Characterization of the Minimal Penalty of a Convex Risk Measure with Applications to Robust Utility Maximization for Lévy Models

Daniel Hernández-Hernández and Leonel Pérez-Hernández

Abstract The minimality of the penalty function associated with a convex risk measure is analyzed in this paper. First, in a general static framework, we provide necessary and sufficient conditions for a penalty function defined in a convex and closed subset of the absolutely continuous measures with respect to some reference measure $\mathbb{P}$ to be minimal on this set. When the probability space supports a Lévy process, we establish results that guarantee the minimality property of a penalty function described in terms of the coefficients associated with the density processes. These results are applied in the solution of the robust utility maximization problem for a market model based on Lévy processes.

Keywords Convex risk measures · Fenchel-Legendre transformation · Minimal penalization · Lévy process · Robust utility maximization

Mathematics Subject Classification 91B30, 46E30

1 Introduction

The definition of coherent risk measure was introduced by Artzner et al. in their fundamental works [1, 2] for finite probability spaces, giving an axiomatic characterization that was extended later by Delbaen [3] to general probability spaces. In the papers mentioned above one of the fundamental axioms was the positive homogeneity, and in further works it was removed, defining the concept of

D. Hernández-Hernández
Centro de Investigación en Matemáticas, Guanajuato, Mexico
e-mail: dher@cimat.mx

L. Pérez-Hernández (✉)
Departamento de Economía y Finanzas, Universidad de Guanajuato, Guanajuato, Mexico

D. Hernández-Hernández et al. (eds.), *XII Symposium of Probability and Stochastic Processes*, Progress in Probability 73,
https://doi.org/10.1007/978-3-319-77643-9_4

convex risk measure introduced by Föllmer and Schied [4, 5], Frittelli and Rosazza Gianin [7, 8] and Heath [10].

This is a rich area that has received a lot of attention and much work has been developed. There exists by now a well established theory in the static and dynamic cases, but there are still many questions unanswered in the static framework that need to be analyzed carefully. The one we focus on in this paper is the characterization of the penalty functions that are minimal for the corresponding static risk measure. Up to now, there are mainly two ways to deal with minimal penalty functions, namely the definition or the biduality relation. With the results presented in this paper we can start with a penalty function, which essentially discriminate models within a convex closed subset of absolutely continuous probability measures with respect to (w.r.t.) the market measure, and then guarantee that it corresponds to the minimal penalty of the corresponding convex risk measure on this subset. This property is, as we will see, closely related with the lower semicontinuity of the penalty function, and the complications to prove this property depend on the structure of the probability space.

We first provide a general framework, within a measurable space with a reference probability measure $\mathbb{P}$, and show necessary and sufficient conditions for a penalty function defined in a convex and closed subset of the absolutely continuous measures with respect to the reference measure to be minimal within this subset. The characterization of the form of the penalty functions that are minimal when the probability space supports a Lévy process is then studied. This requires to characterize the set of absolutely continuous measures for this space, and it is done using results that describe the density process for spaces which support semimartingales with the weak predictable representation property. Roughly speaking, using the weak representation property, every density process splits in two parts, one is related with the continuous local martingale part of the decomposition and the other with the corresponding discontinuous one. It is shown some kind of continuity property for the quadratic variation of a sequence of densities converging in L^1. From this characterization of the densities, a family of penalty functions is proposed, which turned out to be minimal for the risk measures generated by duality.

The previous results are applied to the solution of the robust utility maximization problem. The formulation of this problem, described formally in Sect. 6, is justified by the axiomatic system proposed by Maccheroni et al. [17], which led to utility functionals of the form

$$X \longrightarrow \inf_{\mathbb{Q}\in\mathcal{Q}'} \left\{\mathbb{E}_{\mathbb{Q}}\left[U\left(X\right)\right] + \vartheta\left(\mathbb{Q}\right)\right\}. \tag{1.1}$$

The elements of this display will be described in detail in the last section. For previous works on this direction we refer the interested reader to the works of Quenez [18], Schied [19] and Hernández-Hernández and Schied [11], and references therein.

The paper is organized as follows. Section 2 contains the description of the minimal penalty functions for a general probability space, providing necessary and sufficient conditions, the last one restricted to a subset of equivalent probability measures. Section 3 reports the structure of the densities for a probability space that supports a Lévy processes and the convergence properties needed to prove the lower semicontinuity of the set of penalty functions defined in Sect. 4. In this section we show that these penalty functions are minimal. The description of the market model is presented in Sect. 5, together with the characterization of the equivalent martingale measures and, finally, in the last section we solve the robust utility maximization problem using duality theory.

2 Minimal Penalty Function of Risk Measures Concentrated in $\mathcal{Q}_{\ll}(\mathbb{P})$

Given a penalty function ψ, it is possible to induce a convex risk measure ρ, which in turn has a representation by means of a minimal penalty function ψ_ρ^*. Starting with a penalty function ψ, we give in this section necessary and sufficient conditions in order to guarantee that it is the minimal penalty within the set of absolutely continuous probability measures. We begin recalling briefly some known results from the theory of static risk measures, and then a characterization for minimal penalties is presented.

2.1 Preliminaries from Static Measures of Risk

Let $X : \Omega \to \mathbb{R}$ be a mapping from a set Ω of possible market scenarios, representing the discounted net worth of the position. Uncertainty is represented by the measurable space $(\Omega, \mathcal{F})$, and we denote by $\mathcal{X}$ the linear space of bounded financial positions, including constant functions.

Definition 2.1

(i) The function $\rho : \mathcal{X} \to \mathbb{R}$, quantifying the risk of X, is a *monetary risk measure* if it satisfies the following properties:

$$\text{Monotonicity: If } X \leq Y \text{ then } \rho(X) \geq \rho(Y) \;\; \forall X, Y \in \mathcal{X}. \tag{2.1}$$

$$\text{Translation Invariance: } \rho(X + a) = \rho(X) - a \;\forall a \in \mathbb{R}\; \forall X \in \mathcal{X}. \tag{2.2}$$

(ii) When this function satisfies also the convexity property

$$\rho(\lambda X + (1 - \lambda) Y) \leq \lambda \rho(X) + (1 - \lambda) \rho(Y) \;\; \forall \lambda \in [0, 1] \;\; \forall X, Y \in \mathcal{X}, \tag{2.3}$$

it is said that ρ is a convex risk measure.

(iii) The function ρ is called normalized if $\rho(0) = 0$, and sensitive, with respect to a measure $\mathbb{P}$, when for each $X \in L_+^\infty(\mathbb{P})$ with $\mathbb{P}[X > 0] > 0$ we have that $\rho(-X) > \rho(0)$.

We say that a set function $\mathbb{Q} : \mathcal{F} \to [0, 1]$ is a *probability content* if it is finitely additive and $\mathbb{Q}(\Omega) = 1$. The set of *probability contents* on this measurable space is denoted by $\mathcal{Q}_{cont}$. From the general theory of static convex risk measures [6], we know that any map $\psi : \mathcal{Q}_{cont} \to \mathbb{R} \cup \{+\infty\}$, with $\inf_{\mathbb{Q}\in\mathcal{Q}_{cont}} \psi(\mathbb{Q}) \in \mathbb{R}$, induces a static convex measure of risk as a mapping $\rho : \mathfrak{M}_b \to \mathbb{R}$ given by

$$\rho(X) := \sup_{\mathbb{Q}\in\mathcal{Q}_{cont}} \left\{\mathbb{E}_{\mathbb{Q}}[-X] - \psi(\mathbb{Q})\right\}. \tag{2.4}$$

Here $\mathfrak{M}$ denotes the class of measurable functions and $\mathfrak{M}_b$ the subclass of bounded measurable functions. The function ψ will be referred as a *penalty function*. Föllmer and Schied [5, Theorem 3.2] and Frittelli and Rosazza Gianin [7, Corollary 7] proved that any convex risk measure is essentially of this form.

More precisely, a convex risk measure ρ on the space $\mathfrak{M}_b(\Omega, \mathcal{F})$ has the representation

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}_{cont}} \left\{\mathbb{E}_{\mathbb{Q}}[-X] - \psi_\rho^*(\mathbb{Q})\right\}, \tag{2.5}$$

where

$$\psi_\rho^*(\mathbb{Q}) := \sup_{X\in\mathcal{A}\rho} \mathbb{E}_{\mathbb{Q}}[-X], \tag{2.6}$$

and $\mathcal{A}_\rho := \{X \in \mathfrak{M}_b : \rho(X) \le 0\}$ is the *acceptance set* of ρ.

Remark 2.1 The penalty ψ_ρ^* is called the *minimal penalty function* associated to ρ because, for any other penalty function ψ fulfilling (2.4) , $\psi(\mathbb{Q}) \ge \psi_\rho^*(\mathbb{Q})$, for all $\mathbb{Q} \in \mathcal{Q}_{cont}$. Furthermore, for the minimal penalty function, the next biduality relation is satisfied

$$\psi_\rho^*(\mathbb{Q}) = \sup_{X\in\mathfrak{M}_b(\Omega,\mathcal{F})} \left\{\mathbb{E}_{\mathbb{Q}}[-X] - \rho(X)\right\}, \quad \forall \mathbb{Q} \in \mathcal{Q}_{cont}. \tag{2.7}$$

Let $\mathcal{Q}(\Omega, \mathcal{F})$ be the family of probability measures on the measurable space $(\Omega, \mathcal{F})$. Among the measures of risk, the class of them which representation in (2.5) is concentrated on the set of probability measures $\mathcal{Q} \subset \mathcal{Q}_{cont}$ are of special interest. Recall that a function $I : E \subset \mathbb{R}^\Omega \to \mathbb{R}$ is *sequentially continuous from below (above)* when $\{X_n\}_{n\in\mathbb{N}} \uparrow X \Rightarrow \lim_{n\to\infty} I(X_n) = I(X)$ (respectively $\{X_n\}_{n\in\mathbb{N}} \downarrow X \Rightarrow \lim_{n\to\infty} I(X_n) = I(X)$). Föllmer and Schied [6] proved that any sequentially continuous from below convex measure of risk is concentrated on the set $\mathcal{Q}$. Later, Krätschmer [15, Prop. 3 p. 601] established that the sequential continuity from below is not only a sufficient but also a necessary condition in order to have a representation, by means of the minimal penalty function in terms of probability measures.

We denote by $\mathcal{Q}_{\ll}(\mathbb{P})$ the subclass of absolutely continuous probability measure with respect to $\mathbb{P}$ and by $\mathcal{Q}_{\approx}(\mathbb{P})$ the subclass of equivalent probability measure. Of course, $\mathcal{Q}_{\approx}(\mathbb{P}) \subset \mathcal{Q}_{\ll}(\mathbb{P}) \subset \mathcal{Q}(\Omega, \mathcal{F})$.

Remark 2.2 When a convex risk measures in $\mathcal{X} := L^{\infty}(\mathbb{P})$ satisfies the property

$$\rho(X) = \rho(Y) \text{ if } X = Y \ \mathbb{P}\text{-a.s.} \tag{2.8}$$

and is represented by a penalty function ψ as in (2.4), we have that

$$\mathbb{Q} \in \mathcal{Q}_{cont} \setminus \mathcal{Q}_{cont}^{\ll} \Longrightarrow \psi(\mathbb{Q}) = +\infty, \tag{2.9}$$

where $\mathcal{Q}_{cont}^{\ll}$ is the set of contents absolutely continuous with respect to $\mathbb{P}$; see [6, Lemma 4.30 p. 172].

2.2 Minimal Penalty Functions

In the next sections we will show some of the difficulties that appear to prove the minimality of the penalty function when the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supports a Lévy process. We will also clarify the relevance of this property to get an optimal solution to the robust utility maximization problem in Sect. 6.

In order to establish the results of this section we only need to fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. When we deal with a set of absolutely continuous probability measures $\mathcal{K} \subset \mathcal{Q}_{\ll}(\mathbb{P})$ it is necessary to make reference to some topological concepts, meaning that we are considering the corresponding set of densities and the strong topology in $L^1(\mathbb{P})$. Recall that within a locally convex space, a convex set $\mathcal{K}$ is weakly closed if and only if $\mathcal{K}$ is closed in the original topology [6, Thm A.59].

Lemma 2.1 *Let* $\psi : \mathcal{K} \subset \mathcal{Q}_{\ll}(\mathbb{P}) \rightarrow \mathbb{R} \cup \{+\infty\}$ *be a function with* $\inf_{\mathbb{Q} \in \mathcal{K}} \psi(\mathbb{Q}) \in \mathbb{R}$, *and define the extension* $\psi(\mathbb{Q}) := \infty$ *for each* $\mathbb{Q} \in \mathcal{Q}_{cont} \setminus \mathcal{K}$, *with* $\mathcal{K}$ *a convex closed set. Also, define the function* Ψ, *with domain in* $L^1(\mathbb{P})$, *as*

$$\Psi(D) := \begin{cases} \psi(\mathbb{Q}) \text{ if } D = d\mathbb{Q}/d\mathbb{P} \text{ for } \mathbb{Q} \in \mathcal{K} \\ \infty \text{ otherwise.} \end{cases}$$

Then, for the convex measure of risk $\rho(X) := \sup\limits_{\mathbb{Q} \in \mathcal{Q}_{cont}} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \psi(\mathbb{Q}) \right\}$ *associated with* ψ *the following assertions hold:*

(a) If ρ *has as minimal penalty* ψ_{ρ}^{*} *the function* ψ *(i.e.* $\psi = \psi_{\rho}^{*}$*), then* Ψ *is a proper convex function and lower semicontinuous w.r.t. the (strong)* L^1*-topology or equivalently w.r.t. the weak topology* $\sigma(L^1, L^{\infty})$.

(b) If Ψ is convex and lower semicontinuous w.r.t. the (strong) L^1-topology or equivalently w.r.t. the weak topology $\sigma\left(L^1, L^\infty\right)$, then

$$\psi \mathbf{1}_{\mathcal{Q}_{\ll}(\mathbb{P})} = \psi_\rho^* \mathbf{1}_{\mathcal{Q}_{\ll}(\mathbb{P})}. \tag{2.10}$$

Proof

(a) Recall that $\sigma\left(L^1, L^\infty\right)$ is the coarsest topology on $L^1(\mathbb{P})$ under which every linear operator is continuous, and hence $\Psi_0^X(Z) := \mathbb{E}_{\mathbb{P}}[Z(-X)]$, with $Z \in L^1$, is a continuous function for each $X \in \mathfrak{M}_b(\Omega, \mathcal{F})$ fixed. For $\delta(\mathcal{K}) := \{Z : Z = d\mathbb{Q}/d\mathbb{P} \text{ with } \mathbb{Q} \in \mathcal{K}\}$ we have that

$$\Psi_1^X(Z) := \Psi_0^X(Z)\, \mathbf{1}_{\delta(\mathcal{K})}(Z) + \infty \times \mathbf{1}_{L^1 \setminus \delta(\mathcal{K})}(Z)$$

is clearly lower semicontinuous on $\delta(\mathcal{K})$. For $Z' \in L^1(\mathbb{P}) \setminus \delta(\mathcal{K})$ arbitrary fixed we have from Hahn-Banach's Theorem that there is a continuous lineal functional $l(Z)$ with $l\left(Z'\right) < \inf_{Z \in \delta(\mathcal{K})} l(Z)$. Taking $\varepsilon := \frac{1}{2}\left\{\inf_{Z \in \delta(\mathcal{K})} l(Z) - l\left(Z'\right)\right\}$ we have that the weak open ball $B\left(Z', \varepsilon\right) := \left\{Z \in L^1(\mathbb{P}) : \left|l\left(Z'\right) - l(Z)\right| < \varepsilon\right\}$ satisfies $B\left(Z', \varepsilon\right) \cap \delta(\mathcal{K}) = \varnothing$. Therefore, $\Psi_1^X(Z)$ is weak lower semicontinuous on $L^1(\mathbb{P})$, as well as $\Psi_2^X(Z) := \Psi_1^X(Z) - \rho(X)$. If

$$\psi(\mathbb{Q}) = \psi_\rho^*(\mathbb{Q}) = \sup_{X \in \mathfrak{M}_b(\Omega, \mathcal{F})} \left\{ \int Z(-X)\, d\mathbb{P} - \rho(X) \right\},$$

where $Z := d\mathbb{Q}/d\mathbb{P}$, we have that $\Psi(Z) = \sup_{X \in \mathfrak{M}_b(\Omega, \mathcal{F})} \left\{\Psi_2^X(Z)\right\}$ is the supremum of a family of convex lower semicontinuous functions with respect to the topology $\sigma\left(L^1, L^\infty\right)$, and $\Psi(Z)$ preserves both properties.

(b) For the Fenchel–Legendre transform (conjugate function) $\Psi^* : L^\infty(\mathbb{P}) \longrightarrow \mathbb{R}$ for each $U \in L^\infty(\mathbb{P})$

$$\Psi^*(U) = \sup_{Z \in \delta(\mathcal{K})} \left\{ \int Z U\, d\mathbb{P} - \Psi(Z) \right\} = \sup_{\mathbb{Q} \in \mathcal{Q}_{cont}} \left\{\mathbb{E}_{\mathbb{Q}}[U] - \psi(\mathbb{Q})\right\} \equiv \rho(-U).$$

From the lower semicontinuity of Ψ w.r.t. the weak topology $\sigma\left(L^1, L^\infty\right)$ that $\Psi = \Psi^{**}$. Considering the weak*-topology $\sigma\left(L^\infty(\mathbb{P}), L^1(\mathbb{P})\right)$ for $Z = d\mathbb{Q}/d\mathbb{P}$ we have that

$$\psi(\mathbb{Q}) = \Psi(Z) = \Psi^{**}(Z) = \sup_{U \in L^\infty(\mathbb{P})} \left\{ \int Z(-U)\, d\mathbb{P} - \Psi^*(-U) \right\} = \psi_\rho^*(\mathbb{Q}).$$

□

Remark 2.3

1. As it was pointed out in Remark 2.2, we have that

$$\mathbb{Q} \in \mathcal{Q}_{cont} \setminus \mathcal{Q}_{cont}^{\ll} \Longrightarrow \psi_{\rho}^{*}(\mathbb{Q}) = +\infty = \psi(\mathbb{Q}).$$

   Therefore, under the conditions of Lemma 2.1 (b) the penalty function ψ might differ from ψ_{ρ}^{*} on $\mathcal{Q}_{cont}^{\ll} \setminus \mathcal{Q}_{\ll}$. For instance, the penalty function defined as $\psi(\mathbb{Q}) := \infty \times \mathbf{1}_{\mathcal{Q}_{cont} \setminus \mathcal{Q}_{\ll}(\mathbb{P})}(\mathbb{Q})$ leads to the worst case risk measure $\rho(X) := \sup_{\mathbb{Q} \in \mathcal{Q}_{\ll}(\mathbb{P})} \mathbb{E}_{\mathbb{Q}}[-X]$, which has as minimal penalty the function

$$\psi_{\rho}^{*}(\mathbb{Q}) = \infty \times \mathbf{1}_{\mathcal{Q}_{cont} \setminus \mathcal{Q}_{cont}^{\ll}}(\mathbb{Q}).$$

2. Note that the total variation distance $d_{TV}\left(\mathbb{Q}^1, \mathbb{Q}^2\right) := \sup_{A \in \mathcal{F}} \left|\mathbb{Q}^1[A] - \mathbb{Q}^2[A]\right|$, with $\mathbb{Q}^1, \mathbb{Q}^2 \in \mathcal{Q}_{\ll}$, fulfills that $d_{TV}\left(\mathbb{Q}^1, \mathbb{Q}^2\right) \leq \left\| d\mathbb{Q}^1/d\mathbb{P} - d\mathbb{Q}^2/d\mathbb{P} \right\|_{L^1}$. Therefore, the minimal penalty function is lower semicontinuous in the total variation topology; see Remark 4.16 (b) p. 163 in [6].

3 Fundamentals of Lévy and Semimartingales Processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that $L := \{L_t\}_{t \in \mathbb{R}_+}$ is a Lévy process for this probability space if it is an adapted càdlàg process with independent stationary increments starting at zero. The filtration considered is $\mathbb{F} := \left\{\mathcal{F}_t^{\mathbb{P}}(L)\right\}_{t \in \mathbb{R}_+}$, the completion of its natural filtration, i.e. $\mathcal{F}_t^{\mathbb{P}}(L) := \sigma\{L_s : s \leq t\} \vee \mathcal{N}$ where $\mathcal{N}$ is the σ-algebra generated by all $\mathbb{P}$-null sets. The jump measure of L is denoted by $\mu : \Omega \times (\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_0)) \to \mathbb{N}$ where $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. The dual predictable projection of this measure, also known as its Lévy system, satisfies the relation $\mu^{\mathcal{P}}(dt, dx) = dt \times \nu(dx)$, where $\nu(\cdot) := \mathbb{E}[\mu([0,1] \times \cdot)]$ is the intensity or Lévy measure of L.

The Lévy-Itô decomposition of L is given by

$$L_t = bt + W_t + \int\limits_{[0,t] \times \{0 < |x| \leq 1\}} x d\left\{\mu - \mu^{\mathcal{P}}\right\} + \int\limits_{[0,t] \times \{|x| > 1\}} x\mu(ds, dx). \tag{3.1}$$

It implies that $L^c = W$ is the Wiener process, and hence $[L^c]_t = t$, where $(\cdot)^c$ and $[\,\cdot\,]$ denote the continuous martingale part and the process of quadratic variation of any semimartingale, respectively. For the predictable quadratic variation we use the notation $\langle \cdot \rangle$.

Denote by $\mathcal{V}$ the set of càdlàg, adapted processes with finite variation, and let $\mathcal{V}^+ \subset \mathcal{V}$ be the subset of non-decreasing processes in $\mathcal{V}$ starting at zero. Let $\mathcal{A} \subset \mathcal{V}$ be the class of processes with integrable variation, i.e. $A \in \mathcal{A}$ if and only if $\bigvee_0^\infty A \in$

$L^1(\mathbb{P})$, where $\bigvee_0^t A$ denotes the variation of A over the finite interval $[0,t]$. The subset $\mathcal{A}^+ = \mathcal{A} \cap \mathcal{V}^+$ represents those processes which are also increasing i.e. with non-negative right-continuous increasing trajectories. Furthermore, $\mathcal{A}_{loc}$ (resp. $\mathcal{A}_{loc}^+$) is the collection of adapted processes with locally integrable variation (resp. adapted locally integrable increasing processes). For a càdlàg process X we denote by $X_- := (X_{t-})$ the left hand limit process, where $X_{0-} := X_0$ by convention, and by $\triangle X = (\triangle X_t)$ the jump process $\triangle X_t := X_t - X_{t-}$.

Given an adapted càdlàg semimartingale U, the jump measure and its dual predictable projection (or compensator) are denoted by $\mu_U([0,t]\times A) := \sum_{s\le t} \mathbf{1}_A(\triangle U_s)$ and $\mu_U^{\mathcal{P}}$, respectively. Further, we denote by $\mathcal{P} \subset \mathcal{F}\otimes\mathcal{B}(\mathbb{R}_+)$ the predictable σ-algebra and by $\widetilde{\mathcal{P}} := \mathcal{P}\otimes\mathcal{B}(\mathbb{R}_0)$. With some abuse of notation, we write $\theta_1 \in \widetilde{\mathcal{P}}$ when the function $\theta_1 : \Omega\times\mathbb{R}_+\times\mathbb{R}_0 \to \mathbb{R}$ is $\widetilde{\mathcal{P}}$-measurable and $\theta \in \mathcal{P}$ for predictable processes.

Let

$$\mathcal{L}(U^c) := \Big\{\theta \in \mathcal{P} : \exists \{\tau_n\}_{n\in\mathbb{N}} \text{ sequence of stopping times with } \tau_n \uparrow \infty \text{ and } \mathbb{E}\left[\int_0^{\tau_n} \theta^2 d[U^c]\right] < \infty \ \forall n \in \mathbb{N}\Big\} \tag{3.2}$$

be the class of predictable processes $\theta \in \mathcal{P}$ integrable with respect to U^c in the sense of local martingale, and by

$$\Lambda(U^c) := \left\{\int \theta_0 dU^c : \theta_0 \in \mathcal{L}(U^c)\right\}$$

the linear space of processes which admits a representation as the stochastic integral with respect to U^c. For an integer valued random measure μ' we denote by $\mathcal{G}(\mu')$ the class of functions $\theta_1 : \Omega\times\mathbb{R}_+\times\mathbb{R}_0 \to \mathbb{R}$ satisfying the following conditions:

(i) $\theta_1 \in \widetilde{\mathcal{P}}$,

(ii) $\int_{\mathbb{R}_0} |\theta_1(t,x)| (\mu')^{\mathcal{P}}(\{t\},dx) < \infty \ \forall t > 0$,

(iii) The process

$$\left\{\sqrt{\sum_{s\le t}\left\{\int_{\mathbb{R}_0} \theta_1(s,x)\,\mu'(\{s\},dx) - \int_{\mathbb{R}_0} \theta_1(s,x)\,(\mu')^{\mathcal{P}}(\{s\},dx)\right\}^2}\right\}_{t\in\mathbb{R}_+} \in \mathcal{A}_{loc}^+.$$

The set $\mathcal{G}(\mu')$ represents the domain of the functional $\theta_1 \to \int \theta_1 d\left(\mu' - (\mu')^{\mathcal{P}}\right)$, which assign to θ_1 the unique purely discontinuous local martingale M with

$$\triangle M_t = \int_{\mathbb{R}_0} \theta_1(t,x)\,\mu'(\{t\},dx) - \int_{\mathbb{R}_0} \theta_1(t,x)\left(\mu'\right)^{\mathcal{P}}(\{t\},dx).$$

We use the notation $\int \theta_1 d\left(\mu' - (\mu')^{\mathcal{P}}\right)$ to write the value of this functional in θ_1. It is important to point out that this functional is not, in general, the integral with respect to the difference of two measures. For a detailed exposition on these topics see He et al. [9] or Jacod and Shiryaev [12], which are our basic references.

In particular, for the Lévy process L with jump measure μ,

$$\mathcal{G}(\mu) \equiv \left\{\theta_1 \in \widetilde{\mathcal{P}} : \left\{\sqrt{\sum_{s\leq t} \{\theta_1(s, \triangle L_s)\}^2 \mathbf{1}_{\mathbb{R}_0}(\triangle L_s)}\right\}_{t\in\mathbb{R}_+} \in \mathcal{A}^+_{loc}\right\}, \tag{3.3}$$

since $\mu^{\mathcal{P}}(\{t\} \times A) = 0$, for any Borel set A of $\mathbb{R}_0$.

We say that the semimartingale U has the *weak property of predictable representation* when

$$\mathcal{M}_{loc,0} = \Lambda\left(U^c\right) + \left\{\int \theta_1 d\left(\mu_U - \mu_U^{\mathcal{P}}\right) : \theta_1 \in \mathcal{G}(\mu_U)\right\}, \tag{3.4}$$

where the previous sum is the linear sum of the vector spaces, and $\mathcal{M}_{loc,0}$ is the linear space of local martingales starting at zero.

Let $\mathcal{M}$ and $\mathcal{M}_\infty$ denote the class of càdlàg and càdlàg uniformly integrable martingale respectively. The following lemma is interesting by itself to understand the continuity properties of the quadratic variation for a given convergent sequence of uniformly integrable martingale . It will play a central role in the proof of the lower semicontinuity of the penalization function introduced in Sect. 4. Observe that the assertion of this lemma is valid in a general filtered probability space and not only for the completed natural filtration of the Lévy process introduced above.

Lemma 3.1 *For* $\left\{M^{(n)}\right\}_{n\in\mathbb{N}} \subset \mathcal{M}_\infty$ *and* $M \in \mathcal{M}_\infty$ *the following implication holds*

$$M^{(n)}_\infty \xrightarrow[n\to\infty]{L^1} M_\infty \Longrightarrow \left[M^{(n)} - M\right]_\infty \xrightarrow{\mathbb{P}} 0.$$

Moreover,

$$M^{(n)}_\infty \xrightarrow[n\to\infty]{L^1} M_\infty \Longrightarrow \left[M^{(n)} - M\right]_t \xrightarrow[n\to\infty]{\mathbb{P}} 0 \;\; \forall t.$$

Proof From the L^1 convergence of $M^{(n)}_\infty$ to M_∞, we have that $\{M^{(n)}_\infty\}_{n\in\mathbb{N}} \cup \{M_\infty\}$ is uniformly integrable, which is equivalent to the existence of a convex and increasing function $G : [0, +\infty) \to [0, +\infty)$ such that

$$(i) \quad \lim_{x\to\infty} \frac{G(x)}{x} = \infty,$$

and

$$(ii)\quad \sup_{n\in\mathbb{N}} \mathbb{E}\left[G\left(\left|M_\infty^{(n)}\right|\right)\right] \vee \mathbb{E}\left[G\left(\left|M_\infty\right|\right)\right] < \infty.$$

Now, define the stopping times

$$\tau_k^n := \inf\left\{u > 0 : \sup_{t\leq u}\left|M_t^{(n)} - M_t\right| \geq k\right\}.$$

Observe that the estimation $\sup_{n\in\mathbb{N}} \mathbb{E}\left[G\left(\left|M_{\tau_k^n}^{(n)}\right|\right)\right] \leq \sup_{n\in\mathbb{N}} \mathbb{E}\left[G\left(\left|M_\infty^{(n)}\right|\right)\right]$ implies the uniformly integrability of $\left\{M_{\tau_k^n}^{(n)}\right\}_{n\in\mathbb{N}}$ for each k fixed. Since any uniformly integrable càdlàg martingale is of class $\mathcal{D}$, follows the uniform integrability of $\left\{M_{\tau_k^n}\right\}_{n\in\mathbb{N}}$ for all $k \in \mathbb{N}$, and hence $\left\{\sup_{t\leq\tau_k^n}\left|M_t^{(n)} - M_t\right|\right\}_{n\in\mathbb{N}}$ is uniformly integrable. This and the maximal inequality for supermartingales

$$\begin{aligned}\mathbb{P}\left[\sup_{t\in\mathbb{R}_+}\left|M_t^{(n)} - M_t\right| \geq \varepsilon\right] &\leq \frac{1}{\varepsilon}\left\{\sup_{t\in\mathbb{R}_+}\mathbb{E}\left[\left|M_t^{(n)} - M_t\right|\right]\right\}\\ &\leq \frac{1}{\varepsilon}\mathbb{E}\left[\left|M_\infty^{(n)} - M_\infty\right|\right] \longrightarrow 0,\end{aligned}$$

yields the convergence of $\left\{\sup_{t\leq\tau_k^n}\left|M_t^{(n)} - M_t\right|\right\}_{n\in\mathbb{N}}$ in L^1 to 0. The second Davis' inequality [9, Thm. 10.28] guarantees that, for some constant C,

$$\mathbb{E}\left[\sqrt{\left[M^{(n)} - M\right]_{\tau_k^n}}\right] \leq C\mathbb{E}\left[\sup_{t\leq\tau_k^n}\left|M_t^{(n)} - M_t\right|\right] \underset{n\to\infty}{\longrightarrow} 0 \quad \forall k \in \mathbb{N},$$

and hence $\left[M^{(n)} - M\right]_{\tau_k^n} \xrightarrow[n\to\infty]{\mathbb{P}} 0$ for all $k \in \mathbb{N}$.

Finally, to prove that $\left[M^{(n)} - M\right]_\infty \xrightarrow{\mathbb{P}} 0$ we assume that it is not true, and then $\left[M^{(n)} - M\right]_\infty \overset{\mathbb{P}}{\nrightarrow} 0$ implies that there exist $\varepsilon > 0$ and $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ with

$$d\left(\left[M^{(n_k)} - M\right]_\infty, 0\right) \geq \varepsilon$$

for all $k \in \mathbb{N}$,where $d(X, Y) := \inf\{\varepsilon > 0 : \mathbb{P}[|X - Y| > \varepsilon] \leq \varepsilon\}$ is the Ky Fan metric. We shall denote the subsequence as the original sequence, trying to keep the notation as simple as possible. Using a diagonal argument, a subsequence $\{n_i\}_{i\in\mathbb{N}} \subset \mathbb{N}$ can be chosen, with the property that $d\left(\left[M^{(n_i)} - M\right]_{\tau_k^{n_i}}, 0\right) < \frac{1}{k}$ for all $i \geq k$.

Since

$$\lim_{k\to\infty}\left[M^{(n_i)}-M\right]_{\tau_k^{n_i}}=\left[M^{(n_i)}-M\right]_\infty \quad \mathbb{P}-a.s.,$$

we can find some $k(n_i)\geq i$ such that

$$d\left(\left[M^{(n_i)}-M\right]_{\tau_{k(n_i)}^{n_i}},\left[M^{(n_i)}-M\right]_\infty\right)<\frac{1}{k}.$$

Then, using the estimation

$$\begin{aligned}\mathbb{P}&\left[\left|\left[M^{(n_k)}-M\right]_{\tau_{k(n_k)}^{n_k}}-\left[M^{(n_k)}-M\right]_{\tau_k^{n_k}}\right|>\varepsilon\right]\\ &\leq\mathbb{P}\left[\left\{\sup_{t\in\mathbb{R}_+}\left|M_t^{(n_k)}-M_t\right|\geq k\right\}\right],\end{aligned}$$

it follows that

$$d\left(\left[M^{(n_k)}-M\right]_{\tau_{k(n_k)}^{n_k}},\left[M^{(n_k)}-M\right]_{\tau_k^{n_k}}\right)\underset{k\to\infty}{\longrightarrow}0,$$

which yields a contradiction with $\varepsilon\leq d\left(\left[M^{(n_k)}-M\right]_\infty,0\right)$. Thus, $\left[M^{(n)}-M\right]_\infty \overset{\mathbb{P}}{\to} 0$. The last part of the this lemma follows immediately from the first statement. □

Using the Doob's stopping theorem we can conclude that for $M\in\mathcal{M}_\infty$ and an stopping time τ, that $M^\tau\in\mathcal{M}_\infty$, and therefore it follows as a corollary the following result.

Corollary 3.1 *For $\left\{M^{(n)}\right\}_{n\in\mathbb{N}}\subset\mathcal{M}_\infty$, $M\in\mathcal{M}_\infty$ and τ any stopping time holds*

$$M_\tau^{(n)}\overset{L^1}{\to}M_\tau\Longrightarrow\left[M^{(n)}-M\right]_\tau\overset{\mathbb{P}}{\longrightarrow}0.$$

Proof $\left[\left(M^{(n)}\right)^\tau-M^\tau\right]_\infty=\left[M^{(n)}-M\right]_\infty^\tau=\left[M^{(n)}-M\right]_\tau\overset{\mathbb{P}}{\longrightarrow}0.$ □

3.1 Density Processes

Given an absolutely continuous probability measure $\mathbb{Q}\ll\mathbb{P}$ in a filtered probability space, where a semimartingale with the weak predictable representation property is defined, the structure of the density process has been studied extensively by several authors; see Theorem 14.41 in He et al. [9] or Theorem III.5.19 in Jacod and Shiryaev [12].

Denote by $D_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right]$ the càdlàg version of the density process. For the increasing sequence of stopping times $\tau_n := \inf\left\{t \geq 0 : D_t < \frac{1}{n}\right\}$ $n \geq 1$ and $\tau_0 := \sup_n \tau_n$ we have $D_t(\omega) = 0\ \forall t \geq \tau_0(\omega)$ and $D_t(\omega) > 0\ \forall t < \tau_0(\omega)$, i.e.

$$D = D\mathbf{1}_{[\![0,\tau_0[\![}, \tag{3.5}$$

and the process

$$\frac{1}{D_{s-}}\mathbf{1}_{[\![D_-\neq 0]\!]} \text{ is integrable w.r.t. } D, \tag{3.6}$$

where we abuse of the notation by setting $[\![D_- \neq 0]\!] := \{(\omega,t) \in \Omega \times \mathbb{R}_+ : D_{t-}(\omega) \neq 0\}$. Both conditions (3.5) and (3.6) are necessary and sufficient in order that a semimartingale to be an *exponential semimartigale* [9, Thm. 9.41], i.e. $D = \mathcal{E}(Z)$ the Doléans-Dade exponential of another semimartingale Z. In that case we have

$$\tau_0 = \inf\{t > 0 : D_{t-} = 0 \text{ or } D_t = 0\} = \inf\{t > 0 : \triangle Z_t = -1\}. \tag{3.7}$$

It is well known that the Lévy-processes satisfy the weak property of predictable representation [9], when the completed natural filtration is considered. In the following lemma we present the characterization of the density processes for the case of these processes.

Lemma 3.2 *Given an absolutely continuous probability measure $\mathbb{Q} \ll \mathbb{P}$, there exist coefficients $\theta_0 \in \mathcal{L}(W)$ and $\theta_1 \in \mathcal{G}(\mu)$ such that*

$$\frac{d\mathbb{Q}_t}{d\mathbb{P}_t} = \frac{d\mathbb{Q}_t}{d\mathbb{P}_t}\mathbf{1}_{[\![0,\tau_0[\![} = \mathcal{E}\left(Z^\theta\right)(t), \tag{3.8}$$

where $Z_t^\theta \in \mathcal{M}_{loc}$ is the local martingale given by

$$Z_t^\theta := \int\limits_{]0,t]} \theta_0 dW + \int\limits_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\left(\mu(ds,dx) - ds\,\nu(dx)\right), \tag{3.9}$$

and $\mathcal{E}$ represents the Doleans-Dade exponential of a semimartingale. The coefficients θ_0 and θ_1 are dt-a.s and $\mu_{\mathbb{P}}^{\mathcal{P}}(ds,dx)$-a.s. unique on $[\![0,\tau_0]\!]$ and $[\![0,\tau_0]\!] \times \mathbb{R}_0$ respectively for $\mathbb{P}$-almost all ω. Furthermore, the coefficients can be chosen with $\theta_0 = 0$ on $]\!]\tau_0,\infty[\![$ and $\theta_1 = 0$ on $]\!]\tau_0,\infty[\![\times\mathbb{R}$.

Proof We only address the uniqueness of the coefficients θ_0 and θ_1, because the representation follows from (3.5) and (3.6). Let assume, that we have two possible vectors $\theta := (\theta_0, \theta_1)$ and $\theta' := \left(\theta_0', \theta_1'\right)$ satisfying the representation, i.e.

$$\begin{aligned} D_u \mathbf{1}_{[\![0,\tau_0[\![} &= \int D_{t-} d\{ \int_{]0,t]} \theta_0(s)\, dW_s + \int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,(\mu(ds,dx) - ds\,\nu(dx))\} \\ &= \int D_{t-} d\{ \int_{]0,t]} \theta_0'(s)\, dW_s + \int_{]0,t]\times\mathbb{R}_0} \theta_1'(s,x)\,(\mu(ds,dx) - ds\,\nu(dx))\}, \end{aligned}$$

and thus

$$\begin{aligned} \triangle D_t &= D_{t-} \triangle \left(\int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \right) \\ &= D_{t-} \triangle \left(\int_{]0,t]\times\mathbb{R}_0} \theta_1'(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \right). \end{aligned}$$

Since $D_{t-} > 0$ on $[\![0,\tau_0[\![$, it follows that

$$\triangle \left(\int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \right) = \triangle \left(\int_{]0,t]\times\mathbb{R}_0} \theta_1'(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \right).$$

Since two purely discontinuous local martingales with the same jumps are equal, it follows

$$\begin{aligned} &\int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \\ &\quad = \int_{]0,t]\times\mathbb{R}_0} \widehat{\theta_1}(s,x)\,(\mu(ds,dx) - ds\,\nu(dx)) \end{aligned}$$

and thus

$$\int D_{t-} d\{ \int_{]0,t]} \theta_0(s)\, dW_s\} = \int D_{t-} d\{ \int_{]0,t]} \theta_0'(s)\, dW_s\}.$$

Then,

$$0 = \left[\int D_{s-} d\left\{ \int_{]0,s]} \left(\theta_0'(u) - \theta_0(u)\right) dW_u \right\} \right]_t = \int_{]0,t]} (D_{s-})^2 \left\{\theta_0'(s) - \theta_0(s)\right\}^2 ds$$

and thus $\theta_0' = \theta_0$ dt-$a.s$ on $[\![0,\tau_0]\!]$ for $\mathbb{P}$-almost all ω.

On the other hand,

$$0 = \left\langle \int \left\{ \theta_1' (s,x) - \theta_1 (s,x) \right\} (\mu (ds,dx) - ds\, \nu (dx)) \right\rangle_t$$
$$= \int_{]0,t]\times \mathbb{R}_0} \left\{ \theta_1' (s,x) - \theta_1 (s,x) \right\}^2 \nu (dx)\, ds,$$

implies that $\theta_1 (s,x) = \theta_1' (s,x) \quad \mu_{\mathbb{P}}^{\mathcal{P}} (ds,dx)$-a.s. on $[\![0,\tau_0]\!] \times \mathbb{R}_0$ for $\mathbb{P}$-almost all ω. □

For $\mathbb{Q} \ll \mathbb{P}$ the function $\theta_1 (\omega,t,x)$ described in Lemma 3.2 determines the density of the predictable projection $\mu_{\mathbb{Q}}^{\mathcal{P}} (dt,dx)$ with respect to $\mu_{\mathbb{P}}^{\mathcal{P}} (dt,dx)$ (see He et al. [9] or Jacod and Shiryaev [12]). More precisely, for $B \in (\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_0))$ we have

$$\mu_{\mathbb{Q}}^{\mathcal{P}} (\omega, B) = \int_B (1 + \theta_1 (\omega,t,x))\, \mu_{\mathbb{P}}^{\mathcal{P}} (dt,dx). \tag{3.10}$$

In what follows we restrict ourself to the time interval $[0,T]$, for some $T > 0$ fixed, and we take $\mathcal{F} = \mathcal{F}_T$. The corresponding classes of density processes associated to $\mathcal{Q}_{\ll}(\mathbb{P})$ and $\mathcal{Q}_{\approx}(\mathbb{P})$ are denoted by $\mathcal{D}_{\ll}(\mathbb{P})$ and $\mathcal{D}_{\approx}(\mathbb{P})$, respectively. For instance, in the former case

$$\mathcal{D}_{\ll}(\mathbb{P}) := \left\{ D = \{D_t\}_{t\in[0,T]} : \exists \mathbb{Q} \in \mathcal{Q}_{\ll}(\mathbb{P}) \text{ with } D_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} \right\}, \tag{3.11}$$

and the processes in this set are of the form

$$\begin{aligned} D_t = \exp &\left\{ \int_{]0,t]} \theta_0 dW + \int_{]0,t]\times\mathbb{R}_0} \theta_1 (s,x) (\mu (ds,dx) - \nu (dx)\, ds) \right. \\ &\left. -\tfrac{1}{2} \int_{]0,t]} (\theta_0)^2 ds \right\} \times \\ &\times \exp \left\{ \int_{]0,t]\times\mathbb{R}_0} \{\ln (1 + \theta_1 (s,x)) - \theta_1 (s,x)\}\, \mu (ds,dx) \right\} \end{aligned} \tag{3.12}$$

for $\theta_0 \in \mathcal{L}(W)$ and $\theta_1 \in \mathcal{G}(\mu)$.

The set $\mathcal{D}_{\ll}(\mathbb{P})$ is characterized as follow.

Corollary 3.2 *The process D belongs to $\mathcal{D}_{\ll}(\mathbb{P})$ if and only if there are $\theta_0 \in \mathcal{L}(W)$ and $\theta_1 \in \mathcal{G}(\mu)$ with $\theta_1 \geq -1$ such that $D_t = \mathcal{E}\left(Z^\theta\right)(t)$ $\mathbb{P}$-a.s. $\forall t \in [0,T]$ and $\mathbb{E}_{\mathbb{P}}\left[\mathcal{E}\left(Z^\theta\right)(t)\right] = 1$ $\forall t \geq 0$, where $Z^\theta (t)$ is defined by (3.9).*

Proof The necessity follows from Lemma 3.2. Conversely, let $\theta_0 \in \mathcal{L}(W)$ and $\theta_1 \in \mathcal{G}(\mu)$ be arbitrarily chosen. Since $D_t = \int D_{s-} dZ_s^{\theta} \in \mathcal{M}_{loc}$ is a nonnegative local martingale, it is a supermartingale, with constant expectation from our assumptions. Therefore, it is a martingale, and hence the density process of an absolutely continuous probability measure. □

Since density processes are essentially uniformly integrable martingales, using Lemma 3.1 and Corollary 3.1 the following proposition follows immediately.

Proposition 3.1 *Let* $\left\{\mathbb{Q}^{(n)}\right\}_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{Q}_{\ll}(\mathbb{P})$*, with* $D_T^{(n)} := \left.\frac{d\mathbb{Q}^{(n)}}{d\mathbb{P}}\right|_{\mathcal{F}_T}$ *converging to* $D_T := \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T}$ *in* $L^1(\mathbb{P})$*. For the corresponding density processes* $D_t^{(n)} := \mathbb{E}_{\mathbb{P}}\left[D_T^{(n)} \,|\mathcal{F}_t\right]$ *and* $D_t := \mathbb{E}_{\mathbb{P}}\left[D_T \,|\mathcal{F}_t\right]$*, for* $t \in [0, T]$*, we have*

$$\left[D^{(n)} - D\right]_T \xrightarrow{\mathbb{P}} 0.$$

4 Penalty Functions for Densities

Now, we shall introduce a family of penalty functions for the density processes described in Sect. 3.1, for the absolutely continuous measures $\mathbb{Q} \in \mathcal{Q}_{\ll}(\mathbb{P})$.

Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ and $h_0, h_1 : \mathbb{R} \to \mathbb{R}_+$ be convex functions with $0 = h(0) = h_0(0) = h_1(0)$. Define the penalty function, with τ_0 as in (3.7), by

$$\begin{aligned}\vartheta(\mathbb{Q}) := \mathbb{E}_{\mathbb{Q}}\left[\int_0^{T\wedge\tau_0} h\left(h_0(\theta_0(t)) + \int_{\mathbb{R}_0} \delta(t,x)\, h_1(\theta_1(t,x))\,\nu(dx)\right) dt\right] \mathbf{1}_{\mathcal{Q}_{\ll}}(\mathbb{Q}) \\ +\infty \times \mathbf{1}_{\mathcal{Q}_{cont}\setminus\mathcal{Q}_{\ll}}(\mathbb{Q}),\end{aligned} \tag{4.1}$$

where θ_0, θ_1 are the processes associated to $\mathbb{Q}$ from Lemma 3.2 and $\delta(t,x) : \mathbb{R}_+ \times \mathbb{R}_0 \to \mathbb{R}_+$ is an arbitrary fixed nonnegative function $\delta(t,x) \in \mathcal{G}(\mu)$. Since $\theta_0 \equiv 0$ on $[\![\tau_0, \infty[\![$ and $\theta_1 \equiv 0$ on $[\![\tau_0, \infty[\![\times \mathbb{R}_0$ we have from the conditions imposed to h, h_0, and h_1

$$\begin{aligned}\vartheta(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}\left[\int_0^{T} h\left(h_0(\theta_0(t)) + \int_{\mathbb{R}_0} \delta(t,x)\, h_1(\theta_1(t,x))\,\nu(dx)\right) dt\right] \mathbf{1}_{\mathcal{Q}_{\ll}}(\mathbb{Q}) \\ +\infty \times \mathbf{1}_{\mathcal{Q}_{cont}\setminus\mathcal{Q}_{\ll}}(\mathbb{Q}).\end{aligned} \tag{4.2}$$

Further, define the convex measure of risk

$$\rho(X) := \sup_{\mathbb{Q}\in\mathcal{Q}_{\ll}(\mathbb{P})} \left\{\mathbb{E}_{\mathbb{Q}}[-X] - \vartheta(\mathbb{Q})\right\}. \tag{4.3}$$

Notice that ρ is a normalized and sensitive measure of risk. For each class of probability measures introduced so far, the subclass of those measures with a finite penalization is considered. We will denote by $\mathcal{Q}^{\vartheta}$, $\mathcal{Q}^{\vartheta}_{\ll}(\mathbb{P})$ and $\mathcal{Q}^{\vartheta}_{\approx}(\mathbb{P})$ the corresponding subclasses, i.e.

$$\mathcal{Q}^{\vartheta} := \{\mathbb{Q} \in \mathcal{Q} : \vartheta(\mathbb{Q}) < \infty\}, \ \mathcal{Q}^{\vartheta}_{\ll}(\mathbb{P}) := \mathcal{Q}^{\vartheta} \cap \mathcal{Q}_{\ll}(\mathbb{P}) \text{ and}$$
$$\mathcal{Q}^{\vartheta}_{\approx}(\mathbb{P}) := \mathcal{Q}^{\vartheta} \cap \mathcal{Q}_{\approx}(\mathbb{P}). \tag{4.4}$$

Notice that $\mathcal{Q}^{\vartheta}_{\approx}(\mathbb{P}) \neq \varnothing$.

Next theorem establishes the minimality on $\mathcal{Q}_{\ll}(\mathbb{P})$ of the penalty function introduced above for the risk measure ρ, its proof is based on the sufficient conditions given in Theorem 2.1. This result is relevant to obtain one of the main results of this paper, namely Theorem 6.1.

Theorem 4.1 *The penalty function ϑ defined in (4.2) is equal to the minimal penalty function of the convex risk measure ρ, given by (4.3), on $\mathcal{Q}_{\ll}(\mathbb{P})$, i.e.*

$$\vartheta \mathbf{1}_{\mathcal{Q}_{\ll}(\mathbb{P})} = \psi^{*}_{\rho}\mathbf{1}_{\mathcal{Q}_{\ll}(\mathbb{P})}.$$

Proof From Lemma 2.1 (b), we need to show that the penalization ϑ is proper, convex and that the corresponding identification, defined as $\Theta(Z) := \vartheta(\mathbb{Q})$ if $Z \in \delta\left(\mathcal{Q}_{\ll}(\mathbb{P})\right) := \left\{Z \in L^1(\mathbb{P}) : Z = d\mathbb{Q}/d\mathbb{P} \text{ with } \mathbb{Q} \in \mathcal{Q}_{\ll}(\mathbb{P})\right\}$ and $\Theta(Z) := \infty$ on $L^1 \setminus \delta\left(\mathcal{Q}_{\ll}(\mathbb{P})\right)$, is lower semicontinuous with respect to the strong topology.

First, observe that the function ϑ is proper, since $\vartheta(\mathbb{P}) = 0$. To verify the convexity of ϑ, choose $\mathbb{Q}, \widetilde{\mathbb{Q}} \in \mathcal{Q}^{\vartheta}_{\ll}$ and define $\mathbb{Q}^{\lambda} := \lambda\mathbb{Q} + (1-\lambda)\widetilde{\mathbb{Q}}$, for $\lambda \in [0,1]$. Notice that the corresponding density process can be written as $D^{\lambda} := \dfrac{d\mathbb{Q}^{\lambda}}{d\mathbb{P}} = \lambda D + (1-\lambda)\widetilde{D}$ $\mathbb{P}$-a.s. .

Now, from Lemma 3.2, let (θ_0, θ_1) and $(\widetilde{\theta}_0, \widetilde{\theta}_1)$ be the processes associated to $\mathbb{Q}$ and $\widetilde{\mathbb{Q}}$, respectively, and observe that from

$$D_t = 1 + \int_{[0,t]} D_{s-}\theta_0(s)\, dW_s + \int_{[0,t]\times\mathbb{R}_0} D_{s-}\theta_1(s,x)\, d\left(\mu(ds,dx) - ds\,\nu(dx)\right)$$

and the corresponding expression for $\widetilde{D}$ we have for $\tau^{\lambda}_n := \inf\left\{t \geq 0 : D^{\lambda}_t \leq \frac{1}{n}\right\}$

$$\int_0^{t\wedge\tau^{\lambda}_n} \left(D^{\lambda}_{s-}\right)^{-1} dD^{\lambda}_s = \int_0^{t\wedge\tau^{\lambda}_n} \frac{\lambda D_{s-}\theta_0(s) + (1-\lambda)\widetilde{D}_{s-}\widetilde{\theta}_0(s)}{(\lambda D_{s-} + (1-\lambda)\widetilde{D}_{s-})} dW_s +$$
$$\int_{[0,t\wedge\tau^{\lambda}_n]\times\mathbb{R}_0} \frac{\lambda D_{s-}\theta_1(s,x) + (1-\lambda)\widetilde{D}_{s-}\widetilde{\theta}_1(s,x)}{(\lambda D_{s-} + (1-\lambda)\widetilde{D}_{s-})} d\left(\mu - \mu^{\mathcal{P}}_{\mathbb{P}}\right).$$

The weak predictable representation property of the local martingale $\int_0^{t\wedge\tau_n^\lambda}\left(D_{s-}^\lambda\right)^{-1} dD_s^\lambda$, yield on the other hand

$$\int_0^{t\wedge\tau_n^\lambda}\left(D_{s-}^\lambda\right)^{-1} dD_s^\lambda = \int_0^{t\wedge\tau_n^\lambda}\theta_0^\lambda(s)\,dW_s + \int_{[0,t\wedge\tau_n^\lambda]\times\mathbb{R}_0}\theta_1^\lambda(s,x)\,d\left(\mu-\mu_{\mathbb{P}}^{\mathcal{P}}\right),$$

where identification

$$\theta_0^\lambda(s) = \frac{\lambda D_{s-}\theta_0(s) + (1-\lambda)\widetilde{D}_{s-}\widetilde{\theta}_0(s)}{\left(\lambda D_{s-} + (1-\lambda)\widetilde{D}_{s-}\right)},$$

and

$$\theta_1^\lambda(s,x) = \frac{\lambda D_{s-}\theta_1(s,x) + (1-\lambda)\widetilde{D}_{s-}\widetilde{\theta}_1(s,x)}{\left(\lambda D_{s-} + (1-\lambda)\widetilde{D}_{s-}\right)}.$$

This is possible thanks to the uniqueness of the representation in Lemma 3.2. The convexity follows now from the convexity of h, h_0 and h_1, using the fact that any convex function is continuous in the interior of its domain. More specifically,

$$\begin{aligned}
\vartheta\left(\mathbb{Q}^\lambda\right) \leq{}& \mathbb{E}_{\mathbb{Q}^\lambda}\Bigg[\int_{[0,T]}\frac{\lambda D_s}{(\lambda D_s+(1-\lambda)\widetilde{D}_s)}h\Bigg(h_0(\theta_0(s)) \\
&+\int_{\mathbb{R}_0}\delta(s,x)\,h_1(\theta_1(s,x))\,\nu(dx)\Bigg)ds\Bigg] \\
&+\mathbb{E}_{\mathbb{Q}^\lambda}\Bigg[\int_{[0,T]}\frac{(1-\lambda)\widetilde{D}_s}{(\lambda D_s+(1-\lambda)\widetilde{D}_s)}h\Bigg(h_0\left(\widetilde{\theta}_0(s)\right) \\
&+\int_{\mathbb{R}_0}\delta(s,x)\,h_1(\widetilde{\theta}_1(s,x))\nu(dx)\Bigg)ds\Bigg] \\
={}& \int_{[0,T]}\int_\Omega\frac{\lambda D_s}{\left(\lambda D_s+(1-\lambda)\widetilde{D}_s\right)}h\Bigg(h_0(\theta_0(s)) \\
&+\int_{\mathbb{R}_0}\delta(s,x)\,h_1(\theta_1(s,x))\,\nu(dx)\Bigg) \\
&\times\left(\lambda D_s+(1-\lambda)\widetilde{D}_s\right)\mathbf{1}_{\{\lambda D_s+(1-\lambda)\widetilde{D}_s>0\}}d\mathbb{P}ds \\
&+\int_{[0,T]}\int_\Omega\frac{(1-\lambda)\widetilde{D}_s}{\left(\lambda D_s+(1-\lambda)\widetilde{D}_s\right)}h\Bigg(h_0\left(\widetilde{\theta}_0(s)\right)
\end{aligned}$$

$$+\int_{\mathbb{R}_0}\delta\left(s,x\right)h_1(\widetilde{\theta}_1\left(s,x\right))\nu\left(dx\right)\Bigg)$$
$$\times\left(\lambda D_s+(1-\lambda)\,\widetilde{D}_s\right)\mathbf{1}_{\{\lambda D_s+(1-\lambda)\widetilde{D}_s>0\}}d\mathbb{P}ds$$
$$=\lambda\vartheta\left(\mathbb{Q}\right)+(1-\lambda)\,\vartheta\left(\widetilde{\mathbb{Q}}\right),$$

where we used that

$$\left\{\int_{\mathbb{R}_0}\delta\left(t,x\right)h_1\left(\theta_1\left(t,x\right)\right)\nu\left(dx\right)\right\}_{t\in\mathbb{R}_+}\quad\text{and}\quad\left\{\int_{\mathbb{R}_0}\delta\left(t,x\right)h_1(\widetilde{\theta}_1\left(t,x\right))\nu\left(dx\right)\right\}_{t\in\mathbb{R}_+}$$

are predictable processes.

It remains to prove the lower semicontinuity of Θ. As pointed out earlier, it is enough to consider a sequence of densities $Z^{(n)}:=\frac{d\mathbb{Q}^{(n)}}{d\mathbb{P}}\in\delta\left(\mathcal{Q}_{\ll}\left(\mathbb{P}\right)\right)$ converging in $L^1\left(\mathbb{P}\right)$ to $Z:=\frac{d\mathbb{Q}}{d\mathbb{P}}$. Denote the corresponding density processes by $D^{(n)}$ and D, respectively. In Proposition 3.1 it was verified the convergence in probability to zero of the quadratic variation process

$$\left[D^{(n)}-D\right]_T=\int_0^T\left\{D_{s-}^{(n)}\theta_0^{(n)}\left(s\right)-D_{s-}\theta_0\left(s\right)\right\}^2ds$$
$$+\int_{[0,T]\times\mathbb{R}_0}\left\{D_{s-}^{(n)}\theta_1^{(n)}\left(s,x\right)-D_{s-}\theta_1\left(s,x\right)\right\}^2\mu\left(ds,dx\right).$$

This implies that

$$\left.\begin{array}{c}\int_0^T\left\{D_{s-}^{(n)}\theta_0^{(n)}\left(s\right)-D_{s-}\theta_0\left(s\right)\right\}^2ds\xrightarrow{\mathbb{P}}0,\\ \text{and}\\ \int_{[0,T]\times\mathbb{R}_0}\left\{D_{s-}^{(n)}\theta_1^{(n)}\left(s,x\right)-D_{s-}\theta_1\left(s,x\right)\right\}^2\mu\left(ds,dx\right)\xrightarrow{\mathbb{P}}0.\end{array}\right\}\tag{4.5}$$

Then, for an arbitrary but fixed subsequence, there exists a sub-subsequence such that $\mathbb{P}$-a.s.

$$\left\{D_{s-}^{(n)}\theta_0^{(n)}\left(s\right)-D_{s-}\theta_0\left(s\right)\right\}^2\xrightarrow{L^1(\lambda)}0$$

and

$$\left\{D_{s-}^{(n)}\theta_1^{(n)}\left(s,x\right)-D_{s-}\theta_1\left(s,x\right)\right\}^2\xrightarrow{L^1(\mu)}0,$$

where for simplicity we have denoted the sub-subsequence as the original sequence. Now, we claim that for the former sub-subsequence it also holds that

$$\begin{cases} D_{s-}^{(n)}\theta_0^{(n)}(s) \overset{\lambda\times\mathbb{P}\text{-a.s.}}{\longrightarrow} D_{s-}\theta_0(s), \\ D_{s-}^{(n)}\theta_1^{(n)}(s,x) \overset{\mu\times\mathbb{P}\text{-a.s.}}{\longrightarrow} D_{s-}\theta_1(s,x). \end{cases} \tag{4.6}$$

We present first the arguments for the proof of the second assertion in (4.6). Assuming the opposite, there exists $C \in \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}_0) \otimes \mathcal{F}_T$, with $\mu \times \mathbb{P}[C] > 0$, and such that for each $(s,x,\omega) \in C$

$$\lim_{n\to\infty} \left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 = c \neq 0,$$

or the limit does not exist.

Let $C(\omega) := \{(t,x) \in [0,T] \times \mathbb{R}_0 : (t,x,\omega) \in C\}$ be the ω-section of C. Observe that $B := \{\omega \in \Omega : \mu[C(\omega)] > 0\}$ has positive probability: $\mathbb{P}[B] > 0$.

From (4.5), any arbitrary but fixed subsequence has a sub-subsequence converging $\mathbb{P}$-a.s. Denoting such a sub-subsequence simply by n, we can fix $\omega \in B$ with

$$\begin{aligned} &\int_{C(\omega)} \left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 d\mu(s,x) \\ &\leq \int_{[0,T]\times\mathbb{R}_0} \left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 d\mu(s,x) \underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

and hence $\left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2$ converges in μ-measure to 0 on $C(\omega)$. Again, for any subsequence there is a sub-subsequence converging μ-a.s. to 0. Furthermore, for an arbitrary but fixed $(s,x) \in C(\omega)$, when the limit does not exist

$$\begin{aligned} a &:= \liminf_{n\to\infty} \left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 \\ &\neq \limsup_{n\to\infty} \left\{ D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 =: b, \end{aligned}$$

and we can choose converging subsequences $n(i)$ and $n(j)$ with

$$\lim_{i\to\infty} \left\{ D_{s-}^{n(i)}\theta_1^{n(i)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 = a$$

$$\lim_{j\to\infty} \left\{ D_{s-}^{n(j)}\theta_1^{n(j)}(s,x) - D_{s-}\theta_1(s,x) \right\}^2 = b.$$

From the above argument, there are sub-subsequences $n(i(k))$ and $n(j(k))$ such that

$$a = \lim_{k\to\infty}\left\{D_{s-}^{n(i(k))}\theta_1^{n(i(k))}(s,x) - D_{s-}\theta_1(s,x)\right\}^2 = 0$$

$$b = \lim_{k\to\infty}\left\{D_{s-}^{n(j(k))}\theta_1^{n(j(k))}(s,x) - D_{s-}\theta_1(s,x)\right\}^2 = 0,$$

which is clearly a contradiction.

For the case when

$$\lim_{n\to\infty}\left\{D_{s-}^{(n)}\theta_1^{(n)}(s,x) - D_{s-}\theta_1(s,x)\right\}^2 = c \neq 0,$$

the same argument can be used, and get a subsequence converging to 0, having a contradiction again. Therefore, the second part of our claim in (4.6) holds.

Since $D_{s-}^{(n)}\theta_1^{(n)}(s,x)$, $D_{s-}\theta_1(s,x) \in \mathcal{G}(\mu)$, we have, in particular, that $D_{s-}^{(n)}\theta_1^{(n)}(s,x) \in \widetilde{\mathcal{P}}$ and $D_{s-}\theta_1(s,x) \in \widetilde{\mathcal{P}}$ and hence $C \in \widetilde{\mathcal{P}}$. From the definition of the predictable projection it follows that

$$0 = \mu\times\mathbb{P}[C] = \int_{\Omega}\int_{[0,T]\times\mathbb{R}_0} \mathbf{1}_C(s,\omega)\, d\mu d\mathbb{P} = \int_{\Omega}\int_{[0,T]\times\mathbb{R}_0} \mathbf{1}_C(s,\omega)\, d\mu_{\mathbb{P}}^{\mathcal{P}} d\mathbb{P}$$

$$= \int_{\Omega}\int_{\mathbb{R}_0}\int_{[0,T]} \mathbf{1}_C(s,\omega)\, ds d\nu d\mathbb{P} = \lambda\times\nu\times\mathbb{P}[C],$$

and thus

$$D_{s-}^{(n)}\theta_1^{(n)}(s,x) \overset{\lambda\times\nu\times\mathbb{P}\text{-a.s.}}{\longrightarrow} D_{s-}\theta_1(s,x).$$

Since

$$\int_{\Omega\times[0,T]} \left|D_{t-}^{(n)} - D_{t-}\right| d\mathbb{P}\times dt = \int_{\Omega\times[0,T]} \left|D_t^{(n)} - D_t\right| d\mathbb{P}\times dt \longrightarrow 0,$$

we have that

$$\left\{D_{t-}^{(n)}\right\}_{t\in[0,T]} \overset{L^1(\lambda\times\mathbb{P})}{\longrightarrow} \{D_{t-}\}_{t\in[0,T]} \quad \text{and} \quad \left\{D_t^{(n)}\right\}_{t\in[0,T]} \overset{L^1(\lambda\times\mathbb{P})}{\longrightarrow} \{D_t\}_{t\in[0,T]}.$$

Then, for an arbitrary but fixed subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$, there is a sub-subsequence $\{n_{k_i}\}_{i\in\mathbb{N}} \subset \mathbb{N}$ such that

$$\begin{array}{ccc} D_{t-}^{(n_{k_i})}\theta_1^{(n_{k_i})}(t,x) & \overset{\lambda\times\nu\times\mathbb{P}\text{-a.s.}}{\longrightarrow} & D_{t-}\theta_1(t,x), \\ D_{t-}^{(n_{k_i})} & \overset{\lambda\times\mathbb{P}\text{-a.s.}}{\longrightarrow} & D_{t-}, \\ D_t^{(n_{k_i})} & \overset{\lambda\times\mathbb{P}\text{-a.s.}}{\longrightarrow} & D_t. \end{array}$$

Furthermore, $\mathbb{Q} \ll \mathbb{P}$ implies that $\lambda \times \nu \times \mathbb{Q} \ll \lambda \times \nu \times \mathbb{P}$, and then

$$\begin{array}{ccc} D_{t-}^{(n_{k_i})}\theta_1^{(n_{k_i})}(t,x) & \overset{\lambda\times\nu\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} & D_{t-}\theta_1(t,x), \\ D_{t-}^{(n_{k_i})} & \overset{\lambda\times\nu\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} & D_{t-}, \end{array}$$

and

$$D_t^{(n_{k_i})} \overset{\lambda\times\nu\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} D_t. \tag{4.7}$$

Finally, noting that $\inf D_t > 0$ $\mathbb{Q}$-a.s.

$$\theta_1^{(n_{k_i})}(t,x) \overset{\lambda\times\nu\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} \theta_1(t,x). \tag{4.8}$$

The first assertion in (4.6) can be proved using essentially the same kind of ideas used above for the proof of the second part, concluding that for an arbitrary but fixed subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$, there is a sub-subsequence $\{n_{k_i}\}_{i\in\mathbb{N}} \subset \mathbb{N}$ such that

$$\left\{D_t^{(n_{k_i})}\right\}_{t\in[0,T]} \overset{\lambda\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} \{D_t\}_{t\in[0,T]} \tag{4.9}$$

and

$$\left\{\theta_0^{(n_{k_i})}(t)\right\}_{t\in[0,T]} \overset{\lambda\times\mathbb{Q}\text{-a.s.}}{\longrightarrow} \{\theta_0(t)\}_{t\in[0,T]}. \tag{4.10}$$

We are now ready to finish the proof of the theorem, observing that

$$\begin{aligned} \liminf_{n\to\infty}\vartheta\left(\mathbb{Q}^{(n)}\right) = \liminf_{n\to\infty} \int_{\Omega\times[0,T]} & \left\{h\left(h_0\left(\theta_0^{(n)}(t)\right) + \int_{\mathbb{R}_0}\delta(t,x)\right.\right. \\ & \left.\left. \times h_1\left(\theta_1^{(n)}(t,x)\right)\nu(dx)\right)\right\}\frac{D_t^{(n)}}{D_t}d(\lambda\times\mathbb{Q}). \end{aligned}$$

Let $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ be a subsequence for which the limit inferior is realized. Using (4.7)–(4.10) we can pass to a sub-subsequence $\{n_{k_i}\}_{i\in\mathbb{N}} \subset \mathbb{N}$ and, from the continuity of h, h_0 and h_1, it follows

$$\begin{aligned}
&\liminf_{n\to\infty} \vartheta\left(\mathbb{Q}^{(n)}\right)\\
&\geq \int\limits_{\Omega\times[0,T]} \liminf_{i\to\infty} \left(\left\{h\left(h_0\left(\theta_0^{(n_{k_i})}(t)\right) + \int\limits_{\mathbb{R}_0} \delta(t,x)\, h_1\left(\theta_1^{(n_{k_i})}(t,x)\right)\nu(dx)\right)\right\} \frac{D_t^{(n_{k_i})}}{D_t}\right) d(\lambda\times\mathbb{Q})\\
&\geq \int\limits_{\Omega\times[0,T]} h\left(h_0(\theta_0(t)) + \int_{\mathbb{R}_0} h_1(\theta_1(t,x))\,\nu(dx)\right) d(\lambda\times\mathbb{Q})\\
&= \vartheta(\mathbb{Q}).
\end{aligned}$$

□

5 The Market Model: General Description and Martingale Measures

Let us now consider the stochastic process Y_t with dynamics given by

$$Y_t := \int\limits_{]0,t]} \alpha_s ds + \int\limits_{]0,t]} \beta_s dW_s + \int\limits_{]0,t]\times\mathbb{R}_0} \gamma(s,x)\left(\mu(ds,dx) - \nu(dx)\; ds\right), \qquad (5.1)$$

where α is an adapted process with left continuous paths (càg), β is càdlàg with $\beta \in \mathcal{L}(W)$, and $\gamma \in \mathcal{G}(\mu)$. Throughout we assume that the coefficients α, β and γ fulfill the following conditions:

$$\begin{aligned}
&(A\ 1)\ 0 < c \leq |\beta_t| \quad \forall t \in \mathbb{R}_+ \quad \mathbb{P}\text{-}a.s..\\
&(A\ 2)\ \textstyle\int_0^T \left(\frac{\alpha_u}{\beta_u}\right)^2 du \in \mathfrak{M}_b \text{ i.e. bounded.}\\
&(A\ 3)\ \gamma(t,\triangle L_t) \times \mathbf{1}_{\mathbb{R}_0}(\triangle L_t) \geq -1 \quad \forall t \in \mathbb{R}_+ \quad \mathbb{P}\text{-}a.s..\\
&(A\ 4)\ \left\{\gamma(t,\triangle L_t)\,\mathbf{1}_{\mathbb{R}_0}(\triangle L_t)\right\}_{t\in\mathbb{R}_+} \text{ is a locally bounded process.}
\end{aligned} \qquad (5.2)$$

The market model consists of two assets, one of them is the numéraire, having a strictly positive price. The dynamics of the other risky asset will be modeled as a function of the process Y_t defined above. More specifically, since we are interested in the analysis of problem of robust utility maximization, presented in the next section, the discounted capital process can be written in terms of the wealth invested in this asset, and hence the problem can be written using only the dynamics of the discounted price of this asset. For this reason, throughout we will be concentrated in the dynamics of this price.

The dynamic of the discounted price process S is determined by the process Y as its Doleans-Dade exponential

$$S_t = S_0 \mathcal{E}(Y_t), \tag{5.3}$$

where $\mathcal{E}$ represents the Doleans-Dade exponential of a semimartingale; condition $(A\ 3)$ ensures that the price process is non-negative. This process is an exponential semimartingale if and only if the following two conditions are fulfilled:

$$\begin{aligned} &(i)\ \ S = S\mathbf{1}_{[\![0,\tau[\![}, \text{ for } \tau := \inf\{t > 0 : S_t = 0 \text{ or } S_{t-} = 0\}, \\ &(ii)\ \ \tfrac{1}{S_{t-}}\mathbf{1}_{[\![S_- \neq 0]\!]} \text{ is integrable w.r.t. } S, \end{aligned} \tag{5.4}$$

where $[\![S_- \neq 0]\!] := \{(\omega, t) \in \Omega \times \mathbb{R}_+ : S_{t-}(\omega) \neq 0\}$. The first property in (5.4) is conceptually very appropriate when we are interested in modelling the dynamics of a price process. Recall that a stochastically continuous semimartingale has independent increments if and only if its predictable triplet is non-random. Therefore, in general, the price process S is not a Lévy exponential model, because $[Y^c]_t = \int_0^t (\beta_u)^2\, du$ does not need to be deterministic. However, observe that the price dynamics (5.3) includes Lévy exponential models, for Lévy processes with $|\triangle L_t| \leq 1$.

For the model (5.3) the price process can be written explicitly as

$$\begin{aligned} S_t = S_0 \exp\Bigg\{ &\int_{]0,t]} \alpha_s ds + \int_{]0,t]} \beta_s dW_s + \int_{]0,t]\times\mathbb{R}_0} \gamma(s,x)\left(\mu(ds,dx) - \nu(dx)\ ds\right) \\ &-\tfrac{1}{2}\int_{]0,t]} (\beta_s)^2\, ds \Bigg\} \\ &\times \exp\Bigg\{ \int_{]0,t]\times\mathbb{R}_0} \{\ln(1+\gamma(s,x)) - \gamma(s,x)\}\mu(ds,dx) \Bigg\}. \end{aligned} \tag{5.5}$$

The predictable cádlág process $\{\pi_t\}_{t\in\mathbb{R}_+}$, satisfying the integrability condition $\int_0^t (\pi_s)^2\, ds < \infty$ $\mathbb{P}$-a.s. for all $t \in \mathbb{R}_+$, shall denote the proportion of wealth at time t invested in the risky asset S. For an initial capital x, the discounted wealth $X_t^{x,\pi}$ associated with a self-financing investment strategy (x, π) fulfills the equation

$$X_t^{x,\pi} = x + \int_0^t \frac{X_{u-}^{x,\pi}\pi_u}{S_{u-}} \mathbf{1}_{[\![S_- \neq 0]\!]} dS_u. \tag{5.6}$$

We say that a self-financing strategy (x, π) is *admissible* if the wealth process $X_t^{x,\pi} > 0$ for all $t > 0$. The class of admissible wealth processes with initial wealth less than or equal to x is denoted by $\mathcal{X}(x)$. In what follows we restrict ourself to the time interval $[0, T]$, for some $T > 0$ fixed, and take $\mathcal{F} = \mathcal{F}_T$.

Let us recall briefly the notation introduced in Sect. 3.1. Denote by $\mathcal{Q}_{\ll}(\mathbb{P})$ the subclass of absolutely continuous probability measures with respect to $\mathbb{P}$ and by $\mathcal{Q}_{\approx}(\mathbb{P})$ the subclass of equivalent probability measures. The corresponding classes of density processes associated to $\mathcal{Q}_{\ll}(\mathbb{P})$ and $\mathcal{Q}_{\approx}(\mathbb{P})$ are denoted by $\mathcal{D}_{\ll}(\mathbb{P})$ and $\mathcal{D}_{\approx}(\mathbb{P})$, respectively. The processes in the class $\mathcal{D}_{\ll}(\mathbb{P})$ are of the form

$$\begin{aligned} D_t = \exp & \Bigg\{ \int_{]0,t]} \theta_0 dW + \int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,(\mu(ds,dx) - \nu(dx)\,ds) \\ & -\tfrac{1}{2}\int_{]0,t]} (\theta_0)^2\, ds \Bigg\} \times \\ & \times \exp\Bigg\{ \int_{]0,t]\times\mathbb{R}_0} \{\ln(1+\theta_1(s,x)) - \theta_1(s,x)\}\,\mu(ds,dx) \Bigg\}, \end{aligned} \tag{5.7}$$

for $\theta_0 \in \mathcal{L}(W)$ and $\theta_1 \in \mathcal{G}(\mu)$. If $\int \theta_1(s,x)\,\mu(ds,dx) \in \mathcal{A}_{loc}(\mathbb{P})$ the previous formula can be written as

$$\begin{aligned} D_t = \exp & \Bigg\{ \int_{]0,t]} \theta_0 dW - \frac{1}{2}\int_{]0,t]} (\theta_0(s))^2\, ds \\ & + \int_{]0,t]\times\mathbb{R}_0} \ln(1+\theta_1(s,x))\,\mu(ds,dx) - \int_{]0,t]\times\mathbb{R}_0} \theta_1(s,x)\,\nu(dx)\,ds \Bigg\}. \end{aligned} \tag{5.8}$$

Next result characterizes the class of *equivalent local martingale measures* defined as

$$\mathcal{Q}_{elmm} \equiv \{\mathbb{Q} \in \mathcal{Q}_{\approx}(\mathbb{P}) : \mathcal{X}(1) \subset \mathcal{M}_{loc}(\mathbb{Q})\} = \{\mathbb{Q} \in \mathcal{Q}_{\approx}(\mathbb{P}) : S \in \mathcal{M}_{loc}(\mathbb{Q})\}. \tag{5.9}$$

Observe that $(A\,4)$ is a necessary and sufficient condition for S to be a locally bounded process. This property is crucial in order to obtain the former equality in (5.9). The class of density processes associated with $\mathcal{Q}_{elmm}$ is denoted by $\mathcal{D}_{elmm}(\mathbb{P})$. Kunita [16] gave conditions on the parameters (θ_0, θ_1) of a measure $\mathbb{Q} \in \mathcal{Q}_{\approx}$ in order that it is a local martingale measure for a Lévy exponential model i.e. when $S = \mathcal{E}(L)$. Observe that in this case $\mathcal{Q}_{elmm}(S) = \mathcal{Q}_{elmm}(L)$. Next proposition extends this result, giving conditions on the parameters (θ_0, θ_1) under which an equivalent measure is a local martingale measure for the price model (5.3).

Proposition 5.1 *Given* $\mathbb{Q} \in \mathcal{Q}_{\approx}$, *let* $\theta_0 \in \mathcal{L}(W)$ *and* $\theta_1 \in \mathcal{G}(\mu)$ *be the corresponding processes describing the density processes found in Lemma 3.2.*

Then, the following equivalence holds:

$$\mathbb{Q} \in \mathcal{Q}_{elmm} \Longleftrightarrow \alpha_t + \beta_t \theta_0 (t) + \int_{\mathbb{R}_0} \gamma (t, x) \theta_1 (t, x) \nu (dx) = 0 \; \forall t \geq 0 \quad \mathbb{P}\text{-}a.s. \tag{5.10}$$

6 Robust Utility Maximization

The goal of the economic agent, with an initial capital $x > 0$, will be now to maximize the penalized expected utility from a terminal wealth in the worst case model. Given a penalty function ϑ, this means that the agent seeks to solve the associated robust expected utility problem with value function

$$u (x) := \sup_{X \in \mathcal{X}(x)} \inf_{\mathbb{Q} \in \mathcal{Q}^{\vartheta}_{\ll}(\mathbb{P})} \left\{ \mathbb{E}_{\mathbb{Q}} [U (X_T)] + \vartheta (\mathbb{Q}) \right\}, \tag{6.1}$$

where $\mathcal{Q}^{\vartheta}_{\ll} := \{\mathbb{Q} \ll \mathbb{P} : \vartheta (\mathbb{Q}) < \infty\}$ for a fixed reference measure $\mathbb{P}$; see (4.2). A utility function $U : (0, \infty) \longrightarrow \mathbb{R}$ will be hereafter a strictly increasing, strictly concave, continuously differentiable real function, which satisfies the Inada conditions, namely $U' (0+) = +\infty$ and $U' (\infty-) = 0$.

The Fenchel-Legendre transformation of the function $-U (-x)$ is defined by

$$V (y) = \sup_{x>0} \{U (x) - xy\}, \qquad y > 0. \tag{6.2}$$

This function V is continuously differentiable, decreasing, and strictly convex, satisfying: $V' (0+) = -\infty$, $V' (\infty) = 0$, $V (0+) = U (\infty)$, $V (\infty) = U (0+)$. Further, the biconjugate of U is again U itself, i.e.

$$U (x) = \inf_{y>0} \{V (y) + xy\}, \qquad x > 0.$$

For a fixed prior measure $\mathbb{Q}$, in Kramkov and Schachermayer [13] the dual problem was formulated in terms of the value function

$$v_{\mathbb{Q}} (y) := \inf_{Y \in \mathcal{Y}_{\mathbb{Q}}(y)} \left\{ \mathbb{E}_{\mathbb{Q}} [V (Y_T)] \right\}, \tag{6.3}$$

where

$$\mathcal{Y}_{\mathbb{Q}} (y) := \{Y \geq 0 : \; Y_0 = y, YX \; \mathbb{Q}\text{-supermartingale} \; \forall X \in \mathcal{X} (1)\}. \tag{6.4}$$

A similar problem was studied in [11] for diffusion processes and the logarithmic utility function.

Remark 6.1 To guarantee that the $\mathbb{Q}$-expectations in (6.1) and (6.3) are well defined, we extend the operator $\mathbb{E}_{\mathbb{Q}}[U(\cdot)]$ to $\mathcal{L}^0$, as in Schied [19, p. 111], in the following way

$$\mathbb{E}_{\mathbb{Q}}[X] := \sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbb{Q}}[X \wedge n] = \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}}[X \wedge n] \quad X \in \mathcal{L}^0(\Omega, \mathcal{F}). \tag{6.5}$$

The corresponding dual value function, in the robust setting, is defined by

$$v(y) := \inf_{\mathbb{Q}\in\mathcal{Q}^{\vartheta}_{\ll}} \left\{ v_{\mathbb{Q}}(y) + \vartheta(\mathbb{Q}) \right\}. \tag{6.6}$$

In the rest of this section the connection between the penalty functions (4.1) and the existence of solutions to the penalized robust expected utility problem (6.1) is established. The first step in this direction is to notice that given Theorem 4.1, where the minimality of the penalty function was proved, it is possible to write the primal problem (6.1) as

$$u(x) = \inf_{\mathbb{Q}\in\mathcal{Q}^{\vartheta}_{\ll}(\mathbb{P})} \sup_{X\in\mathcal{X}(x)} \left\{ \mathbb{E}_{\mathbb{Q}}[U(X_T)] + \vartheta(\mathbb{Q}) \right\}.$$

See Schied [19, Theorem 2.3]. Then, based on the duality theory for solving the classical optimal investment problem, the dual problem (6.6) is solved using the analogous sufficient conditions introduced by Kramkov and Schachermayer [13]. More precisely, for the class of utility functions described at the beginning of this section, when

$$v_{\mathbb{Q}}(y) < \infty \quad \text{for all } \mathbb{Q} \in \mathcal{Q}^{\vartheta}_{\approx} \quad \text{and} \quad y > 0, \tag{6.7}$$

where $\mathcal{Q}^{\vartheta}_{\approx} := \{\mathbb{Q} \approx \mathbb{P} : \vartheta(\mathbb{Q}) < \infty\}$ and ϑ is the minimal penalty function of the associated convex measure of risk, we are able to conclude that there exists an optimal solution to the dual problem (6.6), from which we can obtain an optimal solution to (6.1), using Schied [19, Theorems 2.3 and 2.5]. For the proof of the main result of this section, namely Theorem 6.1, we shall verify that these sufficient conditions are satisfied.

6.1 Penalties and Solvability

Let us now introduce the class

$$\mathcal{C} := \left\{ \mathcal{E}\left(Z^{\xi}\right) : \begin{array}{l} \xi := \left(\xi^{(0)}, \xi^{(1)}\right),\ \xi^{(0)} \in \mathcal{L}(W),\ \xi^{(1)} \in \mathcal{G}(\mu),\ \text{with} \\ \alpha_t + \beta_t \xi^{(0)}_t + \int_{\mathbb{R}_0} \gamma(t,x)\, \xi^{(1)}(t,x)\, \nu(dx) = 0 \text{ Lebesgue } \forall t \end{array} \right\}, \tag{6.8}$$

where

$$Z_t^{\xi} := \int\limits_{]0,t]} \xi^{(0)} dW + \int\limits_{]0,t]\times\mathbb{R}_0} \xi^{(1)}(s,x)\left(\mu(ds,dx) - ds\,\nu(dx)\right).$$

Observe that $\mathcal{D}_{elmm}(\mathbb{P}) \subset \mathcal{C} \subset \mathcal{Y}_{\mathbb{P}}(1)$; see (6.4) for the definition of $\mathcal{Y}_{\mathbb{P}}(1)$. This relation between these three sets plays a crucial role in the formulation of the dual problem, even in the non-robust case.

Theorem 6.1 *For $q \in (-\infty, 1)\setminus\{0\}$, let $U(x) := \frac{1}{q}x^q$ be the power utility function, and consider the functions h, h_0 and h_1 as in Sect. 4, satisfying the following conditions:*

$$\begin{aligned}&h(x) \geq \exp\left(\kappa_1 x^2\right) - 1 \text{ where } \kappa_1 := 1 \vee 2\left(2p^2 + p\right)T \text{ and } p := \tfrac{q}{1-q},\\ &h_0(x) \geq |x|,\\ &h_1(x) \geq \tfrac{|x|}{c}, \text{for } c \text{ as in assumption } (A\ 1).\end{aligned}$$

Then, for the penalty function

$$\vartheta_{x^q}(\mathbb{Q}) := \mathbb{E}_{\mathbb{Q}}\left[\int_0^T h\left(h_0(\theta_0(t)) + \int_{\mathbb{R}_0} |\gamma(t,x)|\, h_1(\theta_1(t,x))\,\nu(dx)\right)dt\right],$$

the penalized robust utility maximization problem (6.1) has a solution.

Proof The penalty function ϑ_{x^q} is bounded from below, and by Theorem 4.1 equals on $\mathcal{Q}_{\ll}(\mathbb{P})$ the minimal penalty function of the normalized and sensitive convex measure of risk defined in (4.3). Therefore, we only need to prove that condition (6.7) holds. In order to prove that, fix an arbitrary probability measure $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{x^q}} = \{\mathbb{Q} \approx \mathbb{P} : \vartheta_{x^q}(\mathbb{Q}) < \infty\}$ and let $\theta = (\theta_0, \theta_1)$ be the corresponding coefficients obtained in Lemma 3.2.

(1) In Lemma 4.2, Schied [19] establishes that even for $\mathbb{Q} \in \mathcal{Q}_{\ll}$, with density process D, the next equivalence holds

$$Y \in \mathcal{Y}_{\mathbb{Q}}(y) \Leftrightarrow YD \in \mathcal{Y}_{\mathbb{P}}(y).$$

Therefore, for $\mathbb{Q} \in \mathcal{Q}_{\ll}^{\vartheta_{x^q}}$, with coefficient $\theta = (\theta_0, \theta_1)$, it follows that

$$v_{\mathbb{Q}}(y) = \inf_{Y\in\mathcal{Y}_{\mathbb{P}}(1)}\left\{\mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{Y_T}{D_T^{\mathbb{Q}}}\right)\right]\right\} \leq \inf_{\xi\in\mathcal{C}}\left\{\mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}\left(Z^{\xi}\right)_T}{\mathcal{E}\left(Z^{\theta}\right)_T}\right)\right]\right\}.$$

(2) Define

$$\varepsilon_t := \alpha_t + \beta_t \theta_0(t) + \int_{\mathbb{R}_0} \gamma(t,x)\,\theta_1(t,x)\,\nu(dx),$$

the process involved in the definition of the class $\mathcal{C}$ in (6.8).

When ε_t is identically zero for all $t > 0$, Proposition 5.1 implies that $\mathbb{Q} \in \mathcal{Q}_{elmm}$. However, for $\mathbb{Q} \in \mathcal{Q}_{elmm}$ the constant process $Y \equiv y$ belongs to $\mathcal{Y}_{\mathbb{Q}}(y)$, and it follows that $v_{\mathbb{Q}}(y) < \infty$, for all $y > 0$. In this case the proof is concluded.

If ε is not identically zero, consider $\xi_t^{(0)} := \theta_0(t) - \frac{\varepsilon_t}{\beta_t}$ and $\xi^{(1)} := \theta_1$. Since

$$\left\{ \frac{1}{\beta_t} \int_{\mathbb{R}_0} \gamma(t,x)\,\theta_1(t,x)\,\nu(dx) \right\}_{t\in[0,T]} \in \mathcal{P}$$

and

$$\infty > \vartheta_{x^q}(\mathbb{Q}) \geq \mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left(\frac{1}{\beta_t} \int_{\mathbb{R}_0} \gamma(t,x)\,\theta_1(t,x)\,\nu(dx) \right)^2 dt \right] - T,$$

it follows that $\left\{ \frac{1}{\beta_t} \int_{\mathbb{R}_0} \gamma(t,x)\,\theta_1(t,x)\,\nu(dx) \right\}_{t\in[0,T]} \in \mathcal{L}\left(W'\right)$ for W' a $\mathbb{Q}$-Wiener process and thus also $\xi^{(0)} \in \mathcal{L}\left(W'\right)$. Moreover, for $\xi = \left(\xi^{(0)}, \xi^{(1)}\right)$ we have that $\mathcal{E}\left(Z^{\xi}\right) \in \mathcal{C}$.

Using Girsanov's theorem, we obtain further

$$\frac{\mathcal{E}\left(Z^{\xi}\right)_t}{\mathcal{E}\left(Z^{\theta}\right)_t} = \exp\left\{ \int_{]0,t]} \left(-\frac{\varepsilon_u}{\beta_u} \right) dW'_u - \frac{1}{2} \int_{]0,t]} \left(\frac{\varepsilon_u}{\beta_u} \right)^2 du \right\}.$$

(3) The Cauchy-Bunyakovsky-Schwarz inequality yields

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[V\left(y \frac{\mathcal{E}(Z^{\xi})_T}{\mathcal{E}(Z^{\theta})_T} \right) \right] \\
&\quad = \tfrac{1}{p} y^{-p} \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{ p \int_{]0,T]} \left(\tfrac{\varepsilon_t}{\beta_t} \right) dW' + \tfrac{p}{2} \int_{]0,T]} \left(\tfrac{\varepsilon_t}{\beta_t} \right)^2 dt \right\} \right] \\
&\quad \leq \tfrac{1}{p} y^{-p} \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{ 2p \int_{]0,T]} \left(\tfrac{\varepsilon_t}{\beta_t} \right) dW' - \tfrac{4p^2}{2} \int_{]0,T]} \left(\tfrac{\varepsilon_t}{\beta_t} \right)^2 dt \right\} \right]^{\frac{1}{2}} \\
&\qquad \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{ \left(\tfrac{4p^2}{2} + p \right) \int_{]0,T]} \left(\tfrac{\varepsilon_t}{\beta_t} \right)^2 dt \right\} \right]^{\frac{1}{2}}.
\end{aligned} \tag{6.9}$$

On the other hand, the process

$$\exp\left\{2p\int\limits_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)dW' - \frac{4p^2}{2}\int\limits_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)^2 dt\right\} \in \mathcal{M}_{loc}(\mathbb{Q})$$

is a local $\mathbb{Q}$-martingale and, since it is positive, is a supermartingale. Hence,

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{2p\int\limits_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)dW' - \frac{4p^2}{2}\int\limits_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)^2 dt\right\}\right] \leq 1.$$

Therefore we need only to take care about $\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\left(\frac{4p^2}{2}+p\right)\int_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)^2 dt\right\}\right]$ in order to have the desired integrability. From assumption $(A\ 2)$ we have

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\left(2p^2+p\right)2\int_{]0,T]}\left(\left|\frac{\alpha_t}{\beta_t}\right|\right)^2 dt\right\}\right] < C,$$

and thus

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\left(\frac{4p^2}{2}+p\right)\int_{]0,T]}\left(\frac{\varepsilon_t}{\beta_t}\right)^2 dt\right\}\right] \leq C\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{2\left(2p^2+p\right)\right.\right.$$

$$\left.\left.\times\int\limits_0^T\left(|\theta_0(t)| + \frac{1}{|\beta_t|}\left|\int_{\mathbb{R}_0}\gamma(t,x)\,\theta_1(t,x)\,\nu(dx)\right|\right)^2 dt\right\}\right].$$

Finally, observe that for $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{x^q}}$, using that it has finite penalization $\vartheta_{x^q}(\mathbb{Q}) < \infty$ and Jensen's inequality, we have

$$\infty > \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\frac{\kappa_1}{T}\int\limits_0^T\left(h_0(\theta_0(t)) + \int\limits_{\mathbb{R}_0}|\gamma(t,x)|\,h_1(\theta_1(t,x))\,\nu(dx)\right)^2 dt\right\}\right]$$

$$\geq \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{2\left(2p^2+p\right)\int\limits_0^T\left(|\theta_0(t)| + \frac{1}{|\beta_t|}\left|\int_{\mathbb{R}_0}\gamma(t,x)\right.\right.\right.\right.$$

$$\left.\left.\left.\left.\theta_1(t,x)\,\nu(dx)\right|\right)^2 dt\right\}\right].$$

From the last two displays it follows that the r.h.s. of (6.9) is finite and the theorem follows. □

Next theorem establishes a sufficient condition for the existence of solution to the robust utility maximization problem (6.1) for an arbitrary utility function.

Theorem 6.2 *Suppose that the utility function $\widetilde{U}$ is bounded above by a power utility U, with penalty function ϑ_{x^q} associated to U as in Theorem 6.1. Then, the robust utility maximization problem (6.1) for $\widetilde{U}$ with penalty ϑ_{x^q} has an optimal solution.*

Proof Since $U(x) := \frac{1}{q}x^{-q} \geq \widetilde{U}(x)$ for all $x > 0$, for some $q \in (-\infty, 1) \setminus \{0\}$ the corresponding convex conjugate functions satisfy $V(y) \geq \widetilde{V}(y)$ for each $y > 0$. As it was pointed out in Remark 6.2, we can restrict ourself to the positive part $\widetilde{V}^+(y)$. From Proposition 6.1, we can fix some $Y \in \mathcal{Y}_{\mathbb{Q}}(y)$ such that $\mathbb{E}_{\mathbb{Q}}[V(Y_T)] < \infty$ for any $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{x^q}}$ and $y > 0$, arbitrary, but fixed. Furthermore, the inequality $V(y) \geq \widetilde{V}(y)$ implies that their inverse functions satisfy $\left(V^+\right)^{(-1)}(n) \geq \left(\widetilde{V}^+\right)^{(-1)}(n)$ for all $n \in \mathbb{N}$, and hence

$$\sum_{n=1}^{\infty} \mathbb{Q}\left[Y_T \leq \left(\widetilde{V}^+\right)^{(-1)}(n)\right] \leq \sum_{n=1}^{\infty} \mathbb{Q}\left[Y_T \leq \left(V^+\right)^{(-1)}(n)\right] < \infty.$$

The Moments Lemma ($\mathbb{E}_{\mathbb{Q}}[|X|] < \infty \Leftrightarrow \sum_{n=1}^{\infty} \mathbb{Q}[|X| \geq n] < \infty$) yields $\mathbb{E}_{\mathbb{Q}}\left[\widetilde{V}^+(Y_T)\right] < \infty$, and the assertion follows. □

From the proof of Theorem 6.2 it is clear that the behavior of the convex conjugate function in a neighborhood of zero is fundamental. From this observation we conclude the following.

Corollary 6.1 *Let U be a utility function with convex conjugate V, and ϑ a penalization function such that the robust utility maximization problem (6.1) has a solution. For a utility function $\widetilde{U}$ such that their convex conjugate function $\widetilde{V}$ is majorized in an ε-neighborhood of zero by V, the corresponding utility maximization problem (6.1) has a solution.*

Remark 6.2 When the conjugate convex function V is bounded from above it follows immediately that the penalized robust utility maximization problem (6.1) has a solution for any proper penalty function ϑ. This is the case, for instance, of the power utility function $U(x) := \frac{1}{q}x^q$, for $q \in (-\infty, 0)$, where the convex conjugate function $V(x) = \frac{1}{p}x^{-p} \leq 0$, with $p := \frac{q}{1-q}$.

Next we give an alternative representation of the robust dual value function, introduced in (6.6), in terms of the family $\mathcal{C}$ of stochastic processes.

Theorem 6.3 *For a utility function U satisfying condition (6.7), the dual value function can be written as*

$$\begin{aligned} v(y) &= \inf_{\mathbb{Q}\in\mathcal{Q}_{\approx}^{\vartheta}} \left\{ \inf_{\xi\in\mathcal{C}} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}(Z^{\xi})_T}{D_T^{\mathbb{Q}}}\right)\right]\right\} + \vartheta(\mathbb{Q})\right\} \\ &= \inf_{\mathbb{Q}\in\mathcal{Q}_{\ll}} \left\{ \inf_{\xi\in\mathcal{C}} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}(Z^{\xi})_T}{D_T^{\mathbb{Q}}}\right)\right]\right\} + \vartheta(\mathbb{Q})\right\}. \end{aligned} \tag{6.10}$$

Proof Condition (6.7), together with Lemma 4.4 in [19] and Theorem 2 in [14] , imply the following identity

$$v(y) = \inf_{\mathbb{Q}\in\mathcal{Q}_{\approx}^{\vartheta}} \left\{ \inf_{\widetilde{\mathbb{Q}}\in\mathcal{Q}_{elmm}(\mathbb{Q})} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y d\widetilde{\mathbb{Q}}/d\mathbb{Q}\right)\right]\right\} + \vartheta(\mathbb{Q})\right\}.$$

Since $\mathcal{D}_{elmm}(\mathbb{P}) \subset \mathcal{C}$, we get

$$\begin{aligned} v(y) &\geq \inf_{\mathbb{Q}\in\mathcal{Q}_{\approx}^{\vartheta}} \left\{ \inf_{\xi\in\mathcal{C}} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}(Z^{\xi})_T}{D_T^{\mathbb{Q}}}\right)\right]\right\} + \vartheta(\mathbb{Q})\right\} \\ &\geq \inf_{\mathbb{Q}\in\mathcal{Q}_{\ll}} \left\{ \inf_{\xi\in\mathcal{C}} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}(Z^{\xi})_T}{D_T^{\mathbb{Q}}}\right)\right]\right\} + \vartheta(\mathbb{Q})\right\}. \end{aligned} \tag{6.11}$$

Finally, from Lemma 4.2 in Schied [19] and $\mathcal{C} \subset \mathcal{Y}_{\mathbb{P}}(1)$ follows

$$v_{\mathbb{Q}}(y) \leq \inf_{\xi\in\mathcal{C}} \left\{ \mathbb{E}_{\mathbb{Q}}\left[V\left(y\frac{\mathcal{E}\left(Z^{\xi}\right)_T}{D_T^{\mathbb{Q}}}\right)\right]\right\},$$

and we have the inequalities (6.11) in the other direction, and the result follows. □

6.2 The Logarithmic Utility Case

The existence of solution to the robust problem for the logarithmic utility function $U(x) = \log(x)$ can be obtain using the relation between this utility function and the relative entropy function. Let h, h_0 and h_1 be as in Sect. 4, satisfying also the following growth conditions:

$$\begin{aligned} &h(x) \geq x, \\ &h_0(x) \geq \frac{1}{2}x^2, \\ &h_1(x) \geq \{|x| \vee x\ln(1+x)\}\,\mathbf{1}_{(-1,0)}(x) + x(1+x)\,\mathbf{1}_{\mathbb{R}_+}(x). \end{aligned}$$

Now, define the penalization function

$$\vartheta_{\log}(\mathbb{Q}) := \mathbb{E}_{\mathbb{Q}}\left[\int_0^T h\left(h_0(\theta_0(t)) + \int_{\mathbb{R}_0} h_1(\theta_1(t,x))\,\nu(dx)\right)dt\right]\mathbf{1}_{\mathcal{Q}_{\ll}}(\mathbb{Q}) \\ +\infty \times \mathbf{1}_{\mathcal{Q}_{cont}\setminus\mathcal{Q}_{\ll}}(\mathbb{Q}). \tag{6.12}$$

Remark 6.3 Notice that when $\mathbb{Q} \in \mathcal{Q}_{\ll}^{\vartheta_{\log}}(\mathbb{P})$ with coefficient $\theta = (\theta_0, \theta_1)$ has a finite penalization, the following $\mathbb{Q}$-integrability properties hold:

(6.3.i) $\int_{[0,T]\times\mathbb{R}_0} \theta_1(t,x)\,\mu_{\mathbb{P}}^{\mathcal{P}}(dt,dx) \in \mathcal{L}^1(\mathbb{Q})$

(6.3.ii) $\int_{[0,T]\times\mathbb{R}_0} \{1+\theta_1(t,x)\}\ln(1+\theta_1(t,x))\,\mu_{\mathbb{P}}^{\mathcal{P}}(dt,dx) \in \mathcal{L}^1(\mathbb{Q})$

(6.3.iii) $\int_{[0,T]\times\mathbb{R}_0} \ln(1+\theta_1(s,x))\,\mu(ds,dx) \in \mathcal{L}^1(\mathbb{Q})$

(6.3.iv) $\mathbb{E}_{\mathbb{Q}}\left[\int_{]0,T]\times\mathbb{R}_0} \ln(1+\theta_1)\,d\mu\right] = \mathbb{E}_{\mathbb{Q}}\left[\int_{]0,T]\times\mathbb{R}_0} \{\ln(1+\theta_1)\}(1+\theta_1)\,d\mu_{\mathbb{P}}^{\mathcal{P}}\right]$

In addition, for $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{\log}}(\mathbb{P})$ we have

(6.3.v) $\int_{[0,T]\times\mathbb{R}_0} \theta_1(s,x)\,\mu(ds,dx) < \infty\ \mathbb{P}-a.s.$

For $\mathbb{Q} \in \mathcal{Q}_{\ll}(\mathbb{P})$, the relative entropy function is defined as

$$H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}\left[D_T^{\mathbb{Q}} \log\left(D_T^{\mathbb{Q}}\right)\right].$$

Lemma 6.1 *Given $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{\log}}(\mathbb{P})$, it follows that*

$$H(\mathbb{Q}\,|\mathbb{P}) \le \vartheta_{\log}(\mathbb{Q}).$$

Proof For $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{\log}}(\mathbb{P})$ we have that θ_0 is integrable w.r.t. W' a $\mathbb{Q}$-Wiener process as an square integrable martingale. Further Remark 6.3 implies that

$$H(\mathbb{Q}\,|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\int_0^T (\theta_0)^2\,ds + \int_{]0,T]\times\mathbb{R}_0} \ln(1+\theta_1(s,x))\,\mu(ds,dx) \right. \\ \left. -\int_0^T\int_{\mathbb{R}_0} \theta_1(s,x)\,\nu(dx)\,ds\right]$$

$$\leq \mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left\{\frac{1}{2}(\theta_0)^2 ds + \int_{\mathbb{R}_0} \{\ln(1+\theta_1(s,x))\}\theta_1(s,x)\,\nu(dx)\right\} ds\right]$$
$$\leq \vartheta_{\log}(\mathbb{Q}).$$

□

Using the previous result, the existence of solution to the primal problem (6.1) can be concluded.

Proposition 6.1 *Let $U(x) = \log(x)$ and $\vartheta_{\log}$ be as in (6.12). Then the robust utility maximization problem (6.1) has an optimal solution.*

Proof Again, we only need to verify that condition (6.7) holds. Observe that, for each $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{\log}}(\mathbb{P})$, we have that

$$v_{\mathbb{Q}}(y) \leq \inf_{\xi\in\mathcal{C}}\left\{\mathbb{E}\left[D_T^{\mathbb{Q}}\log\left(\frac{D_T^{\mathbb{Q}}}{\mathcal{E}(Z^{\xi})_T}\right) - \log(y) - 1\right]\right\}.$$

Also, Proposition 5.1 and the Novikov condition yield for $\widetilde{\xi} \in \mathcal{C}$, with $\widetilde{\xi}^{(0)} := -\frac{\alpha_s}{\beta_s}$ and $\widetilde{\xi}^{(1)} := 0$, that $\widetilde{\mathbb{Q}} \in \mathcal{Q}_{elmm}$, where $d\widetilde{\mathbb{Q}}\backslash d\mathbb{P} = D_T^{\widetilde{\xi}} := \mathcal{E}\left(Z^{\widetilde{\xi}}\right)_T$. Further, from Lemma 6.1 we conclude for $\mathbb{Q} \in \mathcal{Q}_{\approx}^{\vartheta_{\log}}(\mathbb{P})$ that

$$\mathbb{E}\left[D_T^{\mathbb{Q}}\log\left(\frac{D_T^{\mathbb{Q}}}{D_T^{\widetilde{\xi}}}\right)\right] = H(\mathbb{Q}\,|\mathbb{P}) + \mathbb{E}_{\mathbb{Q}}\left[\int_0^T \frac{\alpha_s}{\beta_s}\theta_s^{(0)} ds + \frac{1}{2}\int_0^T \left(\frac{\alpha_s}{\beta_s}\right)^2 ds\right] < \infty$$

and the claim follows. □

References

1. P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Thinking coherently. Risk Mag. **10**, 68–71 (1997)
2. P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk. Math. Finance **9**, 203–228 (1999)
3. F. Delbaen, Coherent risk measures on general probability spaces, in *Advances in Finance and Stochastics, Essays in Honor of Dieter Sondermann*, ed. by K. Sandmann, Ph. Schönbucher (Springer, Berlin, 2002), pp. 1–37
4. H. Föllmer, A. Schied, Convex measures of risk and trading constraints. Finance Stoch. **6**, 429–447 (2002)
5. H. Föllmer, A. Schied, Robust preferences and convex risk measures, in *Advances in Finance and Stochastics, Essays in Honor of Dieter Sondermann* (Springer, Berlin, 2002), pp. 39–56
6. H. Föllmer, A. Schied, *Stochastic Finance. An Introduction in Discrete Time*, 2nd edn. de Gruyter Studies in Mathematics, vol. 27 (de Gruyter, Berlin, 2004)

7. M. Frittelli, E. Rosazza Gianin, Putting order in risk measures. J. Bank. Finance **26**, 1473–1486 (2002)
8. M. Frittelli, E. Rosazza Gianin, Dynamic convex risk measures, in *Risk Measures for the 21st Century*, ed. by G. Szegö (Wiley, New York, 2004), pp. 227–248
9. S.W. He, J.G. Wang, J.A. Yan, *Semimartingale Theory and Stochastic Calculus* (Science Press, Beijing, 1992)
10. D. Heath, Back to the future. Plenary Lecture at the First World Congress of the Bachelier Society, Paris (2000)
11. D. Hernández-Hernández, A. Schied, A control approach to robust utility maximization with logarithmic utility and time consistent penalties. Stoch. Process. Appl. **117**, 980–1000 (2007)
12. J. Jacod, A. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd edn. (Springer, Berlin, 2003)
13. D. Kramkov, W. Schachermayer, The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab. **9**, 904–950 (1999)
14. D. Kramkov, W. Schachermayer, Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. Ann. Appl. Probab. **13**, 1504–1516 (2003)
15. V. Krätschmer, Robust representation of convex risk measures by probability measures. Finance Stoch. **9**, 597–608 (2005)
16. H. Kunita, Representation of martingales with jumps and applications to mathematical finance, in *Stochastic Analysis and Related Topics in Kyoto*, ed. by H. Kunita et al. (Mathematical Society of Japan, Tokyo, 2004), pp. 209–232
17. F. Maccheroni, M. Marinacci, A. Rustichini, Ambiguity aversion, robustness and the variational representation of preferences. Econometrica **74**, 1447–1498 (2006)
18. M.-C. Quenez, Optimal portfolio in a multiple-priors model, in *Seminar on Stochastic Analysis, Random Fields and Applications IV*, ed. by R. Dalang, M. Dozzi, F. Russo. Progress in Probability, vol. 58 (Birkhäuser, Basel, 2004), pp. 291–321
19. A. Schied, Optimal investments for risk- and ambiguity-averse preferences: a duality approach. Finance Stoch. **11**, 107–129 (2007)