 0\}}, \quad u \in [t, T].$$

We say that a trading strategy is *admissible* if the corresponding net wealth process is nonnegative. For this, we define the solvency cone

$$\mathcal{S} := \{(b, s) \in \mathbb{R}^2 : b + (1 + \lambda)s > 0, b + (1 - \mu)s > 0\}.$$

With this, whenever $(b, s) \in \overline{\mathcal{S}}$, the investor can liquidate her stock holdings to end up with nonnegative wealth. A trading strategy (L, M) is therefore admissible for an initial position $(b, s) \in \overline{\mathcal{S}}$ if the corresponding pair $(B_{t,b}^{L,M}, S_{t,s}^{L,M})$ takes values in $\overline{\mathcal{S}}$. The set of all admissible trading strategies of this form is denoted by $\mathcal{A}(t, b, s)$.

The objective of the investor is to maximise expected utility of the net terminal wealth after liquidation, i.e.,

$$\mathcal{V}(t, b, s) := \sup_{(L, M) \in \mathcal{A}(t, b, s)} \mathbb{E}[U_p(X_{t,b,s}^{L,M}(T))], \tag{2.1}$$

where the utility function $U_p : (0, \infty) \rightarrow \mathbb{R}$ is defined as

$$U_p(x) := \begin{cases} x^p/p & \text{if } p < 1, p \neq 0, \\ \log x & \text{if } p = 0. \end{cases}$$

We extend U_p to $[0, \infty)$ by setting $U_p(0) := \lim_{x \downarrow 0} U_p(x)$.

2.2 Overview of existing results

As pointed out in Sect. 1, the portfolio problem defined in (2.1) has received considerable interest in the past. In this section, we briefly summarise the results which will be needed in the sequel.

First, it is easy to see that the value function \mathcal{V} defined in (2.1) is finite on $[0, T] \times \mathcal{S}$. More specifically, we have (see Belak et al. [9])

$$U_p(b + \min\{(1 - \mu)s, (1 + \lambda)s\}) \leq \mathcal{V}(t, b, s) \leq \varphi_p(t, b, s),$$

where for $\gamma \in (1 - \mu, 1 + \lambda)$ and $K > 1$, the function $\varphi_p : [0, T] \times \overline{\mathcal{S}} \rightarrow \mathbb{R}$ is given by

$$\varphi_p(t, b, s) := U_p((b + \gamma s)f(t))$$

with $f : [0, T] \rightarrow \mathbb{R}$ given by

$$f(t) := \exp\left(K \frac{1}{2(1-p)} \frac{\alpha^2}{\sigma^2} (T-t)\right).$$

Davis et al. [24] (with some adaptations) and Belak et al. [9, 10] establish that the value function \mathcal{V} is continuous and the unique viscosity solution of the HJB equation

$$0 = \min\{\mathcal{L}^{nt}\mathcal{V}(t, b, s), \mathcal{L}^{buy}\mathcal{V}(t, b, s), \mathcal{L}^{sell}\mathcal{V}(t, b, s)\} \tag{2.2}$$

on $[0, T] \times \mathcal{S}$ with terminal condition

$$\mathcal{V}(T, b, s) = U_p(b + \min\{(1 - \mu)s, (1 + \lambda)s\}), \quad (b, s) \in \overline{\mathcal{S}},$$

and boundary condition

$$\mathcal{V}(t, b, s) = U_p(0), \quad (t, b, s) \in [0, T] \times \partial\mathcal{S}.$$

The differential operators \mathcal{L}^{nt} , \mathcal{L}^{buy} and \mathcal{L}^{sell} in (2.2) are given by

$$\mathcal{L}^{nt}\mathcal{V}(t, b, s) := -\partial_t\mathcal{V}(t, b, s) - \alpha s \partial_s \mathcal{V}(t, b, s) - \frac{1}{2} \sigma^2 s^2 \partial_s^2 \mathcal{V}(t, b, s),$$

$$\mathcal{L}^{buy}\mathcal{V}(t, b, s) := (1 + \lambda) \partial_b \mathcal{V}(t, b, s) - \partial_s \mathcal{V}(t, b, s),$$

$$\mathcal{L}^{sell}\mathcal{V}(t, b, s) := -(1 - \mu) \partial_b \mathcal{V}(t, b, s) + \partial_s \mathcal{V}(t, b, s).$$

The uniqueness of the value function is a consequence of the following comparison principle, obtained in Belak et al. [9, Theorem 4.4].

Theorem 2.1 *Let $u, v : [0, T] \times \overline{\mathcal{S}} \rightarrow \overline{\mathbb{R}}$ and fix $\varepsilon > 0$. Assume that u is an upper semicontinuous viscosity subsolution and v is a lower semicontinuous viscosity supersolution of (2.2) such that*

$$U_p(b + \min\{(1 - \mu)s, (1 + \lambda)s\}) \leq u(t, b, s), v(t, b, s) \leq \varphi_p(t, b, s).$$

If $u(T, b, s) \leq v(T, b + \varepsilon, s)$ and $u(t, b, s) \leq U_p(0)$ for every $(b, s) \in \partial\mathcal{S}$, then $u(t, b, s) \leq v(t, b + \varepsilon, s)$ on $[0, T] \times \overline{\mathcal{S}}$.

It is expected that the operators \mathcal{L}^{nt} , \mathcal{L}^{buy} and $\mathcal{L}^{\text{sell}}$ determine the optimal strategies. To be more precise, given a smooth solution v of (2.2), we define

$$\begin{aligned} \mathcal{R}^{\text{buy}}(v) &:= \{(t, b, s) \in [0, T) \times \mathcal{S} : \mathcal{L}^{\text{buy}}v(t, b, s) = 0\}, \\ \mathcal{R}^{\text{sell}}(v) &:= \{(t, b, s) \in [0, T) \times \mathcal{S} : \mathcal{L}^{\text{sell}}v(t, b, s) = 0\}, \\ \mathcal{R}^{\text{nt}}(v) &:= ([0, T) \times \mathcal{S}) \setminus (\mathcal{R}^{\text{buy}} \cup \mathcal{R}^{\text{sell}}). \end{aligned}$$

We expect that the optimal strategy keeps the process (B, S) inside the no-trading region $\mathcal{R}^{\text{nt}}(v)$ by reflecting (B, S) at the boundary $\partial\mathcal{R}^{\text{nt}}(v)$. Note that we do not include the terminal time T in the definition of the trading regions since we require the investor to liquidate her holdings in the stock at time T , and we do not include the boundary $\partial\mathcal{S}$ since the only admissible, and hence optimal, strategy on $\partial\mathcal{S}$ is to instantly close the stock position and refrain from further trading; see Lemma 3.1 below.

While with the previous result, the value function \mathcal{V} is completely characterised (and can be computed numerically), it does not suffice to construct and verify the optimal strategies. It is therefore necessary to study the HJB equation in more detail for the existence of a regular solution.

We denote by \mathcal{S}_0 and $\overline{\mathcal{S}}_0$ the restrictions to positive stock holdings of the solvency region and of its closure, respectively, i.e.,

$$\mathcal{S}_0 := \{(b, s) \in \mathcal{S} : s > 0\} \quad \text{and} \quad \overline{\mathcal{S}}_0 := \{(b, s) \in \overline{\mathcal{S}} : s > 0\}.$$

Dai and Yi [21, Theorem 5.1] show that the HJB equation admits a classical solution on the restricted solvency region:

Theorem 2.2 *There exists a continuous function $V : [0, T) \times \overline{\mathcal{S}}_0 \rightarrow \mathbb{R}$ such that $V \in C^{1,2}([0, T) \times \mathcal{S}_0) \setminus F$ and $\partial_t V \leq 0$ which solves the HJB equation (2.2) in the classical sense. Here, the set F is given by*

$$F := \begin{cases} \emptyset & \text{if } \pi_M < 1, \\ \{(t, b, s) \in [0, T) \times \mathcal{S}_0 : b = 0\} & \text{if } \pi_M = 1, \\ \{(t, b, s) \in [0, T) \times \mathcal{S}_0 : b = 0, t = t^{\text{up}}\} & \text{if } \pi_M > 1, \end{cases} \quad (2.3)$$

where

$$t^{\text{up}} := T - \frac{\log(1 + \lambda) - \log(1 - \mu)}{\alpha - (1 - p)\sigma^2} \quad (2.4)$$

and $\pi_M := \alpha / ((1 - p)\sigma^2)$ denotes the Merton fraction.

Let us emphasise here that the combination of the fact that \mathcal{V} is a viscosity solution of the HJB equation on $[0, T) \times \mathcal{S}$, the fact that V is a classical solution of the HJB equation on $[0, T) \times \mathcal{S}_0$, and the uniqueness result implied by the comparison principle in Theorem 2.1 do not imply that $\mathcal{V} = V$ on $[0, T) \times \mathcal{S}_0$; some additional work is necessary to arrive at this conclusion. Indeed, while it is not too difficult to show that V is also a viscosity solution on $[0, T) \times \mathcal{S}_0$ (which is in particular immediate

outside the set F), one would need to extend V to a continuous viscosity solution defined on the entire state space $[0, T) \times \mathcal{S}$ to apply the comparison theorem. Verifying that such an extension exists, however, requires additional work including a careful study of the behaviour of V (and its partial derivatives) as $s \downarrow 0$. While we believe that it is possible to follow this direct approach, we shall nevertheless take a different route which has the advantage of additionally verifying optimality of our candidate optimal trading strategies. These candidate strategies are defined in terms of trading regions implied by the classical solution V . More precisely, Theorem 2.2 allows us to define the reduced trading regions

$$\begin{aligned} \mathcal{R}_0^{\text{buy}} &:= \{(t, b, s) \in [0, T) \times \mathcal{S}_0 : \mathcal{L}^{\text{buy}} V(t, b, s) = 0\}, \\ \mathcal{R}_0^{\text{sell}} &:= \{(t, b, s) \in [0, T) \times \mathcal{S}_0 : \mathcal{L}^{\text{sell}} V(t, b, s) = 0\}, \\ \mathcal{R}_0^{\text{nt}} &:= ([0, T) \times \mathcal{S}_0) \setminus (\mathcal{R}_0^{\text{buy}} \cup \mathcal{R}_0^{\text{sell}}). \end{aligned}$$

Note that we must have $\mathcal{L}^{\text{nt}} V = 0$ on $\mathcal{R}_0^{\text{nt}}$. In order to construct the optimal strategies, it is important to determine the geometry of these sets and the location of the boundaries between them. Dai and Yi [21, Theorems 4.3, 4.5 and 4.7] provide the following characterisation of these free boundaries.

Theorem 2.3 1) *There exist nonnegative, nonincreasing functions $\underline{\pi} : [0, T) \rightarrow \mathbb{R}$ and $\overline{\pi} : [0, T) \rightarrow \mathbb{R}$ with $\underline{\pi}(t) < \overline{\pi}(t)$ for all $t \in [0, T)$ such that*

$$\begin{aligned} \mathcal{R}_0^{\text{buy}} &= \left\{ (t, b, s) \in [0, T) \times \mathcal{S}_0 : \frac{s}{b+s} \leq \underline{\pi}(t) \right\}, \\ \mathcal{R}_0^{\text{sell}} &= \left\{ (t, b, s) \in [0, T) \times \mathcal{S}_0 : \frac{s}{b+s} \geq \overline{\pi}(t) \right\}, \\ \mathcal{R}_0^{\text{nt}} &= \left\{ (t, b, s) \in [0, T) \times \mathcal{S}_0 : \underline{\pi}(t) < \frac{s}{b+s} < \overline{\pi}(t) \right\}. \end{aligned}$$

Moreover, V is of class C^∞ on $\mathcal{R}_0^{\text{nt}}$.

2) *The function $\underline{\pi}$ is continuous and satisfies*

$$\underline{\pi}(t) \begin{cases} < 1 & \text{if } \pi_M \leq 1, \\ > 1 & \text{if } \pi_M > 1, t < t^{\text{up}}, \\ = 1 & \text{if } \pi_M > 1, t = t^{\text{up}}, \\ < 1 & \text{if } \pi_M > 1, t > t^{\text{up}}, \end{cases}$$

where t^{up} is defined in (2.4). Furthermore, $\underline{\pi}(t) = 0$ for $t \in [t^{\text{down}}, T)$, where

$$t^{\text{down}} := T - \frac{\log(1 + \lambda) - \log(1 - \mu)}{\alpha}.$$

3) It holds that

$$\bar{\pi}(t) \begin{cases} < 1 & \text{if } \pi_M < 1, \\ = 1 & \text{if } \pi_M = 1, \\ > 1 & \text{if } \pi_M > 1, \end{cases}$$

and $\bar{\pi} \in C^\infty([0, T])$.

Remark 2.4 A close inspection of the results in Dai and Yi [21] implies that

$$\inf_{t \in [0, T]} |\bar{\pi}(t) - \underline{\pi}(t)| > 0.$$

Indeed, Dai and Yi [21, Sect. 5] show that the HJB equation (2.2) can be transformed into a double obstacle problem with obstacles given by $1/(x + 1 + \lambda)$ (determining $\underline{\pi}$) and $1/(x + 1 - \mu)$ (determining $\bar{\pi}$), respectively. Since V is continuous and the distance between the obstacles is strictly positive, this implies that the above infimum is also strictly positive.

Figures 1–3 below sketch the different scenarios for the location of the free boundaries. Note that t^{up} is the time point at which the lower free boundary is equal to one, i.e., $\underline{\pi}(t^{\text{up}}) = 1$ (this may only happen if $\pi_M > 1$), and t^{down} is the time point from which onwards the lower free boundary is equal to zero, i.e., $\underline{\pi}(t) = 0$ for all $t \in [t^{\text{down}}, T]$.

For obvious reasons, we refer to $\underline{\pi}$ and $\bar{\pi}$ as the buy and sell boundary, respectively. If our conjecture that the buy and sell boundaries characterise the optimal strategies is indeed correct (which will be rigorously proved in Sect. 4), Theorem 2.3 has the following implications:

1) If $\pi_M < 1$ (cf. Fig. 1), i.e., if borrowing is not optimal in the absence of costs, then it is also not optimal in the presence of costs.

2) If $\pi_M = 1$ (cf. Fig. 2), i.e., if it is optimal to invest all money in the stock in the absence of transaction costs, then two cases must be distinguished in the presence of costs. If the initial position of the investor is such that $b \leq 0$, then the bond position is closed and all money is kept in the stock (since $\bar{\pi} = 1$). However, if the initial position is such that $b > 0$, then it is not optimal to close the bond position. This is because we force the investor to close the stock position at the terminal time T , and hence it is too expensive to first buy shares of the stock at the initial time just to liquidate the stock position once the investment horizon is reached.

Fig. 1 The trading regions for $\pi_M < 1$

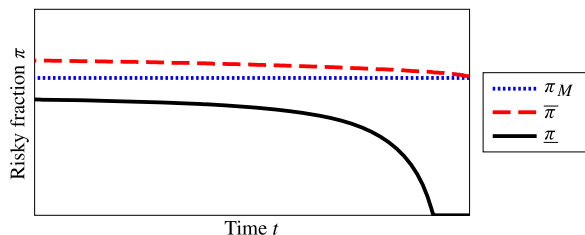


Fig. 2 The trading regions for $\pi_M = 1$

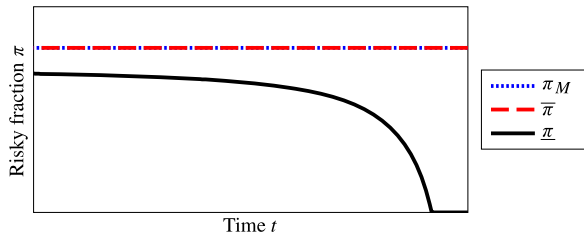
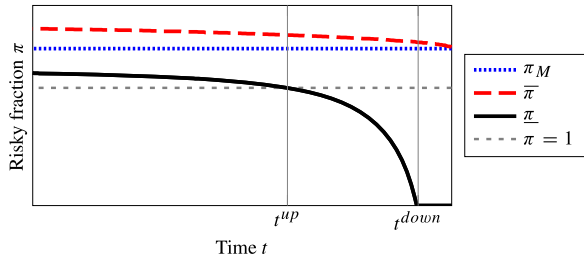


Fig. 3 The trading regions for $\pi_M > 1$



3) If $\pi_M > 1$ (cf. Fig. 3), i.e., if borrowing is optimal without costs, we need to distinguish three cases. Since the investor never switches from borrowing to no-borrowing or vice versa after the initial transaction ($\bar{\pi} > 1$ and $\underline{\pi}$ is nonincreasing), the initial transaction determines whether borrowing or no-borrowing is optimal:

$t^{up} > 0$. In this case, borrowing is optimal since $\underline{\pi}(0) > 1$.

$t^{up} = 0$. If the initial position is such that $b < 0$, then borrowing is optimal. Otherwise the investor invests all her wealth in the stock (since $\underline{\pi}(0) = 1$).

$t^{up} < 0$. In this case, borrowing is optimal if $b < 0$ and no-borrowing is optimal if $b \geq 0$. This is because $\underline{\pi}(t) < 1 < \bar{\pi}(t)$ for all $t \in [0, T)$.

4) In any case, as soon as $t \geq t^{down}$, the investor refrains from buying shares of the stock since $\underline{\pi}(t) = 0$.

5) If the initial position (b, s) is such that $s \leq 0$, then whenever $\underline{\pi}(t) = 0$, it is optimal to liquidate the stock position and refrain from further trading. Whenever $\underline{\pi}(t) > 0$, the investor performs an initial transaction which takes her position to the boundary of the no-trading region. This is proved in Sect. 4, but intuitively this behaviour is clear: Since the excess return α is positive and the investor has to liquidate her stock holdings at time T , it should never be optimal to have a short position in the stock.

3 Construction of the optimal strategies

We proceed with the construction of the candidate optimal strategies. For this, we fix an arbitrary initial datum $(t_0, b_0, s_0) \in [0, T) \times \bar{S}$. We first observe that if $(b_0, s_0) \in \partial S$, then the only admissible and hence optimal strategy is to immediately close the position and refrain from further trading. The proof follows as in Shreve and Soner [37, Remark 2.1].

Lemma 3.1 *Let $(b_0, s_0) \in \partial S$. Then the only admissible strategy is to instantly jump to the position $(0, 0)$ and remain there.*

In what follows, we may thus assume $(b_0, s_0) \in \mathcal{S}$. For the construction of the optimal strategy, we need to prove the existence of nondecreasing processes $L^* = (L^*(u))_{u \in [t, T]}$ and $M^* = (M^*(u))_{u \in [t, T]}$ which turn the controlled wealth process $(B^*, S^*) := (B_{t_0, b_0}^{L^*, M^*}, S_{t_0, s_0}^{L^*, M^*})$ into a diffusion reflected at the boundary of $\mathcal{R}_0^{\text{nt}}$. We first observe that we can without loss of generality assume that the initial position (t_0, b_0, s_0) is an element of the closure of the no-trading region $\mathcal{R}_0^{\text{nt}}$. Indeed, if $(t_0, b_0, s_0) \notin \overline{\mathcal{R}_0^{\text{nt}}}$, then we can find (b^*, s^*) and (minimal) $\ell, m \geq 0$ such that $(t_0, b^*, s^*) \in \partial \mathcal{R}_0^{\text{nt}}$ and

$$b^* = b_0 - (1 + \lambda)\ell + (1 - \mu)m, \quad s^* = s_0 + \ell - m.$$

With this, if (L^*, M^*) is the candidate optimal strategy for (t_0, b^*, s^*) , the pair $(L^* + \ell, M^* + m)$ is the candidate optimal strategy for (t_0, b_0, s_0) . In other words, by a suitable initial transaction, we can always ensure that we start within the closure of the no-trading region.

Next, we observe that in the following, we can rule out all cases in which the investor liquidates either the bond or the stock position at time t_0 and refrains from further transactions. Comparing with Figs. 1–3 and recalling that we may assume $(t_0, b_0, s_0) \in \overline{\mathcal{R}_0^{\text{nt}}}$, these cases are

- (a) π_M arbitrary, $t_0 \geq t^{\text{down}}$ and $s_0 = 0$;
- (b) $\pi_M = 1$ with $s_0 > 0$ and $b_0 = 0$;
- (c) $\pi_M > 1$ with $s_0 > 0$, $b_0 = 0$, and $t_0 \geq t^{\text{up}}$;
- (d) $\pi_M > 1$ with $s_0 = 0$, $b_0 > 0$, and $t_0 = t^{\text{up}}$.

The remaining cases are given by

- (e) $\pi_M < 1$ with $s_0, b_0 > 0$;
- (f) $\pi_M = 1$ with $s_0, b_0 > 0$;
- (g) $\pi_M > 1$ with $s_0, b_0 > 0$ and $t_0 > t^{\text{up}}$;
- (h) $\pi_M > 1$ with $s_0 > 0$, $b_0 < 0$.

The cases (e)–(g) are no-borrowing cases, whereas we expect borrowing to be optimal in the case (h). It turns out that for the construction of the reflected diffusions, it is advantageous to consider the change of variables s/b in the no-borrowing case and $s/(-b)$ in the borrowing case. For this, we define

$$\mathcal{S}_+ := \{(b, s) \in \mathcal{S} : b > 0, s > 0\}, \quad \mathcal{S}_- := \{(b, s) \in \mathcal{S} : b < 0, s > 0\}.$$

In the sequel, we work on the reduced state space $[0, T] \times \mathcal{S}_+$ in the no-borrowing cases (e)–(g) and $[0, T] \times \mathcal{S}_-$ in the borrowing case (h).

3.1 Construction in the no-borrowing case (e)

The main idea for the construction of the optimal strategy is to find a suitable transformation of the state space so that the problem of constructing an obliquely reflected diffusion in an unbounded and time-dependent cone simplifies to normal reflection in a time-dependent interval. The transformation and construction is based on ideas

from Gerhold et al. [26], and hence we keep the exposition to a minimum. We restrict ourselves to the case $p < 1$, $p \neq 0$ (power utility) and remark that the construction for the case $p = 0$ (log utility) follows similarly.

In the situation of case (e), i.e., $\pi_M < 1$ with $s_0, b_0 > 0$, we first observe that $V \in C^{1,2}([0, T] \times \mathcal{S}_0)$ by Theorem 2.2. We define

$$\ell(t) := \frac{\pi(t)}{1 - \pi(t)} \quad \text{and} \quad u(t) := \frac{\bar{\pi}(t)}{1 - \bar{\pi}(t)}, \tag{3.1}$$

and note that ℓ and u constitute the buy and sell boundaries under the change of variables $(b, s) \mapsto s/b$, i.e.,

$$\mathcal{R}_0^{\text{nt}} = \left\{ (t, b, s) \in [0, T] \times \bar{\mathcal{S}} : \ell(t) < \frac{s}{b} < u(t) \right\}.$$

By Theorem 2.3, we see that $\ell < u$, $\ell \in C([0, T])$ and $u \in C^\infty([0, T])$.

On the set $[0, T] \times \mathcal{S}_+$ we consider the transformation

$$V(t, b, s) = b^p \exp\left(-p \int_{\log \frac{s}{bu(t)}}^0 w(t, y) \, dy\right) = b^p \exp\left(-p \int_x^0 w(t, y) \, dy\right),$$

where

$$x = x(t, b, s) := \log \frac{s}{bu(t)}.$$

With this and using that V satisfies $\mathcal{L}^{\text{buy}} V \geq 0$ and $\mathcal{L}^{\text{sell}} V \geq 0$, we see that w satisfies

$$1 - \mu \leq \frac{w(t, x)}{u(t)(1 - w(t, x))e^x} \leq 1 + \lambda, \tag{3.2}$$

and equality holds if and only if $\mathcal{L}^{\text{sell}} V = 0$ or $\mathcal{L}^{\text{buy}} V = 0$, respectively. Moreover, since $\mathcal{L}^{\text{nt}} V = 0$ whenever $\mathcal{L}^{\text{buy}} V > 0$ and $\mathcal{L}^{\text{sell}} V > 0$, we have

$$\begin{aligned} 0 &= \int_x^0 \partial_t w(t, y) \, dy - \left(\alpha - \frac{1}{2}\sigma^2 - \frac{u'(t)}{u(t)}\right) w(t, x) \\ &\quad - \frac{1}{2}p\sigma^2 w(t, x)^2 - \frac{1}{2}\sigma^2 \partial_x w(t, x) \end{aligned}$$

whenever $w/(u(1 - w)e^x) \notin \{1 - \mu, 1 + \lambda\}$. Taking the derivative with respect to x in the last equation, we obtain

$$\begin{aligned} \frac{1}{2}\sigma^2 \partial_x^2 w(t, x) &= -\partial_t w(t, x) - \left(\alpha - \frac{1}{2}\sigma^2 - \frac{u'(t)}{u(t)}\right) \partial_x w(t, x) \\ &\quad - p\sigma^2 w(t, x) \partial_x w(t, x). \end{aligned}$$

Consider again the fraction in (3.2), i.e.,

$$f(t, x) := \frac{w(t, x)}{u(t)(1 - w(t, x))e^x}.$$

Since by (3.1) the points $x = 0$ and $x = \log(\ell(t)/u(t))$ constitute the boundary points of the no-trading region in the new variables, we must have

$$f(t, x) = 1 - \mu \quad \text{if } x \geq 0, \tag{3.3}$$

$$f(t, x) = 1 + \lambda \quad \text{if } x \leq \log \frac{\ell(t)}{u(t)}. \tag{3.4}$$

Moreover, note that for the point $x = \log(\ell(t)/u(t))$, these considerations are only valid for $t \in [0, t^{\text{down}})$ since otherwise $\log(\ell(t)/u(t)) = -\infty$.

Remark 3.2 We have

$$f(t, x) = \frac{w(t, x)}{u(t)(1 - w(t, x))e^x} \in [1 - \mu, 1 + \lambda]$$

and $f(t, x) \in \{1 - \mu, 1 + \lambda\}$ inside the buy and sell regions. This suggests that

$$f(t, X^*(t))P^1(t) \quad \text{with} \quad X^*(t) = \log \frac{S^*(t)}{B^*(t)u(t)}$$

(where (B^*, S^*) is the optimally controlled portfolio process) is the shadow price in our problem. This can be confirmed as in Gerhold et al. [26].

The next step is to construct a reflected diffusion in the time-dependent interval $[\log(\ell/u), 0]$.

Lemma 3.3 *There exist a process $\Psi = (\Psi(t))_{t \in [t_0, T]}$ and nondecreasing processes $L = (L(t))_{t \in [t_0, T]}$ and $M = (M(t))_{t \in [t_0, T]}$ such that L is constant on $[t^{\text{down}}, T)$ and*

$$d\Psi(t) = \left(\alpha - \frac{1}{2}\sigma^2 - \frac{u'(t)}{u(t)} \right) dt + \sigma dW(t) + dL(t) - dM(t), \tag{3.5}$$

with

$$\Psi(t_0) = \log \frac{s_0}{b_0 u(t_0)},$$

and such that Ψ is a diffusion reflected on the boundaries of the time-dependent interval $[\log(\ell/u), 0]$.

Proof This follows from Słomiński and Wojciechowski [38, Theorem 3.3] together with Remark 2.4. □

Let us now define a process $N = (N(t))_{t \in [t_0, T]}$ by $N(t_0) = s_0/P^1(t_0)$ and, for $t \in [t_0, T)$,

$$dN(t) = N(t) \left(1 - w \left(t, \log \frac{\ell(t)}{u(t)} \right) \right) dL(t) - N(t) (1 - w(t, 0)) dM(t). \tag{3.6}$$

Remark 3.4 Comparing with Gerhold et al. [26], we interpret ue^Ψ as the optimal stock-to-bond ratio and N as the optimal cumulative number of shares of the stock bought up to time t . Furthermore, again as in Gerhold et al. [26], the function w can be interpreted as the optimal risky fraction, and hence $1 - w$ coincides with the optimal fraction of wealth invested into the bond which gives a nice interpretation for all the terms occurring in (3.6).

With this, we now have all the tools at hand to construct the optimal strategies. Let us define a process $S^* = (S^*(t))_{t \in [t_0, T]}$ through $S^*(t) := N(t)P^1(t)$. Then $S^*(t_0) = N(t_0)P^1(t_0) = s_0$ and

$$\begin{aligned} dS^*(t) &= \alpha S^*(t) dt + \sigma S^*(t) dW(t) \\ &\quad + S^*(t) \left(1 - w \left(t, \log \frac{\ell(t)}{u(t)} \right) \right) dL(t) \\ &\quad - S^*(t) (1 - w(t, 0)) dM(t), \quad t \in [t_0, T]. \end{aligned}$$

Also define $B^* = (B^*(t))_{t \in [t_0, T]}$ by $B^*(t) := S^*(t)e^{-\Psi(t)}/u(t)$ so that $B^*(t_0) = S^*(t_0)/(u(t_0)e^{\Psi(t_0)}) = b_0$ and

$$dB^*(t) = -w \left(t, \log \frac{\ell(t)}{u(t)} \right) B^*(t) dL(t) + w(t, 0) B^*(t) dM(t), \quad t \in [t_0, T].$$

Using the definition of B^* , (3.3) and (3.4), we see that

$$w(t, 0) B^*(t) = \frac{w(t, 0) S^*(t)}{u(t) e^0} = (1 - \mu) S^*(t) (1 - w(t, 0)),$$

and similarly

$$\begin{aligned} w \left(t, \log \frac{\ell(t)}{u(t)} \right) B^*(t) &= \frac{w(t, \log(\ell(t)/u(t))) S^*(t)}{u(t) e^{\log(\ell(t)/u(t))}} \\ &= (1 + \lambda) S(t)^* \left(1 - w \left(t, \log \frac{\ell(t)}{u(t)} \right) \right). \end{aligned}$$

So in total, the dynamics of B^* simplify to

$$\begin{aligned} dB^*(t) &= -(1 + \lambda) S^*(t) \left(1 - w \left(t, \log \frac{\ell(t)}{u(t)} \right) \right) dL(t) \\ &\quad + (1 - \mu) S^*(t) (1 - w(t, 0)) dM(t), \quad t \in [t_0, T]. \end{aligned}$$

Hence, if we define

$$\begin{aligned} dL^*(t) &= N(t)P^1(t) \left(1 - w \left(t, \log \frac{\ell(t)}{u(t)} \right) \right) dL(t), \quad t \in [t_0, T], \\ dM^*(t) &= N(t)P^1(t) (1 - w(t, 0)) dM(t), \quad t \in [t_0, T], \end{aligned}$$

with $L^*(t_0) = 0, M^*(t_0) = 0$ and if we set (liquidation at the terminal time)

$$\begin{aligned} L^*(T) &= L^*(T-) + S^*(T-) \mathbb{1}_{\{S^*(T-) < 0\}}, \\ M^*(T) &= M^*(T-) + S^*(T-) \mathbb{1}_{\{S^*(T-) > 0\}}, \\ S^*(T) &= 0, \\ B^*(T) &= B^*(T-) + \min\{(1 - \mu)S^*(T-), (1 + \lambda)S^*(T-)\}, \end{aligned}$$

then $(L^*, M^*) \in \mathcal{A}(t_0, b_0, s_0)$, and (B^*, S^*) is a diffusion reflected at $\partial \mathcal{R}_0^{\text{int}}$.

3.2 Construction in the other cases

The construction in the other cases (f)–(h) follows in a similar way as the construction in the case (e). Let us outline the critical differences.

Assume first that we are in one of the no-borrowing cases (f) or (g). That is, we either have $\pi_M = 1$ with $s_0, b_0 > 0$, or $\pi_M > 1$ with $s_0, b_0 > 0$ and $t_0 > t^{\text{up}}$. While the construction here is similar to the case (e), we have to be more careful since the upper boundary in terms of the transformation s/b is now equal to infinity which does not allow us to consider the transformation $x = \log(s/(bu(t)))$. However, since the upper boundary is now equal to infinity, we deal with one-sided reflection which simplifies matters again (we never have to sell shares of the stock!). As before, we define the lower boundary as $\ell(t) := \underline{\pi}(t)/(1 - \underline{\pi}(t))$ and consider the slightly different transformation

$$V(t, b, s) = b^p \exp\left(-p \int_{\log(s/b)}^0 w(t, y) dy\right).$$

Setting $x = \log(s/b)$ and arguing in a similar fashion as before, the existence of the candidate optimal strategy follows. Note, however, that the process Ψ now has to be constructed without the $u'(t)/u(t)$ term in its drift (see (3.5)).

Let us now turn to the borrowing case (h). That is, assume $\pi_M > 1, s_0 > 0, b_0 < 0$ and $t_0 \in [0, T)$. Since $b_0 < 0$ and we want the optimally controlled bond wealth B^* to satisfy $B^* < 0$, we have to consider a different transformation. More precisely, we consider the transformation $s/(-b)$ instead. We first define the trading boundaries to be

$$\ell(t) := -\frac{\bar{\pi}(t)}{1 - \bar{\pi}(t)} \quad \text{and} \quad u(t) := -\frac{\underline{\pi}(t)}{1 - \underline{\pi}(t)}.$$

Theorem 2.3 implies $0 < \ell(t) < u(t) \leq \infty, \ell \in C^\infty([0, T))$ and $u \in C([0, T))$ (note, however, that in this case the lower boundary ℓ is defined by means of $\bar{\pi}$ instead of $\underline{\pi}$). The transformation of the function V is then chosen to be

$$V(t, b, s) = (-b)^p \exp\left(-p \int_0^{\log \frac{s}{-b(t)}} w(t, y) dy\right),$$

where we restrict V to the set \mathcal{S}_- . This leads to similar calculations as in the case (e), but with the lower boundary ℓ in place of the upper boundary u in the drift term of the process Ψ (see (3.5)).

4 Verification and value function regularity

We now proceed by verifying that the strategies constructed in the previous section are indeed optimal. Since V is only defined on $[0, T] \times \overline{\mathcal{S}}_0$, classical verification arguments are difficult (it will turn out that \mathcal{V} is not sufficiently regular everywhere, see Theorem 4.14). Instead, we adapt the approach introduced in Christensen [14] for impulse control problems to our setting. The idea is to show that the value function is the pointwise minimum of a suitable set of superharmonic functions.

More precisely, let \mathbb{H} be the set of all continuous functions $h : [0, T] \times \overline{\mathcal{S}} \rightarrow \overline{\mathbb{R}}$ satisfying the following properties:

(i) $h(T, b, s) \geq \mathcal{V}(T, b, s)$ on $\{T\} \times \overline{\mathcal{S}}$.

(ii) h is *nonincreasing in the direction of transactions*, i.e., for all $(t, b, s) \in [0, T] \times \overline{\mathcal{S}}$ and $\ell, m \geq 0$ with $(b - (1 + \lambda)\ell + (1 - \mu)m, s + \ell - m) \in \overline{\mathcal{S}}$, it holds that

$$h(t, b, s) \geq h(t, b - (1 + \lambda)\ell + (1 - \mu)m, s + \ell - m).$$

(iii) h is space-time superharmonic with respect to the uncontrolled portfolio process. More precisely, denote by $(B^0, S^0) = (B_{t,b}^0, S_{t,s}^0)$ the wealth process corresponding to the strategy $L \equiv M \equiv 0$ and let ϑ be the first hitting time of $\partial\mathcal{S}$. Then h is called *space-time superharmonic* if

$$h(t, b, s) \geq \mathbb{E}[h(\tau \wedge \vartheta, B_{t,b}^0(\tau \wedge \vartheta), S_{t,s}^0(\tau \wedge \vartheta))]$$

for every $[t, T]$ -valued stopping time τ .

(iv) h satisfies the bounds

$$U_p(b + \min\{(1 - \mu)s, (1 + \lambda)s\}) \leq h(t, b, s) \leq \varphi_p(t, b, s).$$

We expect that \mathcal{V} is the pointwise minimum of the elements of \mathbb{H} . If this is true, we can prove the optimality of (L^*, M^*) as follows:

- 1) Show that every $h \in \mathbb{H}$ dominates \mathcal{V} .
- 2) Define the function

$$h_0(t, b, s) := \mathbb{E}[U_p(X_{t,b,s}^{L^*, M^*}(T))]$$

and show that $h_0 \in \mathbb{H}$.

It then follows that $\mathcal{V} \leq h_0$, but $h_0 \leq \mathcal{V}$ since (L^*, M^*) is admissible. Hence $h_0 = \mathcal{V}$ and (L^*, M^*) is optimal.

The advantage of this approach is that h_0 (and hence \mathcal{V}) need not be of class $C^{1,2}$ everywhere for the above argument to work. Indeed, in order to show that h_0 satisfies the monotonicity property (ii), we only require piecewise continuous differentiability in certain spatial directions for each fixed time point t , and to verify the superharmonicity property (iii), we only require some regularity along the paths of the uncontrolled state process, which naturally avoid the points of degeneracy of the infinitesimal generator \mathcal{L}^{nt} (and hence the points on the b - and s -axes where regularity fails).

In Lemma 4.1 below, we show that every $h \in \mathbb{H}$ dominates \mathcal{V} . Then we proceed by analysing the regularity of h_0 and use these results to show that h_0 is superharmonic (Proposition 4.11) and nonincreasing in the direction of transactions (Proposition 4.12). The optimality of (L^*, M^*) then follows in Theorem 4.13.

Lemma 4.1 *Let $h \in \mathbb{H}$. Then $\mathcal{V} \leq h$.*

Proof We show that h is a viscosity supersolution of the HJB equation (2.2). By Theorem 2.1 (comparison principle), it then follows that for every $\varepsilon > 0$, we have $\mathcal{V}(t, b, s) \leq h(t, b + \varepsilon, s)$ everywhere, and by the continuity of h , we can send $\varepsilon \downarrow 0$ to conclude.

Let us therefore fix $(t_0, b_0, s_0) \in [0, T) \times \mathcal{S}$ and let $\varphi \in C^{1,2}([0, T) \times \mathcal{S})$ be such that $\varphi \leq h$ and $\varphi(t_0, b_0, s_0) = h(t_0, b_0, s_0)$. We have to show that

$$\min\{\mathcal{L}^{\text{nt}}\varphi(t_0, b_0, s_0), \mathcal{L}^{\text{buy}}\varphi(t_0, b_0, s_0), \mathcal{L}^{\text{sell}}\varphi(t_0, b_0, s_0)\} \geq 0.$$

Let $\ell > 0$ be such that $(b_0 - (1 + \lambda)\ell, s_0 + \ell) \in \mathcal{S}$. Then

$$\begin{aligned} &\varphi(t_0, b_0, s_0) - \varphi(t_0, b_0 - (1 + \lambda)\ell, s_0 + \ell) \\ &\geq h(t_0, b_0, s_0) - h(t_0, b_0 - (1 + \lambda)\ell, s_0 + \ell) \geq 0 \end{aligned}$$

since $\varphi(t_0, b_0, s_0) = h(t_0, b_0, s_0)$, $\varphi \leq h$ and h is nonincreasing in the direction of transactions. Now divide by ℓ and send $\ell \downarrow 0$ to obtain $\mathcal{L}^{\text{buy}}\varphi(t_0, b_0, s_0) \geq 0$. By similar arguments, we have $\mathcal{L}^{\text{sell}}\varphi(t_0, b_0, s_0) \geq 0$, and hence it only remains to show that $\mathcal{L}^{\text{nt}}\varphi(t_0, b_0, s_0) \geq 0$.

Suppose that on the contrary, we have $\mathcal{L}^{\text{nt}}\varphi(t_0, b_0, s_0) < 0$. Then there exist $\varepsilon, \delta > 0$ such that $t_0 + \varepsilon < T$, $\overline{B}_\varepsilon(b_0, s_0) \subseteq \mathcal{S}$ and $\mathcal{L}^{\text{nt}}\varphi(t, b, s) < -\delta$ for all $(t, b, s) \in [t_0, t_0 + \varepsilon] \times \overline{B}_\varepsilon(b_0, s_0)$. Now define the stopping time

$$\tau_\varepsilon := \inf \{u \geq t_0 : (B_{t_0, b_0}^0(u), S_{t_0, s_0}^0(u)) \notin \overline{B}_\varepsilon(b_0, s_0)\} \wedge (t_0 + \varepsilon).$$

Since h is space-time superharmonic and by Itô’s formula, we have

$$\begin{aligned} \varphi(t_0, b_0, s_0) &= h(t_0, b_0, s_0) \\ &\geq \mathbb{E}[h(\tau_\varepsilon, B_{t_0, b_0}^0(\tau_\varepsilon), S_{t_0, s_0}^0(\tau_\varepsilon))] \\ &\geq \mathbb{E}[\varphi(\tau_\varepsilon, B_{t_0, b_0}^0(\tau_\varepsilon), S_{t_0, s_0}^0(\tau_\varepsilon))] \\ &= \varphi(t_0, b_0, s_0) - \mathbb{E}\left[\int_{t_0}^{\tau_\varepsilon} \mathcal{L}^{\text{nt}}\varphi(u, B_{t_0, b_0}^0(u), S_{t_0, s_0}^0(u)) \, du\right], \end{aligned}$$

i.e., $\mathbb{E}[\int_{t_0}^{\tau_\varepsilon} \mathcal{L}^{\text{nt}}\varphi(u, B_{t_0, b_0}^0(u), S_{t_0, s_0}^0(u)) \, du] \geq 0$. This, however, must imply that

$$\max_{u \in [t_0, t_0 + \varepsilon], (b, s) \in \overline{B}_\varepsilon(b_0, s_0)} \mathcal{L}^{\text{nt}}\varphi(u, b, s) \geq 0.$$

Sending $\varepsilon \downarrow 0$ implies that $\mathcal{L}^{\text{nt}}\varphi(t_0, b_0, s_0) \geq 0$ which is a contradiction. □

Remark 4.2 From the superharmonicity and the monotonicity in the direction of transactions, it is easily seen that $h(\cdot, B^{L,M}, S^{L,M})$ is a supermartingale for any $h \in \mathbb{H}$ and any piecewise constant trading strategy (L, M) and thus

$$h(t, b, s) \geq \mathbb{E}[U_p(X_{t,b,s}^{L,M}(T))] \text{ for any piecewise constant } (L, M) \in \mathcal{A}(t, b, s).$$

In the context of the closely related stochastic Perron method, one studies instead the set of all functions \bar{h} for which $\bar{h}(\cdot, B^{L,M}, S^{L,M})$ is a supermartingale for *all* $(L, M) \in \mathcal{A}(t, b, s)$ and shows that the pointwise infimum of this set coincides with the value function as a means of obtaining the viscosity characterisation of \mathcal{V} .

In Sect. 3 and Lemma 3.1, we have constructed candidate optimal strategies $(L^*, M^*) = (L_{t,b,s}^*(u), M_{t,b,s}^*(u))_{u \in [t,T]}$ for every $(t, b, s) \in [0, T] \times \bar{\mathcal{S}}$. Moreover, it is obvious that a candidate optimal strategy $(L_{T,b,s}^*(u), M_{T,b,s}^*(u))$ is the strategy which merely liquidates the stock position s . This allows us to define the function

$$h_0(t, b, s) := \mathbb{E}[U_p(X_{t,b,s}^{L^*,M^*}(T))], \quad (t, b, s) \in [0, T] \times \bar{\mathcal{S}}. \tag{4.1}$$

Our next aim is to show that $h_0 \in \mathbb{H}$ and hence $h_0 = \mathcal{V}$ and (L^*, M^*) is optimal. As a first step, we show that h_0 coincides with V on $[0, T] \times \mathcal{S}_0$.

Proposition 4.3 *The function h_0 defined in (4.1) coincides with the classical solution V of the HJB on the reduced state space $[0, T] \times \mathcal{S}_0$.*

Proof Let $(t, b, s) \in [0, T] \times \mathcal{S}_0$. If (t, b, s) is such that we are in one of the liquidation cases (a), (b) or (c), then direct computations reveal that $h_0(t, b, s) = V(t, b, s)$ since V is explicitly known at these points (cf. Dai and Yi [21, Proposition 3.2]). For example, assume that $p = 0, \pi_M > 1$ and $(t, b, s) = (t^{up}, 0, s)$. Then Dai and Yi [21, Proposition 3.2] show that

$$\begin{aligned} V(t, b, s) &= \log s + \log(1 - \mu) + \left(\alpha - \frac{1}{2}\sigma^2\right)(T - t) \\ &= \mathbb{E}[\log((1 - \mu)S_{t,s}^{L^*,M^*}(T))] \\ &= \mathbb{E}[U_0(b + (1 - \mu)S_{t,s}^{L^*,M^*}(T))] = h_0(t, b, s). \end{aligned}$$

We therefore exclude these cases in the sequel. For ease of notation, we denote the controlled processes $(B_{t,b,s}^{L^*,M^*}, S_{t,b,s}^{L^*,M^*})$ by (B^*, S^*) .

First, let us remark that $(t, b - (1 + \lambda)L^*(t) + (1 - \mu)M^*(t), s + L^*(t) - M^*(t))$ is contained in \mathcal{S}_0 since $(t, b, s) \in \mathcal{S}_0$ and $\mathcal{R}_0^{int} \subseteq [0, T] \times \mathcal{S}_0$. By the fundamental theorem of calculus for line integrals, we thus have

$$V(t, b, s) = V(t, b - (1 + \lambda)L^*(t) + (1 - \mu)M^*(t), s + L^*(t) - M^*(t))$$

since $dL^*(t) \neq 0$ only if $(t, b, s) \in \mathcal{R}_0^{buy}$ and $dM^*(t) \neq 0$ only if $(t, b, s) \in \mathcal{R}_0^{sell}$. For all $n \in \mathbb{N}$, define a stopping time

$$\tau_n := \inf \left\{ u \geq t : \int_t^u |\sigma S^*(r) \partial_s V(r, B^*(r), S^*(r))|^2 dr \geq n \right\} \wedge T.$$

Note that $(\cdot, B^*, S^*) \in ([0, T] \times \mathcal{S}_0) \setminus F$ after the initial transaction since we have ruled out the liquidation cases. Thus, using that V is in $C^{1,2}([0, T] \times \mathcal{S}_0) \setminus F$ and solves the HJB equation, we obtain

$$\begin{aligned} V(t, b, s) &= V(t, b - (1 + \lambda)L^*(t) + (1 - \mu)M^*(t), s + L^*(t) - M^*(t)) \\ &= V(\tau_n, B^*(\tau_n), S^*(\tau_n)) - \int_t^{\tau_n} \sigma S^*(u) \partial_s V(u, B^*(u), S^*(u)) dW(u). \end{aligned}$$

Taking expectations on both sides shows by the definition of τ_n that

$$V(t, b, s) = \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n))].$$

We are left with showing that

$$\lim_{n \rightarrow \infty} \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n))] = \mathbb{E}[U_p(X_{t,b,s}^{L^*,M^*}(T))] = h_0(t, b, s). \tag{4.2}$$

We first note that there exist constants $C_1, C_2 > 0$ such that

$$C_1 \mathcal{V}(t, b, s) \leq V(t, b, s) \leq C_2 \varphi_p(t, b, s) \quad \text{for all } (t, b, s) \in \overline{\mathcal{R}_0^{\text{nt}}}. \tag{4.3}$$

To see this, assume that $p \neq 0$ (the case $p = 0$ is similar). Since V, \mathcal{V} and φ_p are homogeneous of order p , we can write

$$\begin{aligned} V(t, b, s) &= (b + s)^p V(t, 1 - \pi, \pi), \\ \mathcal{V}(t, b, s) &= (b + s)^p \mathcal{V}(t, 1 - \pi, \pi), \\ \varphi_p(t, b, s) &= (b + s)^p \varphi_p(t, 1 - \pi, \pi), \end{aligned}$$

where $\pi = \pi(b, s) := s/(b + s)$. But $\pi(b, s)$ is bounded on $\overline{\mathcal{R}_0^{\text{nt}}}$, and hence so are $V(t, 1 - \pi, \pi), \mathcal{V}(t, 1 - \pi, \pi)$ and $\varphi_p(t, 1 - \pi, \pi)$, from which we infer (4.3).

Case 1: $p \in (0, 1)$. We claim that $(V(\tau_n, B^*(\tau_n), S^*(\tau_n)))_{n \in \mathbb{N}}$ is uniformly integrable, in which case (4.2) holds. Let $\varepsilon > 0$ be such that $p(1 + \varepsilon) < 1$. Then

$$\begin{aligned} 0 &\leq \mathbb{E}[|V(\tau_n, B^*(\tau_n), S^*(\tau_n))|^{1+\varepsilon}] \\ &\leq (C_2)^{1+\varepsilon} \mathbb{E}[|\varphi_p(\tau_n, B^*(\tau_n), S^*(\tau_n))|^{1+\varepsilon}] \\ &= (C_2)^{1+\varepsilon} \frac{1+\varepsilon}{p^\varepsilon} \mathbb{E}[\varphi_{p(1+\varepsilon)}(\tau_n, B^*(\tau_n), S^*(\tau_n))] \\ &\leq (C_2)^{1+\varepsilon} \frac{1+\varepsilon}{p^\varepsilon} \varphi_{p(1+\varepsilon)}(t, b, s). \end{aligned}$$

Here, we have used that $U_p = (1 + \varepsilon)p^{-\varepsilon} U_{p(1+\varepsilon)}$ for the equality in the third line, and the fact that $\varphi_{p(1+\varepsilon)}(\cdot, B^*, S^*)$ is a supermartingale (compare with the proof of Lemma 2.2 in Belak et al. [9]) to arrive at the last inequality.

Case 2: $p < 0$. We write

$$\begin{aligned} &\mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n))] \\ &= \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n < T\}}] + \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n = T\}}]. \end{aligned}$$

Since monotone convergence gives

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n=T\}}] &= \lim_{n \rightarrow \infty} \mathbb{E}[U_p(X_{t,b,s}^{L^*,M^*}(T)) \mathbb{1}_{\{\tau_n=T\}}] \\ &= \mathbb{E}[U_p(X_{t,b,s}^{L^*,M^*}(T))], \end{aligned}$$

we only have to show that $\lim_{n \rightarrow \infty} \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n < T\}}] = 0$. But this follows from admissibility of (L^*, M^*) and monotone convergence since

$$\begin{aligned} 0 &\geq \lim_{n \rightarrow \infty} \mathbb{E}[V(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n < T\}}] \\ &\geq C_1 \lim_{n \rightarrow \infty} \mathbb{E}[\mathcal{V}(\tau_n, B^*(\tau_n), S^*(\tau_n)) \mathbb{1}_{\{\tau_n < T\}}] \\ &\geq C_1 \lim_{n \rightarrow \infty} \mathbb{E}[U_p(X_{t,b,s}^{L^*,M^*}(T)) \mathbb{1}_{\{\tau_n < T\}}] = 0. \end{aligned}$$

Case 3: $p = 0$. This follows in a similar fashion as in Cases 1 and 2 by splitting $V(\tau_n, B^*(\tau_n), S^*(\tau_n))$ into its positive and negative parts and using that we have the estimate $-x^{-p}/p \leq \log x \leq x^p/p$ for every $p \in (0, 1)$. □

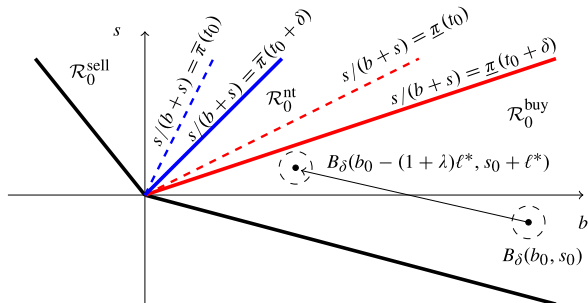
Proposition 4.3 proves the regularity of h_0 on $[0, T) \times \mathcal{S}_0$. The following lemmas investigate the regularity of h_0 for $s \leq 0$.

Lemma 4.4 *It holds that $h_0 \in C^{1,2}([0, t^{\text{down}}) \times (\mathcal{S} \setminus \mathcal{S}_0))$. Moreover, for every $(t, b, s) \in [0, t^{\text{down}}) \times (\mathcal{S} \setminus \mathcal{S}_0)$ and $\ell^* > 0$ with $(t, b - (1 + \lambda)\ell^*, s + \ell^*) \in \mathcal{R}_0^{\text{buy}}$ and $s + \ell^* > 0$, we have*

$$\begin{aligned} \partial_t h_0(t, b, s) &= \partial_t V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*, s+\ell^*)}, \\ \partial_b h_0(t, b, s) &= \partial_b V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*, s+\ell^*)}, \\ \partial_s h_0(t, b, s) &= \partial_s V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*, s+\ell^*)}, \\ \partial_s^2 h_0(t, b, s) &= \partial_s^2 V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*, s+\ell^*)}. \end{aligned}$$

Proof The idea is to bootstrap the regularity of h_0 for nonpositive s from the regularity of $h_0 = V$ for positive s inside the buy region $\mathcal{R}_0^{\text{buy}}$; see Fig. 4.

Fig. 4 Bootstrapping the regularity of h_0



Fix $(t_0, b_0, s_0) \in [0, T] \times (\mathcal{S} \setminus \mathcal{S}_0)$ (in particular $s_0 \leq 0$) with $t_0 < t^{\text{down}}$ and let $\delta > 0$ be such that $t_0 + \delta < t^{\text{down}}$ and $B_\delta(b_0, s_0) \subseteq \mathcal{S}$. By making δ smaller if necessary, we may furthermore assume that $[t_0, t_0 + \delta) \times B_\delta(b_0, s_0) \subseteq \mathcal{R}_0^{\text{buy}}$. Now, for every $(t, b, s) \in [t_0, t_0 + \delta) \times B_\delta(b_0, s_0)$, there exists some $\ell_0 > 0$ such that $(b - (1 + \lambda)\ell_0, s + \ell_0) \in \partial\mathcal{R}_0^{\text{nt}} \cap \partial\mathcal{R}_0^{\text{buy}}$ and $s + \ell_0 > 0$ (since $t < t^{\text{down}}$ and hence $\underline{\pi}(t) > 0$ by Theorem 2.3). Moreover, by the construction of h_0 ,

$$h_0(t, b, s) = h_0(t, b - (1 + \lambda)\ell, s + \ell)$$

for every $\ell \in [0, \ell_0]$. By the monotonicity of the buy boundary, making δ even smaller if necessary, we can therefore find some $\ell^* \in (0, \ell_0)$ such that the interval $[t_0, t_0 + \delta) \times B_\delta(b_0 - (1 + \lambda)\ell^*, s_0 + \ell^*)$ is contained in the interior of $\mathcal{R}_0^{\text{buy}}$ and $s > 0$ for all $(b, s) \in B_\delta(b_0 - (1 + \lambda)\ell^*, s_0 + \ell^*)$. Note that by construction and Proposition 4.3, we have

$$h_0(t, b, s) = h_0(t, b - (1 + \lambda)\ell^*, s + \ell^*) = V(t, b - (1 + \lambda)\ell^*, s + \ell^*)$$

for all $(t, b, s) \in [t_0, t_0 + \delta) \times B_\delta(b_0, s_0)$. Since

$$V \in C^{1,2}\left([t_0, t_0 + \delta) \times B_\delta(b_0 - (1 + \lambda)\ell^*, s_0 + \ell^*)\right),$$

the result follows. □

Proposition 4.5 *The function h_0 satisfies*

$$\mathcal{L}^{\text{nt}}h_0(t, b, s) \geq 0, \quad \mathcal{L}^{\text{buy}}h_0(t, b, s) = 0 \quad \text{and} \quad \mathcal{L}^{\text{sell}}h_0(t, b, s) > 0$$

in the classical sense on $[0, t^{\text{down}}) \times (\mathcal{S} \setminus \mathcal{S}_0)$.

Proof It follows immediately from Lemma 4.4 for all $(t, b, s) \in [0, t^{\text{down}}) \times (\mathcal{S} \setminus \mathcal{S}_0)$ that

$$\mathcal{L}^{\text{buy}}h_0(t, b, s) = \mathcal{L}^{\text{buy}}V(t, b - (1 + \lambda)\ell^*, s + \ell^*) = 0, \tag{4.4}$$

$$\mathcal{L}^{\text{sell}}h_0(t, b, s) = \mathcal{L}^{\text{sell}}V(t, b - (1 + \lambda)\ell^*, s + \ell^*) > 0, \tag{4.5}$$

for a suitable choice of ℓ^* . From (4.4), we obtain

$$\partial_s h_0(t, b, s) = (1 + \lambda)\partial_b h_0(t, b, s).$$

Plugging this into (4.5) yields

$$(1 - \mu)\partial_b h_0(t, b, s) < \partial_s h_0(t, b, s) = (1 + \lambda)\partial_b h_0(t, b, s)$$

which implies $\partial_b h_0(t, b, s) > 0$ and thus $\partial_s h_0(t, b, s) > 0$ for all $(b, s) \in (\mathcal{S} \setminus \mathcal{S}_0)$. It only remains to show that

$$\mathcal{L}^{\text{nt}}h_0(t, b, s) \geq 0.$$

Case 1: $s = 0$. Fix some $\ell^* > 0$ such that $(t, b, \ell^*) \in \mathcal{R}_0^{\text{buy}}$ so that

$$\partial_t h_0(t, b, 0) = \partial_t V(t, b - (1 + \lambda)\ell^*, \ell^*) \leq 0$$

by Theorem 2.2. Therefore,

$$\mathcal{L}^{\text{nt}} h_0(t, b, 0) = -\partial_t h_0(t, b, 0) \geq 0. \tag{4.6}$$

Case 2: $s < 0$. For some suitable $\ell^* > 0$, we have

$$\begin{aligned} \partial_t h_0(t, b, s) &= \partial_t V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*,s+\ell^*)} = \partial_t h_0(t, b - (1 + \lambda)s, 0), \\ \partial_s h_0(t, b, s) &= \partial_s V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*,s+\ell^*)} = \partial_s h_0(t, b - (1 + \lambda)s, 0), \\ \partial_s^2 h_0(t, b, s) &= \partial_s^2 V(t, b, s)|_{(b,s)=(b-(1+\lambda)\ell^*,s+\ell^*)} = \partial_s^2 h_0(t, b - (1 + \lambda)s, 0). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathcal{L}^{\text{nt}} h_0(t, b, s) &= -\partial_t h_0(t, b, s) - \alpha s \partial_s h_0(t, b, s) - \frac{1}{2} \sigma^2 s^2 \partial_s^2 h_0(t, b, s) \\ &= -\partial_t h_0(t, b - (1 + \lambda)s, 0) - \alpha s \partial_s h_0(t, b - (1 + \lambda)s, 0) \\ &\quad - \frac{1}{2} \sigma^2 s^2 \partial_s^2 h_0(t, b - (1 + \lambda)s, 0). \end{aligned}$$

By (4.6), we have

$$-\partial_t h_0(t, b - (1 + \lambda)s, 0) = \mathcal{L}^{\text{nt}} h_0(t, b - (1 + \lambda)s, 0) \geq 0.$$

Moreover, since $s < 0$ and $\partial_s h_0(t, b - (1 + \lambda)s, 0) > 0$, we have

$$-\alpha s \partial_s h_0(t, b - (1 + \lambda)s, 0) > 0,$$

and since $\partial_s^2 h_0(t, b - (1 + \lambda)s, 0) \leq 0$ (V is concave; see Dai and Yi [21, Remark 4.2]), we see that

$$-\frac{1}{2} \sigma^2 s^2 \partial_s^2 h_0(t, b - (1 + \lambda)s, 0) \geq 0.$$

Putting the pieces together, we obtain $\mathcal{L}^{\text{nt}} h_0(t, b, s) > 0$. □

We have similar statements for the case $t \geq t^{\text{down}}$ with $s < 0$; the proofs are however significantly easier.

Lemma 4.6 *We have $h_0 \in C^\infty([t^{\text{down}}, T) \times (\mathcal{S} \setminus \{(b, s) \in \mathcal{S} : s \geq 0\}))$, and h_0 is given explicitly as*

$$h_0(t, b, s) = U_p(b + (1 + \lambda)s)$$

on $[t^{\text{down}}, T) \times \mathcal{S} \setminus \{(b, s) \in \mathcal{S} : s > 0\}$.

Proof This is an immediate consequence of the definition of h_0 and (L^*, M^*) . Indeed, if $(t, b, s) \in [t^{\text{down}}, T) \times (\mathcal{S} \setminus \{(b, s) \in \mathcal{S} : s > 0\})$, then (L^*, M^*) is such that the stock position is immediately liquidated and the investor refrains from further trading. That is,

$$\begin{aligned} h_0(t, b, s) &= h_0(t, b + (1 + \lambda)s, 0) \\ &= \mathbb{E}[U_p(B_{t, b+(1+\lambda)s}^{L^*, M^*}(T) + 0)] = U_p(b + (1 + \lambda)s), \end{aligned}$$

from which the assertion of the lemma follows. □

Proposition 4.7 *The function h_0 satisfies*

$$\mathcal{L}^{\text{nt}}h_0(t, b, s) > 0, \quad \mathcal{L}^{\text{buy}}h_0(t, b, s) = 0 \quad \text{and} \quad \mathcal{L}^{\text{sell}}h_0(t, b, s) > 0$$

in the classical sense on $[t^{\text{down}}, T) \times (\mathcal{S} \setminus \{(b, s) \in \mathcal{S} : s \geq 0\})$.

Proof This follows directly since $h_0(t, b, s) = U_p(b + (1 + \lambda)s)$ by Lemma 4.6. □

Corollary 4.8 *For every $t \in [t^{\text{down}}, T)$, we have $h_0(t, \cdot) \in C(\bar{\mathcal{S}})$.*

Proof It suffices to show that $(b, s) \mapsto h_0(t, b, s)$ is continuous at $(b, 0)$ for every $b > 0$. By Lemma 4.6, $h_0(t, b, 0)$ is given explicitly as $h_0(t, b, 0) = U_p(b)$. Moreover, for $s > 0$, $h_0(t, b, s) = U_p(b + (1 - \mu)s)$ and $\partial_t h_0 \leq 0$ and thus

$$h_0(t, b, s) \geq h_0(T, b, s) \geq h_0(t, b, 0) = U_p(b) \quad \text{for all } s > 0,$$

which implies that h_0 is at least lower semi-continuous at $(b, 0)$ for every $b > 0$. By Theorems 2.2 and 2.3, $h_0 = V$ is $C^{1,2}$ and satisfies

$$\partial_t h_0(t, b, s) = -\alpha s \partial_s h_0(t, b, s) - \frac{1}{2} \sigma^2 s^2 \partial_s^2 h_0(t, b, s) \geq -\alpha s \partial_s h_0(t, b, s) \quad (4.7)$$

for $t \in [t^{\text{down}}, T)$, $b > 0$ and $s > 0$ sufficiently small (so that (t, b, s) is in the no-trading region) since $\partial_s^2 h_0 \leq 0$. An argument as in Lemma 3.4 in Dai and Yi [21] shows that $\partial_s h_0(t, \cdot)$ is bounded from above on $(b - \delta, b + \delta) \times (0, \bar{s}]$ for every $0 < \delta < b$ and $\bar{s} > 0$, uniformly in $t \in [t^{\text{down}}, T)$. Thus (4.7) implies that

$$\liminf_{\hat{b} \rightarrow b, s \downarrow 0} \partial_t h_0(t, \hat{b}, s) \geq 0 \quad \text{uniformly in } t \in [t^{\text{down}}, T).$$

But since $\partial_t h_0 \leq 0$, we must have

$$\lim_{\hat{b} \rightarrow b, s \downarrow 0} \partial_t h_0(t, \hat{b}, s) = 0 \quad \text{uniformly in } t \in [t^{\text{down}}, T). \quad (4.8)$$

Now choose $(b_k, s_k)_{k \in \mathbb{N}}$ with $b_k, s_k > 0$ and $(b_k, s_k) \rightarrow (b, 0)$. Then (4.8) implies that for every $\varepsilon > 0$, there exists some $K \in \mathbb{N}$ such that

$$h_0(t, b_k, s_k) \leq h_0(T, b_k, s_k) + \varepsilon = U_p(b_k + (1 - \mu)s_k) + \varepsilon \quad \text{for all } k \geq K,$$

from which we conclude that

$$\limsup_{k \rightarrow \infty} h_0(t, b_k, s_k) \leq \lim_{k \rightarrow \infty} U_p(b_k + (1 - \mu)s_k) = U_p(b) = h_0(t, b, 0),$$

i.e., h_0 is also upper semi-continuous at $(b, 0)$. □

Remark 4.9 We cannot expect more regularity of h_0 at $s = 0$. Indeed, taking the derivative with respect to s in the equality in (4.7) and formally sending $s > 0$ to zero shows that $\partial_s h_0(t, b, 0+)$ solves

$$\begin{aligned} 0 &= -\partial_t \partial_s h_0(t, b, 0+) - \alpha \partial_s h_0(t, b, 0+), \\ h_0(T, b, 0+) &= \partial_s U_p(b + (1 + \lambda)s)|_{s=0}, \end{aligned}$$

i.e., $\partial_s h_0(t, b, 0+) = e^{\alpha(T-t)} \partial_s U_p(b + (1 + \lambda)s)|_{s=0}$. This is in contrast to

$$\partial_s h_0(t, b, 0-) = \partial_s U_p(b + (1 + \lambda)s)|_{s=0},$$

suggesting that $\partial_s h_0$ is not continuous at $s = 0$.

We are now ready to prove that $h_0 \in \mathbb{H}$. By construction, we already know that $h_0 \leq \mathcal{V} \leq \varphi_p$. Moreover, it is clear from the above analysis that

$$h_0(t, b, s) \geq U_p(b + \min\{(1 - \mu)s, (1 + \lambda)s\})$$

since this is clearly satisfied for $t = T$ and h_0 is nonincreasing in t .

We proceed in three steps. First we show that h_0 is continuous, then we show that h_0 is superharmonic, and finally we show that h_0 is nonincreasing in the direction of transactions.

Proposition 4.10 *The function h_0 is continuous on $[0, T] \times \overline{\mathcal{S}}$.*

Proof By Proposition 4.3, Lemmas 4.4 and 4.6 and Corollary 4.8, it only remains to prove that h_0 is continuous in $(t^{\text{down}}, b, 0)$ for every $b \geq 0$. Moreover, by Lemma 4.6 and Corollary 4.8, we can restrict attention to sequences $(t_n, b_n, s_n)_{n \in \mathbb{N}}$ converging to $(t^{\text{down}}, b, 0)$ with $t_n < t^{\text{down}}$ and $s_n > 0$ for all $n \in \mathbb{N}$. Next, we observe that every such sequence is eventually contained in $\mathcal{R}_0^{\text{buy}} \cup \mathcal{R}_0^{\text{nt}}$ since $s_n \rightarrow 0$ and $\overline{\pi} > 0$, and we may furthermore assume $b_n > 0$ for all $n \in \mathbb{N}$. Now, if $(t_n, b_n, s_n) \in \mathcal{R}_0^{\text{buy}}$, there is $\ell_n > 0$ with $(t_n, b_n - (1 + \lambda)\ell_n, s_n + \ell_n) \in \partial \mathcal{R}_0^{\text{nt}}$ and

$$h_0(t_n, b_n, s_n) = h_0(t_n, b_n - (1 + \lambda)\ell_n, s_n + \ell_n),$$

which implies that we may restrict further to sequences in $\overline{\mathcal{R}_0^{\text{nt}}}$. But because we have $\mathcal{L}^{\text{nt}} h_0(t, b, s) = \mathcal{L}^{\text{nt}} V(t, b, s) = 0$ for all $(t, b, s) \in \overline{\mathcal{R}_0^{\text{nt}}}$ with $s > 0$, we can argue as in the proof of Corollary 4.8. □

Proposition 4.11 *The function h_0 is superharmonic.*

Proof Fix $(t, b, s) \in [0, T) \times \mathcal{S}$, let τ be a $[t, T]$ -valued stopping time and ϑ the first exit time of the uncontrolled portfolio process $(B_{t,b}^0, S_{t,s}^0)$ from \mathcal{S} .

Case 1: $s > 0$. In this case, $S_{t,s}^0(u) > 0$ for all $u \in [t, \tau \wedge \vartheta]$ and hence $h_0 = V$ is $C^{1,2}$ along the paths of $(u, B^0(u), S^0(u))$ by Proposition 4.3. Let $\varepsilon > 0$ and denote by ϑ^ε the first exit time of $(B_{t,b}^0 + \varepsilon, S_{t,s}^0)$ from \mathcal{S} . Then clearly $\vartheta^\varepsilon \geq \vartheta$. For every $n \in \mathbb{N}$, let us define a stopping time

$$\tau_n := \inf \left\{ u \geq t : \int_t^u |\sigma S_{t,s}^0(r) \partial_s h_0(r, B_{t,b}^0(r) + \varepsilon, S_{t,s}^0(r))|^2 dr \geq n \right\} \wedge \tau \wedge \vartheta.$$

An application of Itô’s formula shows that

$$\begin{aligned} h_0(t, b + \varepsilon, s) &= h_0(\tau_n, B_{t,b}^0(\tau_n) + \varepsilon, S_{t,s}^0(\tau_n)) \\ &\quad + \int_t^{\tau_n} \mathcal{L}^{\text{nt}} h_0(u, B_{t,b}^0(u) + \varepsilon, S_{t,s}^0(u)) du \\ &\quad + \int_t^{\tau_n} \sigma S_{t,s}^0(u) \partial_s h_0(u, B_{t,b}^0(u) + \varepsilon, S_{t,s}^0(u)) dW(u). \end{aligned}$$

Taking expectations yields

$$\begin{aligned} h_0(t, b + \varepsilon, s) &= \mathbb{E} \left[h_0(\tau_n, B_{t,b}^0(\tau_n) + \varepsilon, S_{t,s}^0(\tau_n)) \right. \\ &\quad \left. + \int_t^{\tau_n} \mathcal{L}^{\text{nt}} h_0(u, B_{t,b}^0(u) + \varepsilon, S_{t,s}^0(u)) du \right]. \end{aligned}$$

By Proposition 4.3, h_0 is a classical solution of the HJB on $[0, T] \times \mathcal{S}_0$, and so we see that

$$\int_t^{\tau_n} \mathcal{L}^{\text{nt}} h_0(u, B_{t,b}^0(u), S_{t,s}^0(u)) du \geq 0$$

and hence

$$h_0(t, b + \varepsilon, s) \geq \mathbb{E}[h_0(\tau_n, B_{t,b}^0(\tau_n) + \varepsilon, S_{t,s}^0(\tau_n))]. \tag{4.9}$$

Next, since $\partial_t h_0 \leq 0$ and $h_0(T, b, s) = V(T, b, s) = U_p(b + (1 - \mu)s)$, we have $h_0(t, b, s) \geq U_p(b + (1 - \mu)s)$ and hence

$$h_0(\tau_n, B_{t,b}^0(\tau_n) + \varepsilon, S_{t,s}^0(\tau_n)) \geq U_p(\varepsilon).$$

We can therefore send $n \rightarrow \infty$ in (4.9), use Fatou’s lemma and $\partial_b h_0 \geq 0$ to obtain

$$\begin{aligned} h_0(t, b + \varepsilon, s) &\geq \mathbb{E}[h_0(\tau \wedge \vartheta, B_{t,b}^0(\tau \wedge \vartheta) + \varepsilon, S_{t,s}^0(\tau \wedge \vartheta))] \\ &\geq \mathbb{E}[h_0(\tau \wedge \vartheta, B_{t,b}^0(\tau \wedge \vartheta), S_{t,s}^0(\tau \wedge \vartheta))]. \end{aligned}$$

Sending $\varepsilon \downarrow 0$ then shows that h is superharmonic.

Case 2: $s \leq 0$ and $t \geq t^{\text{down}}$. For simplicity, we write $\tilde{\tau} := \tau \wedge \vartheta$. If $t \geq t^{\text{down}}$, then (L^*, M^*) performs an initial transaction from (b, s) to $(b + (1 + \lambda)s, 0)$ so that

$$h_0(t, b, s) = h(t, b + (1 + \lambda)s, 0) = U_p(b + (1 + \lambda)s). \tag{4.10}$$

On the other hand, by the same arguments and Jensen’s inequality, we have

$$\begin{aligned} \mathbb{E}[h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))] &= \mathbb{E}[h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}) + (1 + \lambda)S_{t,s}^0(\tilde{\tau}), 0)] \\ &= \mathbb{E}[U_p(B_{t,b}^0(\tilde{\tau}) + (1 + \lambda)S_{t,s}^0(\tilde{\tau}))] \\ &\leq U_p(\mathbb{E}[B_{t,b}^0(\tilde{\tau}) + (1 + \lambda)S_{t,s}^0(\tilde{\tau})]) \\ &= U_p(b + (1 + \lambda)\mathbb{E}[S_{t,s}^0(\tilde{\tau})]). \end{aligned}$$

Now since $S_{t,s}^0$ is a supermartingale for every $s \leq 0$, it follows that $\mathbb{E}[S_{t,s}^0(\tilde{\tau})] \leq s$ and hence

$$\mathbb{E}[h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))] \leq U_p(b + (1 + \lambda)s) = h_0(t, b, s)$$

by the monotonicity of U_p and (4.10).

Case 3: $s \leq 0$ and $t < t^{\text{down}}$. We have

$$\begin{aligned} h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau})) &= h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))\mathbb{1}_{\{\tilde{\tau} < t^{\text{down}}\}} \\ &\quad + h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))\mathbb{1}_{\{\tilde{\tau} \geq t^{\text{down}}\}}. \end{aligned}$$

On $\{\tilde{\tau} \geq t^{\text{down}}\}$, we have as before that

$$\mathbb{E}[h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))] \leq U_p(b + (1 + \lambda)s) = h_0(t, b, s);$$

so we may without loss of generality assume that $\tilde{\tau} < t^{\text{down}}$. However, since we know by Lemma 4.4 and Proposition 4.5 that h_0 is $C^{1,2}$ and satisfies the HJB in the classical sense, we obtain

$$h(t, b, s) \geq \mathbb{E}[h_0(\tilde{\tau}, B_{t,b}^0(\tilde{\tau}), S_{t,s}^0(\tilde{\tau}))]$$

as in the case $s > 0$. □

Proposition 4.12 *The function h_0 is nonincreasing in the direction of transactions.*

Proof Fix $(t, b, s) \in [0, T] \times \bar{\mathcal{S}}$ and let $\ell, m \geq 0$ be such that

$$(b - (1 + \lambda)\ell + (1 - \mu)m, s + \ell - m) \in \bar{\mathcal{S}}.$$

We have to show that

$$h_0(t, b, s) \geq h_0(t, b - (1 + \lambda)\ell + (1 - \mu)m, s + \ell - m).$$

However, by Proposition 4.10, $h(t, \cdot)$ is continuous and satisfies

$$\mathcal{L}^{\text{buy}}h_0(t, b, s) \geq 0 \quad \text{and} \quad \mathcal{L}^{\text{sell}}h_0(t, b, s) \geq 0$$

for every $(t, b, s) \in ([0, T) \times \mathcal{S}) \setminus \{(t, b, s) \in [0, T) \times \mathcal{S} : t \geq t^{\text{down}}, s = 0\}$ and $(t, b, s) \notin F$ (defined in (2.3)) by Propositions 4.3, 4.5 and 4.7. Therefore, by the fundamental theorem of calculus for line integrals, we immediately obtain the claim. \square

Combining Propositions 4.10–4.12 proves the optimality of (L^*, M^*) .

Theorem 4.13 *We have $h_0 \in \mathbb{H}$, and thus $h_0 = \mathcal{V}$ and (L^*, M^*) is optimal.*

Since $h_0 = \mathcal{V}$, we furthermore have the following regularity result.

Theorem 4.14 *The value function \mathcal{V} is continuous everywhere and (at least) of class $C^{1,2}$ except for possibly the points (t, b, s) for which one of the following statements is true:*

- 1) $b = 0$ and (t, b, s) is on the buy boundary.
 - 2) $\pi_M = 1$ and $b = 0$.
 - 3) $t = t^{\text{down}}$ and $s \leq 0$. However, $\mathcal{V}(t^{\text{down}}, \cdot) \in C^2(\mathcal{S} \setminus \{(b, s) \in \mathcal{S} : s = 0\})$.
 - 4) $t \geq t^{\text{down}}$ and $s = 0$. However, $\mathcal{V}(\cdot, b, 0) \in C^\infty((t^{\text{down}}, T))$ for all $b \geq 0$.
- Moreover, \mathcal{V} is of class C^∞ on $\mathcal{R}_0^{\text{nt}}$.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Akian, M., Menaldi, J.L., Sulem, A.: On an investment-consumption model with transaction costs. *SIAM J. Control Optim.* **34**, 329–364 (1996)
2. Akian, M., S equier, P., Sulem, A.: A finite horizon multidimensional portfolio selection problem with singular transactions. In: Proceedings of the 34th IEEE Conference on Decision and Control, vol. 3, pp. 2193–2198. IEEE Press New York (1995)
3. Akian, M., Sulem, A., Taksar, M.I.: Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logarithmic utility. *Math. Finance* **11**, 153–188 (2001)
4. Bayraktar, E., Sirbu, M.: Stochastic Perron's method and verification without smoothness using viscosity comparison: the linear case. *Proc. Am. Math. Soc.* **140**, 3645–3654 (2012)
5. Bayraktar, E., Sirbu, M.: Stochastic Perron's method for Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* **51**, 4274–4294 (2013)
6. Bayraktar, E., Sirbu, M.: Stochastic Perron's method and verification without smoothness using viscosity comparison: obstacle problems and Dynkin games. *Proc. Am. Math. Soc.* **142**, 1399–1412 (2014)
7. Bayraktar, E., Zhang, Y.: Stochastic Perron's method for the probability of lifetime ruin problem under transaction costs. *SIAM J. Control Optim.* **53**, 91–113 (2015)
8. Belak, C., Christensen, S., Seifried, F.T.: A general verification result for stochastic impulse control problems. *SIAM J. Control Optim.* **55**, 627–649 (2017)
9. Belak, C., Menkens, O., Sass, J.: On the uniqueness of unbounded viscosity solutions arising in an optimal terminal wealth problem with transaction costs. *SIAM J. Control Optim.* **53**, 2878–2897 (2015)
10. Belak, C., Menkens, O., Sass, J.: Worst-case portfolio optimization with proportional transaction costs. *Stochastics* **87**, 623–663 (2015)
11. Bichuch, M.: Asymptotic analysis for optimal investment in finite time with transaction costs. *SIAM J. Financ. Math.* **3**, 433–458 (2012)
12. Chen, Y., Dai, M., Zhao, K.: Finite horizon optimal investment and consumption with CARA utility and proportional transaction costs. In: Zhang, T., Zhou, X. (eds.) *Stochastic Analysis and Applications to Finance*. Interdisciplinary Mathematical Sciences, pp. 39–54. World Scientific, Hackensack (2012)

13. Choi, J.H., Sîrbu, M., Žitković, G.: Shadow prices and well-posedness in the problem of optimal investment and consumption with transaction costs. *SIAM J. Control Optim.* **51**, 4414–4449 (2013)
14. Christensen, S.: On the solution of general impulse control problems using superharmonic functions. *Stoch. Process. Appl.* **124**, 709–729 (2014)
15. Constantinides, G.M., Magill, M.J.P.: Portfolio selection with transaction costs. *J. Econ. Theory* **13**, 245–263 (1976)
16. Czichowsky, C., Schachermayer, W.: Duality theory for portfolio optimisation under transaction costs. *Ann. Appl. Probab.* **26**, 1888–1941 (2016)
17. Czichowsky, C., Schachermayer, W.: Strong supermartingales and limits of nonnegative martingales. *Ann. Probab.* **44**, 171–205 (2016)
18. Czichowsky, C., Schachermayer, W.: Portfolio optimisation beyond semimartingales: shadow prices and fractional Brownian motion. *Ann. Appl. Probab.* **27**, 1414–1451 (2017)
19. Czichowsky, C., Schachermayer, W., Yang, J.: Shadow prices for continuous processes. *Math. Finance* **27**, 623–658 (2017)
20. Dai, M., Jiang, L.S., Li, P.F., Yi, F.H.: Finite horizon optimal investment and consumption with transaction costs. *SIAM J. Control Optim.* **48**, 1134–1154 (2009)
21. Dai, M., Yi, F.H.: Finite-horizon optimal investment with transaction costs: a parabolic double obstacle problem. *J. Differ. Equ.* **246**, 1445–1469 (2009)
22. Dai, M., Zhong, Y.: Penalty methods for continuous-time portfolio selection with proportional transaction costs. *J. Comput. Finance* **13**, 1–31 (2010)
23. Davis, M.H.A., Norman, A.R.: Portfolio selection with transaction costs. *Math. Oper. Res.* **15**, 676–713 (1990)
24. Davis, M.H.A., Panas, V.G., Zariphopoulou, T.: European option pricing with transaction costs. *SIAM J. Control Optim.* **31**, 470–493 (1993)
25. de Vallière, D., Kabanov, Y., Lépinette, E.: Consumption–investment problem with transaction costs for Lévy-driven price processes. *Finance Stoch.* **20**, 705–740 (2016)
26. Gerhold, S., Guasoni, P., Muhle-Karbe, J., Schachermayer, W.: Transaction costs, trading volume, and the liquidity premium. *Finance Stoch.* **18**, 1–37 (2014)
27. Gerhold, S., Muhle-Karbe, J., Schachermayer, W.: Asymptotics and duality for the Davis and Norman problem. *Stochastics* **84**, 625–641 (2012)
28. Gerhold, S., Muhle-Karbe, J., Schachermayer, W.: The dual optimizer for the growth-optimal portfolio under transaction costs. *Finance Stoch.* **17**, 325–354 (2013)
29. Herczegh, A., Prokaj, V.: Shadow price in the power utility case. *Ann. Appl. Probab.* **25**, 2671–2707 (2015)
30. Herzog, R., Kunisch, K., Sass, J.: Primal–dual methods for the computation of trading regions under proportional transaction costs. *Math. Methods Oper. Res.* **77**, 101–130 (2013)
31. Janeček, K., Shreve, S.E.: Asymptotic analysis for optimal investment and consumption with transaction costs. *Finance Stoch.* **8**, 181–206 (2004)
32. Kabanov, Y., Klüppelberg, C.: A geometric approach to portfolio optimization in models with transaction costs. *Finance Stoch.* **8**, 207–227 (2004)
33. Kallsen, J., Li, S.: Portfolio optimization under small transaction costs: a convex duality approach (2013). Preprint, Available online at [arXiv:1309.3479](https://arxiv.org/abs/1309.3479)
34. Kallsen, J., Muhle-Karbe, J.: On using shadow prices in portfolio optimization with transaction costs. *Ann. Appl. Probab.* **20**, 1341–1358 (2010)
35. Kallsen, J., Muhle-Karbe, J.: The general structure of optimal investment and consumption with small transaction costs. *Math. Finance* **27**, 659–703 (2017)
36. Kunisch, K., Sass, J.: Trading regions under proportional transaction costs. In: Waldmann, K.-H., Stocker, U.M. (eds.) *Operations Research Proceedings, Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), September 6–8, 2006* pp. 563–568. Springer, Berlin (2007)
37. Shreve, S.E., Soner, H.M.: Optimal investment and consumption with transaction costs. *Ann. Appl. Probab.* **4**, 609–692 (1994)
38. Słomiński, L., Wojciechowski, T.: Stochastic differential equations with time-dependent reflecting barriers. *Stochastics* **85**, 27–47 (2013)