Finance Stoch (2016) 20:99–121
DOI 10.1007/s00780-015-0274-y

Facelifting in utility maximization

Kasper Larsen[1] · Halil Mete Soner[2] · Gordan Žitković[3]

Received: 9 April 2014 / Published online: 17 September 2015

Abstract We establish the existence and characterization of a primal and a dual facelift—discontinuity of the value function at the terminal time—for utility maximization in incomplete semimartingale-driven financial markets. Unlike in the lower and upper hedging problems, and somewhat unexpectedly, a facelift turns out to exist in utility maximization despite strict convexity in the objective function. In addition to discussing our results in their natural, Markovian environment, we also use them to show that the dual optimizer cannot be found in the set of countably additive (martingale) measures in a wide variety of situations.

Keywords Boundary layer · Convex analysis · Convex duality · Facelift · Financial mathematics · Incomplete markets · Markov processes · Utility maximization · Unspanned endowment

Mathematics Subject Classification (2010) 91G10 · 91G80 · 60K35

JEL Classification C61 · G11

During the preparation of this work, the first author has been supported by the National Science Foundation under Grant No. DMS-1411809 (2014–2017). The second author has been supported by the Swiss National Foundation through grant SNF 200021_153555 and by the Swiss Finance Institute. The third author has been supported by the National Science Foundation under Grants No. DMS-0706947 (2010–2015) and No. DMS-1107465 (2012–2017). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

✉ K. Larsen
kasperl@andrew.cmu.edu

H.M. Soner
mete.soner@math.ethz.ch

G. Žitković
gordanz@math.utexas.edu

[1] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[2] Department of Mathematics, ETH, 8092 Zürich, Switzerland

[3] Department of Mathematics, University of Texas at Austin, Austin, TX 78712, USA

1 Introduction

Valuation, or pricing, is one of the central problems in mathematical finance. Its goal is to assign a monetary value to a contingent claim based on the economic principles of supply and demand. When the claim is liquidly traded in a financial market, the only meaningful notion of price is the one at which the claim is trading. When the claim is not traded, but is replicable in an arbitrage-free market, its unique price is determined by the no-arbitrage principle. The case of a nonreplicable claim is the most complex one. Here, the no-arbitrage principle alone does not suffice, and additional economic input is needed. This input usually comes in the form of a risk profile of the agents involved in the transaction. In extreme cases, one comes up with the notions of upper and lower hedging prices, whereas in between those, the pricing procedure typically involves a solution of a utility maximization problem.

Utility maximization problems arise in contexts such as optimal investment and equilibrium problems. In fact, they play a central role in mathematical finance and financial economics. This fact is quite evident from the range of literature both in mathematics as well as economics and finance that treats them. Instead of providing a list of the most important references, we simply point the reader to the monograph [12] and the references therein for a thorough literature review from the inception of the subject to 1998. The more recent history, at least as far as the relevance to the present paper is concerned, can be found in the papers [11] and [5], where the problem is treated in great mathematical generality.

Both in pricing and utility maximization, there is a significant jump in mathematical and conceptual difficulty as one transitions from complete to incomplete models. In pricing, it is well known that the upper (and lower) hedging price of a nonreplicable claim cannot be expressed as the expectation under a (local, σ-) martingale measure (see [6, Theorem 5.16]). In other words, when viewed as a linear optimization problem over the set of martingale measures, the value of the upper hedging problem is attained only when a suitable relaxation is introduced. This relaxation almost always (implicitly or explicitly) involves a closure in the weak* topology and the passage from countably additive to merely finitely additive measures. This phenomenon is well understood not only from the functional-analytic, but also from the control-theoretic and analytic points of view. Indeed, stochastic target problems (introduced in [21]; see [22, Chap. 8] for an overview and further references) provide an approach using the related partial differential equations. We also understand that the passage from countable to finite additivity corresponds, loosely speaking, to the lack of weak compactness of minimizing sequences, and that it often leads to a discontinuity in the problem's value function at the terminal time. In mathematical finance, this naturally leads to a "facelifting" procedure where one upper-hedges (or lower-hedges) a contingent claim by (perfectly) hedging another contingent claim whose payoff is an upper majorant of the original payoff in a specific class (see, e.g., [2, 3, 19–21, 7], and [4]). Let us note that in the present paper, the term facelift refers to a formally different concept. Namely, it pertains to the discontinuity of the value function in an optimization problem at the terminal time. This is related—but not identical—to the terminology used, for example, in [7], where "facelift" refers to the modified replicable claim, which is perfectly replicated in order to superhedge the original claim. However, in a typical Markovian setting, such a claim will be given as a function of the state variables, whose difference from the original payoff is exactly the size of the discontinuity in the value function (superreplication price). Yet other, slightly different interpretations can be found in the literature, but they all seem to fall within the conceptual ballpark of our use.

The literature on nonlinear problems such as utility maximization and utility-based pricing is much narrower in scope. In this context, one must also distinguish between the need for relaxation and the existence of a facelift. Whereas, as we show below, the existence of a facelift is related to the nonexistence of a minimizer without an appropriate relaxation in most cases of interest, the opposite implication does not hold. In fact, the only known cases in the literature where it is shown that a relaxation is necessary (in [11] and [8]) do not come with a facelift (as can be deduced from our main theorem). Moreover, they appear in non-Markovian settings, are constructed using heavy functional-analytic machinery (the Rosenthal subsequence splitting lemma in [8], e.g.) and do not involve a nonreplicable random endowment.

When a nonreplicable random endowment is present—which is invariably the case when one wants to consider pricing approaches other than marginal utility-based pricing (such as indifference or conditional marginal pricing)—no answer can be found in the existing literature. Indeed, the papers [5, 13], and others treat such problems theoretically, and both pose and solve a class of dual utility maximization problems over an appropriate relaxation of the set of (σ- or local) martingale measures. Intuitively, this relaxation is helpful because of the interplay between the strong intertemporal admissibility constraints on the trading strategies and the sudden appearance of the endowment at the terminal date. Nevertheless, no rigorous proof of the necessity of such an enlargement has been given in the literature before. One can speculate that one reason why such results do not exist is because we never expected to see a facelift in such problems. Therefore, by a somewhat perverted logic, we did not expect a finitely additive relaxation to be truly necessary, except in pathological cases. After all, the objective function is strictly convex—there are no "flat parts" to produce infinite Hamiltonians and the related explosion in control, which leads to the emergence of a facelift (see, e.g., [15, Sect. 4.3.2] for an accessible treatment). Indeed, if one tries to apply the "exploding Hamiltonian" test to virtually any Markovian incarnation of a (primal or dual) utility maximization problem with a random endowment, the results will be inconclusive—the Hamiltonian never explodes. It came consequently as a great surprise to us when we discovered that the "Hamiltonian test" is impotent in this case and that the facelift appears virtually generically. Moreover, there is no need for pathology at all. As we explain in our illustrative Sect. 2, in what one can quite confidently call the "simplest nontrivial incomplete utility maximization problem with nonreplicable random endowment," the facelift invariably appears. Moreover, in many setups, every time it appears, one can show that the corresponding dual problem does not admit a minimizer in the class of countably additive measures. This fact is not only of theoretical value; it has important implications for the numerical treatment of the problem. Indeed, the lack of countable additivity of the dual optimizer points to a particularly unwieldy type of unboundedness—or, worse, the nonexistence—of the optimal dual control.

After the aforementioned illustrative example in Sect. 2, we turn to a general semimartingale model of a financial market in Sect. 3 and analyze the asymptotic behavior of the value function of the dual utility maximization problem with random endowment as the time horizon shrinks to 0. While keeping the same underlying market structure, we let the random endowment vary with the horizon in a rather general fashion. We show here that the limiting value of the value function exists under minimal conditions on the inputs, compute its value explicitly, and argue that it often differs from the limiting value of the objective, that is, that a facelift exists.

The choice of the shrinking time horizon—as opposed to the one of Sect. 2, where the current time gets closer and closer to the horizon—is made here for mathematical convenience. Whereas it may be of interest in it own right when one wants to study utility maximization on very short horizons, our main concern is to understand how the value function of the (dual) utility maximization problem behaves close to maturity. In Markovian models, as described in Sect. 3.6, the two views can be reconciled by observing that various control problems corresponding to the same value of the state variable, but varying values of the time parameter, can be coupled on the same probability space. In this way, the study of the "forward" convergence of value functions can be aided by the natural RCLL properties of trajectories of canonical Markov processes and the abstract results of Sect. 3.

In Sect. 4, we take up a related problem and show that under mild conditions on the random endowment, the objective function in the dual utility maximization problem can be replaced by a smaller function without changing its value. We use this to show that if the random endowment is nonreplicable and its negative admits a unique minimal (smallest) replicable majorant, then the dual utility maximization problem cannot have a solution among the countably additive measures.

Section 5 is devoted to an in-depth study of the only nonstandard assumption made in our main theorem in Sect. 3, namely the existence of the so-called germ price. Therein, two general sufficient conditions are given, and concrete examples where they hold are described.

2 An illustrative example

Before we develop a theory in a general semimartingale market model, the purpose of this section is to show that a facelift—together with all of its repercussions such as nonattainment in the class of countably additive martingale measures—already appears in the simplest of models and is not a "cooked-up" consequence of a pathological choice of the modeling framework.

2.1 The market model

For a given time horizon $T>0$, we let $(B_t)_{t\in[0,T]}$ and $(W_t)_{t\in[0,T]}$ be two independent Brownian motions on $(\Omega,\mathcal{F},\mathbb{P})$, and $\mathbb{F}=(\mathcal{F}_t)_{t\in[0,T]}$ the standard augmentation of their natural filtration $\mathbb{F}^{B,W}$. The financial market model consists of a *money-market account* $(S^{(0)}_t)_{t\in[0,T]}$ and a *risky security* $(S_t)_{t\in[0,T]}$. For simplicity, we assume a zero interest rate, that is, $S^{(0)}\equiv 1$, and we model S by the geometric Brownian motion

$$dS_t = S_t(\mu\,dt+\sigma\,dB_t),\qquad S_0:=1. \tag{2.1}$$

Assuming throughout that $\sigma > 0$ and $\mu \neq 0$, we set $\lambda = \mu/\sigma$ and interpret λ as the *market price of risk*. Except for W's presence in the filtration, this model follows the usual Black–Scholes–Samuelson paradigm; the Brownian motion W will play a role in the dynamics of the random endowment that we describe below.

2.2 Trading and admissibility

The investor's initial wealth is denoted by x; at time $t \in [0, T]$, he/she holds π_t shares of the stock S. The usual self-financing condition dictates that the agent's total wealth admits the dynamics

$$X_t^{x,\pi} = x + \int_0^t S_u \pi_u (\mu\, du + \sigma\, dB_u), \quad t \in [0, T]. \tag{2.2}$$

To ensure that the integral is well defined, we require that $\int_0^T \pi_u^2\, du < \infty$ a.s. When additionally, there exists a constant a such that $X_t^{x,\pi} \geq -a$ for all $t \in [0, T]$, $\mathbb{P}$-a.s., we call π *admissible* and write $\pi \in \mathcal{A}$.

2.3 Preferences and random endowment

For the purposes of this example, we model the agent's preferences by a utility function of the "power" type but also note that all statements in this section remain true for a much larger class:

$$U(x) := \frac{1}{p} x^p \quad \text{for } p \in (-\infty, 1) \setminus \{0\} \quad \text{or} \quad U(x) := \log x \quad \text{for } p = 0.$$

For definiteness, we set $U(x) = -\infty$ for $x < 0$ (and at $x = 0$ for $p \leq 0$).

In addition to the investment opportunities provided by the financial market given by $(S^{(0)}, S)$, the investor receives a lump-sum payment (a random endowment, stochastic income, etc.) at time T of the form $\varphi(\eta_T)$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a bounded continuous function, and $\eta_t := \eta_0 + W_t$ for $t \in [0, T]$. We note that the endowment $\varphi(\eta_T)$ cannot be replicated by trading in $(S^{(0)}, S)$ as soon as φ is not a constant function. On the other hand, more and more information about its value is gathered by the agent as t goes to T, so it cannot be treated as an independent random variable either.

2.4 The primal problem

Keeping track of the time horizon $T > 0$, the initial wealth $x \in \mathbb{R}$, and the initial value $\eta_0 \in \mathbb{R}$ of the process η, we pose the following optimization problem faced by a rational agent with the characteristics described above:

$$u(T, \eta_0, x) := \sup_{\pi \in \mathcal{A}} \mathbb{E}\big[U\big(X_T^{x,\pi} + \varphi(\eta_T)\big)\big].$$

Let $x_c = x_c(T, \eta_0) \in \mathbb{R}$ be such that we have $u(T, \eta_0, x) = -\infty$ for $x < x_c$ and $u(T, \eta_0, x) > -\infty$ for $x > x_c$. In [5], it is shown that $-x_c$ coincides with the superreplication cost of $-\varphi(\eta_T)$. In our case, thanks to the fact that $\eta_t = \eta_0 + W_t$, we have $x_c = -\inf \varphi$, independently of η_0 and $T > 0$.

2.5 The dual problem

Let $\mathcal{M}$ denote the set of all $\mathbb{P}$-equivalent probability measures $\mathbb{Q}$ on $\mathcal{F}_T$ for which S defined by (2.1) is a $\mathbb{Q}$-martingale. In our simple model, the structure of $\mathcal{M}$ is well known and completely described. Indeed, a probability measure $\mathbb{Q}$ is in $\mathcal{M}$ if and only if its Radon–Nikodým derivative on $\mathcal{F}_T$ is given by $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T$, where Z is an exponential martingale of the (differential) form

$$dZ_t^\nu = -Z_t^\nu(\lambda\, dB_t + \nu_t\, dW_t), \quad Z_0 = 1, \tag{2.3}$$

for some progressively measurable process ν with $\int_0^T \nu_u^2\, du < \infty$ a.s.

With the dual utility function given by $V(z) := \sup_{x>0}(U(x) - xz)$ for $z > 0$, we define the value function of the *dual problem* by

$$v(T, \eta_0, z) := \inf_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}\left[V\left(z \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] + z\mathbb{E}^{\mathbb{Q}}\left[\varphi(\eta_T) \right] \right) \tag{2.4}$$

for $z > 0$, $\eta_0 \in \mathbb{R}$, and $T > 0$.

Remark 2.1 To ensure the existence of a minimizer, the authors in [5] identify $\mathcal{M}$ with a subset of $\mathbb{L}^1_+(\mathbb{P})$ and embed it naturally into the bidual $\mathrm{ba}(\mathbb{P}) := \mathbb{L}^\infty(\mathbb{P})^* \supseteq \mathbb{L}^1(\mathbb{P})$. With the weak*-closure of $\mathcal{M}$ in $\mathrm{ba}(\mathbb{P})$ denoted by $\overline{\mathcal{M}}_T$ and the dual pairing between $\mathbb{L}^\infty(\mathbb{P})$ and $\mathrm{ba}(\mathbb{P})$ by $\langle\cdot,\cdot\rangle$, the (relaxed) dual value function is then defined by

$$\tilde{v}(T, \eta_0, z) := \inf_{\mathbb{Q}\in\overline{\mathcal{M}}_T} \left(\mathbb{E}\left[V\left(z \frac{d\mathbb{Q}^r}{d\mathbb{P}} \right) \right] + z\langle \mathbb{Q}, \varphi(\eta_T) \rangle \right), \tag{2.5}$$

where $\mathbb{Q}^r \in \mathbb{L}^1(\mathbb{P})$ denotes the regular part in the Yosida–Hewitt decomposition $\mathbb{Q} = \mathbb{Q}^r + \mathbb{Q}^s$ of $\mathbb{Q} \in \mathrm{ba}_+(\mathbb{P})$. It is shown in [5] that the dual minimizer $\hat{\mathbb{Q}} = \hat{\mathbb{Q}}^{T,\eta,z}$ is always attained in $\overline{\mathcal{M}}_T$.

Theorem 2.10 in [14] states that $v = \tilde{v}$, that is, that the finitely additive relaxation is unnecessary if one is interested in the value function alone. This allows us to work with random variables $\frac{d\mathbb{Q}}{d\mathbb{P}} \in \mathbb{L}^1_+(\mathbb{P})$ in the sequel, instead of finitely additive measures and their regular parts as needed in (2.5).

2.6 A naive approach via HJB

If we were to approach the utility maximization problem via the formal dynamic programming principle, then we should start by embedding it into a family of problems starting at $t \in [0, T]$, with the terminal time T, and depending additionally on the states x and η. Thanks to the Markovian structure and without loss of generality, instead of varying the initial time t, we use the "time-to-go" variable $T - t$. Moreover, we abuse the notation and denote this variable simply by T, giving it an alternative interpretation of the (varying) time horizon. In this way, all the effects in the regime

$t \sim T$ show up at $T \sim 0$. The formal HJB equation is now given by

$$\begin{cases} u_T = \sup_{\pi \in \mathbb{R}} \mathcal{L}^\pi u, \\ u(0, \eta, x) = U\big(x + \varphi(\eta)\big). \end{cases} \tag{2.6}$$

Here $u_T := \frac{\partial}{\partial T} u$, and $\mathcal{L}^\pi$ is the (controlled) formal infinitesimal generator of the process $(X^{x,\pi}, \eta)$. With $X^{x,\pi}$ defined by (2.2) and $\eta_t^{\eta_0} := \eta_0 + W_t$, this generator becomes

$$\mathcal{L}^\pi u := \mu \pi u_x + \frac{1}{2}\sigma^2 \pi^2 u_{xx} + \frac{1}{2} u_{\eta\eta}.$$

Similarly, the HJB equation for the dual value function formally reads as

$$\begin{cases} v_T = \inf_{\nu \in \mathbb{R}} \mathcal{N}^\nu v, \\ v(0, \eta, z) = V(z) + z\varphi(\eta), \end{cases}$$

where the dynamics (2.3) for Z^ν produces the generator

$$\mathcal{N}^\nu v = \frac{1}{2} z^2 (\lambda^2 + \nu^2) v_{zz} + \frac{1}{2} v_{\eta\eta} - z\nu v_{z\eta}.$$

The seemingly natural choice for the primal domain $\mathcal{D}_u$ and the interpretation of the initial condition in (2.6) are

(1′) $\mathcal{D}_u = \{(T, \eta, x) \in [0, \infty) \times \mathbb{R} \times \mathbb{R} : x + \varphi(\eta) > 0\}$, and
(2′) $\lim_{T \downarrow 0} u(T, \eta, x) = U(x + \varphi(\eta))$.

Similarly, the dual domain $\mathcal{D}_v$ and the initial condition for the dual problem are expected to be

(3′) $\mathcal{D}_v = \{(T, \eta, z) \in [0, \infty) \times \mathbb{R} \times \mathbb{R} : z > 0\}$, and
(4′) $\lim_{T \downarrow 0} v(T, \eta, z) = V(z) + z\varphi(\eta)$.

It turns out, however, that …

2.7 … the naive approach is not always the right one

In the remainder of the paper, we show in much greater generality that the prescriptions (1′)–(4′) do not fully correspond to reality. Even in the simple Black–Scholes-type model (2.1) with utilities of power type, the value functions behave quite differently. If we set

$$\operatorname{dom} u := \operatorname{int}\{(T, \eta, x) \in [0, \infty) \times \mathbb{R} \times \mathbb{R} : u(T, \eta, x) \in (-\infty, \infty)\}$$

and

$$\operatorname{dom} v := \operatorname{int}\{(T, \eta, z) \in [0, \infty) \times \mathbb{R} \times (0, \infty) : v(T, \eta, z) \in (-\infty, \infty)\},$$

then we have the following result (a special case of Theorem 3.5 below).

Proposition 2.2 *In the setting of the current section, we have*

1. $\operatorname{dom} u = \{(T, \eta, x) \in (0,\infty) \times \mathbb{R} \times \mathbb{R} : x > -\inf \varphi\}$.
2. $\lim_{T\downarrow 0} u(T, \eta, x) = \begin{cases} U(x+\varphi(\eta)), & x \geq -\inf\varphi, \\ -\infty, & x < -\inf\varphi. \end{cases}$
3. $\operatorname{dom} v = \{(T, \eta, z) \in (0,\infty) \times \mathbb{R} \times \mathbb{R} : z > 0\}$.
4. $\lim_{T\downarrow 0} v(T, \eta, z) = \underline{V}(\eta, z)$, *where*

$$\underline{V}(\eta, z) := \begin{cases} V(z) + z\varphi(\eta), & z < z_c, \\ V(z_c) + z_c\varphi(\eta) + (z - z_c)\inf\varphi, & z \geq z_c, \end{cases}$$

and $z_c(\eta) := U'(\varphi(\eta) - \inf\varphi) \in (0,\infty]$, *so that* $V'(z_c(\eta)) + \varphi(\eta) = \inf\varphi$.

Remark 2.3 Proposition 2.2 states that both the primal and the dual value functions exhibit a *facelift* phenomenon:

1. In the primal case, the facelift "cuts off" a part of the domain and leads to an effective initial condition for which the Inada conditions fail. Indeed, we have $U'(x+\varphi(\eta)) \not\to \infty$ as $x \to -\inf\varphi$, for all η, unless φ is constant. On the other hand, as soon as $T > 0$, we have $\frac{\partial}{\partial x} u(T, x, \eta) \to \infty$ as $x \to -\inf\varphi$, for all η.
2. The situation with the dual problem appears even more severe. Even though the effective domain turns out to be exactly as expected, the limiting value $\underline{V}$ of v differs from (4′) in the previous section. Indeed, unless φ is constant, we have $\underline{V}(z, \eta) < V(z) + z\varphi(\eta)$ for $z > z_c(\eta)$.
3. If one tries to apply the usual "Hamiltonian" test (as in, e.g., [15, Sect. 4.3.2]), no facelift—as in Proposition 2.2, (4)—will be detected. Indeed, the dual Hamiltonian acts on a $G : (0,\infty) \times \mathbb{R} \to \mathbb{R}$ as

$$\inf_{\nu\in\mathbb{R}} \mathcal{N}^\nu G(z, \eta) = \frac{1}{2} z^2 \lambda^2 G_{zz} + \frac{1}{2} G_{\eta\eta} - \frac{1}{2} \frac{G_{z\eta}^2}{G_{zz}}.$$

   Because $V_{zz} > 0$ (assuming that φ is smooth enough), this infimum is finite for $G(z, \eta) = V(z) + z\varphi(\eta)$. The case of a general φ leads to the same conclusion but requires a viscosity interpretation; so we do not go into details here.

One important consequence of the facelift is the following (the proof is a combination of Remark 4.1, Corollary 4.3, Proposition 4.4).

Proposition 2.4 *In the setting of the current section, let* $(T, \eta_0) \in (0,\infty) \times \mathbb{R}$, *and let a nonconstant, bounded, and continuous function* $\varphi : \mathbb{R} \to \mathbb{R}$ *be given. Then for all large-enough* $z > 0$, *the problem* (2.4) *does not admit a minimizer in* $\mathcal{M}$.

If additionally $\mathbb{E}[z_c(\eta_0 + W_T)] < \infty$, *then the previous statement holds for all* $z > 0$.

Proposition 2.4 shows that even in the simplest of incomplete continuous-time financial models, the set of countably additive martingale measures $\mathcal{M}$ is not big enough to host the dual optimizer, as soon as the random endowment is unspanned (nonreplicable). A suitable relaxation (e.g., to the set of finitely additive martingale measures) is therefore truly needed.

From [5] we know that the problem (2.4) always admits a minimizer $\hat{\mathbb{Q}}$ in the weak*-closure of $\mathcal{M}$ in ba($\mathbb{P}$). When $\hat{\mathbb{Q}} \notin \mathcal{M}$, both components in the Yosida–Hewitt decomposition $\hat{\mathbb{Q}} = \hat{\mathbb{Q}}^r + \hat{\mathbb{Q}}^s$ are nontrivial. To see this, we recall that Theorem 3.2(ii) in [5] produces the first-order condition

$$U'\big(X_T^{x,\hat{\pi}} + \varphi(\eta_T)\big) = y\frac{d\hat{\mathbb{Q}}^r}{d\mathbb{P}},$$

where $y = y(x) > 0$ is the corresponding Lagrange multiplier. Since $U'(\cdot) > 0$, we necessarily have $\frac{d\hat{\mathbb{Q}}^r}{d\mathbb{P}} > 0$ $\mathbb{P}$-almost surely. However, closed-form expressions for $\hat{\mathbb{Q}}^r$ and $\hat{\mathbb{Q}}^s$ in the setting of Proposition 2.4 remain unavailable.

3 A general market model

We start by describing a general semimartingale financial model that serves as setting for our (abstract) result. It is built on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ that satisfies the usual conditions of right-continuity and completeness; we assume in addition that $\mathcal{F}_0$ is $\mathbb{P}$-trivial. The choice of the constant 1 as the time horizon is arbitrary; it simply indicates that only the values of the ingredients in a neighborhood of 0 are of interest.

3.1 The asset price model

Let $(S_t)_{t\in[0,1]}$ be an $(\mathcal{F}_t)_{t\in[0,1]}$-adapted, RCLL semimartingale that satisfies the following assumption (see [6] for the definition and an in-depth discussion of the concept of σ-martingale):

(**A1**) The set of σ-martingale measures

$$\mathcal{M}_\sigma^e := \{\mathbb{Q} \sim \mathbb{P} : S \text{ is a } \sigma\text{- martingale under } \mathbb{Q}\}$$

is nonempty.

As shown in [6, Theorem 1.1], Assumption (**A1**) is equivalent to the no-arbitrage condition NFLVR (see [6] for the details).

In the context of utility maximization, it is easier to use a mild modification $\mathcal{M}$ of the set $\mathcal{M}_\sigma^e$, which is defined as follows. Let the *admissible set* $\mathcal{A}$ consist of all $\mathbb{F}$-predictable S-integrable processes π such that $\int_0^\cdot \pi_u\, dS_u$ is a.s. uniformly bounded from below. We also define the set of gains processes $\mathcal{X}$ by

$$\mathcal{X} := \left\{ \int_0^\cdot \pi_u\, dS_u : \pi \in \mathcal{A} \right\}.$$

Thanks to a result of Ansel and Stricker (see [1, Corollaire 3.5])—or the σ-martingale property—each $X \in \mathcal{X}$ is a $\mathbb{Q}$-local martingale (and therefore a supermartingale) for any $\mathbb{Q} \in \mathcal{M}^e_\sigma$. Therefore, the set

$$\mathcal{M} := \{\mathbb{Q} \sim \mathbb{P} : X \text{ is a } \mathbb{Q}\text{-supermartingale for each } X \in \mathcal{X}\}$$

includes $\mathcal{M}^e_\sigma$. The difference is often not very significant since Proposition 4.7 in [6] states that $\mathcal{M}^e_\sigma$ is dense in $\mathcal{M}$ in the total-variation norm.

3.2 The utility function and its dual

Let U be a reasonably elastic utility function, that is, a function $U : (0,\infty) \to \mathbb{R}$ with the following properties:

$$(\mathbf{A2}) \begin{cases} (1)\ U \text{ is strictly concave, } C^1\text{, and strictly increasing on } (0,\infty). \\ (2)\ \lim_{x\searrow 0} U'(x) = +\infty \text{ and } \lim_{x\to\infty} U'(x) = 0. \\ (3)\ \lim_{x\to\infty} \frac{xU'(x)}{U(x)} < 1 \text{ if } \sup_{x>0} U(x) > 0. \end{cases}$$

We extend U to $(-\infty, 0]$ by $U(x) := -\infty$ for $x < 0$ and $U(0) := \lim_{x\searrow 0} U(x)$. The conjugate (dual utility function) $V : (0,\infty) \to (-\infty,\infty)$ of U is defined by

$$V(z) := \sup_{x\in\mathbb{R}} \big(U(x) - xz\big) \quad \text{for } z > 0.$$

3.3 Value functions

Given a bounded adapted and RCLL process $(\varphi_t)_{t\in[0,1]}$, for $x \in \mathbb{R}$ and $T \in [0,1]$, we set

$$u(T,x) := \sup_{\pi\in\mathcal{A}} \mathbb{E}\left[U\left(x + \int_0^T \pi_u\, dS_u + \varphi_T\right)\right] \tag{3.1}$$

with the usual convention that $\mathbb{E}[\xi] = -\infty$ as soon as $\mathbb{E}[\xi^-] = \infty$. We call u the *(primal) value function* and note that $u(0,x) = U(x + \varphi_0)$ for all $x \in \mathbb{R}$.

Remark 3.1 As in Sect. 2, we use the (slightly nonstandard) notation T for the time variable to stress the fact that in our principal interpretation, it plays the role of the *time-to-go*. This is also done to avoid the possible confusion with the usual interpretation of the parameter t as the current time with the time-to-go being given by $T - t$. We continue using the variable t as the generic “dummy” time parameter for stochastic processes.

For $\mathbb{Q} \in \mathcal{M}$, we let the density process $(Z^{\mathbb{Q}}_t)_{t\in[0,T]}$ be an RCLL version of

$$Z^{\mathbb{Q}}_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right], \quad t \in [0,1].$$

We also introduce $\mathcal{Z} := \{Z^{\mathbb{Q}} : \mathbb{Q} \in \mathcal{M}\}$ and define the *dual value function* $v : [0,1] \times (0,\infty) \to (-\infty,\infty]$ by

$$v(T,z) := \inf_{Z\in\mathcal{Z}} \mathbb{E}[V(zZ_T) + zZ_T\varphi_T]. \tag{3.2}$$

Remark 3.2 [5] show that the conjugate to the primal value function (3.1) equals the expression on the right-hand side of (3.2), but with $\mathcal{Z}$ replaced (in a suitable manner, see Remark 2.1) by its weak*-closure in ba($\mathbb{P}$). Theorem 2.10 in [14] shows that such a relaxation is not necessary and that infimizing over $\mathcal{Z}$ yields the same value function.

In order to have a nontrivial dual problem, we also ask for finiteness of its value function on its entire domain:

(**A3**) For all $T \leq 1$ and $z > 0$, we have $v(T, z) < \infty$.

Remark 3.3 Thanks to the convexity of V and the reasonable asymptotic elasticity condition (3) in Assumption (**A2**) (see [11, Lemma 6.3] for details), Assumption (**A3**) is equivalent to the existence of $\mathbb{Q} \in \mathcal{M}$ such that $V^+(zZ_1^{\mathbb{Q}}) \in \mathbb{L}^1$ for *some* $z > 0$. In that case, moreover, we have $V(zZ_1^{\mathbb{Q}}) \in \mathbb{L}^1$ for *all* $z > 0$.

Remark 3.3 and the convexity of V guarantee that the set

$$\mathcal{Z}^V(z) := \{Z \in \mathcal{Z} : \mathbb{E}[V^+(zZ_1^{\mathbb{Q}})] < \infty\} =: \mathcal{Z}^V$$

is independent of $z > 0$, nonempty, and enjoys the property that

$$v(T, z) = \inf_{Z \in \mathcal{Z}^V} \mathbb{E}\left[V(zZ_T) + zZ_T \varphi_T\right] \quad \text{for all } T \in (0, 1] \text{ and } z > 0.$$

The set of all corresponding $\mathbb{Q} \in \mathcal{M}$ is denoted by $\mathcal{M}^V$, and its elements are referred to as *V-finite*.

3.4 The lower hedging germ price

With $\mathcal{S}_T$ denoting the set of all $[0, T]$-valued stopping times, we define the *lower American germ price* of $(\varphi_t)_{t \in [0,1]}$ by

$$\Phi^A := \lim_{T \searrow 0} \Phi_T^A, \quad \text{where } \Phi_T^A := \inf_{Z \in \mathcal{Z}, \tau \in \mathcal{S}_T} \mathbb{E}[Z_\tau \varphi_\tau].$$

The *European* counterpart is defined as

$$\Phi^E := \limsup_{T \searrow 0} \Phi_T^E, \quad \text{where } \Phi_T^E := \inf_{Z \in \mathcal{Z}} \mathbb{E}[Z_T \varphi_T].$$

We assume that the two limits are equal:

(**A4**) $\Phi^A = \Phi^E$,

and we denote the common value by Φ and call it the (*lower hedging*) *germ price*.

Remark 3.4 (1) Even though Assumption (**A4**) is not as standard as, for example, (**A2**) or (**A3**), it is in fact quite mild and is satisfied in a wide variety of cases. Section 5 is devoted to sufficient conditions and examples related to Assumption (**A4**). We observe right away, however, that the following three properties follow directly from it:

$$\Phi \leq \varphi_0, \quad \Phi_T^A \nearrow \Phi, \quad \text{and} \quad \Phi_T^E \to \Phi \quad \text{a.s. as } T \searrow 0.$$

(2) By the density of $\mathcal{M}_\sigma^e$ in $\mathcal{M}$, referred to in Sect. 3.1, the infima in the definitions of Φ^A and Φ^E can be taken over the smaller set of densities Z of all σ-martingale measures $\mathcal{M}_\sigma^e$.

(3) Kabanov and Stricker have shown (see [9, Corollary 1.3]) that the infima in the definitions of lower hedging prices can be taken over the set of all V-finite measures as long as V satisfies a mild growth condition. Such a condition is satisfied in our case thanks to the assumption of reasonable asymptotic elasticity in (**A2**), part (3) and Lemma 6.3 in [11]; so

$$\Phi_T^E = \inf_{Z \in \mathcal{Z}^V} \mathbb{E}[Z_T \varphi_T] \quad \text{for all } T > 0. \tag{3.3}$$

This will be useful in the proof of Lemma 3.7 below.

3.5 The form of the facelift and the main theorem

Given two nonnegative constants φ, ψ, let $z \mapsto \underline{V}(z; \varphi, \psi)$ denote the largest convex function below $z \mapsto V(z) + \varphi z$ such that $\underline{V}(z; \varphi, \psi) - z\psi$ is nonincreasing. This function is given by

$$\underline{V}(z; \varphi, \psi) = \sup_{x > -\psi} \big(U(x + \varphi) - xz\big)$$

or, equivalently, by

$$\underline{V}(z; \varphi, \psi) = \begin{cases} V(z) + \varphi z, & z \leq z_c, \\ V(z_c) + \varphi z_c + \psi(z - z_c), & z > z_c. \end{cases}$$

Here z_c is the (unique) solution to $V'(z_c) + \varphi = \psi$ when it exists, and $z_c = +\infty$ otherwise. The special case where $\varphi = \varphi_0$ and $\psi = \Phi$ appears in our main theorem below, so we give it its own notation as

$$\underline{V}(z) := \underline{V}(z; \varphi_0, \Phi).$$

We are now ready to state and prove the central result of this section; it identifies explicitly the shape of the facelift in both the primal and the dual problem. The proof is given in Sect. 3.7 below.

Theorem 3.5 *Under Assumptions* (**A1**)–(**A4**), *we have*

$$\begin{aligned} \lim_{T \searrow 0} u(T, x) &= \begin{cases} U(x + \varphi_0), & x > -\Phi, \\ -\infty, & x < -\Phi, \end{cases} \\ \lim_{T \searrow 0} v(T, z) &= \underline{V}(z) \quad \textit{for all } z > 0. \end{aligned} \tag{3.4}$$

3.6 The true home of Theorem 3.5

One can argue that the natural home for our facelifting result of Theorem 3.5 lies in a class of interconnected optimization problems in a Markovian setting. Indeed, we should *not* like to adopt the somewhat unnatural interpretation of its result in the sense of the asymptotic behavior of the dual value function as the time horizon shrinks to 0 (with the random endowment somehow depending on it). Rather, we should like to think of the time as getting closer to the maturity and the function v as a section of the entire time-dependent value function, in the spirit of the dynamic programming principle. The way to pass from one framework to the other is rather simple: when the dynamics of the underlying state process is homogeneous, one can couple the problems corresponding to the same value of the state, but with varying times, on the same probability space as follows.

Let $\mathfrak{F}$ be a nonempty Hausdorff LCCB (locally compact with a countable base) and therefore a Polish topological space (Euclidean or discrete). For a nonempty G_δ (in particular, open or closed) subset $\mathfrak{S}$ of $\mathbb{R}^d$ for some $d \in \mathbb{N}$, the product $E = \mathfrak{S} \times \mathfrak{F}$ is Hausdorff LCCB and Polish. We work exclusively on the canonical path space $\Omega = D_E[0, \infty)$ consisting of all E-valued RCLL (right-continuous with left limits) paths on $[0, \infty)$, with the σ-algebra $\mathcal{F}$ generated by all coordinate maps.

The coordinate process is denoted by η, and its components by

1. $S = (S^1, \ldots, S^d)$ that is $\mathfrak{S}$-valued (modeling a risky actively traded asset), and
2. F that is $\mathfrak{F}$-valued (modeling a nontraded factor).

The "physical" dynamics of η will be described via a strong Markov family $(\mathbb{P}^\eta)_{\eta\in E}$ of probability measures on D_E. Let $\mathbb{F}^0$ be the (raw) filtration on D_E, generated by the coordinate maps, and let $\mathcal{F}^\eta_t$ be the $\mathbb{P}^\eta$-completion of $\mathcal{F}^0_t$. Thanks to Blumenthal's 0–1 law, $(\mathcal{F}^\eta_t)_{t\geq 0}$ is right-continuous and satisfies the $(\mathbb{P}^\eta)_{\eta\in E}$-usual conditions (see [18, Chap. 3, § 3] for details).

To be able to use Theorem 3.5 under each $\mathbb{P}^\eta$, we impose (**A1**)–(**A4**) on each probability space $(\Omega, (\mathcal{F}^\eta_t)_{t\in[0,1]}, \mathcal{F}, \mathbb{P}^\eta)$; the sets $\mathcal{M}^\eta$ and $\mathcal{Z}^\eta$ are simply the η-parameterized versions of the eponymous objects defined earlier in this section. Similarly, the admissible set depends on $\eta \in E$, and the family is denoted by $(\mathcal{A}^\eta)_{\eta\in E}$. We work with the utility function (and its dual) that satisfy the conditions of Assumption (**A2**). Given a time horizon $T \in (0, 1]$ and $t \in [0, T]$, we define the primal value function

$$u(t, \eta, x) := \sup_{\pi\in\mathcal{A}^\eta} \mathbb{E}^\eta\left[U\left(x + \int_0^{T-t} \pi_u\, dS_u + \varphi(\eta_{T-t})\right)\right],$$

where φ is a bounded and continuous function on E. Similarly, the dual value function is given by

$$v(t, \eta, z) := \inf_{Z\in\mathcal{Z}^\eta} \mathbb{E}^\eta[V(zZ_{T-t}) + zZ_{T-t}\varphi(\eta_{T-t})].$$

Under mild additional conditions on S (it will, e.g., suffice that it is either bounded from below or that its jumps are bounded from below), we have the following version of the dynamic programming principle (see Theorem 3.17 in [23]):

$$v(t, \eta, z) = \inf_{Z\in\mathcal{Z}^\eta} \mathbb{E}^\eta\left[v(\tau, \eta_\tau, zZ_\tau)\right], \quad v(T, \eta, z) = V(z) + z\varphi(\eta), \tag{3.5}$$

for any random time τ of the form $\tau = t + \sigma$, where σ with values in $[0, T - t]$ is an $\mathbb{F}^\eta$-stopping time. It is also shown in [23] that the function v is (jointly) universally measurable, so that the expectation on the right-hand side of (3.5) is well defined. As shown in the last paragraph of Sect. 3.4 in [23], the idea of the proof of Lemma 3.6 below can be used to establish the dynamic programming principle for the primal problem as well.

Equation (3.5) often serves as an analytic description of the value function. In continuous time, it is usually infinitesimalized into a PDE and studied, together with its terminal condition, as a nonlinear Cauchy problem. As already mentioned in Sect. 2, in our case a facelift (boundary-layer) phenomenon appears, and this terminal condition comes in a nonstandard form. Indeed, Theorem 3.5 in the present setting becomes

$$v(t, \eta, z) \longrightarrow \underline{V}\big(z; \varphi(\eta), \Phi(\eta)\big), \tag{3.6}$$

where $\Phi(\eta)$ is as in Sect. 3.4, with the dependence on η emphasized. We conjecture that (3.5) and (3.6) suffice to characterize the value function v in a wide class of models (possibly via a PDE approach) but do not pursue this interesting question in the present paper.

3.7 A proof of Theorem 3.5

We split the proof of our main Theorem 3.5 into lemmas and start from a statement that allows us to focus completely on the dual problem.

Lemma 3.6 *Under Assumptions* (**A1**)–(**A4**), *the first equality in* (3.4) *follows from the second one.*

Proof Suppose that $\lim_{T\searrow 0} v(T, z) = \underline{V}(z)$ for all $z > 0$. The conjugate relationship between the primal and the dual value functions

$$u(T, x) = \inf_{z>0} \big(v(T, z) + xz\big) \quad \text{for } x \in \mathbb{R}, T \in (0, 1],$$

established in [5] and further extended in [14], allows us to apply the tools of classical convex analysis. Indeed, the assumed pointwise convergence of the function v transfers directly to the convex conjugate in the interior of its effective domain (see Theorem 11.34 in [17]). One only needs to check that the limiting function for the primal value function in (3.4) and the function $\underline{V}$ are convex conjugates of each other. □

We focus now exclusively on the dual problem and examine the asymptotic behavior of the function v in the large-z regime.

Lemma 3.7 *Under the Assumptions* (**A1**)–(**A3**), *for all* $T \in (0, 1]$, *the function* $z \to v(T, z)$ *is convex, and*

$$\lim_{z\to\infty} \frac{1}{z} v(T, z) = \Phi_T^E.$$

Proof Convexity of $v(T,\cdot)$ follows from convexity of $V(T,\cdot)$ and $\mathcal{Z}$. For the second statement, we fix $T \in (0,1]$, pick an arbitrary $\varepsilon > 0$, and note that for all $Z \in \mathcal{Z}$,

$$\frac{1}{z}\mathbb{E}[V(zZ_T) + zZ_T\varphi_T] \ge \frac{1}{z}U(\varepsilon) + \mathbb{E}[Z_T(\varphi_T - \varepsilon)].$$

Passing to the infimum over all $Z \in \mathcal{Z}^V$ and using the result in (3.3), we get

$$\frac{1}{z}v(T,z) \ge \frac{1}{z}U(\varepsilon) + \Phi_T^E - \varepsilon, \quad \text{and so } \liminf_{z\to\infty} \frac{1}{z}v(T,z) \ge \Phi_T^E.$$

On the other hand, by the monotone convergence theorem we have

$$\lim_{z\to\infty} \frac{1}{z}\mathbb{E}[V(zZ_T)] = 0 \quad \text{for } Z \in \mathcal{Z}^V.$$

Therefore, for $Z \in \mathcal{Z}$, we have

$$\limsup_{z\to\infty} \frac{1}{z}v(T,z) \le \limsup_{z\to\infty} \frac{1}{z}\mathbb{E}[V(zZ_T)] + \mathbb{E}[Z_T\varphi_T] = \mathbb{E}[Z_T\varphi_T].$$

To complete the proof, it suffices to infimize over all $Z \in \mathcal{Z}^V$. □

We define $v(0_+,z) := \liminf_{T\searrow 0} v(T,z)$ and $v(0^+,z) := \limsup_{T\searrow 0} v(T,z)$.

Lemma 3.8 *Under Assumptions* (**A1**)–(**A4**), $v(0^+,z) \le \underline{V}(z)$ *for all* $z > 0$.

Proof By Lemma 3.7 the function $z \mapsto v(T,z) - z\Phi_T^E$ is convex and nonincreasing for all $T \in (0,1]$. Therefore, so is the function $z \mapsto v(0^+,z) - z\Phi$. Indeed, $\Phi = \lim_{T\searrow 0} \Phi_T^E$, and both convexity and the nonincreasing property are preserved by the limit superior operator. On the other hand, for $z > 0$ and $Z \in \mathcal{Z}^V$, the process

$$t \mapsto V(zZ_t), \quad t \in [0,1],$$

is a uniformly integrable RCLL submartingale. Therefore, $\mathbb{E}[V(zZ_T)] \to V(z)$ as $T \to 0$. Thus,

$$v(0^+,z) \le \limsup_{T\searrow 0} \mathbb{E}[V(zZ_T) + zZ_T\varphi_T] = V(z) + z\varphi_0.$$

It remains to use the definition of $\underline{V}$. □

Lemma 3.9 *Under Assumptions* (**A1**)–(**A4**), $v(0_+,z) \ge \underline{V}(z)$ *for all* $z > 0$.

Proof For $T \in (0,1]$ and $t \in [0,1]$, we set

$$X_t = \operatorname{esssup}_{Z\in\mathcal{Z},\tau\in[t,T]} \mathbb{E}\left[-\frac{Z_T}{Z_t}\varphi_\tau \middle| \mathcal{F}_t\right] \quad \text{for } t \le T \quad \text{and} \quad X_t = X_T \text{ for } t > T.$$

By Proposition 4.3 in [10], $(X_t)_{t\in[0,1]}$ admits an RCLL version, and the process $(Z_tX_t)_{t\in[0,1]}$ is a supermartingale for each $Z \in \mathcal{Z}$. Also, we have $X_t + \varphi_t \geq 0$ for all $t \leq T$ and $X_0 = -\Phi_T^A$. For $x > 0$, Fenchel's inequality produces

$$V(zZ_t) + zZ_t\varphi_t \geq U(x + X_t + \varphi_t) - zZ_t(x + X_t) \quad \text{a.s.}, \ Z \in \mathcal{Z}.$$

By taking expectations we find

$$\begin{aligned}\mathbb{E}[V(zZ_t) + zZ_t\varphi_t] &\geq \mathbb{E}[U(x + X_t + \varphi_t)] - zx - z\mathbb{E}[Z_tX_t] \\ &\geq \mathbb{E}[U(x + X_t + \varphi_t)] - zx - zX_0 \\ &= \mathbb{E}[U(x + X_t + \varphi_t)] - zx + z\Phi_t^A,\end{aligned}$$

where the second inequality follows from the supermartingale property of ZX. Since $x > 0$, we can use Fatou's lemma to see that

$$\begin{aligned}v(0+, z) &\geq \liminf_{t\searrow 0} \mathbb{E}[U(x + X_t + \varphi_t)] - z(x - \Phi) \\ &\geq \mathbb{E}\Big[U\Big(x + \liminf_{t\searrow 0}(X_t + \varphi_t)\Big)\Big] - z(x - \Phi) \\ &= U(x + \varphi_0 - \Phi_T^A) - z(x - \Phi),\end{aligned}$$

where the last equality follows from the right-continuity of X and φ. It remains to let $T \searrow 0$ and then maximize over all $x > 0$. □

4 A modified objective

Our next result states that the seemingly local effect of a facelift is sometimes felt far away from it as well. We adopt the setting of Sect. 3, with Assumptions (**A1**)–(**A3**) in place, but do not assume (**A4**). Since the results in this section are not asymptotic in nature, we chose and fix a time horizon $T > 0$ and replace the time set $[0, 1]$ from Sect. 3 by the generic $[0, T]$.

As a preparation for our result on the modified objective, we define the set

$$\mathcal{C} := \left\{x + \int_0^T \pi_u \, dS_u : x \in \mathbb{R}, \pi \in \mathcal{A}\right\}.$$

The following property for the variable φ_T will be crucial in the sequel:

(**B1**) There exists a random variable $\underline{\varphi}_T \in \mathcal{C}$ such that $X + \underline{\varphi}_T \geq 0$ a.s. whenever $X \in \mathcal{C}$ and $X + \varphi_T \geq 0$ a.s.

Remark 4.1 One can construct one-period examples on a three-element probability space where (**B1**) fails. Nevertheless, there are plenty of cases when it always holds.

For example, in [2], (**B1**) is shown to hold in a related problem. In particular, in the setting of Sect. 2, we have

$$x+\int_0^t \pi_u dS_u \geq \mathbb{E}^{\mathbb{Q}}\left[x+\int_0^T \pi_u dS_u \middle| \mathcal{F}_t\right] \geq -\mathbb{E}^{\mathbb{Q}}[\varphi(\eta_0+W_T)|\mathcal{F}_t].$$

By optimizing over $\mathbb{Q}\in\mathcal{M}$ we then find

$$x+\int_0^t \pi_u dS_u \geq -\inf\varphi.$$

Consequently, in the setting of Sect. 2, we have $\underline{\varphi}_T=\inf\varphi$.

Theorem 4.2 *Suppose Assumptions* (**A1**)–(**A3**) *and* (**B1**) *hold. Then for all* $z\in(0,\infty)$ *and* $T\in(0,1]$, *we have the representation*

$$v(T,z)=\inf_{Z\in\mathcal{Z}}\mathbb{E}[\underline{V}(zZ_T;\varphi_T,\underline{\varphi}_T)]. \tag{4.1}$$

Proof Let $(T,z)\mapsto\underline{v}(T,z)$ denote the function defined by the right-hand side of (4.1). Since $\underline{V}(z;\varphi_T,\underline{\varphi}_T)\leq V(z)+z\varphi_T$ for all $z>0$ a.s., we clearly have $\underline{v}\leq v$. To prove the converse inequality, we pick $x\in\mathbb{R}$ and $\pi\in\mathcal{A}$ satisfying $\mathbb{E}[U(x+\int_0^T\pi_u\,dS_u+\varphi_T)]>-\infty$. That implies

$$x+\int_0^T \pi_u\,dS_u+\varphi_T\geq 0 \quad \text{a.s.}$$

Therefore, there exists $\underline{\varphi}_T\in\mathcal{C}$ such that

$$x+\int_0^T \pi_u\,dS_u+\underline{\varphi}_T\geq 0.$$

This produces

$$\mathbb{E}\left[U\left(x+\int_0^T \pi_u\,dS_u+\varphi_T\right)\right]=\mathbb{E}\left[\underline{U}\left(x+\int_0^T \pi_u\,dS_u+\varphi_T;\underline{\varphi}_T\right)\right],$$

where we have introduced

$$\underline{U}(x;\underline{\varphi}):=\begin{cases}U(x), & x>-\underline{\varphi},\\ -\infty & \text{otherwise.}\end{cases}$$

We have

$$\sup_{x\in\mathbb{R}}\left(\underline{U}(x+\varphi)-xz\right)=\sup_{x>-\underline{\varphi}}\left(U(x+\varphi)-xz\right)=\underline{V}(z;\varphi,\underline{\varphi}),$$

which, together with the supermartingale property of $Z_t(x+\int_0^t\pi_u dS_u)$ for each $Z\in\mathcal{Z}$, produces

$$\mathbb{E}\left[U\left(x+\int_0^T \pi_u\,dS_u+\varphi_T\right)\right]-xz\leq\mathbb{E}[\underline{V}(zZ_T;\varphi_T,\underline{\varphi}_T)]$$

for all $z > 0$. This in turn implies that

$$u(T, x) - xz \leq \underline{v}(T, z).$$

The claim now follows by using the conjugacy of the primal and dual value functions, as established in [5] and extended in [14] (see Remark 2.1 for details). □

The result of Theorem 4.2 has an interesting consequence:

Corollary 4.3 *Suppose that the conditions of Theorem 4.2 hold and* $\mathbb{P}[\underline{\varphi}_T \neq \varphi_T] > 0$. *Then the dual problem* (3.2) does not *admit a minimizer* $Z \in \mathcal{Z}$ *for all* $z > 0$ *large enough.*

Proof Since $\underline{\varphi}_T \neq \varphi_T$ with positive probability, there exists a constant $z_0 \in (0, \infty)$ such that

$$\mathbb{P}[V(z) + z\varphi_T > \underline{V}(z; \varphi_T, \underline{\varphi}_T)] > 0 \quad \text{for } z \geq z_0. \tag{4.2}$$

Additionally, we have the trivial inequality $V(z) + z\varphi_T \geq \underline{V}(z; \varphi_T, \underline{\varphi}_T)$ a.s. for all $z \in (0, \infty)$. Suppose now that $\hat{Z} = \hat{Z}(z)$ is the dual minimizer, that is, the minimizer in (3.2), corresponding to $z \geq z_0$. By Theorem 4.2, it must also be a minimizer for the right-hand side of (4.1), and it must have the property that

$$V(zZ_T) + z\hat{Z}_T\varphi_T = \underline{V}(z\hat{Z}_T; \varphi_T, \underline{\varphi}_T) \quad \text{a.s.}$$

The inequality in (4.2), however, implies that

$$\mathbb{P}[z\hat{Z}_T < z_0] = 1,$$

which is in contradiction with $z \geq z_0$ and $\mathbb{E}[\hat{Z}_T] = 1$. □

In the model of Sect. 2, one can improve Corollary 4.3 and show nonattainment for any $z > 0$, provided that φ does not stay "too close" to its minimum:

Proposition 4.4 *In the setting of Sect.* 2, *assume that*

$$\mathbb{E}\big[U'\big(\varphi(\eta_0 + W_T) - \inf\varphi\big)\big] < \infty. \tag{4.3}$$

Then the dual problem (2.4) *at* (T, η_0) *does not admit a minimizer in* $\mathcal{Z}$ *for any* $z > 0$.

Proof Given (4.3), we assume that there exist $z > 0$ and $\hat{Z} \in \mathcal{Z}$ that attain the infimum in (3.2). As in the proof of Corollary 4.3, this implies that $z\hat{Z}_T \leq Y$ a.s., where $Y := U'(\varphi(\eta_0 + W_T) - \inf\varphi)$. Thanks to the special structure of the set $\mathcal{Z}$ in the model of Sect. 2, there exists a predictable and W-integrable process $(\hat{\nu}_t)_{t\in[0,T]}$ such that

$$\hat{Z}_T = \mathcal{E}(-\lambda \cdot B)_T \hat{H}_T, \qquad \hat{H}_t := \mathcal{E}(\hat{\nu} \cdot W)_t.$$

We define the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ as the usual augmentation of

$$\mathcal{G}_t^{\mathrm{raw}} := \sigma(B_u, W_s; u \le T, s \le t), \quad t \in [0,T].$$

The process W is a $\mathcal{G}$-Brownian motion, and $\hat{\nu}$ is $\mathcal{G}$-predictable; so $\hat{H}$ is a $\mathcal{G}$-local martingale, and in particular, we have $\mathbb{E}[\hat{H}_T|\mathcal{G}_0] \le 1$. Therefore,

$$e^{\lambda B_T - \frac{1}{2}\lambda^2 T} \ge e^{\lambda B_T - \frac{1}{2}\lambda^2 T}\mathbb{E}[\hat{H}_T|\mathcal{G}_0] = \mathbb{E}[\hat{Z}_T|\mathcal{G}_0],$$

where the inequality is in fact an a.s. equality since both sides have expectation one. Using the fact that Y is independent of $\mathcal{G}_0$, we conclude that

$$e^{\lambda B_T - \frac{1}{2}\lambda^2 T} \le \frac{1}{z}\mathbb{E}[Y] < \infty \quad \text{a.s.},$$

which is a contradiction with the fact (derived from the assumption that $\mu \ne 0$) that the distribution of the left-hand side has support $(0,\infty)$. □

5 Sufficient conditions for (A4)

Condition (**A4**) in Sect. 3 plays a major role in the proof of Theorem 3.5 and guarantees that the process $(\varphi_t)_{t\in[0,1]}$ does not oscillate too much as $t \searrow 0$. Clearly, it (or a version of it) must be imposed; indeed, the very form of the facelift depends on the value (and existence) of the limiting germ price Φ. We present here two sufficient conditions for its validity, which apply to a wide variety of situations often encountered in mathematical finance.

5.1 Complete markets

In the case of a complete market, we have the following:

Proposition 5.1 *If $\mathcal{M} = \{\mathbb{Q}\}$ for some $\mathbb{Q} \sim \mathbb{P}$, then* (**A4**) *holds with*

$$\Phi^E = \Phi^A = \varphi_0.$$

Proof It suffices to note that by the dominated convergence theorem and the RCLL assumption we have

$$\mathbb{E}^{\mathbb{Q}}\Big[\inf_{t\in[0,T]}\varphi_t\Big] \longrightarrow \varphi_0 \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}[\varphi_T] \longrightarrow \varphi_0 \quad \text{as } T \searrow 0.$$ □

5.2 Sufficient controllability

Our second sufficient condition assumes that there exists a process $(\eta_t)_{t\in[0,1]}$ with values in some topological space E such that $\varphi_t = \varphi(\eta_t)$, $t \in [0,1]$, for some continuous and bounded function $\varphi: E \to \mathbb{R}$.

We start with a general condition, phrased as a lemma, which heuristically says that (**A4**) holds if $(\eta_t)_{t\in[0,1]}$ can be well controlled toward any point in E, within any positive amount of time. To state it, for each $T \in (0,1]$, we define the set d_T of $\mathbb{Q}$-distributions of η_t as $\mathbb{Q}$ ranges through $\mathcal{M}$ and $t \in (0,T]$.

Lemma 5.2 *Suppose that φ is a bounded and continuous function and that for each $\eta \in E$ and each $T > 0$, there exists a sequence $(\mu_n)_{n\in\mathbb{N}}$ in d_T such that $\mu_n \Rightarrow \delta_\eta$ weakly. Then* (**A4**) *holds with*

$$\Phi^E = \Phi^A = \inf\varphi. \tag{5.1}$$

Proof Since φ is continuous and bounded, the assumptions imply that

$$\inf_{\mu\in d_T} \int \varphi(\eta)\, d\mu(\eta) = \inf\varphi.$$

Therefore, $\Phi_T^E = \inf\varphi$ for $T \in (0,1]$. Recalling that $\inf\varphi \le \Phi_T^A \le \Phi_T^E$ for all $T \in (0,1]$ by construction, we conclude that (5.1) holds. □

Next, we describe a large class of models with $E := \mathbb{R}^d$ to which Lemma 5.2 applies. We start by fixing a filtered probability space with a filtration $\mathbb{F}$ satisfying the usual conditions. Let $\mathcal{S}^g$ denote the set of all $\mathbb{R}^d$-valued semimartingales R with $R_0 = 0$ for which there exist

1. a semimartingale decomposition $R = M + F$ into a local martingale M and a finite variation process F, and
2. a (deterministic) function $g_R : [0,1] \to [0,\infty)$ with $\lim_{T\searrow 0} g_R(T) = 0$,

such that with $|F|$ denoting the total variation process of F and $[M,M]$ the quadratic variation process of M, we have

$$|F|_T + [M,M]_T \le g_R(T) \quad \text{a.s. for all } T \in [0,1].$$

We note that a posteriori membership in $\mathcal{S}^g$ immediately makes any semimartingale special and we can (and do) talk about its unique canonical decomposition without ambiguity.

Remark 5.3 An example of an element in the class $\mathcal{S}^g$ is a process of the form

$$R_t = \int_0^t \alpha_u\, du + \int_0^t \beta_u\, dB_u + \int_0^t \gamma_u\, dN_u, \quad t \in [0,1],$$

where α, β, and γ are uniformly bounded predictable processes valued in, respectively, $\mathbb{R}^d$, $\mathbb{R}^{d\times d}$, and $\mathbb{R}^d$; B is a d-dimensional Brownian motion, and N is a d-dimensional Poisson process.

The class $\mathcal{S}^g$ is important in our setting because it admits moment estimates uniformly over all equivalent measure changes that preserve the semimartingale decomposition. The next result follows directly from the Burkholder–Davis–Gundy inequalities (see Theorem 48 in [16]), and we skip the proof.

Lemma 5.4 *For each $R \in \mathcal{S}^g$ with canonical decomposition $R = M + F$, there exists a function $h_R : [0,1] \to (0,\infty)$ with $h_R(t) \to 0$ as $t \searrow 0$ such that*

$$\mathbb{E}^{\mathbb{Q}}[|R_t|] \le h_R(t) \quad \textit{for all } t \in [0,1],$$

for any $\mathbb{Q} \sim \mathbb{P}$ such that M is a $\mathbb{Q}$-local martingale.

Theorem 5.5 *Suppose that the $\mathbb{R}^m$-valued process S and the $\mathbb{R}^d$-valued factor process η are semimartingales that satisfy the following assumptions*:

(1) *There exists a $\mathbb{P}$-equivalent measure $\mathbb{Q}^0$ such that S is a local martingale.*
(2) *The process $(\eta_t)_{t\in[0,1]}$ is of the form*

$$\eta_t = \eta_0 + \int_0^t \beta_u \, dW_u + R_t, \quad t \in [0,1],$$

where W is a Brownian motion strongly orthogonal to S and to Z^0 (the density process of $\mathbb{Q}^0$ with respect to $\mathbb{P}$), β is a bounded predictable process whose absolute value is bounded away from 0, and R is in $\mathcal{S}^g$ with canonical decomposition $R = M + F$, where M is strongly orthogonal to W. Then condition (**A4**) *holds, and $\Phi^E = \Phi^A = \inf \varphi$.*

Proof We start by constructing a large enough subfamily of the family of local martingale measures. For a bounded predictable process $(\nu_t)_{t\in[0,1]}$, we define the processes

$$H_t^\nu := \mathcal{E}(\nu \cdot W)_t, \quad Z_t^\nu := Z_t^0 H_t^\nu, \quad t \in [0,1].$$

The strong orthogonality between W and Z^0 and the continuity of W ensure that $[W, Z^0] \equiv 0$; hence, Z^ν is a local martingale. To see that it is a martingale, we note that

$$\mathbb{E}[Z_T^\nu] = \mathbb{E}[Z_T^0 H_T^\nu] = \mathbb{E}^{\mathbb{Q}^0}[H_T^\nu].$$

Since W remains a Brownian motion under $\mathbb{Q}^0$ and since ν is bounded, we can use Novikov's condition to get $\mathbb{E}^{\mathbb{Q}^0}[H_T^\nu] = 1$, from which the martingale property follows. We can then define $\frac{d\mathbb{Q}^\nu}{d\mathbb{P}} := Z_T^\nu$.

We fix a constant $\eta \in \mathbb{R}$. For $n \in \mathbb{N}$, we define the bounded process

$$\nu_t^n := \begin{cases} \frac{n(\eta - \eta_0)}{\beta_t}, & t \le 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\eta_0 + \int_0^t \beta_u \nu_u^n \, du = \eta$ for $t \ge 1/n$, and

$$\eta_{1/n} = \eta_0 + R_{1/n} + \int_0^{1/n} \beta_u \, dW_u = \eta + R_{1/n} + \int_0^{1/n} \beta_u \, dW_u^n,$$

where $W_t^n := W_t - \int_0^t \nu_u^n \, du$. Thanks to the orthogonality assumption, the local martingale part M of R is a $\mathbb{Q}^{\nu^n}$-local martingale, and so $\mathbb{Q}^{\nu_n}$ and R satisfy the conditions of Lemma 5.4. Moreover, W^n is a $\mathbb{Q}^{\nu_n}$-Brownian motion; so

$$\mathbb{E}^{\mathbb{Q}^{\nu^n}}[|\eta_{1/n} - \eta|] \leq C\left(\mathbb{E}^{\mathbb{Q}^{\nu^n}}[|R_{1/n}|] + \mathbb{E}^{\mathbb{Q}^{\nu^n}}\left[\left|\int_0^{1/n} \beta_u \, d\tilde{W}_u\right|\right]\right)$$

for some constant C. The right-hand side is bounded from above by a linear combination of $h_R(1/n)$ and $1/n$; so it converges to 0 as $n \to \infty$. Therefore, we have

$$\int_{\mathbb{R}^d} |x - \eta| \mu_n(dx) \longrightarrow 0 \quad \text{as } n \to \infty,$$

where μ_n denotes the distribution of $\eta_{1/n}$ under $\mathbb{Q}^{\nu^n}$. Consequently, $\mu_n \Rightarrow \delta_\eta$, and Lemma 5.2 can be applied. □

Remark 5.6 A family of examples of models that satisfy Assumption (1) of Theorem 5.5 is furnished by processes of the form

$$dS_t := \mu_t \, dt + \sigma_t \, dB_t + dJ_t, \qquad S_0 \in \mathbb{R},$$

where B is an $\mathbb{F}$-Brownian motion independent of W, J is an $\mathbb{F}$-local martingale (possibly with jumps) that is strongly orthogonal to B, and μ and σ are predictable processes. To apply Theorem 5.5, it suffices to note that the process $Z^0 := \mathcal{E}(-(\mu/\sigma) \cdot B)$ is a strictly positive martingale whenever μ/σ is sufficiently integrable.

Acknowledgements The authors would like to thank Mihai Sîrbu, Kim Weston, and the two anonymous referees for their many constructive comments.

References

1. Ansel, J.-P., Stricker, C.: Couverture des actifs contingents et prix maximum. Ann. Inst. Henri Poincaré Probab. Stat. **30**, 303–315 (1994)
2. Broadie, M., Cvitanić, J., Soner, H.M.: Optimal replication of contingent claims under portfolio constraints. Rev. Financ. Stud. **11**, 59–79 (1998)
3. Bouchard, B., Touzi, N.: Explicit solution to the multivariate super-replication problem under transaction costs. Ann. Appl. Probab. **10**, 685–708 (2000)
4. Chassagneux, J.-F., Elie, R., Kharroubi, I.: When terminal facelift enforces delta constraints. Finance Stoch. **19**, 329–362 (2015)
5. Cvitanić, J., Schachermayer, W., Wang, H.: Utility maximization in incomplete markets with random endowment. Finance Stoch. **5**, 237–259 (2001)
6. Delbaen, F., Schachermayer, W.: The fundamental theorem of asset pricing for unbounded stochastic processes. Math. Ann. **312**, 215–250 (1998)
7. Guasoni, P., Rásonyi, M., Schachermayer, W.: Consistent price systems and face-lifting pricing under transaction costs. Ann. Appl. Probab. **18**, 491–520 (2008)
8. Hugonnier, J., Kramkov, D., Schachermayer, W.: On utility based pricing of contingent claims in incomplete markets. Math. Finance **15**, 203–212 (2005)
9. Kabanov, Y., Stricker, C.: On equivalent martingale measures with bounded densities. In: Azéma, J., Émery, M., Ledoux, M., Yor, M. (eds.) Séminaire de Probabilités XXXV. Lecture Notes in Math., vol. 1755, pp. 139–148 (2001)

10. Kramkov, D.: Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. Probab. Theory Relat. Fields **105**, 459–479 (1996)
11. Kramkov, D., Schachermayer, W.: The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab. **9**, 904–950 (1999)
12. Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance. Springer, Berlin (1998)
13. Karatzas, I., Žitković, G.: Optimal consumption from investment and random endowment in incomplete semimartingale markets. Ann. Probab. **31**, 1821–1858 (2003)
14. Larsen, K., Žitković, G.: Utility maximization under convex portfolio constraints. Ann. Appl. Probab. **23**, 665–692 (2013)
15. Pham, H.: Continuous-Time Stochastic Control and Optimization with Financial Applications. Springer, Berlin (2009)
16. Protter, P.: Stochastic Integration and Differential Equations, 2nd edn. Applications of Mathematics (New York), vol. 21. Springer, Berlin (2004)
17. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317. Springer, Berlin (1998)
18. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293. Springer, Berlin (1999)
19. Schmock, U., Shreve, S.E., Wystup, U.: Valuation of exotic options under shortselling constraints. Finance Stoch. **6**, 143–172 (2002)
20. Soner, H.M., Touzi, N.: Superreplication under gamma constraints. SIAM J. Control Optim. **39**, 73–96 (2000)
21. Soner, H.M., Touzi, N.: Dynamic programming for stochastic target problems and geometric flows. J. Eur. Math. Soc. **4**, 201–236 (2002)
22. Touzi, N.: Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. Fields Institute Monographs, vol. 29. Springer, New York (2013)
23. Žitković, G.: Dynamic programming for controlled Markov families: abstractly and over martingale measures. SIAM J. Control Optim. **52**, 1597–1621 (2014)