Positivity (2023) 27:40
https://doi.org/10.1007/s11117-023-00993-7

Positivity

Convex duality for partial hedging of American options: continuous price processes

Ari-Pekka Perkkiö[1] · Erick Treviño-Aguilar[2]

Received: 8 February 2022 / Accepted: 3 May 2023 / Published online: 21 May 2023

Abstract

Partial hedging of American options is an interesting minimax problem and in this paper we establish its dual problem that concerns only maximization. The case of a continuous price process is considered under a general incomplete market. Our construction of a duality requires a careful preparation in order to define the dual domain with a compactness property. A key step is an extension of linear functionals preserving norm and positivity.

Keywords American options · Partial hedging · Convex duality · Normal integrands

Mathematics Subject Classification 46N10 · 60G40 · 91G20

1 Introduction

American options are financial contracts in which the owner has the right to exercise his option at any time in a specific period. From the point of view of the seller derives the problem of partial hedging. If the market model is complete, that is, the model admits a unique martingale measure, the option can be perfectly replicated and the solution is provided by the Snell envelope of the American option with respect to the unique martingale measure; see Bensoussan [3] and Karatzas [13]. However, in incomplete markets, there is in general no perfect replication. In this case, the seller may decide to take some risk and it is meaningful to design portfolios generating partial hedgings which given a budget constraint, minimize the *shortfall risk*: The positive gap between the option's payoff and the capital generated by the selected portfolio. To quantify shortfall risk, a loss function is applied and then, the expected value

✉ Erick Treviño-Aguilar
erick.trevino@im.unam.mx

1 Department of Mathematics, Ludwig Maximilians Universität München, Theresienstr. 39, 80333 Munich, Germany

2 Instituto de Matemáticas, Unidad Cuernavaca, Universidad Nacional Autónoma de México (UNAM), Cuernavaca Morelos, México

with respect to the market model is calculated. This point of view was systematically developed for European options by Föllmer and Leukert [10]. An analogous approach for American options was studied by Mulinacci [15].

The existence of optimal partial hedging strategies for American options is known from the above cited references. However, so far, optimality conditions characterizing optimal strategies are still lacking. Indeed, partial hedging of American options is a difficult problem. It involves the minimization over a class of stochastic processes which must satisfy some constraints and the functional to be minimized requires to take the supremum over a class of stopping times. This means that partial hedging of American options lead to a minimax problem over constrained classes and optimality conditions are elusive. From known experience, convex duality is always a promising approach for a solution and our goal in this paper is to advance convex duality for the problem of partial hedging of American options. Under mild conditions we establish a dual problem which concerns only maximization; see equation (26) below. Hence, for the dual problem optimality conditions are accessible and will be studied in future work. Our construction of a duality requires a careful preparation in order to define the dual domain with a compactness property allowing the application of Sion's minimax theorem. A key step is a careful extension of linear functionals preserving norm and positivity. Another key step is to reduce the form of the dual problem and to this end, we apply an interchange rule for processes developed in [16, Theorem 11].

After this introduction the paper is organized as follows. In Sect. 2 we fix notation. In Sect. 3, we present the problem of partial hedging of American options in incomplete markets and continuous time. In Sect. 4, we prepare all necessary preliminaries to establish the dual problem of partial hedging, the main result is stated in Theorem 12. The proof is distributed in Sects. 5 and 6.

2 Notations and setting

In this section we introduce basic notation and a few concepts from convex and functional analysis that will be necessary. For $\mathbb{T} > 0$ a positive real number, we fix a stochastic base $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq \mathbb{T}}, \mathbb{P})$. We assume that the filtration $\mathbb{F}$ satisfies the usual conditions of right continuity and completeness. We assume that the probability measure $\mathbb{P}$ is $0-1$ on $\mathcal{F}_0$. Expectation with respect to $\mathbb{P}$ is simply denoted by $E[\cdot]$. Equality of random variables always means $\mathbb{P}$-a.s. equality. Let $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ and $\bar{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\}$ where $\mathbb{R}_+$ denotes the non-negative real numbers. The family of $\mathbb{F}$-stopping times with values in the interval $[0, \mathbb{T}]$ is denoted by $\mathcal{T}$. We use the french abreviations càdlàg for right continuous with left limits, càglàd for left-continuous with right limits, and càg for left-continuous. The optional σ-algebra defined on the product space $\Omega \times [0, \mathbb{T}]$ is generated by the family of $\mathbb{F}$-adapted, càdlàg stochastic processes seen as functions on the product space. We denote this σ-algebra by $\mathcal{O}$; see e.g., Jacod and Shiryaev [12, Section 1 c] for basic properties of this σ-algebra defined in the product space $\Omega \times [0, \infty)$. The predictable σ-algebra defined on the set $\Omega \times [0, \mathbb{T}]$ is generated by the family of $\mathbb{F}$-adapted, càg stochastic processes. This σ-algebra is usually denoted by $\mathcal{P}$; see e.g., Jacod and Shiryaev [12, Section 2 a]. Let L^∞ be the Banach space of random variables which are essentially bounded. Recall that a stochastic process x is said to be of class(D) if the family of random variables $\{\mathsf{x}_\tau | \tau \in \mathcal{T}\}$ is uniformly integrable.

In order to develop convex duality for our problem of partial hedging, a crucial step will be to apply Sion's minimax theorem [21]. This result applies in a setting of topological vector spaces. To prepare this analytical environment we introduce three Banach spaces of stochastic processes. The topological vector spaces for the minimax theorem will be defined in terms of those Banach spaces and their topological duals.

Let $\mathcal{R}$ be the space of $\mathbb{F}$-adapted processes x of class(D) with càdlàg paths and with

$$\|\mathsf{x}\|_{\mathcal{R}} := \sup_{\tau \in \mathcal{T}} E[|\mathsf{x}_\tau|] < \infty.$$

The norm $\|\cdot\|_{\mathcal{R}}$ makes $\mathcal{R}$ a Banach space; see [7, Definition in paragraph 20 and Theorem 22, Chapter VI, pp. 89–90]. Let $\mathcal{R}_\infty$ be the family of processes in $\mathcal{R}$ with the further property

$$\|\mathsf{x}\|_{\mathcal{R}_\infty} := \left\|\mathsf{x}^*\right\|_{L^\infty} < \infty, \tag{1}$$

where

$$\mathsf{x}^* := \sup_{s \in [0, \mathbb{T}]} |\mathsf{x}_s| \,. \tag{2}$$

We denote by $\mathcal{R}_\infty^+$ the non negative elements of $\mathcal{R}_\infty$. We denote by $\mathcal{R}_1$ the stochastic processes in $\mathcal{R}$ such that

$$\|\mathsf{x}\|_{\mathcal{R}_1} := E[\mathsf{x}^*] < \infty. \tag{3}$$

The norm $\|\cdot\|_{\mathcal{R}_1}$ makes $\mathcal{R}_1$ a Banach space which is studied by Bismut [4].

The Banach structure of the spaces $\mathcal{R}$ and $\mathcal{R}_1$ are important for our duality results. At the same level of importance are their order structure, indeed both are Riesz spaces as defined in [1]. For the specific result where the Riesz structure is crucial see the proof of Lemma 7 below.

2.1 Quasi-randomized stopping times

We denote by $\mathcal{R}_1^*$ the topological dual space of $\mathcal{R}_1$. The Proposition 1.3 of Bismut [4] identifies $\mathcal{R}_1^*$ as follows. For an element $\Phi \in \mathcal{R}_1^*$ there exists a pair $(\phi, \tilde{\phi})$ of right-continuous, $\mathbb{F}$-adapted processes ϕ and $\tilde{\phi}$ with

$$\langle \Phi, \mathsf{x} \rangle = E\left[\int_{[0,\mathbb{T}]} \mathsf{x}_t \, d\phi_t + \int_{[0,\mathbb{T}]} \mathsf{x}_{t-} \, d\tilde{\phi}_t\right] \tag{4a}$$

$$\text{integrable variation, } \phi \text{ not necessarily zero at } t = 0, \tag{4b}$$

$$\tilde{\phi} \text{ a predictable process with } \tilde{\phi}_0 = 0, \text{ and} \tag{4c}$$

$$\text{concentrated in a countable family of predictable stopping times.} \tag{4d}$$

We next specify a class $\overline{\mathsf{K}}$ of elements of $\mathcal{R}_1^*$ which in agreement with Definition 1.3 of Bismut [4] we call the class of *quasi-randomized stopping times*. A linear functional $\Phi \in \mathcal{R}_1^*$ is an element of $\overline{\mathsf{K}}$ if in the representation (4a), ϕ and $\tilde{\phi}$ are non decreasing, non negative and $\phi_{\mathbb{T}} + \tilde{\phi}_{\mathbb{T}} = 1$. We denote by K the elements of $\overline{\mathsf{K}}$ such that the process $\tilde{\phi}$ in the representation (4a) is equal to zero while $\phi_{\mathbb{T}} = 1$.

Remark 1 Typical elements of $\overline{\mathsf{K}}$ will be denoted by Φ and Ψ and the non-decreasing processes in the representation (4a) are denoted by $(\phi, \tilde{\phi})$ and $(\psi, \tilde{\psi})$, respectively.

Remark 2 A stopping time $\tau \in \mathcal{T}$ can be identified with an element of K having associated the non-decreasing process $1_{\{\tau \le t\}}$ in the representation (4a). Moreover, for nonnegative $\mathsf{x} \in \mathcal{R}_1$ we have

$$\sup_{\tau \in \mathcal{T}} E[\mathsf{x}_\tau] = \sup_{\Phi \in \overline{\mathsf{K}}} \langle \Phi, \mathsf{x} \rangle , \tag{5}$$

see Bismut [4, Proposition 1.4].

Elements in $\overline{\mathsf{K}}$ acting by the representation (4a) define continuous linear functionals on $\mathcal{R}$ and $\mathcal{R}_\infty$ and (5) holds true. These properties are necessary for the functional analysis framework that we use in order to establish duality for our main problem. These are the statements of Proposition 2 below, the essence is Lemma 1 following next. Note however that Lemma 1 does not require the property of being class(D) and we estate and prove in this generality.

Lemma 1 *Let x be a càdlàg process with $\sup_{\tau \in \mathcal{T}} E[|\mathsf{x}_\tau|] < \infty$. For $\Phi \in \overline{K}$ we have*

$$\langle \Phi, |x| \rangle \le \|x\|_{\mathcal{R}} .$$

Proof Take first $\Phi \in \mathsf{K}$. We model the proof as in [12, Lemma 3.12, Section I.3]. Let $c_s := \inf\{t \mid \phi_t \ge s\}$. We have

$$\int_0^1 E\left[|\mathsf{x}_{c_s}| 1_{\{c_s < \infty\}}\right] ds \le \sup_{\tau \in \mathcal{T}} E[|\mathsf{x}_\tau|] < \infty.$$

Hence, by Tonelli's theorem [5, Thm. 3.4.5] we see that $|\mathsf{x}_{c_s}|$ is integrable with respect to $E[\int \cdot\, ds]$ and we apply Fubini's theorem [5, Thm. 3.4.4] to get

$$\begin{aligned} \int_0^1 E\left[|\mathsf{x}_{c_s}| 1_{\{c_s < \infty\}}\right] ds &= E\left[\int_0^1 |\mathsf{x}_{c_s}| 1_{\{c_s < \infty\}} ds\right] \\ &= E\left[\int_0^{\mathbb{T}} |\mathsf{x}_s| d\phi_s\right], \end{aligned}$$

where in the second equality we applied the change of variable formula [11, p 29, equation (1)] (see also [18, Proposition (4.9) p. 8]).

Now take a general $\Phi = (\phi, \tilde{\phi}) \in \overline{\mathsf{K}}$. Let $\tilde{U}$ be the Snell envelope of $|\mathsf{x}|$ in the generality of [14, Theorem 4] in that $|\mathsf{x}|$ is not necessarily of class(D). Now consider the Mertens decomposition $\tilde{U} = U - I$ where U is a right-continuous supermartingale and I is a left-continuous non-decreasing process with $I_0 = 0$; see [14, p. 51]. Take the Doob-Meyer decomposition $U = U_0 + M - C$ where M is a local martingale with $M_0 = 0$ and C is a non decreasing predictable process with $E[C_{\mathbb{T}}] < \infty$; see [17, Thm. III.3.13, p.115]. We have

$$\begin{aligned}\langle \Phi, |\mathsf{x}| \rangle &\leq E\left[\int_{[0,\mathbb{T}]} U_0 + M_t d\phi_t + \int_{[0,\mathbb{T}]} U_0 + M_{t-} d\tilde{\phi}_t\right] \\ &= E\left[\int_{[0,\mathbb{T}]} (U_0 + M_t) d(\phi_t + \tilde{\phi}_t)\right]\end{aligned}$$

where in the equality M replaces M_- since $\tilde{\phi}$ is predictable; see [12, Lemma 2.27, p.22]. On considering $\phi + \tilde{\phi}$ as an element of K we obtain from the first part that

$$\langle \Phi, |\mathsf{x}| \rangle \leq \|U_0 + M\|_{\mathcal{R}} = U_0.$$

This proves the lemma since $U_0 = \sup_{\tau \in \mathcal{T}} E[|\mathsf{x}_\tau|]$. □

We obtain the following result as direct consequence of Lemma 1.

Proposition 2 *A quasi-randomized stopping time defines through (4a) a continuous linear functional on all of $\mathcal{R}$. Furthermore, the equality (5) holds true for non negative processes in $\mathcal{R}$. Similar claims hold true for $\mathcal{R}_\infty$.*

A useful consequence of Proposition 2 is an identification of $\overline{\mathsf{K}}$ as a family of functionals in the topological dual of $\mathcal{R}$.

Proposition 3 *Let $\Phi \in \mathcal{R}^*$ be a positive linear functional of unitary operator norm and with $\langle \Phi, 1 \rangle = 1$. The restriction of Φ to $\mathcal{R}_1$ has a unique extension to $\mathcal{R}$. In particular, there exists a pair $(\phi, \tilde{\phi})$ of right-continuous processes with $\phi_{\mathbb{T}} + \tilde{\phi}_{\mathbb{T}} = 1$ such that the representation (4a) holds true in $\mathcal{R}$.*

Proof Let $\tilde{\Phi}$ be the restriction of Φ to $\mathcal{R}_1$. It defines an element of $\mathcal{R}_1^*$ since for $\mathsf{x} \in \mathcal{R}_1$

$$\left|\left\langle \tilde{\Phi}, \mathsf{x} \right\rangle\right| \leq \|\mathsf{x}\|_{\mathcal{R}} \leq \|\mathsf{x}\|_{\mathcal{R}_1}.$$

Hence, there exists a pair $(\phi, \tilde{\phi})$ such that the representation (4a) holds true in $\mathcal{R}_1$ for $\tilde{\Phi}$. According to Proposition 2, the representation (4a) extends also to elements of $\mathcal{R}$ while preserving the norm. The proof will be complete after showing that this extension coincides with Φ. This is a density argument. For $\mathsf{x} \in \mathcal{R}$ and $\lambda > 0$, let $T = \inf\{t \in [0, \mathbb{T}] \mid |\mathsf{x}_t| > \lambda\}$. We have that

$$P(\{T < \infty\}) \leq P(\{\mathsf{x}^* > \lambda\}) \leq \frac{1}{\lambda} \|\mathsf{x}\|_{\mathcal{R}}$$

where the second inequality follows from [7, equation (21.2), p.90]. The process $\tilde{\mathsf{x}} := (-\lambda) \vee \mathsf{x}^T \wedge \lambda$ is an element of $\mathcal{R}_1$. It converges to x in $\mathcal{R}$ (here we use the property of class(D)). Indeed:

$$\sup_{\tau\in\mathcal{T}} E[|\mathsf{x}_\tau - \tilde{\mathsf{x}}_\tau|] \leq \sup_{\tau\in\mathcal{T}} E[|\mathsf{x}_\tau| + |\mathsf{x}_T|; T < \infty, \tau \geq T].$$

Hence, by uniform integrability

$$\lim_{\lambda\to\infty} \sup_{\tau\in\mathcal{T}} E[|\mathsf{x}_\tau - \tilde{\mathsf{x}}_\tau|] = 0.$$

□

3 Partial hedging

We model the discounted price of an asset in a financial market by an $\mathbb{F}$-adapted semi-martingale $S := \{S_t\}_{0\leq t\leq \mathbb{T}}$ whose trajectories are continuous. We assume the market model is arbitrage-free in the sense that the set of equivalent martingale measures

$$\mathcal{M} = \{\mathbb{P}^* \sim \mathbb{P} \mid S \text{ is a local-martingale under } \mathbb{P}^*\}, \tag{6}$$

is nonempty. For equivalence of the property $\mathcal{M} \neq \emptyset$ and the property of No Free Lunch with Vanishing Risk we refer to Delbaen and Schachermayer [6]. We denote by $E_{\mathbb{P}^*}[\cdot]$ the corresponding $\mathbb{P}^*$-expectation. For a non negative adapted process v we make use of the notation

$$\pi(\mathsf{v}) := \sup_{\mathbb{P}\in\mathcal{M}} \sup_{\tau\in\mathcal{T}} E_{\mathbb{P}}[\mathsf{v}_\tau]\,. \tag{7}$$

We call to $\pi(\mathsf{v})$ the superhedging cost of v. For a motivation of this concept see Föllmer and Kramkov [9] and references.

We fix a càdlàg non-negative $\mathbb{F}$-adapted process H that will satisfy integrability conditions introduced below; see (9) and (10). The process H represents the discounted payoff of an American option. Fix $p > 1$ and let $l : \mathbb{R} \to \mathbb{R}_+$ be the function defined by $l(x) = \frac{1}{p}(x_+)^p$ where x_+ denotes the positive part of x. Note that l is a convex increasing function with $l'(0) = 0$ and

$$\lim_{x\to\infty} l'(x) = \infty. \tag{8}$$

The *seller* of the American option H has a non-negative initial capital c and might reduce risk by holding another position v. The maximal expected *shortfall risk* weighted by the function l is defined by

$$\sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - \mathsf{v}_\tau)].$$

We assume that the superhedging cost $\pi(\mathsf{H})$ of the American option H is finite:

$$\pi(\mathsf{H}) < \infty. \tag{9}$$

Moreover, we also assume that for some $p' > p$

$$\sup_{\tau \in \mathcal{T}} E[(\mathsf{H}_\tau)^{p'}] < \infty. \tag{10}$$

Note that (10) implies that $l(\mathsf{H}) \in \mathcal{R}$. Let $\mathcal{V}_c$ denote the family defined by

$$\mathcal{V}_c := \{\text{non negative } \mathsf{v} \in \mathcal{R}_\infty \mid \pi(\mathsf{v}) \leq c\}. \tag{11}$$

It will be convenient to denote by $\mathcal{V}_c^\#$ the elements of $\mathcal{V}_c$ where the inequality (11) is indeed an equality. Thus

$$\mathcal{V}_c^\# := \{\mathsf{v} \in \mathcal{V}_c \mid \pi(\mathsf{v}) = c\}. \tag{12}$$

Let

$$P(c) := \inf_{\mathsf{v} \in \mathcal{V}_c} \sup_{\tau \in \mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}_\tau\right)\right]. \tag{13}$$

We clearly have that

$$0 \leq P(c) < \infty, \text{ for } c \in [0, \infty). \tag{14}$$

The next result says that we can equivalently define the value $P(c)$ in the class $\mathcal{V}_c^\#$. The proof is obvious and we omit it.

Lemma 4 *The value $P(c)$ defined in equation (13) can equivalently be defined as*

$$P(c) = \inf_{\mathsf{v} \in \mathcal{V}_c^\#} \sup_{\tau \in \mathcal{T}} E\left[l\left(H_\tau - \mathsf{v}_\tau\right)\right]. \tag{15}$$

The following result is a consequence to (5) which holds true in $\mathcal{R}$ due to Proposition 2. We use the notation of (4a).

Lemma 5 *For $c > 0$ the value $P(c)$ is invariant if we take $\overline{K}$ instead of $\mathcal{T}$:*

$$P(c) = \inf_{\mathsf{v} \in \mathcal{V}_c^\#} \sup_{\Phi \in \overline{K}} \langle \Phi, l\left(H - \mathsf{v}\right)\rangle. \tag{16}$$

We defined $P(c)$ in terms of the class of processes $\mathcal{V}_c$ which are bounded processes. In general we do not expect the infimum to be attained in $\mathcal{V}_c$. However, this class is convenient for technical reasons in the convex duality approach and the value $P(c)$ does not change if we consider more in general non-negative processes with $\pi(\mathsf{v}) = c$. This last claim is a corollary to the next proposition. Note that we do not require v to be an element of $\mathcal{R}$.

Proposition 6 *Let v be a non-negative process. For $n \in \mathbb{N}$ let $\mathsf{v}^n := \mathsf{v} \wedge n$. Then*

$$\left\| l(H_\tau - \mathsf{v}_\tau^n) - l(H_\tau - \mathsf{v}_\tau) \right\|_{\mathcal{R}} \leq \sqrt[\alpha]{\left\| H^{p'} \right\|_{\mathcal{R}}} \sqrt[\beta]{\frac{1}{n}} \left\| H \right\|_{\mathcal{R}}$$

where $\alpha := \frac{p'}{p}$ *and* $\beta = \frac{p'}{p'-p} > 1$ *so that* $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Proof For $\tau \in \mathcal{T}$ we have

$$l(\mathsf{H}_\tau - \mathsf{v}_\tau^n) - l(\mathsf{H}_\tau - \mathsf{v}_\tau) = \begin{cases} l(\mathsf{H}_\tau - n) & \text{if } n < \mathsf{H}_\tau \le \mathsf{v}_\tau \\ l'(\mathsf{H}_\tau - \eta)(\mathsf{v}_\tau - \mathsf{v}_\tau^n) & \text{if } n < \mathsf{v}_\tau \le \mathsf{H}_\tau \\ 0 & \text{otherwise} \end{cases}$$

where η is a random variable with $\mathsf{v}_\tau^n \le \eta \le \mathsf{v}_\tau$. Hence

$$l(\mathsf{H}_\tau - \mathsf{v}_\tau^n) - l(\mathsf{H}_\tau - \mathsf{v}_\tau) \le \begin{cases} \mathsf{H}_\tau^p & \text{if } n < \mathsf{H}_\tau \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Let $\alpha := \frac{p'}{p}$ and $\beta = \frac{p'}{p'-p}$. Then

$$E[\mathsf{H}_\tau^p 1_{\{\mathsf{H}_\tau > n\}}] \le \sqrt[\alpha]{E[\mathsf{H}_\tau^{p'}]}\sqrt[\beta]{P[\mathsf{H}_\tau > n]} \le \sqrt[\alpha]{E[\mathsf{H}_\tau^{p'}]}\sqrt[\beta]{\frac{1}{n}\,\|\mathsf{H}\|_{\mathcal{R}}}$$

where the second inequality holds true due to the estimation in [7, (21.2)]. Thus

$$E\left[l(\mathsf{H}_\tau - \mathsf{v}_\tau^n) - l(\mathsf{H}_\tau - \mathsf{v}_\tau)\right] \le \sqrt[\alpha]{\left\|\mathsf{H}^{p'}\right\|_{\mathcal{R}}}\sqrt[\beta]{\frac{1}{n}\,\|\mathsf{H}\|_{\mathcal{R}}}.$$

This inequality proves the proposition. ◻

Note that $\mathsf{H} \in \mathcal{R}$ due to (10). Hence $P(c) = 0$ for $c \in [\pi(\mathsf{H}), +\infty)$, due to Proposition 6. Moreover, in this case there exists $\mathsf{v} \in \mathcal{R}$ with $\pi(\mathsf{v}) = \pi(\mathsf{H})$ which can be written as a stochastic integral with respect to the price process of the underlying assets. In this generality, the existence of such stochastic integral is a consequence to the celebrated optional decomposition theorem; see Föllmer and Kramkov [9].

4 Convex duality

Consider $\mathcal{R}$ and $\mathcal{R}_1$ as topological vector spaces with respect to the topology generated by their respective norms. Let $\mathbb{X}$ be the locally convex space defined as the direct sum $\mathcal{R} \oplus \mathcal{R}_1$; see [20, p. 55]. In the topological dual $\mathbb{X}^*$ (weak dual in the terminology of [20]) we consider the weak-star topology $\sigma(\mathbb{X}^*, \mathbb{X})$. Note that $\mathbb{X}^*$ is a locally convex topological vector space with respect to $\sigma(\mathbb{X}^*, \mathbb{X})$; see [20, p. 52].

In order to establish the dual problem to $P(\cdot)$ the first task to do is to define a specific class of elements in $\mathbb{X}^*$. This is done below in subsection 4.1 where dual variables are defined. In subsection 4.3 we present our main result on convex duality that establishes the dual problem of $P(\cdot)$; this is Theorem 12. However, before we can formulate this result a list of definitions are necessary and will take almost all of this section. The proof of Theorem 12 requires a minimax result for a Lagrangian functional. This is precisely the purpose of Sect. 5 below.

4.1 Dual variables

The price process S is a semimartingale and the stochastic integral makes sense. For a predictable process ξ, integrable with respect to S, we denote by $\int \xi dS$ the integral.

In this section the next family of processes will be crucial.

Definition 1 Let

$$\mathcal{X} := \left\{ c + \int \xi dS \mid c \in \mathbb{R}, \xi \text{ predictable and integrable with respect to } S \right\} \cap \mathcal{R}_\infty.$$

let $\tilde{\mathscr{M}}$ be the class of linear functionals on $\mathcal{X}$ constructed in the following way. For $\tau \in \mathcal{T}$ and a martingale measure $Q \in \mathcal{M}$ we define a functional Γ for $X \in \mathcal{X}$ by

$$\langle \Gamma, X \rangle := E_Q [X_\tau]. \tag{18}$$

This definition is the initial step in our construction of the dual domain. Indeed, we are interested in elements of the dual $\mathcal{R}_1^*$ extending such functionals Γ's. Actually, we will consider only positive extensions with operator norm bounded by one.

Definition 2 We denote by $\mathscr{M}$ the positive continuous linear functionals defined on $\mathcal{R}_1$ with operator norm bounded by one and such that their restriction to $\mathcal{X}$ can be obtained as in equation (18).

Linear extensions preserving positivity of linear forms is an important topic; see e.g., [1] for a systematic presentation. For our purposes we need the following result. The lemma is essentially an extension result. We cannot use Proposition 2 since here a change of probability measure appears in the definition of the linear form. Instead, the proof boils down to an application of a suitable version of the Hahn-Banach theorem.

Lemma 7 *The class $\mathscr{M}$ is non empty and every element of $\mathcal{M}$ is represented. More precisely, every functional in* (18) *can be extented to a continuous linear functional in $\mathcal{R}_1$ preserving the norm and positivity.*

Proof The proof is based on [1, Theorem 1.27] with the following details. In [1, Theorem 1.27] take $E := \mathcal{R}_1$, $F := \mathbb{R}$, $G := \mathcal{X}$, and define $p(\mathsf{x}) := \sup_{\tau \in \mathcal{T}} E[\mathsf{x}_\tau]$ together with $q(\mathsf{x}) := p(\mathsf{x}^+)$. Furthermore, for $\mathsf{x} \in G$, define $T(\mathsf{x})$ by (18), hence $T(\mathsf{x}) := E_Q[\mathsf{x}_\tau]$. Clearly E and F are Riesz spaces with F Dedekind complete. Moreover, T is a positive operator dominated from above by p on G.

Although the linear subspace G is not a Riesz subspace, the inequality $T(\mathsf{x}) \le q(\mathsf{x})$, for $\mathsf{x} \in G$ holds true. Hence, T extends to a positive operator R on all of E by the implication (3) $\implies$ (1) in [1, Theorem 1.27]. The positive operator R satisfies $R(\mathsf{x}) \le q(\mathsf{x})$ for $\mathsf{x} \in E$, hence, it has a norm less or equal than one. □

Proposition 8 *Let $\overline{\mathscr{M}^{co}}$ be the $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-closed convex hull of $\mathscr{M}$. Then, the family $\overline{\mathscr{M}^{co}}$ is $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-compact. Moreover, any element of $\overline{\mathscr{M}^{co}}$ admits a representation* (4a)*, although not necessarily normalized.*

Proof The closed unit ball $B = \{\mathsf{x}^* \in \mathcal{R}_1^* \mid \|\mathsf{x}^*\|_{\mathcal{R}_1^*} \leq 1\}$ is $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-compact due to Banach-Alaoglu theorem in the form [20, Corollary 4.3 p.84].

The family $\overline{\mathscr{M}^{co}}$ is obviously convex and $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-closed. Moreover, it is a subset of B since $\mathscr{M}^{co} \subset B$. Thus, $\overline{\mathscr{M}^{co}}$ is $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-compact.

The second claim of the lemma follows from [4, Proposition 1.3]. □

Now that we obtained the compact set $\overline{\mathscr{M}^{co}}$ through the convex and closed hull of $\mathscr{M}$, we must verify that we did not enlarged too much. This is the purpose of the next proposition. This is the only part where we explicitly use the fact that the price process S is continuous in order to guarantee that all the stochastic integrals of S are locally bounded.

Proposition 9 *For non negative $\mathsf{v} \in \mathcal{R}_\infty$ we have*

$$\pi(\mathsf{v}) = \sup_{\Gamma \in \overline{\mathscr{M}^{co}}} \langle \Gamma, \mathsf{v} \rangle .$$

Proof The inequality $\leq$ is clear. Now we prove equality.

For $\Gamma_0 \in \overline{\mathscr{M}^{co}}$ it holds true that for non-negative $\mathsf{v} \in \mathcal{R}_\infty$

$$\langle \Gamma_0, \mathsf{v} \rangle \leq \sup_{\Gamma \in \mathscr{M}^{co}} \langle \Gamma, \mathsf{v} \rangle = \sup_{\Gamma \in \mathscr{M}} \langle \Gamma, \mathsf{v} \rangle ,$$

since Γ_0 is in the $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$-closure of $\mathscr{M}^{co}$. Hence, it is sufficient to show that

$$\sup_{\Gamma \in \mathscr{M}} \langle \Gamma, \mathsf{v} \rangle \leq \pi(\mathsf{v}).$$

The key issue is that we only have control for a functional $\Gamma \in \mathscr{M}$ in the subspace $\mathcal{X}$ while v is not necessarily in this space. To couple with this, we apply the optional decomposition theorem [8] in order to obtain a predictable S-integrable process ξ such that for $Y := \pi(\mathsf{v}) + \int \xi dS$ it holds true that $Y \geq \mathsf{v}$. Here again the problem is that we do not know that $Y \in \mathcal{X}$. However, Y is locally bounded and it can be approximated with elements of $\mathcal{X}$, while Γ has the necessary continuity properties on $\mathcal{R}_1$. We provide the details.

Let Q be an equivalent local martingale measure of S and τ be a stopping time such that for $X \in \mathcal{X}$ the functional $\Gamma \in \mathcal{M}$ is given by $\langle \Gamma, X \rangle = E_Q[X_\tau]$. Let $\{\rho_n\}_{n\in\mathbb{N}}$ be a localizing sequence of stopping times such that $\mathbb{P}(\rho_n \geq \mathbb{T}) \nearrow 1$, and $Y^{\rho_n} := \{Y_{t\wedge\rho_n}\}_{t\in[0,\mathbb{T}]}$ is bounded. Note that the first property is stronger than the usual condition $\lim_{n\to\infty} \rho_n = \mathbb{T}$. In this case, $Y^{\rho_n} \in \mathcal{X}$ and we have

$$\left\langle \Gamma, Y^{\rho_n} \right\rangle = E_Q[Y_\tau^{\rho_n}] \leq \pi(\mathsf{v}).$$

The functional Γ is positive and therefore $\langle \Gamma, \mathsf{v}^{\rho_n} \rangle \leq \langle \Gamma, Y^{\rho_n} \rangle$. The continuity of Γ yields

$$\langle \Gamma, \mathsf{v} \rangle = \lim_{n\to\infty} \left\langle \Gamma, \mathsf{v}^{\rho_n} \right\rangle \leq \pi(\mathsf{v}),$$

since the sequence v^{ρ_n} converges to v on $\mathcal{R}_1$. Indeed, $|\mathsf{v} - \mathsf{v}^{\rho_n}| = 0$ on $\{\rho_n \geq \mathbb{T}\}$ and $|\mathsf{v} - \mathsf{v}^{\rho_n}|^* \leq 2\,|\mathsf{v}|_{L^\infty}$. □

Now we are ready to define the dual domain for the dual problem of $P(\cdot)$.

Definition 3 With a slight abuse of notation, let $\overline{\mathsf{K}}$ be the class of quasi-randomized stopping times introduced in Sect. 2.1, now considered as functionals in $\mathcal{R}$ by the identification in Proposition 3. Let $\mathfrak{M} \subset \mathbb{X}^*$ be defined by the canonical embedding of $\overline{\mathsf{K}} \times \overline{\mathscr{M}^{co}}$ into $\mathbb{X}^*$. This means that $\mathfrak{M}$ consists of linear functionals $T : \mathbb{X} \to \mathbb{R}$ constructed in the following way. For $(\Phi, \Gamma) \in \overline{\mathsf{K}} \times \overline{\mathscr{M}^{co}}$ let T be the linear functional defined by $T(x, y) := \langle \Phi, x\rangle + \langle \Gamma, y\rangle$ for $(x, y) \in \mathcal{R} \times \mathcal{R}_1$. It holds true that $T \in \mathbb{X}^*$; see [20, Theorem 4.3 p.137].

Proposition 10 *The weak dual* $\mathbb{X}^*$ *is algebraically isomorphic with* $\mathcal{R}^* \times \mathcal{R}_1^*$ *and* $\sigma(\mathbb{X}^*, \mathbb{X}) = \sigma(\mathcal{R}^*, \mathcal{R}) \times \sigma(\mathcal{R}_1^*, \mathcal{R}_1)$. *Moreover, the set* $\mathfrak{M}$ *is* $\sigma(\mathbb{X}^*, \mathbb{X})$*-compact.*

Proof The first and second claims about $\mathbb{X}^*$ hold true due to [20, Corollary 1 p. 138]. The family $\mathfrak{M}$ is compact since it is the product of sets that are compact in their respective spaces. Specifically, $\overline{\mathsf{K}}$ is compact with respect to $\sigma(\mathcal{R}^*, \mathcal{R})$ by the identification in Proposition 3, by following the steps in the proof of [4, Theorem 1.1]. The family $\overline{\mathscr{M}^{co}}$ is compact with respect to $\sigma(\mathcal{R}_1^*, \mathcal{R}_1)$ due to Proposition 8. □

4.2 Normal integrands

We recall in this part the concept of a convex normal integrand since it is necessary in the definition of the dual problem of $P(\cdot)$. Let $(\Xi, \sigma(\Xi))$ be a measurable space and T a topological space. Recall that a set valued mapping (or correspondence) $S : \Xi \mapsto T$ is measurable if the inverse image $S^{-1}(O) := \{\xi \in \Xi \mid S(\xi) \cap O \neq \emptyset\}$ of every open set O is $\sigma(\Xi)$-measurable. The following concept will be central to the development in the paper. A function $\mathbf{h} : \Xi \times T \to \overline{\mathbb{R}}$ is a *normal integrand* if its epigraphical mapping defined by

$$\operatorname{epi}\mathbf{h}(\xi) := \{(u, \alpha) \in T \times \mathbb{R} \mid \mathbf{h}(\xi, u) \leq \alpha\},$$

is closed-valued and $\sigma(\Xi)$-measurable. Additionally, if T carries a linear structure and $\mathbf{h}$ is convex in its second argument, $\mathbf{h}$ is a *normal convex integrand*.

We will consider the special case $\Xi = \Omega \times [0, \mathbb{T}]$ with σ-algebras $\mathcal{O}$ or $\mathcal{P}$. In this case, we say that $\mathbf{h}$ is optional (resp. predictable) if epi $\mathbf{h}$ is $\mathcal{O}$ (resp. $\mathcal{P}$) measurable.

4.3 The dual problem

In Definition 3 we presented the class $\mathfrak{M}$ whose elements are going to be dual variables for the dual problem of $P(\cdot)$ and now we introduce functionals acting on them. These preliminary definitions account for the dual framework in the duality formulated at the end of this paragraph; see Theorem 12.

We define a function $L : \mathbb{R}^4 \to \bar{\mathbb{R}}$ by

$$L(h, v, a, \lambda) := \begin{cases} al(h - v) + \lambda v & \text{if } (h, v, a, \lambda) \in \mathbb{R}^4_+, \\ \infty & \text{otherwise.} \end{cases} \tag{19}$$

Remark 3 We will constantly use the average of $\Phi \in \overline{\mathsf{K}}$ and $\Gamma \in \overline{\mathscr{M}^{co}}$ which we denote by $\Theta = (\theta, \tilde{\theta})$. Thus, if $\Phi = (\phi, \tilde{\phi})$ and $\Gamma = (\gamma, \tilde{\gamma})$ then $\theta = \frac{\phi+\gamma}{2}$ and $\tilde{\theta} = \frac{\tilde{\phi}+\tilde{\gamma}}{2}$. The density process of ϕ (resp. $\tilde{\phi}$) with respect to θ (resp. $\tilde{\theta}$) is denoted by $D\phi$ (resp. $D\tilde{\phi}$). Analogously $D\gamma$ (resp. $D\tilde{\gamma}$) is the density process of γ (resp. $\tilde{\gamma}$) with respect to θ (resp. $\tilde{\theta}$).

For $\mu = (\Phi, \Gamma) \in \mathfrak{M}$ and $\lambda > 0$ we define, with the notation in Remark 3, the convex normal integrands $\mathbf{h}^{\mu,\lambda}$ and $\tilde{\mathbf{h}}^{\mu,\lambda}$, respectively, by

$$\mathbf{h}^{\mu,\lambda}(\omega, t, v) := L\left(\mathsf{H}_t(\omega), v, D\phi_t(\omega), \lambda D\gamma_t(\omega)\right), \tag{20}$$

$$\tilde{\mathbf{h}}^{\mu,\lambda}(\omega, t, v) := L\left(\mathsf{H}_{t-}(\omega), v, D\tilde{\phi}_t(\omega), \lambda D\tilde{\gamma}_t(\omega)\right). \tag{21}$$

Let $u : \mathbb{R}^3 \to \mathbb{R}_+$ be defined by

$$u(h, a, \lambda) := \min_{v\in\mathbb{R}} L(h, v; a, \lambda). \tag{22}$$

We define the normal integrands:

$$\mathbf{u}^{\mu,\lambda}(\omega, t) := u(\mathsf{H}_t(\omega), D\phi_t(\omega), \lambda D\gamma_t(\omega)), \tag{23}$$

$$\tilde{\mathbf{u}}^{\mu,\lambda}(\omega, t) := u(\mathsf{H}_{t-}(\omega), D\tilde{\phi}_t(\omega), \lambda D\tilde{\gamma}_t(\omega)). \tag{24}$$

We define $\mathsf{J}^* : \mathfrak{M} \times \mathbb{R}_+ \to \bar{\mathbb{R}}$ by

$$\mathsf{J}^*(\mu, \lambda) := \int \mathbb{P}(d\omega) \left[\int_{[0,\mathbb{T}]} \mathbf{u}^{\mu,\lambda}(\omega, t)\theta(\omega, dt) + \int_{[0,\mathbb{T}]} \tilde{\mathbf{u}}^{\mu,\lambda}(\omega, t)\tilde{\theta}(\omega, dt) \right]. \tag{25}$$

We also define a ‘dual value function’ by

$$D(\lambda) := \sup_{\mu\in\mathfrak{M}} \mathsf{J}^*(\mu, \lambda). \tag{26}$$

We define the ‘Lagrangian’ $\Lambda : \mathcal{R}_\infty \times \mathfrak{M} \times \mathbb{R}_+ \to \bar{\mathbb{R}}$ by

$$\begin{aligned} \Lambda(\mathsf{v}; \mu, \lambda) &:= \delta_{\mathcal{R}^+_\infty}(\mathsf{v}) + \langle \Phi, l(\mathsf{H} - \mathsf{v}) \rangle + \lambda \langle \Gamma, \mathsf{v} \rangle \\ &= \delta_{\mathcal{R}^+_\infty}(\mathsf{v}) + \int_\Omega \mathbb{P}(d\omega) \quad \left[\int_{[0,\mathbb{T}]} \mathbf{h}^{\mu,\lambda}(\omega, t, \mathsf{v}(\omega, t))\theta(\omega, dt) \right. \\ &\quad \left. + \int_{[0,\mathbb{T}]} \tilde{\mathbf{h}}^{\mu,\lambda}(\omega, t, \mathsf{v}(\omega, t))\tilde{\theta}(\omega, dt) \right], \end{aligned} \tag{27}$$

here $\delta_{\mathcal{R}^+_\infty}(\mathsf{v})$ is equal to zero if v is an element of $\mathcal{R}^+_\infty$, otherwise it is equal to ∞.

Remark 4 The definition of Λ introduces a "hard constraint" in the process v. Another possibility is to consider the functional in (27) without $\delta_{\mathcal{R}^+_\infty}()$. This second definition also introduces constraints through the definition of L but only up to the functional μ. Note however that the definition (27) is more natural and will be important in the proof of Proposition 11 below.

The proof of the next proposition is based on the interchange rule [16, Theorem 11] for convex integrals of càdlàg processes. This proposition provides a crucial step for the proof of Theorem 12 below.

Proposition 11 *For $\mu \in \mathfrak{M}$, and $\lambda > 0$ fixed, we have*

$$J^*(\mu, \lambda) = \inf_{\mathsf{v} \in \mathcal{R}_\infty} \Lambda(\mathsf{v}; \mu, \lambda) = \inf_{\mathsf{v} \in \mathcal{R}^+_\infty} \Lambda(\mathsf{v}; \mu, \lambda). \tag{28}$$

Proof We do an application of the interchange rule [16, Theorem 11] , and for this, we verify its conditions. For unexplained notation see [16].

We start with [16, Assumption 3] . The process θ seen as a random measure is clearly optional. The function $\mathbf{h}^{\mu,\lambda}$ must be an optional convex normal integrand. Convexity follows directly from the convexity of l. The set-valued mapping epi $\mathbf{h}$ is closed-valued due to the continuity of the function L. Now we check measurability with respect to $\mathcal{O}$ let O be an open set of $\mathbb{R}^2$ and let $\tilde{O} := O \cap \mathbb{Q}^2$. We have

$$(\operatorname{epi} \mathbf{h}^{\mu,\lambda})^{-1}(O) = (\operatorname{epi} \mathbf{h}^{\mu,\lambda})^{-1}(\tilde{O}),$$

due to the continuity of L. The right-hand side of the equality is a countable union of sets of the form

$$\{(\omega, t) \mid L\left(\mathsf{H}_t(\omega), v, D\phi_t(\omega), \lambda D\gamma_t(\omega)\right) \le \alpha\}, \text{ for } (v, \alpha) \in \tilde{O},$$

which clearly is optional.

[16, Assumption 3] requires that the mapping $(t, v) \to \mathbf{h}^{\mu,\lambda}(\omega, t, v)$ satisfy [16, Assumption 1] for $\omega \notin N$ and N a null event. To verify this condition we must introduce a few notation. For $\omega \in \Omega$ let $S_t(\omega) = \{v \in \mathbb{R} \mid v \in \operatorname{cl} \operatorname{dom} \mathbf{h}^{\mu,\lambda}(\omega, t, \cdot)\}$ where $\operatorname{dom} \mathbf{h}^{\mu,\lambda}(\omega, t, \cdot) = \{v \mid \mathbf{h}^{\mu,\lambda}(\omega, t, v) < \infty\}$. It is clear that $S_t(\omega) = \mathbb{R}_+$. Let moreover, $D(S)(\omega)$ denote the class of right continuous, with finite left limits functions y in the interval $[0, \mathbb{T}]$ which are selections of $S(\omega)$, so $y(t) \in S_t(\omega)$ for all $t \in [0, \mathbb{T}]$. Note that $D(S)(\omega)$ consists of all non negative right continuous, with finite left limits functions. For the first part of [16, Assumption 1] take $q \in \mathbb{Q}_+$ and define the constant function $y^q(t) := q$. Then $y^q \in D(S)(\omega)$ and $S_t(\omega) = \operatorname{cl}\{y^q(t)\}$ for all $t \in [0, T]$. The second part of [16, Assumption 1] holds easily as well, but requires to introduce the following notation. Let $B(\omega) = \{y \in D(S)(\omega) \mid \int_{[0,\mathbb{T}]} \mathbf{h}^{\mu,\lambda}(\omega, t, y(t))\theta(\omega, dt) < \infty\}$. It is easy to see that $B(\omega) = D(S)(\omega)$, which implies the second part of [16, Assumption 1]. The final part of [16, Assumption 3], holds true trivially since $\mathbf{h}^{\mu,\lambda}$ is non negative. The verification of [16, Assumption 3] is now complete. The verification of [16, Assumption 4] is similar. We only verify the first part of [16, Assumption 5], since the verification of the second part is similar. Take $\mathsf{v} \in \mathcal{R}_\infty$ with

$$\int \mathbb{P}(d\omega)\left[\int_{[0,\mathbb{T}]} \mathbf{h}^{\mu,\lambda}(\omega, t, \mathsf{v}(\omega, t))\theta(\omega, dt)\right] < \infty.$$

It suffices to show that

$$\int \mathbb{P}(d\omega)\left[\int_{[0,\mathbb{T}]} \tilde{\mathbf{h}}^{\mu,\lambda}(\omega, t, \mathsf{v}_-(\omega, t))\tilde{\theta}(\omega, dt)\right] < \infty.$$

This follows from

$$\int \mathbb{P}(d\omega)\left[\int_{[0,\mathbb{T}]} \tilde{\mathbf{h}}^{\mu,\lambda}(\omega, t, \mathsf{v}_-(\omega, t))\tilde{\theta}(\omega, dt)\right] \le \|l(\mathsf{H})\|_{\mathcal{R}} + \lambda \|\mathsf{v}\|_{\mathcal{R}_\infty}.$$

The conditions of [16, Theorem 11] have been established. As a consequence we see that

$$\inf_{\mathsf{v}\in\mathcal{R}^\infty} \Lambda(\mathsf{v}; \mu, \lambda) = E\left[\int_{[0,\mathbb{T}]} \inf_{v\in\mathbb{R}} \mathbf{h}^{\mu,\lambda}(v)d\theta\right] + E\left[\int_{[0,\mathbb{T}]} \inf_{v\in\mathbb{R}} \tilde{\mathbf{h}}^{\mu,\lambda}(v)d\tilde{\theta}\right].$$

To conclude the proof, note that

$$E\left[\int_{[0,\mathbb{T}]} \inf_{v\in\mathbb{R}} \mathbf{h}^{\mu,\lambda}(v)d\theta\right] + E\left[\int_{[0,\mathbb{T}]} \inf_{v\in\mathbb{R}} \tilde{\mathbf{h}}^{\mu,\lambda}(v)d\tilde{\theta}\right] = \mathsf{J}^*(\mu, \lambda),$$

and this yields the last equality in (28). □

The proof of the next theorem is distributed in Sects. 5 and 6. The results in Appendix 7 are also necessary and of independent interest.

Theorem 12 *The function $P(\cdot)$ is convex and continuous. The functions D and P are conjugate to each other:*

$$D(\lambda) = \inf_{c>0}\{P(c) + \lambda c\} \tag{29}$$

$$P(c) = \sup_{\lambda>0}\{D(\lambda) - \lambda c\}. \tag{30}$$

Moreover, for $\lambda > 0$ fixed, there exists $\mu^* \in \mathfrak{M}$ such that $D(\lambda) = \mathsf{J}^*(\mu^*, \lambda)$.

Remark 5 We justify here the expression “The functions D and P are conjugate to each other”. For the concepts of proper and closed convex functions see [19]. Recall that the Legendre-Fenchel transform of a function f is given by

$$f^*(y) = \sup_x\{xy - f(x)\}$$

where x and y are elements of an Euclidean space. Let

$$\tilde{P}(x) := \begin{cases} P(x) & \text{for } x \geq 0 \\ +\infty & otherwise. \end{cases}$$

It is clear that $\tilde{P}$ is a proper function. It is convex and closed by Theorem 12. It is easy to see that

$$\tilde{P}^*(\lambda) = \begin{cases} \sup_{x\geq 0}\{\lambda x - \tilde{P}(x)\} & \text{for } \lambda \leq 0 \\ +\infty & otherwise. \end{cases}$$

Then $D(\lambda) = -\tilde{P}^*(-\lambda)$ for $\lambda > 0$.

5 A minimax identity

In this section we work with Λ, defined in (27). More precisely, we work with the restriction of Λ to the set $\mathcal{R}^+_\infty$, and we continue with the same notation Λ. Our main goal here is to prove a minimax identity which is going to be crucial for the proof of Theorem 12. To this end, we start with the necessary preliminaries in order to apply Sion's minimax theorem [21].

Lemma 13 *For $\lambda > 0$ fixed, $(\mathsf{v}; \mu) \to \Lambda(\mathsf{v}; \mu, \lambda)$ is convex in v and concave in $\mu \in \mathfrak{M}$.*

Proof It is clear that Λ is convex in v by the convexity of l. It is linear in μ so it is concave. □

Lemma 14 *For $\mu \in \mathfrak{M}$, $\lambda > 0$ fixed, $\Lambda(\cdot, \mu, \lambda)$ is lower-semicontinuous with respect to the norm in $\mathcal{R}_\infty$.*

Proof For $\alpha \in \mathbb{R}$ the lower level set

$$A := \{\mathsf{v} \in \mathcal{R}^+_\infty \mid \Lambda(\mathsf{v}; \mu, \lambda) \leq \alpha\}$$

is closed with respect to the norm of $\mathcal{R}_\infty$. Indeed, let $\{\mathsf{v}^n\}_{n\in\mathbb{N}}$ be a sequence in A converging to v with respect to the norm in $\mathcal{R}_\infty$. Note that v must be an element of $\mathcal{R}^+_\infty$. Let μ be given by the pair (Φ, Γ). We claim that $l(\mathsf{H} - \mathsf{v}^n)$ converges to $l(\mathsf{H} - \mathsf{v})$ in $\mathcal{R}$. Then

$$\lim_{n\to\infty} \left\langle \Phi, l(\mathsf{H} - \mathsf{v}^n) \right\rangle = \langle \Phi, l(\mathsf{H} - \mathsf{v}) \rangle ,$$

since Φ is a continuous linear functional in $\mathcal{R}$. Analogously

$$\lim_{n\to\infty} \left\langle \Gamma, \lambda\mathsf{v}^n \right\rangle = \langle \Gamma, \lambda\mathsf{v} \rangle ,$$

since v^n converges to v also with respect to the norm in $\mathcal{R}_1$ and $\Gamma \in \mathcal{R}_1^*$. As a consequence $\mathsf{v} \in A$.

Now we verify the claim. Take $\epsilon > 0$. We have for $\tau \in \mathcal{T}$

$$E[|l(\mathsf{H}_\tau - \mathsf{v}_\tau) - l(\mathsf{H}_\tau - \mathsf{v}_\tau^n)|] \leq l(\epsilon) + E[l'(\mathsf{H}_\tau)|\mathsf{v}_\tau^n - \mathsf{v}_\tau|1_{\{\mathsf{H}_\tau \geq \epsilon\}}].$$

Let $\alpha := \frac{p'}{p}$ and $\beta = \frac{p'}{p'-p}$. Similarly to the arguments in the proof of Proposition 6, we have

$$\begin{aligned} E[l'(\mathsf{H}_\tau)|\mathsf{v}_\tau^n - \mathsf{v}_\tau|1_{\{\mathsf{H}_\tau \geq \epsilon\}}] = E\left[\mathsf{H}_\tau^p \frac{|\mathsf{v}_\tau^n - \mathsf{v}_\tau|}{\mathsf{H}_\tau} 1_{\{\mathsf{H}_\tau \geq \epsilon\}}\right] &\leq \frac{1}{\epsilon} E\left[\mathsf{H}_\tau^p |\mathsf{v}_\tau^n - \mathsf{v}_\tau| 1_{\{\mathsf{H}_\tau \geq \epsilon\}}\right] \\ &\leq \frac{1}{\epsilon} \sqrt[\alpha]{E[\mathsf{H}_\tau^{p'}]} \sqrt[\beta]{E\left[|\mathsf{v}_\tau^n - \mathsf{v}_\tau|^\beta\right]} \\ &\leq \frac{1}{\epsilon} \sqrt[\alpha]{\left\|\mathsf{H}^{p'}\right\|_{\mathcal{R}}} \left\|\mathsf{v}^n - \mathsf{v}\right\|_{\mathcal{R}_\infty}. \end{aligned}$$

Hence

$$\left\|l(\mathsf{H} - \mathsf{v}) - l(\mathsf{H} - \mathsf{v}^n)\right\|_{\mathcal{R}} \leq l(\epsilon) + \frac{1}{\epsilon} \sqrt[\alpha]{\|\mathsf{H}\|_{\mathcal{R}}^{p'}} \left\|\mathsf{v}^n - \mathsf{v}\right\|_{\mathcal{R}_\infty}.$$

This last inequality proves the claim. The proof of the lemma is complete. □

Lemma 15 *For $\mathsf{v} \in \mathcal{R}_\infty^+$, $\lambda > 0$ fixed, $\Lambda(\mathsf{v}; \cdot, \lambda)$ is $\sigma(\mathbb{X}^*, \mathbb{X})$-continuous.*

Proof For $\mathsf{v} \in \mathcal{R}_\infty^+$ and $\lambda > 0$ fixed the functional $\Lambda(\mathsf{v}, \cdot, \lambda)$ is an evaluation at $(l(\mathsf{H} - \mathsf{v}), \lambda\mathsf{v}) \in \mathbb{X}$. Thus, from the definition of the weak star topology it follows continuity. □

After the preliminary results, the next proposition follows from Sion's minimax theorem [21].

Proposition 16 *For $\lambda > 0$ fixed, the following equality holds true*

$$\inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \max_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda) = \max_{\mu \in \mathfrak{M}} \inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \Lambda(\mathsf{v}; \mu, \lambda).$$

The next result shows how the minimax equality in Proposition 16 is connected to the value function $P(\cdot)$ and how it traces the way to duality.

Proposition 17 *For λ fixed, we have*

$$\inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \max_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda) = \inf_{c > 0} \{P(c) + \lambda c\}.$$

Proof It is clear that

$$\inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \sup_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda) = \inf_{c > 0} \inf_{\mathsf{v} \in \mathcal{V}_c^\#} \sup_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda).$$

For $\mathsf{v} \in \mathcal{V}_c^{\#}$ fixed, we have

$$\sup_{\Gamma \in \overline{\mathscr{M}^{co}}} \langle \Gamma, \mathsf{v} \rangle = c,$$

due to Proposition 9. Then

$$\begin{aligned}\sup_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda) &= \sup_{\Phi \in \overline{\mathsf{K}}} \langle \Phi, l(\mathsf{H} - \mathsf{v}) \rangle + \lambda \sup_{\Gamma \in \overline{\mathscr{M}^{co}}} \langle \Gamma, \mathsf{v} \rangle \\ &= \sup_{\Phi \in \overline{\mathsf{K}}} \langle \Phi, l(\mathsf{H} - \mathsf{v}) \rangle + \lambda c.\end{aligned}$$

Therefore,

$$\inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \sup_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda) = \inf_{c>0} \{P(c) + \lambda c\},$$

due to Lemma 4. □

6 Conclusion of the proof of Theorem 12

We are ready to conclude the proof of Theorem 12.

Proof For λ fixed we have

$$\begin{aligned} D(\lambda) &= \sup_{\mu \in \mathfrak{M}} \inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \Lambda(\mathsf{v}; \mu, \lambda) && (31)\\ &= \inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \Lambda(\mathsf{v}; \mu^*, \lambda), \text{ for some } \mu^* \in \mathfrak{M}, && (32)\\ &= \mathsf{J}^*(\lambda, \mu^*), && (33)\end{aligned}$$

where (31) follows from the interchange rule in Proposition 11, (32) from the minimax equality in Proposition 16 and (33) follows again from the interchange rule in Proposition 11.

We also have by the minimax equality of Proposition 16 that

$$D(\lambda) = \inf_{\mathsf{v} \in \mathcal{R}_\infty^+} \sup_{\mu \in \mathfrak{M}} \Lambda(\mathsf{v}; \mu, \lambda).$$

This equality together with Proposition 17 yields

$$D(\lambda) = \inf_{c>0} \{P(c) + \lambda c\}.$$

Thus, D is conjugate to P. The function P is a convex continuous function; see Lemma 18 and Proposition 20. Then, the function P is conjugate to D as well; see Rockafellar [19, Theorems 7.1 and 12.2]. □

7 A Convexity and continuity of the value function $P(\cdot)$

Lemma 18 *The value function $P(\cdot)$ is convex.*

Proof Take $\alpha \in (0, 1)$ and $c^1, c^2 \in \mathbb{R}_+$. For $\epsilon > 0$ there exists $\mathsf{v}^i \in \mathcal{V}_{c^i}$ such that

$$\sup_{\tau \in \mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^i_\tau\right)\right] \le P(c^i) + \epsilon, \text{ for } i = 1, 2,$$

due to (14). Let $c^3 := \alpha c^1 + (1-\alpha)c^2$ and $\mathsf{v}^3 := \alpha \mathsf{v}^1 + (1-\alpha)\mathsf{v}^2$. It is clear that $\mathsf{v}^3 \in \mathcal{V}_{c^3}$. We have

$$\begin{aligned} P(c^3) &\le \sup_{\tau \in \mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^3_\tau\right)\right] \\ &\le \alpha \sup_{\tau \in \mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^1_\tau\right)\right] + (1-\alpha) \sup_{\tau \in \mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^2_\tau\right)\right] \\ &\le \alpha P(c^1) + (1-\alpha) P(c^2) + \epsilon, \end{aligned}$$

where in the second inequality we have used the convexity of the function l. Taking the limit as ϵ decreases to zero, we obtain the convexity of the function $P(\cdot)$. □

The convexity property of $P(\cdot)$ yields local Lipschitz continuity on $(0, \pi(\mathsf{H}))$, see Rockafellar [19, Theorem 10.4]. In Proposition 20 below we prove continuity up to the boundary. We need the preliminary Lemma 19 below which is a result due to [22, Lemma 8], but here we need a different formulation. We need a preliminary definition.

Definition 4 Let Y^n be a sequence of positive processes. Then, the sequence Fatou converges to a positive process Y if there is a countable dense subset $\mathscr{T}$ of $[0, \mathbb{T}]$ such that

$$\begin{aligned} Y_t &= \liminf_{s \searrow t, s \in \mathscr{T}} \liminf_{n \to \infty} Y^n_s \\ &= \limsup_{s \searrow t, s \in \mathscr{T}} \limsup_{n \to \infty} Y^n_s \end{aligned}$$

Lemma 19 requires a strong uniform integrability condition but it changes lim inf in [22, Lemma 8] to lim which indeed is useful.

Lemma 19 *(Lemma 8 in [22]) Let Y^n be a sequence of non negative martingales Fatou converging to a process Y. Assume that the sequence is uniformly bounded by a constant k in that $\|Y^n\|_{\mathcal{R}_\infty} \le k$. Then, there exists a countable set $K \subset [0, \mathbb{T})$ such that for $t \in [0, \mathbb{T}] \backslash K$ we have $Y_t = \lim_{n \to \infty} Y^n_t$ a.s.*

Proof [22, Lemma 8] yields the existence of a set $K^1 \subset [0, \mathbb{T})$ such that for $t \in [0, \mathbb{T}] \backslash K^1$ we have $Y_t = \liminf_{n \to \infty} Y^n_t$ a.s. Now define $Z^n_t := k - Y^n_t$. It is clear that Z^n is a non negative martingale. It is also clear that Z^n Fatou converges to $k - Y$. Hence, there exists $K^2 \subset [0, \mathbb{T})$ such that for $t \in [0, \mathbb{T}] \backslash K^2$ we have $Z_t = \liminf_{n \to \infty} Z^n_t$ a.s.

which implies $Y_t = \limsup_{n\to\infty} Y_t^n$ a.s. To conclude the proof define $K = K^1 \cup K^2$. □

Proposition 20 *The function $P(\cdot)$ is continuous in the interval* $[0, \pi(H)]$.

Proof We only need to prove continuity for $c \in \{0, \pi(\mathsf{H})\}$.

We start with $c = 0$. Let $\{c_n\}_{n\in\mathbb{N}}$ be a sequence decreasing to zero. Let $\tilde{v}(n)$ be a non negative process with $\pi(\tilde{v}(n)) = c_n$ and such that $P(c_n) + \frac{1}{n} > \sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - \tilde{v}_\tau(n))]$. Without loss of generality, we assume that $\tilde{v}(n) = c_n + \int \xi^n dS$ and then it is a $\mathbb{P}^*$-local martingale for each n and $\mathbb{P}^* \in \mathcal{M}$; see eg., [2, Corollary 3.5].

For $m \in \mathbb{N}$ we define $\tilde{w}(n, m) := \tilde{v}(n) \wedge m$. The processes $\tilde{w}(n, m)$ continue to be non negative $\mathbb{P}^*$-local martingales and being uniformly bounded by m they are martingales. According to Proposition 6 for m sufficiently large we also have

$$P(c_n) + \frac{1}{n} > \sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - \tilde{w}_\tau(n, m))].$$

There are convex combinations $w(n, m) \in conv\,\langle \tilde{w}(n, m), \tilde{w}(n+1, m), \ldots\rangle$ Fatou-converging to a non-negative càdlàg process $w(m)$; see [9, Lemma 5.2]. In this case, there exists a dense countable set $D \subset [0, \mathbb{T}]$ such that for $t \in D$ we have $w_t(m) = \lim_{n\to\infty} w_t(n, m)$ due to Lemma 19. In particular, for a stopping time τ taking a finite number of values in D and $\mathbb{P}^* \in \mathcal{M}$

$$E_{\mathbb{P}^*}[w_\tau(m)] \le \liminf_{n\to\infty} E_{\mathbb{P}^*}[w_\tau(n, m)] \le \liminf_{n\to\infty} c_n = 0.$$

As a consequence

$$\pi(w(m)) = 0.$$

Thus,

$$P(0) \le \sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - w_\tau(m))]. \tag{34}$$

Let $\widehat{\tau}$ be a stopping time taking a finite number of values in D and ϵ-optimal in the following sense

$$\sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - w_\tau(m))] \le \epsilon + E[l(\mathsf{H}_{\widehat{\tau}} - w_{\widehat{\tau}}(m))].$$

For $\widehat{\tau}$ we have

$$\begin{aligned} E[l(\mathsf{H}_{\widehat{\tau}} - w_{\widehat{\tau}}(m))] &= E[\lim_{n\to\infty} l(\mathsf{H}_{\widehat{\tau}} - w_{\widehat{\tau}}(n, m))] \\ &\le \liminf_{n\to\infty} E[l(\mathsf{H}_{\widehat{\tau}} - w_{\widehat{\tau}}(n, m))] \\ &\le \liminf_{n\to\infty} P(c_n), \end{aligned}$$

where in the first equality we have used that $w_t(m) = \lim_{n\to\infty} w_t(n,m)$ for $t \in D$. Hence

$$\sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau - w_\tau(m))] \leq \epsilon + \liminf_{n\to\infty} P(c_n). \tag{35}$$

As a consequence we see that

$$P(0) \leq \liminf_{n\to\infty} P(c_n),$$

due to equations (34) and (35). This yields the continuity of P at zero.

For $c_0 = \pi(\mathsf{H})$ note that $P(c_0) = 0$, since we can take $\mathsf{v} = \mathsf{H}$. Let $\{c_n\}_{n\in\mathbb{N}}$ be a sequence increasing to c_0. Let $\mathsf{v}^n := \frac{c_n}{c_0}\mathsf{H}$. Then $\mathsf{v}^n \in \mathcal{V}_{c_n}$ and

$$0 = P(c_0) \leq P(c_n) \leq \sup_{\tau\in\mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^n_\tau\right)\right].$$

For any stopping time τ

$$0 \leq E[l(\mathsf{H}_\tau - \mathsf{v}^n_\tau)] \leq \left(1 - \frac{c_n}{c_0}\right) E[l(\mathsf{H}_\tau)],$$

due to the convexity of l. As a consequence

$$\sup_{\tau\in\mathcal{T}} E\left[l\left(\mathsf{H}_\tau - \mathsf{v}^n_\tau\right)\right] \leq \left(1 - \frac{c_n}{c_0}\right) \sup_{\tau\in\mathcal{T}} E[l(\mathsf{H}_\tau)].$$

We have proved the continuity of P at $\pi(\mathsf{H})$. □

References

1. Aliprantis, C.D., Burkinshaw, O.: Positive Operators. Springer, New York (2006)
2. Ansel, J.P., Stricker, C.: Couverture des actifs contingents et prix maximum. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 30(2):303–315, (1994)
3. Bensoussan, A.: On the theory of option pricing. Acta Appl. Math. **2**(2), 139–158 (1984)
4. Bismut, J.M.: Temps d'arrêt optimal, quasi-temps d'arrêt et retournement du temps. Ann. Probab. **7**(6), 933–964 (1979)
5. Bogachev, V.I.: Measure theory, vol. I. Springer-Verlag, Berlin (2007)

6. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**(3), 463–520 (1994)
7. Dellacherie, C., Meyer, P.A.: Probabilities and Potential B. Chapters V to VIII. North-Holland, New York (1982)
8. Föllmer, H., Kabanov, Y.M.: Optional decomposition and Lagrange multipliers. Financ. Stoch. **2**(1), 69–81 (1998)
9. Föllmer, H., Kramkov, D.O.: Optional decompositions under constraints. Probab. Theory Relat. Fields **109**(1), 1–25 (1997)
10. Föllmer, H., Leukert, P.: Efficient hedging: cost versus shortfall risk. Finance Stochast. **4**(2), 117–146 (2000)
11. Jacod, J.: Calcul Stochastique et Problèmes de Martingales, 1st edn. Springer, New York (1979)
12. Jacod, J., Shiryayev, A.N.: Limit theorems for stochastic processes. A comprehensive studies in mathematics, vol. 288, 2nd edn. Springer, Berlin (2002)
13. Karatzas, I.: On the pricing of American options. Appl. Math. Optim. **17**(1), 37–60 (1988)
14. Mertens, J.-F.: Théorie des processus stochastiques généraux applications aux surmartingales. Z. Wahrscheinlichkeitstheorie Verwandte Geb **22**, 45–68 (1972)
15. Mulinacci, S.: The efficient hedging problem for American options. Finance Stochast. **15**(2), 365–397 (2011)
16. Perkkiö, A. P., Trevino Aguilar, E.: Convex integral functionals of càdlàg processes. Submitted to Stochastic Processes and their Applications (2022)
17. Protter, P.: Stochastic integration and differential equations. In Stochastic modelling and applied probability, volume 21. Springer, Berlin (2005)
18. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, New York (2005)
19. Rockafellar, R. T.: Convex Analysis. Princeton Landmarks in Mathematics. Princeton University Press, (1970)
20. Schaefer, H.H., Wolff, M.P.: Topological Vector Spaces, 2nd edn. Springer, New York (1999)
21. Sion, M.: On general minimax theorems. Pac. J. Math. **8**, 171–176 (1958)
22. Žitković, G.: A filtered version of the Bipolar Theorem of Brannath and Schachermayer. J. Theor. Probab. **15**(1), 41–61 (2002)

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