The Annals of Applied Probability
2002, Vol. 12, No. 1, 143–172

MINIMIZING SHORTFALL RISK AND APPLICATIONS TO FINANCE AND INSURANCE PROBLEMS

BY HUYÊN PHAM

Université Paris 7

We consider a controlled process governed by $X^{x,\theta} = x + \int \theta \, dS + H^\theta$, where S is a semimartingale, Θ the set of control processes θ is a convex subset of $L(S)$ and $\{H^\theta : \theta \in \Theta\}$ is a concave family of adapted processes with finite variation. We study the problem of minimizing the shortfall risk defined as the expectation of the shortfall $(B - X_T^{x,\theta})_+$ weighted by some loss function, where B is a given nonnegative measurable random variable. Such a criterion has been introduced by Föllmer and Leukert [*Finance Stoch.* **4** (1999) 117–146] motivated by a hedging problem in an incomplete financial market context: $\Theta = L(S)$ and $H^\theta \equiv 0$. Using change of measures and optional decomposition under constraints, we state an existence result to this optimization problem and show some qualitative properties of the associated value function. A verification theorem in terms of a dual control problem is established which is used to obtain a quantitative description of the solution. Finally, we give some applications to hedging problems in constrained portfolios, large investor and reinsurance models.

1. Introduction. In this paper, we study the following stochastic control problem. Let S be an $\mathbb{R}^m$-valued semimartingale on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, P)$. We denote by $L(S)$ the set of $\mathbb{R}^m$-valued predictable processes integrable with respect to S. We prescribe a convex subset Θ of $L(S)$ and a concave family of adapted processes $\{H^\theta : \theta \in \Theta\}$ with finite variation. We consider the controlled process,

$$X_t^{x,\theta} = x + \int_0^t \theta_u \, dS_u + H_t^\theta, \qquad 0 \le t \le T.$$

Given $x \in \mathbb{R}$, a control $\theta \in \Theta$ is said to be admissible, and we denote $\theta \in \mathcal{A}(x)$, if the state process $X^{x,\theta}$ satisfies an arbitrary uniform lower bound and $X_T^{x,\theta} \ge 0$. Given a nonnegative $\mathcal{F}_T$-measurable random variable B, and a loss function l, that is, a nondecreasing and convex function on $\mathbb{R}_+$, we then consider the stochastic optimization problem:

$$\text{minimize } E[l(B - X_T^{x,\theta})_+] \quad \text{overall } \theta \in \mathcal{A}(x).$$

In the case where S is a diffusion process, $\Theta = L(S)$, $H^\theta \equiv 0$, Kulldorf (1993) and Heath (1995) have studied this problem for the loss function $l(x) = \mathbb{1}_{x>0}$

Received September 2000; revised February 2001.

AMS 2000 subject classifications. 93E20, 60G44, 60H05, 60H30, 90A46.

Key words and phrases. Shortfall risk minimization, semimartingales, optional decomposition under constraints, duality theory, finance and insurance.

and for B a constant. Motivated by a hedging problem in an incomplete financial market context, Föllmer and Leukert (1999, 2000) have extended the approach of the last cited authors. In their model, the semimartingale process S is the (discounted) price of risky assets, $\theta \in L(S)$ is an unconstrained portfolio strategy, $H^\theta \equiv 0$, $X^{x,\theta}$ is the wealth process required to be nonnegative and B is interpreted as a nonnegative contingent claim to be hedged. It is well known that the extreme approach of the superhedging criterion, which consists in finding an initial wealth x and a portfolio strategy θ such that $X_T^{x,\theta} \geq B$ or equivalently $(B - X_T^{x,\theta})_+ = 0$, a.s., may lead to a very expensive initial cost x. Therefore, Föllmer and Leukert introduced the concept of quantile hedging and more generally the shortfall risk criterion which consists in minimizing over strategies θ the expectation of the shortfall $(B - X_T^{x,\theta})_+$ weighted by some loss function, given an initial wealth x. Related works also include Cvitanić and Karatzas (1999), Cvitanić (2000), Spivak and Cvitanić (1999) in a model driven by a Brownian motion and Pham (2000) for lower partial moments $l(x) = x^p/p$ in a general discrete time setting.

Here, we shall consider a general semimartingale S and a convex subset Θ of $L(S)$, which leads to extensions of results in constrained portfolios models with possible jumps in price process. In addition to the stochastic integral term, we prescribe an adapted process H^θ. In the simplest case where H^θ does not depend on θ, this term can be interpreted as a labor income or random endowment; see Cuoco (1997), El Karoui and Jeanblanc (1998) or Cvitanić, Schachermayer and Wang (2000). In the general case, the term H^θ arises in large investor and reinsurance models and can be interpreted as a cost of intervention on the gain process $\int \theta \, dS$.

The modern approach for solving these control problems uses probabilistic methods rather than P.D.E. methods via the Bellman equation. This allows relaxing the assumption of Markov state process required in the P.D.E. approach. Another advantage is that one can prove existence results more simply and under weaker assumptions. Finally, in some cases, one can derive explicit solutions which are not easily obtained by P.D.E. methods. The main probabilistic tool used in the papers of Kulldorf (1993) and Heath (1995) is the martingale representation theorem. Föllmer and Leukert (1999, 2000) combine optional decomposition for supermartingales of El Karoui and Quenez (1995) and Kramkov (1996) and the Neyman–Pearson technique. In our context, we shall use the general optional decomposition under constraints of Föllmer and Kramkov (1997). Touzi (2000) also uses such a decomposition in the particular case of marked point processes S arising in an insurance model.

A crucial assumption required in this optional decomposition theorem is that the set $\{X^{0,\theta} : \theta \in \Theta\}$ be closed for the semimartingale topology. We devote a complete section to checking this assumption in different general models which can be used for applications in finance and insurance. In our general setting, the Neyman–Pearson lemma cannot be applied. We shall rather apply a convex duality approach in order to derive a quantitative description of the solution to the minimization problem. The main point is to state a convexity property on the upper variation term arising from the optional decomposition under constraints.

The paper is organized as follows. Section 2 describes the notations and general assumptions of the model and formulates the control problem precisely. In Section 3, we introduce auxiliary probability measures and upper variation processes which are used in Section 4 to provide a dual characterization of constrained controlled processes by means of the optional decomposition of Föllmer and Kramkov (1997). In Section 5, we prove existence of a solution to the minimization problem. In Section 6, we give an explicit description of the solution in terms of the solution to a dual control problem obtained by a convex duality approach. Section 7 is concerned with a verification in different classes of models of the closure property required in the optional decomposition. We conclude the paper in Section 8 with some applications: we consider the shortfall risk problem in the Black–Scholes model with short-sales constraints, the case of price pressure in the Black–Scholes model and finally a reinsurance problem. Some proofs are found in the Appendix.

2. Formulation of the control problem. We consider an $\mathbb{R}^m$-valued semimartingale S on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions of right-continuity and completeness. $T > 0$ is a fixed finite time horizon, and we assume that $\mathcal{F}_0$ is trivial and $\mathcal{F}_T = \mathcal{F}$. We denote by $L(S)$ the set of all predictable $\mathbb{R}^m$-valued processes integrable with respect to S. We consider a subset Θ of $L(S)$ containing the zero element and convex in the following sense: for any predictable process ζ valued in [0, 1] and for all $\theta^1, \theta^2 \in \Theta$, we have $\zeta\theta^1 + (1-\zeta)\theta^2 \in \Theta$. We consider a family $\{H^\theta\colon \theta \in \Theta\}$ of adapted processes with finite variation, with initial value 0 and such that the process H^0 is bounded. In the following, we shall denote by $\mathcal{I}$ the set of all nondecreasing adapted processes with initial value 0, and we shall assume the following concavity property.

(Hc) For any predictable process ζ valued in [0, 1] and for all $\theta^1, \theta^2 \in \Theta$, we have

$$H^{\zeta\theta^1+(1-\zeta)\theta^2} - \int \zeta\, dH^{\theta^1} - \int (1-\zeta)\, dH^{\theta^2} \in \mathcal{I}.$$

We say that the family $\{H^\theta : \theta \in \Theta\}$ is linear if the particular case of (Hc) is satisfied.

(Hl) For any predictable process ζ valued in [0, 1] and for all $\theta^1, \theta^2 \in \Theta$, we have

$$H^{\zeta\theta^1+(1-\zeta)\theta^2} - \int \zeta\, dH^{\theta^1} - \int (1-\zeta)\, dH^{\theta^2} \equiv 0.$$

Given $x \in \mathbb{R}$ and $\theta \in \Theta$, we consider the controlled process,

$$X_t^{x,\theta} = x + \int_0^t \theta_u\, dS_u + H_t^\theta, \qquad 0 \le t \le T.$$

We shall make the following closure property assumption. Under condition (Hc), we shall assume that:

(CPc) The family $\{X^{0,\theta} - C : \theta \in \Theta, C \in \mathcal{I}\}$ is closed for the semimartingale topology.

We recall that the semimartingale topology is associated to the Emery distance between two semimartingales X^1 and X^2 defined as

$$D_E(X^1, X^2) = \sup_{|\theta| \le 1} \left(\sum_{n \ge 1} 2^{-n} E\left[\left| \int_0^{T \wedge n} \theta_t \, d(X_t^1 - X_t^2) \right| \wedge 1 \right] \right), \tag{2.1}$$

where the supremum is taken over all predictable processes θ is bounded by 1. For this metric, the space of semimartingales is complete. Under condition (Hl), we shall assume that:

(CPl) The family $\{X^{0,\theta} : \theta \in \Theta\}$ is closed for the semimartingale topology.

These concave and closure properties are crucial in our approach and are inspired by Föllmer and Kramkov (1997). Section 7 is devoted to a verification of these conditions in different models.

Given $x \in \mathbb{R}$, we say that a control $\theta \in \Theta$ is admissible, and we denote $\theta \in \mathcal{A}(x)$, if the following conditions are satisfied:

$$X_t^{x,\theta} \ge d \quad \text{a.s.} \quad \forall t \in [0, T] \text{ for some } d \in \mathbb{R}, \tag{2.2}$$

$$X_T^{x,\theta} \ge 0 \quad \text{a.s.} \tag{2.3}$$

Consider now a nonnegative $\mathcal{F}_T$-measurable random variable B and introduce a loss function l, that is, a nondecreasing and convex function defined on $\mathbb{R}_+$ with $l(0) = 0$ and $l(y) > 0$ for $y > 0$. Given $x \in \mathbb{R}$ and $\theta \in \mathcal{A}(x)$, the shortfall is defined by $(B - X_t^{x,\theta})_+$. Our objective is to minimize over admissible controls the expectation of the shortfall, weighted by the loss function l. We shall then study the stochastic optimization problem,

$$\inf_{\theta \in \mathcal{A}(x)} E[l(B - X_t^{x,\theta})_+], \qquad x \in \mathbb{R}. \tag{$\mathbb{P}(x)$}$$

3. Auxiliary probability measures and upper variation processes. We first recall some general definitions introduced in Föllmer and Kramkov (1997). Let $\mathcal{Y}$ be a family of semimartingales, locally bounded from below, with initial value 0 and containing the constant process 0. The family $\mathcal{Y}$ is called predictably convex if for any $Y^1, Y^2 \in \mathcal{Y}$ and for any predictable process ζ valued in [0, 1], we have

$\int \zeta \, dY^1 + \int (1-\zeta)\, dY^2 \in \mathcal{Y}$. The set $\mathcal{P}(\mathcal{Y})$ is the class of all probability measures $Q \sim P$ with the following property: there exists $A \in \mathcal{I}_p$, set of nondecreasing predictable processes with $A_0 = 0$, such that

$$Y - A \text{ is a } Q\text{-local supermartingale for any } Y \in \mathcal{Y}. \tag{3.1}$$

An upper variation process of $\mathcal{Y}$ under $Q \in \mathcal{P}(\mathcal{Y})$ is an element $A^{\mathcal{Y}}(Q)$ in $\mathcal{I}_p$ satisfying (3.1) and such that $A - A^{\mathcal{Y}}(Q) \in \mathcal{I}$ for any $A \in \mathcal{I}_p$ satisfying (3.1). The set $\mathcal{P}(\mathcal{Y})$ and the upper variation process were introduced in Föllmer and Kramkov (1997) in order to develop optional versions under constraints of the Doob–Meyer decomposition. A crucial assumption required in their theorem is the following closure property.

(CP) If $(Y^n)_n$ is a sequence in $\mathcal{Y}$ which is uniformly bounded from below and converges in the semimartingale topology to Y then we have $Y \in \mathcal{Y}$.

In our context, we consider the following family:

$$\mathcal{S} = \{X^{0,\theta} - H^0 : \theta \in \Theta_{\text{loc}}\},$$

where Θ_{loc} is the set of processes $\theta \in \Theta$ such that $X^{0,\theta} - H^0$ is locally bounded from below. $\mathcal{S}$ is a family of semimartingales which are locally bounded from below, with initial value 0, and containing the constant process 0 attained for $\theta \equiv 0$. We also introduce the family $\mathcal{S}'$ of processes in the form $X - C$, for $X \in \mathcal{S}$ and $C \in \mathcal{I}$.

The following result is easy to prove.

LEMMA 3.1. *We have*

$$\mathcal{P}(\mathcal{S}) = \mathcal{P}(\mathcal{S}') := \mathcal{P}.$$

Under condition (Hc), *the set $\mathcal{S}'$ is predictably convex and the upper variation processes of $\mathcal{S}$ and $\mathcal{S}'$ under $Q \in \mathcal{P}$ exist*, *are unique and are equal*,

$$A^{\mathcal{S}}(Q) = A^{\mathcal{S}'}(Q) := A(Q) \qquad \forall\, Q \in \mathcal{P}.$$

Moreover, *under the particular case* (Hl), the set $\mathcal{S}$ is predictably convex.

PROOF. Let $Q \in \mathcal{P}(\mathcal{S})$ and $A \in \mathcal{I}_p$ such that $Y - A$ is a Q-local supermartingale for any $Y \in \mathcal{S}$. Then, we obviously have that

$$Y - C - A \text{ is a } Q\text{-local supermartingale for any } Y \in \mathcal{S} \text{ and } C \in \mathcal{I},$$

which shows that $\mathcal{P}(\mathcal{S}) \subset \mathcal{P}(\mathcal{S}')$. Conversely, by choosing $C \equiv 0$ in (3.2), we obtain that $\mathcal{P}(\mathcal{S}') \subset \mathcal{P}(\mathcal{S})$ and then that these two sets are equal.

Let $U^1 = X^{0,\theta^1} - H^0 - C^1$ and $U^2 = X^{0,\theta^2} - H^0 - C^2$ be two elements of $\mathcal{S}'$ and ζ a predictable process valued in [0, 1]. Then, a straightforward calculation shows that

$$\int \zeta \, dU^1 + \int (1-\zeta)\, dU^2 = X^{0,\theta} - H^0 - C, \tag{3.2}$$

where

$$\theta = \zeta\theta^1 + (1-\zeta)\theta^2,$$
$$C = \int \zeta\, dC^1 + \int (1-\zeta)\, dC^2 + H^\theta - \int \zeta\, dH^{\theta^1} - \int (1-\zeta)\, dH^{\theta^2}.$$

By the convexity property on the set Θ, we have $\theta \in \Theta$ and by the concavity property (Hc), we have $C \in \mathcal{I}$. Since U^1 and U^2 are locally bounded from below and ζ is bounded, we deduce from (3.3) that $X^{0,\theta} - H^0$ is locally bounded from below. Therefore $\theta \in \Theta_{\text{loc}}$ and so $\int \zeta\, dU^1 + \int (1-\zeta)\, dU^2$ lies in $\mathcal{S}'$.

Since $\mathcal{S}'$ is predictably convex, it follows by Lemma 2.1 of Föllmer and Kramkov (1997) that the upper variation process $A^{\mathcal{S}'}(Q)$ of $\mathcal{S}'$ under some fixed $Q \in \mathcal{P}(\mathcal{S}') = \mathcal{P}(\mathcal{S})$ exists and is unique. Hence, $Y - C - A^{\mathcal{S}'}(Q)$ is a Q-local supermartingale for any $Y \in \mathcal{S}$. In particular for $C \equiv 0$, we have that

$$Y - A^{\mathcal{S}'}(Q) \text{ is a } Q\text{-local supermartingale for any } Y \in \mathcal{S}.$$

Now, let $A \in \mathcal{I}_p$ such that $Y - A$ is a Q-local supermartingale for any $Y \in \mathcal{S}$. Then, $Y - C - A$ is a Q-local supermartingale for any $Y \in \mathcal{S}$ and $C \in \mathcal{I}$ and so by definition of the upper variation process $A^{\mathcal{S}'}(Q)$, we have $A - A^{\mathcal{S}'}(Q) \in \mathcal{I}_p$. This shows that $A^{\mathcal{S}'}(Q)$ is the upper variation process of $\mathcal{S}$ under Q.

Finally, by similar arguments, the predictable convexity property of $\mathcal{S}$ is easily checked under condition (Hl). □

4. Dual characterization of constrained controlled processes. Throughout this section and the two following, we assume that $\mathcal{P} \neq \emptyset$ and that conditions (Hc) and (CPc) or (HI) and (CPl) hold.

We say that a nonnegative $\mathcal{F}_T$-measurable random variable X is dominated by a controlled process if there exist $u_0 \in \mathbb{R}$ and $\theta \in \mathcal{A}(u_0)$ such that

$$X \le X_T^{u_0,\theta}, \quad P \text{ a.s.}$$

The following result provides a dual characterization of dominated random variables. We shall also assume that

$$E^Q[A_T(Q)] < \infty \qquad \forall\, Q \in \mathcal{P}, \tag{4.1}$$

and that the process

$$\operatorname*{ess\,inf}_{Q\in\mathcal{P}} E^Q[A_T(Q)|\mathcal{F}_t] \quad \text{is bounded in } (t,\omega). \tag{4.2}$$

These conditions are satisfied in all examples of the last section.

THEOREM 4.1. *Assume that* (4.1) *and* (4.2) *hold and let X be a nonnegative $\mathcal{F}_T$-measurable random variable.*

(i) *Given* $x \in \mathbb{R}$, *there exists* $\theta \in \mathcal{A}(x)$ *such that* $X \leq X_T^{x,\theta}$ *a.s. if and only if*

$$v_0(X) := \sup_{Q \in \mathcal{P}} E^Q[X - H_T^0 - A_T(Q)] \leq x. \tag{4.3}$$

(ii) *Suppose that* $v_0(X) < \infty$ *and denote by* $\theta^X \in \mathcal{A}(v_0(X))$ *the control process s.t.* $X \leq X_T^{v_0(X), \theta^X}$, *a.s. If there exists* $\widehat{Q} \in \mathcal{P}$ *such that* $v_0(X) = E^{\widehat{Q}}[X - H_T^0 - A_T(\widehat{Q})]$, *then we have actually* $X = X_T^{v_0(X), \theta^X}$, *and the associated controlled process is given by*

$$X_t^{v_0(X), \theta^X} = E^{\widehat{Q}}[X - H_T^0 + H_t^0 - A_T(\widehat{Q}) + A_t(\widehat{Q}) \mid \mathcal{F}_t], \qquad 0 \leq t \leq T.$$

This theorem shows that $v_0(X)$ is the least initial value allowing the domination of X by an admissible control. In the terminology of mathematical finance, $v_0(X)$ is called superreplication cost of X and the associated control θ^X is a superhedging strategy of X.

PROOF OF THEOREM 4.1. (i) *Necessary condition.* Let $\theta \in \mathcal{A}(x)$ such that $X \leq X_T^{x,\theta}$, a.s. By condition (4.1), the nondecreasing feature of $A(Q)$ and since the process $X^{x,\theta} - H^0$ is bounded from below, we deduce by Fatou's lemma that the Q-local supermartingale $X^{x,\theta} - H^0 - A(Q)$ is actually a supermartingale under $Q \in \mathcal{P}$. Therefore, we have

$$E^Q[X - H_T^0 - A_T(Q)] \leq E^Q[X_T^{x,\theta} - H_T^0 - A_T(Q)] \leq x$$

for any $Q \in \mathcal{P}$ and so (4.3).

Sufficient condition. Consider the bounded from below $\mathcal{F}_T$-measurable random variable $X_0 = X - H_T^0$. By (4.3), we have

$$\sup_{Q \in \mathcal{P}} E^Q[X^0 - A_T(Q)] = v_0(X) \leq x < \infty.$$

Then by similar arguments as in Lemma A.1 in Föllmer and Kramkov (1997), there exists a RCLL version of the process

$$V_t = \operatorname*{ess\,sup}_{Q \in \mathcal{P}} E^Q[X^0 - A_T(Q) + A_t(Q) \mid \mathcal{F}_t], \qquad 0 \leq t \leq T,$$

and for any $Q \in \mathcal{P}$, the process $V - A(Q)$ is a Q-local supermartingale. Moreover, by (4.2), the process V is bounded from below. It is clear that under condition (CPc) [resp. (CPl)], the family $\mathcal{S}'$ (resp. $\mathcal{S}$) satisfies the closure property (CP). Therefore, by Lemma 3.1, the optional decomposition under constraints of Föllmer and Kramkov (1997) (see their Theorem 3.1) can be applied so that the process V admits a decomposition

$$V = v_0(X) + U - \widetilde{C},$$

where $U \in \mathcal{S}'$ (resp. $\mathcal{S}$) and $\widetilde{C}$ is an (optional) nondecreasing process, $\widetilde{C}_0 = 0$. Therefore, there exists $\theta \in \Theta_{\mathrm{loc}}$ such that

$$V_t \le v_0(X) + X_t^{0,\theta} - H_t^0 \le X_t^{x,\theta} - H_t^0, \qquad 0 \le t \le T. \tag{4.4}$$

Since the process V is bounded from below, we deduce by (4.4) that $X^{x,\theta}$ is also bounded from below so that (2.2) is satisfied. Moreover, inequality (4.4) for $t = T$ shows that

$$V_T = X - H_T^0 \le X_T^{x,\theta} - H_T^0,$$

and so $0 \le X \le X_T^{x,\theta}$, P a.s., with $\theta \in \mathcal{A}(x)$.

(ii) Suppose that $v_0(X) < \infty$ and $v_0(X) = E^{\widehat{Q}}[X - H_T^0 - A_T(\widehat{Q})]$ for some $\widehat{Q} \in \mathcal{P}$. By the supermartingale property of $X^{v_0(X),\theta^X} - H^0 - A(\widehat{Q})$ under $\widehat{Q}$, we then have

$$\begin{aligned} E^{\widehat{Q}}[X_T^{v_0(X),\theta^X} - X] &= E^{\widehat{Q}}\big[X_T^{v_0(X),\theta^X} - H_T^0 - A_T(\widehat{Q}) - \big(X - H_T^0 - A_T(\widehat{Q})\big)\big] \\ &\le v_0(X) - E^{\widehat{Q}}[X - H_T^0 - A_T(\widehat{Q})] = 0. \end{aligned}$$

Since $\widehat{Q} \sim P$, this proves that $X_T^{v_0(X),\theta^X} = X$, P a.s. Moreover, since $E^{\widehat{Q}} \times [X_T^{v_0(X),\theta^X} - H_T^0 - A_T(\widehat{Q})] = v_0(X)$, we conclude that the $\widehat{Q}$-supermartingale $X^{v_0(X),\theta^X} - H^0 - A(\widehat{Q})$ is actually a $\widehat{Q}$-martingale which ends the proof. $\square$

As an immediate consequence of the last theorem, we have a necessary and sufficient condition in terms of the set $\mathcal{P}$ in order for the set of admissible controls to be nonempty.

COROLLARY 4.1. *For all $x \in R$, $\mathcal{A}(x) \neq \varnothing$ if and only if*

$$v_0(0) = \sup_{Q \in \mathcal{P}} E^Q[-H_T^0 - A_T(Q)] \le x. \tag{4.5}$$

PROOF. Suppose that $\mathcal{A}(x) \neq \varnothing$. Then there exists $\theta \in \mathcal{A}(x)$ such that $X_T^{x,\theta} \ge X = 0$. By Theorem 4.1, we then have $v_0(X) \le x$. Conversely, suppose that $v_0(0) \le x$. Then by Theorem 4.1, there exists $\theta \in \mathcal{A}(x)$ such that $0 \le X_T^{x,\theta}$ and in particular $\mathcal{A}(x)$ is nonempty. $\square$

5. Existence to the minimization problem. By using the dual characterization of constrained controlled processes in the previous section, we shall reduce the initial dynamic control problem into a static optimization problem. For any $x \in \mathbb{R}$, we denote

$$\begin{aligned} \mathcal{C}(x) = \Big\{ & X\ \mathcal{F}_T\text{-measurable} : 0 \le X \le B\ P\text{-a.s.} \\ & \text{and } v_0(X) = \sup_{Q \in \mathcal{P}} E^Q[X - H_T^0 - A_T(Q)] \le x \Big\}. \end{aligned}$$

Notice that $\mathcal{C}(x)$ is nonempty if and only if $x \geq v_0(0)$. We consider the static problem,

$$(\mathbb{S}(x)) \qquad J(x) = \inf_{X \in \mathcal{C}(x)} E[l(B - X)], \qquad x \geq v_0(0).$$

The following main result proves the existence of a solution to the dynamic problem $(\mathbb{P}(x))$ and relates it to the solution of the static problem $(\mathbb{S}(x))$. It also provides some qualitative properties of the associated value function.

THEOREM 5.1. *Assume that* (4.1) *and* (4.2) *hold and that* $l(B) \in L^1(P)$.

(i) *For any* $x \geq v_0(0)$, *there exists* $X^*(x) \in \mathcal{C}(x)$ *solution of* $(\mathbb{S}(x))$ *and* B *is solution of* $(\mathbb{S}(x))$ *for* $x \geq v_0(B)$. *Moreover*, *if* l *is strictly convex*, *any two such solutions coincide* P*-a.s.*

(ii) *The function* J *is nonincreasing and convex on* $[v_0(0), \infty)$, *strictly decreasing on* $[v_0(0), v_0(B)]$ *and equal to zero on* $[v_0(B), \infty)$. *For any* $x \in [v_0(0), v_0(B)]$, *we have*

$$\sup_{Q \in \mathcal{P}} E^Q[X^*(x) - H_T^0 - A_T(Q)] = x. \tag{5.1}$$

Moreover, *if* l *is strictly convex*, *then* J *is strictly convex on* $[v_0(0), v_0(B)]$.

(iii) *For any* $x \geq v_0(0)$, *there exists* $\theta^*(x) \in \mathcal{A}(x)$ *such that* $X^*(x) \leq X_T^{x,\theta^*(x)}$, P*-a.s.*, *and* $\theta^*(x)$ *is solution to the dynamic problem* $(\mathbb{P}(x))$. *Moreover*, *we have*

$$J(x) = \inf_{\theta \in \mathcal{A}(x)} E\big[l(B - X_T^{x,\theta})_+\big] \qquad \forall\, x \geq v_0(0). \tag{5.2}$$

PROOF. (i) Let $x \geq v_0(0)$ and $(X^n)_n \in \mathcal{C}(x)$ be a minimizing sequence for the problem $(\mathbb{S}(x))$; that is,

$$\lim_{n \to \infty} E[l(B - X^n)] = \inf_{X \in \mathcal{C}(x)} E[l(B - X)].$$

Since $X^n \geq 0$, then by Lemma A.1.1 of Delbaen and Schachermayer (1994), there exists a sequence of $\mathcal{F}_T$-measurable random variables $\widehat{X}^n \in \mathrm{conv}(X^n, X^{n+1}, \ldots)$ such that $\widehat{X}^n$ converges almost surely to $X^*(x)$ $\mathcal{F}_T$-measurable. Since $0 \leq \widehat{X}^n \leq B$, we deduce that $0 \leq X^*(x) \leq B$. By Fatou's lemma, we have for all $Q \in \mathcal{P}$,

$$\begin{aligned} E^Q[X^*(x) - H_T^0 - A_T(Q)] &\leq \liminf_{n \to \infty} E^Q[\widehat{X}^n - H_T^0 - A_T(Q)] \\ &= \liminf_{n \to \infty} E^Q[X^n - H_T^0 - A_T(Q)] \\ &\leq x, \end{aligned}$$

hence $X^*(x) \in \mathcal{C}(x)$. Now, since l is convex and $l(B) \in L^1(P)$, we have by the dominated convergence theorem,

$$\begin{aligned} \inf_{X \in \mathcal{C}(x)} E[l(B - X)] &= \lim_{n \to \infty} E[l(B - X^n)] \\ &\geq \lim_{n \to \infty} E[l(B - \widehat{X}^n)] \\ &= E[l(B - X^*(x))], \end{aligned}$$

which proves that $X^*(x)$ solves $(\mathbb{S}(x))$. Now, suppose that $x \geq v_0(B)$. Then $B \in \mathcal{C}(x)$ and is obviously a solution to $(\mathbb{S}(x))$, and in this case $J(x) = 0$.

Let X^1 and X^2 be two solutions of $(\mathbb{S}(x))$ and $\varepsilon \in (0, 1)$. Set $X^\varepsilon = (1-\varepsilon)X^1 + \varepsilon X^2 \in \mathcal{C}(x)$. By convexity of function l, we have

$$E[l(B - X^\varepsilon)] \leq (1-\varepsilon)E[l(B - X^1)] + \epsilon E[l(B - X^2)] \tag{5.3}$$

$$= \inf_{X \in \mathcal{C}(x)} E[l(B - X)]. \tag{5.4}$$

Suppose that $P[X^1 \neq X^2] > 0$. Then by the strict convexity of l, we should have strict inequality in (5.3), which is a contradiction with (5.4).

(ii) Let $v_0(0) \leq x_1 \leq x_2$. Since $\mathcal{C}(x_1) \subset \mathcal{C}(x_2)$, we deduce that $J(x_2) \leq J(x_1)$ and so J is nonincreasing on $[v_0(0), \infty)$. Notice also that $(X^*(x_1) + X^*(x_2))/2 \in \mathcal{C}((x_1 + x_2)/2)$. Then, by convexity of function l, we have

$$\begin{aligned} J\left(\frac{x_1 + x_2}{2}\right) &\leq E\left[l\left(B - \frac{X^*(x_1) + X^*(x_2)}{2}\right)\right] \\ &\leq \frac{1}{2}E\big[l\big(B - X^*(x_1)\big)\big] + \frac{1}{2}E\big[l\big(B - X^*(x_2)\big)\big] \\ &= \frac{1}{2}J(x_1) + \frac{1}{2}J(x_2), \end{aligned}$$

which proves the convexity of J on $[v_0(0), \infty)$. We have already seen that $J(x) = 0$ for $x \geq v_0(B)$. To end the proof of assertion (ii), we now suppose that $v_0(B) > v_0(0)$ (otherwise there is nothing else to prove). First, notice that since l is a nonnegative function, cancelling only on 0, it follows that $J(x) = 0$ if and only if $X^*(x) = B$ which implies that $x \geq v_0(B)$. Therefore, for all $b \leq x < v_0(B)$, we have $J(x) > 0$. Let us check that J is strictly decreasing on $[v_0(0), v_0(B)]$. On the contrary, there would exist $v_0(0) \leq x_1 < x_2 < v_0(B)$ such that $J(x_1) = J(x_2)$. Then, there exists $\alpha \in (0, 1)$ such that $x_2 = \alpha x_1 + (1-\alpha)v_0(B)$. By convexity of function J, we should have

$$J(x_2) \leq \alpha J(x_1) + (1-\alpha)J\big(v_0(B)\big) = \alpha J(x_1).$$

Since $J(x_1) = J(x_2) > 0$, this would imply that $\alpha > 1$, a contradiction. Let us now prove (5.1). On the contrary, we should have $b \leq \tilde{x} := v_0(X^*(x)) < x$. Then $X^*(x) \in \mathcal{C}(\tilde{x})$ and so $J(\tilde{x}) \leq E[l(B - X^*(x))] = J(x)$, a contradiction with the fact that J is strictly decreasing on $[v_0(0), v_0(B)]$. Let $v_0(0) \leq x_1 < x_2 \leq v_0(B)$. We have $(X^*(x_1) + X^*(x_2))/2 \in \mathcal{C}((x_1 + x_2)/2)$. Moreover, since $0 < J(x_2) < J(x_1)$, we have $X^*(x_1) \neq X^*(x_2)$. Then, by the strict convexity of function l, we obtain

$$\begin{aligned} J\left(\frac{x_1 + x_2}{2}\right) &\leq E\left[l\left(B - \frac{X^*(x_1) + X^*(x_2)}{2}\right)\right] \\ &< \frac{1}{2}E\big[l\big(B - X^*(x_1)\big)\big] + \frac{1}{2}E\big[l\big(B - X^*(x_2)\big)\big] \\ &= \frac{1}{2}J(x_1) + \frac{1}{2}J(x_2), \end{aligned}$$

which proves the strict convexity of J on $[v_0(0), v_0(B)]$.

(iii) Fix some $x \geq v_0(0)$. Let $\theta \in \mathcal{A}(x)$ and set $X = B \wedge X_T^{x,\theta} = B - (B - X_T^{x,\theta})_+$. Then $0 \leq X \leq B$. Since $X \leq X_T^{x,\theta}$, we have $v_0(X) \leq x$, by Theorem 4.1, and so $X \in \mathcal{C}(x)$. We then have

$$E\big[l(B - X_T^{x,\theta})_+\big] = E[l(B - X)] \geq J(x),$$

and then

$$\inf_{\theta \in \mathcal{A}(x)} E\big[l(B - X_T^{x,\theta})_+\big] \geq J(x). \tag{5.5}$$

Conversely, let $X \in \mathcal{C}(x)$. We then have $v_0(X) \leq x < \infty$. We deduce by Theorem 4.1 that there exists $\theta \in \mathcal{A}(x)$ such that

$$X_T^{x,\theta} \geq X, \quad P\text{-a.s.}, \tag{5.6}$$

and therefore, recalling that $X \leq B$,

$$(B - X_T^{x,\theta})_+ \leq B - X, \quad P\text{-a.s.} \tag{5.7}$$

Now, since the function l is nondecreasing, we obtain

$$E\big[l(B - X_T^{x,\theta})_+\big] \leq E[l(B - X)],$$

which, combined with (5.5), proves (5.2). Finally, as in (5.6) and (5.7), there exists an admissible control $\theta^*(x) \in \mathcal{A}(x)$ such that $X^*(x) \leq X_T^{x,\theta^*(x)}$, P-a.s., and we have

$$\big(B - X_T^{x,\theta^*(x)}\big)_+ \leq B - X^*(x) \quad P\text{-a.s.},$$

hence,

$$E\big[l\big(B - X_T^{x,\theta^*(x)}\big)_+\big] \leq E\big[l\big(B - X^*(x)\big)\big] = J(x),$$

which proves that $\theta^*(x)$ solves problem $(\mathbb{P}(x))$. $\square$

6. Convex duality approach and structure of the solution. The purpose of this section is to provide a quantitative description of $X^*(x)$ and $\theta^*(x)$ solutions of the optimization problems $(\mathbb{P}(x))$ and $(\mathbb{S}(x))$, and of the associated value function $J(x)$. In Föllmer and Leukert (1999, 2000), the Neyman–Pearson lemma approach is emphasized in order to describe the structure of the solution to the static problem. Here, we have an upper variation term due to the constraints set Θ and the process H^θ and therefore, the Neyman–Pearson technique cannot be applied. We shall use a convex duality approach.

We assume that the function $l \in C^1(0, \infty)$, the derivative l' is strictly increasing with $l'(0^+) = 0$ and $l'(\infty) = \infty$. We denote by $I = (l')^{-1}$ the inverse function of l'. Starting from the state-dependent convex function $0 \le x \le B \mapsto l(B-x)$, we consider its stochastic Fenchel–Legendre transform,

$$\begin{aligned}\widetilde{L}(y,\omega) &= \max_{0\le x\le B}[-l(B-x)-xy]\\ &= -l\big(B\wedge I(y)\big) - y\big(B-I(y)\big)_+, \qquad y\ge 0.\end{aligned}$$

We now consider the dual control problem,

$$(\mathcal{D}(y)) \qquad \widetilde{J}(y) = \inf_{Q\in P} E\bigg[\widetilde{L}\bigg(y\frac{dQ}{dP},\omega\bigg) + y\frac{dQ}{dP}\big(A_T(Q)+H_T^0\big)\bigg], \qquad y\ge 0.$$

It is straightforward to see that $\widetilde{J}$ is convex $[0,\infty)$.

In contrast with the problems of maximizing expected utility of terminal wealth where existence to the dual problem is established in order to prove existence of an optimal portfolio choice [see Kramkov and Schachermayer (1999) and Cvitanić, Schachermayer and Wang (2000)] since we have already proved the existence of the minimization problem ($\mathbb{P}(x)$), we do not focus here on a general existence result for the dual problem ($\mathcal{D}(y)$). Our object is to provide a description of the solution to the problem ($\mathbb{P}(x)$) in function of a solution to the problem ($\mathcal{D}(y)$) when it exists. This can be viewed as a verification theorem expressed in terms of the dual control problem. In a Markovian context, this is an alternative to the usual verification theorem of stochastic control problems expressed in terms of the value function of the primal problem. In the last section, we shall see that the dual control problem leads to a standard Bellman equation where the existence of a smooth solution can be proved simply.

Recall that for $x \ge v_0(B)$, $B \in \mathcal{C}(x)$ is solution to the static problem ($\mathbb{S}(x)$). We then consider in the sequel the case $v_0(0) < x < v_0(B) < \infty$.

THEOREM 6.1. *Assume that* (4.1) *and* (4.2) *hold*, $l(B-d_T)\in L^1(P)$ *and* $v_0(0) < v_0(B) < \infty$. *Suppose that for all* $y>0$, *there exists a solution* $Q^*(y)\in\mathcal{P}$ *to problem* ($\mathcal{D}(y)$). *Then*:

(i) $\widetilde{J}$ *is differentiable on* $(0,\infty)$ *with derivative*

$$\widetilde{J}'(y) = -E^{Q^*(y)}\bigg[\bigg(B - I\bigg(y\frac{dQ^*(y)}{dP}\bigg)\bigg)_+ - H_T^0 - A_T\big(Q^*(y)\big)\bigg], \tag{6.1}$$

for all $y>0$.

(ii) *Let* $v_0(0) < x < v_0(B)$. *Then, there exists* $\hat{y}>0$ *that attains the infimum in* $\inf_{y>0}\{\widetilde{J}(y)+xy\}$, *and we have*

$$\widetilde{J}'(\hat{y}) = -x.$$

The unique solution of ($\mathbb{S}(x)$) *is then given by*

$$X^*(x) = \bigg(B - I\bigg(\hat{y}\frac{dQ^*(\hat{y})}{dP}\bigg)\bigg)_+.$$

There exists $\theta^*(x) \in \mathcal{A}(x)$ *such that* $X^*(x) = X_T^{x,\theta^*(x)}$, P-*a.s.*, *and* $\theta^*(x)$ *is solution to* $(\mathbb{P}(x))$. *Moreover*, *we have*

$$X_t^{x,\theta^*(x)} = E^{Q^*(\hat{y})}\big[X^*(x) - H_T^0 + H_t^0 - A_T\big(Q^*(\hat{y})\big) + A_t\big(Q^*(\hat{y})\big) \,\big|\, \mathcal{F}_t\big],$$
$$0 \le t \le T.$$

(iii) *We have the duality relation*

$$J(x) = \max_{y \ge 0}[-\widetilde{J}(y) - xy] \qquad \forall\, x > v_0(0).$$

The proof of this theorem is based on classical arguments of convex duality [see, e.g., Kramkov and Schachermayer (1997)] with modifications arising from the upper variation term. More precisely, we need the following general result.

LEMMA 6.1. *Assume that* (4.1) *holds*. *Then the set* $\mathcal{P}$ *is convex and the function*

$$\mathcal{P} \to \mathbb{R}_+,$$
$$Q \mapsto E^Q[A_T(Q)]$$

is convex.

PROOF. Let $Q^1, Q^2 \in \mathcal{P}$, Z^1, Z^2 their density processes, $\alpha \in [0,1]$ and denote by $Q \sim P$ the probability measure $Q = \alpha Q^1 + (1-\alpha)Q^2$ and by Z its density process. Consider the process $A^Q \in \mathcal{I}_p$ defined by

$$A_t^Q = \alpha \int_0^t \frac{Z_u^1}{Z_u}\, dA_u(Q^1) + (1-\alpha)\int_0^t \frac{Z_u^2}{Z_u}\, dA_u(Q^2), \qquad 0 \le t \le T.$$

Fix $0 \le u \le t \le T$. We have, for $i = 1, 2$,

$$\begin{aligned} &E\bigg[Z_t \int_0^t \frac{Z_v^i}{Z_v}\, dA_v(Q^i) \,\bigg|\, \mathcal{F}_u\bigg] \\ &\quad = Z_u \int_0^u \frac{Z_v^i}{Z_v}\, dA_v(Q^i) + E\bigg[Z_t \int_u^t \frac{Z_v^i}{Z_v}\, dA_v(Q^i) \,\bigg|\, \mathcal{F}_u\bigg] \\ &\quad = Z_u \int_0^u \frac{Z_v^i}{Z_v}\, dA_v(Q^i) + E\bigg[\int_u^t Z_v^i\, dA_v(Q^i) \,\bigg|\, \mathcal{F}_u\bigg] \qquad (6.2) \\ &\quad = Z_u \int_0^u \frac{Z_v^i}{Z_v}\, dA_v(Q^i) + E\big[Z_t^i\big(A_t(Q^i) - A_u(Q^i)\big) \,\big|\, \mathcal{F}_u\big] \\ &\quad = Z_u \int_0^u \frac{Z_v^i}{Z_v}\, dA_v(Q^i) - Z_u^i A_u(Q^i) + E[Z_t^i A_t(Q^i) \mid \mathcal{F}_u] \end{aligned}$$

where we used the properties that Z is a P-martingale, Bayes formula and law of iterated conditional expectations. Now, fix some $U \in \mathcal{S}'$. For simplicity, we assume that U is bounded from below. The general case follows by localization arguments. By writing that

$$Z_t(U_t - A_t^Q) = \alpha\Big[Z_t^1 U_t - Z_t \int_0^t \frac{Z_v^1}{Z_v}\, dA_v(Q^1)\Big] + (1-\alpha)\Big[Z_t^2 U_t - Z_t \int_0^t \frac{Z_v^2}{Z_v}\, dA_v(Q^2)\Big],$$

and using relations (6.2) for $i = 1, 2,$ we obtain by the supermartingale property of $Z^i(U - A(Q^i))$ under P,

$$\begin{aligned}
E\big[Z_t(U_t - A_t^Q) \mid \mathcal{F}_u\big] &= \alpha E\big[Z_t^1\big(U_t - A_t(Q^1)\big) \,\big|\, \mathcal{F}_u\big] \\
&\quad + (1-\alpha) E\big[Z_t^2\big(U_t - A_t(Q^2)\big) \,\big|\, \mathcal{F}_u\big] \\
&\quad - \alpha Z_u \int_0^u \frac{Z_v^1}{Z_v}\, dA_v(Q^1) - (1-\alpha) Z_u \int_0^u \frac{Z_v^2}{Z_v}\, dA_v(Q^2) \\
&\quad + \alpha Z_u^1 A_u(Q^1) + (1-\alpha) Z_u^2 A_u(Q^2) \\
&\le \alpha Z_u^1\big(U_u - A_u(Q^1)\big) + (1-\alpha) Z_u^2\big(U_u - A_u(Q^2)\big) \\
&\quad - Z_u A_u^Q + \alpha Z_u^1 A_u(Q^1) + (1-\alpha) Z_u^2 A_u(Q^2) \\
&= Z_u(U_u - A_u^Q).
\end{aligned}$$

This proves the supermartingale property under P of $Z(U - A^Q)$ and so the supermartingale property under Q of $U - A^Q$. Therefore $Q \in \mathcal{P}$ and by definition of the upper variation process, we have $A^Q - A(Q) \in \mathcal{I}_p$ and so $A_T(Q) \le A_T^Q$. Moreover, applying (6.2) for $u = 0$ and $t = T$, we have

$$\begin{aligned}
E^Q[A_T(Q)] &\le E[Z_T A_T^Q] \\
&= \alpha E[Z_T^1 A_T(Q^1)] + (1-\alpha) E[Z_T^2 A_T(Q^2)] \\
&= \alpha E^{Q^1}[A_T(Q^1)] + (1-\alpha) E^{Q^2}[A_T(Q^2)]
\end{aligned}$$

which proves the convexity of function $Q \in \mathcal{P} \mapsto E^Q[A_T(Q)]$. $\square$

Given Lemma 6.1, the proof of Theorem 6.1 is standard and is then reported in the Appendix.

7. Closure property. This section is concerned with the verification of the concave and closure properties formulated in Section 2, in different classes of models useful for applications in finance and insurance.

7.1. *Constrained portfolios.* In this model, we suppose that Θ is closed in $L(S)$ with respect to the metric $d_E(\theta^1, \theta^2) = D_E(\int \theta^1 \, dS, \int \theta^2 \, dS)$ where D_E is the Emery distance in (2.1) defining the semimartingale topology. We consider $H^\theta \equiv 0$ so that the controlled process $X^{x,\theta} = x + \int \theta \, dS$ is interpreted as a wealth process of a constrained portfolio strategy θ in a financial market model of price process S. For the case where $\Theta = L(S)$, this corresponds to the incomplete market model. When $\Theta = \{\theta \in L(S) : \theta^i \geq 0, \; 1 \leq i \leq l\}$, this corresponds to a no short-selling model on the first l assets.

Obviously, the linear condition (Hl) is satisfied. The closure property (CPl) follows from Mémin's theorem (1980), which states that the space of stochastic integrals is closed for the semimartingale topology.

7.2. *Labor income model.* We consider the same model as in the previous paragraph, but we prescribe in addition an adapted process (p_t) such that $\int_0^T |p_t| \, dt < \infty$, a.s. We consider then for all $\theta \in \Theta$, the adapted process with finite variation,

$$H_t^\theta = \int_0^t p_u \, du, \qquad 0 \leq t \leq T.$$

The process p represents an income rate per unit of time. Again, the linear condition (Hl) is obvious. The closure property (CPl) follows from the invariance of the Emery distance in (2.1) under translation.

7.3. *Concave model.* We assume that S is a continuous semimartingale with canonical decomposition,

$$S_t = S_0 + M_t + A_t, \qquad t \in [0, T].$$

We denote by $\langle M \rangle$ the sharp bracket process of M. We shall assume that $\langle M \rangle$ is absolutely continuous with respect to the Lebesgue measure on $[0, T]$ and we define the predictable $m \times m$-matrix valued process $\sigma = (\sigma_t)_{0 \leq t \leq T}$ by

$$\langle M \rangle_t = \int_0^t \sigma_u \, du, \qquad t \in [0, T]. \tag{7.1}$$

We assume, following the terminology of Schweizer (1994), that S satisfies the structure condition, in the sense that there exists a predictable $\mathbb{R}^m$-valued process $\lambda = (\lambda_t)_{0 \leq t \leq T}$ such that

$$A_t = \int_0^t d\langle M \rangle_u \lambda_u \, du = \int_0^t \sigma_u \lambda_u \, du, \qquad t \in [0, T]. \tag{7.2}$$

We assume that

$$\int_0^T \lambda_t' \sigma_t \lambda_t \, dt \quad \text{is bounded,} \tag{7.3}$$

where $'$ stands for the transposition. We refer to Schweizer (1994) for an interpretation of $\int_0^T \lambda_t' \sigma_t \lambda_t \, dt$ as a mean-variance tradeoff. Finally, we shall make the additional nondegeneracy assumption:

$$\sigma_t \text{ is definite positive a.s. } \forall\, t \in [0, T]. \tag{7.4}$$

We are given a closed convex set K of $\mathbb{R}^m$ containing 0 and a measurable function $h(t, z)$ of $[0, T] \times K$ into $\mathbb{R}$ such that the function $h(t, \cdot)$ is concave in $z \in K$ and satisfies the property: there exists $k \geq 0$, $\forall z_1, z_2 \in K$, $\forall \alpha \in [0, 1]$, $\forall t \in [0, T]$,

$$\big|h\big(t, \alpha z_1 + (1-\alpha) z_2\big)\big| \leq k\big(1 + |h(t, z_1)| + |h(t, z_2)|\big). \tag{7.5}$$

We also assume that the function $h(t, \cdot)$ is Lipschitz in $z \in K$ uniformly in $t \in [0, T]$ and either (i) K is bounded or (ii) the $m \times m$ matrix σ_t is uniformly elliptic a.s. for all $t \in [0, T]$: there exists $c > 0$ such that $\xi' \sigma_t^c \xi \geq c|\xi|^2$, a.s., $\forall \xi \in \mathbb{R}^m$, $\forall\, t \in [0, T]$.

We consider $\Theta = \{\theta \in L(S) : \int_0^T |h(t, \theta_t)| \, dt < \infty$, a.s., and $\theta_t \in K$, a.s. $\forall\, t \in [0, T]\}$. We prescribe then for all $\theta \in \Theta$, the adapted process with finite variation,

$$H_t^\theta = \int_0^t h(u, \theta_u) \, du, \qquad 0 \leq t \leq T.$$

Such a concave model is used in applications for large investor models [see Cuoco and Cvitanić (1998)] and insurance models [see Touzi (2000)].

The following lemma states the concave and closure properties (Hc) and (CPc).

LEMMA 7.1. (i) *The family* $\{H^\theta = \int h(t, \theta_t) \, dt : \theta \in \Theta\}$ *satisfies* (Hc).

(ii) *The set* $\{X^{0,\theta} - C = \int \theta \, dS + \int h(t, \theta_t) \, dt - C : \theta \in \Theta,\ \mathcal{C} \in \mathcal{I}\}$ *is closed for the semimartingale topology*.

For the proof see the Appendix.

8. Applications.

8.1. *Short-sales constraints in the Black–Scholes model.* We consider the standard Black–Scholes model where the underlying stock price process is given by a geometric Brownian motion,

$$dS_t = S_t(\mu \, dt + \sigma \, dW_t), \qquad S_0 = s_0,$$

where $\mu \geq 0$, $\sigma > 0$ and W is a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ equipped with the filtration generated by W. For simplicity, we set the interest rate to zero. The wealth process starting from an initial capital $x \in \mathbb{R}$ and a trading strategy $\theta \in L(S)$, number of units invested in the stock, is then written as

$$X_t^{x,\theta} = x + \int_0^t \theta_u \, dS_u = x + \int_0^t \theta_u S_u(\mu \, du + \sigma \, dW_u).$$

We impose no short-sales constraints on trading strategies,

$$\theta_t \in K = [0, \infty) \quad \text{a.s. } \forall t \in [0, T].$$

This model is a particular case of the one considered in Section 7.2. We have $\Theta = \Theta_{\text{loc}} = \{(\theta_t)_{0 \le t \le T}$ adapted process, $\int_0^T \theta_t^2 \, dt < \infty$ and $\theta_t \ge 0$, a.s., $0 \le t \le T\}$. It is well known from the martingale representation theorem for martingales that all probability measures $Q \sim P$ have a Radon–Nikodym density of the form $dQ/dP = Z_T^\rho$, where

$$dZ_t^\rho = -\rho_t Z_t^\rho \, dW_t, \qquad Z_0^\rho = 1,$$

and $\rho \in \mathcal{D} = \{(\rho_t)_{0 \le t \le T}$ adapted process, $\int_0^T \rho_t^2 \, dt < \infty$ and $E[Z_T^\rho] = 1\}$. Since $\mathcal{S} = \{X^{0,\theta} : \theta \in \Theta_{\text{loc}}\}$ is a cone of semimartingales, $\mathcal{P}$ is the set of all probability measures $Q \sim P$ such that $X^{0,\theta}$ is a Q-local supermartingale for any $\theta \in \Theta_{\text{loc}}$ and the upper variation process of $Q \in \mathcal{P}$ is zero. By Girsanov's theorem, we obtain

$$\mathcal{P} = \{P^\rho = Z_T^\rho \cdot P : \rho \in \mathcal{D}_C\}$$

with

$$\mathcal{D}_C = \left\{\rho \in \mathcal{D} : \rho_t \ge \frac{\mu}{\sigma} \text{ a.s., } 0 \le t \le T\right\}.$$

We consider the problem of minimizing the shortfall risk of an European call option $B = g(S_T) = (S_T - \kappa)_+$ and for a loss function l satisfying the assumptions of Section 6. Notice that from Corollary 4.1, the set of admissible controls $\mathcal{A}(x)$ is nonempty iff $x \ge v_0(0) = 0$. We fix now some initial wealth $x \ge 0$. The dynamic version of the dual control problem $(\mathcal{D}(y))$ is

$$\widetilde{J}(t, s, y) = \inf_{\rho \in \mathcal{D}_C} E\left[\widetilde{L}\left(y \frac{Z_T^\rho}{Z_t^\rho}, s \frac{S_T}{S_t}\right)\right], \qquad (t, s, y) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+, \tag{8.1}$$

where $\widetilde{L}(y, s) = -l(g(s) \wedge I(y)) - y(g(s) - I(y))_+$. We have $\widetilde{J}(y) = \widetilde{J}(0, s_0, y)$. The Bellman equation associated to this control problem is

$$-\frac{\partial v}{\partial t} - \mu s \frac{\partial v}{\partial s} - \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} + \sup_{r \ge \mu/\sigma}\left[-\frac{1}{2} r^2 y^2 \frac{\partial^2 v}{\partial y^2} + \sigma r s y \frac{\partial^2 v}{\partial s \partial y}\right] = 0, \tag{8.2}$$

together with the terminal boundary condition $v(T, s, y) = \widetilde{L}(y, s)$. Define now the positive constant process in $\mathcal{D}_c$ by

$$r^* = \frac{\mu}{\sigma}.$$

The new proposition shows that r^* is the optimal control for (8.1) or $(\mathcal{D}(y))$.

PROPOSITION 8.1. *The value function of the dual control problem is equal to* $\widetilde{V}$ *where*

$$(8.3)\qquad \widetilde{V}(t,s,y) = E\left[\widetilde{L}\left(y\frac{Z_T^{r^*}}{Z_T^{r^*}}, s\frac{S_T}{S_t}\right)\right], \qquad (t,s,y) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R}_+.$$

PROOF. We only sketch the main arguments of the proof. Using classical regularity results for parabolic linear P.D.E. and the Feynman–Kac formula [see, e.g., Friedman (1975)], we have that the function $\widetilde{V}$ lies in $C^{1,2}([0,T], \mathbb{R}_+ \times \mathbb{R}_+)$ and is solution of

$$(8.4)\qquad -\frac{\partial v}{\partial t} - \mu s\frac{\partial v}{\partial s} - \frac{1}{2}\sigma^2 s^2\frac{\partial^2 v}{\partial s^2} - \frac{1}{2}r^{*2}y^2\frac{\partial^2 v}{\partial y^2} + \sigma r^* s y\frac{\partial^2 v}{\partial s \partial y} = 0,$$

with the boundary condition $\widetilde{V}(T,s,y) = \widetilde{L}(y,s)$. Moreover, using the dominated convergence theorem, it is easy to check that one can differentiate $\widetilde{V}$ with respect to y inside the expectation operator in (8.3), so that for all $(t,s,y) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R}_+$,

$$\frac{\partial \widetilde{V}}{\partial y}(t,s,y) = -E\left\{\frac{Z_T^{r^*}}{Z_T^{r^*}}\left[\left(g\left(s\frac{S_T}{S_t}\right) - I\left(y\frac{Z_T^{r^*}}{Z_T^{r^*}}\right)\right)_+\right]\right\}.$$

From this last expression, we deduce that $\frac{\partial^2 \widetilde{V}}{\partial y^2} \geq 0$ and $\frac{\partial^2 \widetilde{V}}{\partial s \partial y} \leq 0$. Recalling that $\mu \geq 0$, we then see that $\widetilde{V}$ is solution to the Bellman equation (8.2). Now, by the dynamic programming principle, one proves by standard arguments [see, e.g., Fleming and Soner (1993)] that the lower semicontinuous envelope of $\widetilde{J}$ is a viscosity supersolution of the Bellman equation (8.2) and then by the maximum principle, $\widetilde{V} \leq \widetilde{J}$. Since the converse inequality is obvious by definition of $\widetilde{J}$, this shows that $\widetilde{V} = \widetilde{J}$ and r^* is the optimal associated (constant) control. $\square$

Proposition 8.1 shows that the solution to the dual problem does not depend on y and is equal to $Q^*(y) = P^{\mu/\sigma}$. This is the unique martingale measure of the unconstrained Black–Scholes model. Such a model was considered by Föllmer and Leukert (1999). Since their optimal strategy satisfies a posteriori the constraint $\theta_t^*(x) \geq 0$, then it is also the solution to the no short-sales constraints model.

8.2. *Price pressure in the Black–Scholes model.* We consider a variation of the Black–Scholes model where the drift of the underlying price process is affected by the investor's strategy for the following:

$$d\overline{S_t} = \overline{S_t}\big(\bar{\mu}(\theta_t)\,dt + \sigma\,dW_t\big),$$

where $\bar{\mu}(z) = \mu + f(z)$ and

$$f(z) = \begin{cases} -\dfrac{az}{|z|}, & \text{if } z \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

for some $\mu \in \mathbb{R}$, $\sigma > 0$ and $a > 0$. This means that buying stock depresses its expected return while shorting it increases the expected return. This example is taken from Cuoco and Cvitanić (1998). The wealth process starting from an initial wealth x and a strategy θ, amount invested in the stock, is then be written as

$$X_t^{x,\theta} = x + \int_0^t \frac{\theta_u}{\overline{S}_u} d\overline{S}_u = x + \int_0^t \theta_u(\mu\,du + \sigma\,dW_u) + \int_0^t \theta_u f(\theta_u)\,du.$$

We assume no constraints on trading strategies, $K = \mathbb{R}$. This is a particular case of the model considered in Section 7.3 with the function h equal to $h(t, z) = zf(z)$. The family $\mathcal{S} = \{X^{0,\theta} : \theta \in L(W)\}$ is a cone of semimartingales so that $\mathcal{P}$ is the set of probability measures $Q \sim P$ such that $X^{0,\theta}$ is a Q-local supermartingale for any $\theta \in L(W)$ and the upper variation process is zero. Keeping the same notations as in the previous paragraph, we deduce by Girsanov's theorem that the dynamics of $X^{0,\theta}$ under P^ρ, $\rho \in \mathcal{D}$, is

$$dX_t^{x,\theta} = \theta_t[f(\theta_t) + \mu - \sigma\rho_t]\,dt + \theta_t\sigma\,dW_t^\rho,$$

where $W^\rho = W + \int \rho\,dt$ is a P^ρ-Brownian motion. It follows that $P^\rho \in \mathcal{P}$ if and only if

$$\sup_{z\in\mathbb{R}}[zf(z) + z(\mu - \sigma\rho_t)] < \infty \quad \text{a.s.,}\ 0 \le t \le T,$$

or equivalently $|\mu - \sigma\rho_t| \le a$ a.s., $0 \le t \le T$. We obtain then the following characterization of $\mathcal{P}$:

$$\mathcal{P} = \{P^\rho = Z_T^\rho \cdot P : \rho \in \mathcal{D}_a\},$$

where

$$\mathcal{D}_a = \left\{\rho \in \mathcal{D} : \frac{\mu - a}{\sigma} \le \rho_t \le \frac{\mu + a}{\sigma} \text{ a.s., } 0 \le t \le T\right\}.$$

We consider the problem of minimizing the shortfall risk of a constant B in $\mathbb{R}$ (e.g., a riskless asset) for a loss function l satisfying assumptions of Section 6. Notice that from Corollary 4.1, the set of admissible trading strategies $\mathcal{A}(x)$ is nonempty iff $x \ge v_0(0) = 0$. The dynamic version of the dual control problem $(\mathcal{D}(y))$ is

$$\widetilde{J}(t, y) = \inf_{\rho\in\mathcal{D}_a} E\left[\overline{L}\left(y\frac{Z_T^\rho}{Z_T^\rho}\right)\right], \qquad (t, y) \in [0, T] \times \mathbb{R}_+, \tag{8.5}$$

where $\overline{L}(y) = -l(B \wedge I(y)) - y(B - I(y))_+$. We have $\widetilde{J}(y) = \widetilde{J}(0, y)$. The Bellman equation associated to this control problem is

$$-\frac{\partial v}{\partial t} + \sup_{\frac{\mu-a}{\sigma} \le r \le \frac{\mu+a}{\sigma}}\left[-\frac{1}{2}r^2y^2\frac{\partial^2 v}{\partial y^2}\right] = 0,$$

together with the terminal boundary condition $v(T, y) = \widetilde{L}(y)$. By similar arguments as in the proof of Proposition 8.1 (using convexity of the function $\widetilde{J}$ in y), one proves that the optimal control of (8.5) is given by the constant process,

$$r^* = \begin{cases} \dfrac{\mu - a}{\sigma}, & \text{if } \mu > a, \\ 0, & \text{if } \mu \le a. \end{cases}$$

8.3. *A reinsurance problem.* We consider an insurance company which reinsures a fraction $1 - \theta_t$ of the incoming claims. The times of arrival of the claims are modelled by a Poisson process (N_t) with constant intensity π and the magnitude of the incoming claims is constant, equal to $\delta \ge 0$. The premium rate per unit of time received by the company is a constant $\alpha \ge 0$ and the premium rate per unit of time paid by the company to the reinsurer is $\beta \ge \alpha$. We shall consider here a fair premium reinsurance rate, that is, $\beta = \delta\pi$. The risk process of the insurance company is governed by

$$\begin{aligned} dX_t^{x,\theta} &= [\alpha - \beta(1 - \theta_t)]\,dt - \theta_t \delta\, dN_t, \\ X_0^{x,\theta} &= x. \end{aligned}$$

The reinsurance trading strategy is constrained to remain in $K = [0, 1]$. This is a particular case of the model in Section 7.3 with the function h equal to $h(t, z) = \alpha - \beta(1 - z)$. Assuming that the probability space $(\Omega, \mathcal{F}, P)$ is equipped with the filtration generated by the Poisson process, it is well known from the martingale representation theorem for random measures [see, e.g., Brémaud (1981)] that all probability measures $Q \sim P$ have a Radon–Nikodym density of the form $dQ/dP = Z_T^\rho$ where

$$dZ_t^\rho = (\rho_t - 1) Z_t^\rho - d\widetilde{N}_t, \qquad Z_0^\rho = 1,$$

$\widetilde{N} = N - \int \pi\, dt$ is the P-compensated martingale of N and $\rho \in \mathcal{D} = \{(\rho_t)_{0 \le t \le T}$ predictable process: $\rho_t > 0$, a.s., $0 \le t \le T$, $\int_0^T |\ln \rho_t| + \rho_t dt < \infty$ and $E[Z_T^\rho] = 1\}$. By Girsanov's theorem, the dynamics of $X^{0,\theta}$ under $P^\rho := Z_T^\rho \cdot P$ is

$$dX_t^{0,\theta} = [\alpha - \beta + \beta\theta_t(1 - \rho_t)]\,dt - \theta_t \delta\, d\widetilde{N}_t^\rho,$$

where $\widetilde{N}_t^\rho = N_t - \int \rho\pi\, dt$ is the P^ρ-compensated martingale of N. The set $\mathcal{S}$ is equal to $\{(X^{0,\theta} - (\alpha - \beta)t)_{0 \le t \le T} : \theta \in \Theta_{\mathrm{loc}}\}$. The predictable compensator of $(X_t^{0,\theta} - (\alpha - \beta)t)_{0 \le t \le T}, \theta \in \Theta_{\mathrm{loc}}$, under P^ρ, $\rho \in \mathcal{D}$, is

$$A_t^{p,\theta} = \int_0^t \beta\theta_u(1 - \rho_u)\,du, \qquad 0 \le t \le T.$$

We then deduce that $\mathcal{P} = \{P^\rho : \rho \in \mathcal{D}\}$ and the upper variation process of P^ρ is

$$A_t(P^\rho) = \int_0^t \beta(1 - \rho_u)_+\,du, \qquad 0 \le t \le T.$$

We consider the problem of minimizing the shortfall risk of a constant B in $\mathbb{R}$ interpreted as a benchmark upper constant level for a loss function l satisfying assumptions of Section 6. By Corollary 4.1, the set of admissible reinsurance strategies $\mathcal{A}(x)$ is nonempty iff $x \geq b = (\beta - \alpha)T$. We fix now some initial reserve $x \geq b$. The dynamic version of the dual control problem $(\mathcal{D}(y))$ is then written as

$$\widetilde{J}(t, y) = \inf_{\rho \in \mathcal{D}} E\left[\widetilde{L}\left(y \frac{Z_T^\rho}{Z_T^\rho}\right) + y \int_t^T \frac{Z_u^\rho}{Z_t^\rho} (\beta(1 - \rho_u)_+ + \alpha - \beta)\, du\right], \tag{8.6}$$

for all $(t, y) \in [0, T] \times \mathbb{R}_+$, where $\widetilde{L}(y) = -l(B \wedge I(y)) - y(B - I(y))_+$. We have $\widetilde{J}(y) = \widetilde{J}(0, y)$. The Bellman equation associated to this control problem is

$$\begin{aligned} &-\frac{\partial v}{\partial t} + y(\beta - \alpha) \\ &\quad + \sup_{r > 0}\left[-\left(v(t, ry) - v(t, y) - (r-1)y\frac{\partial v}{\partial y}(t, y)\right)\pi - y\beta(1 - r)_+\right] = 0, \end{aligned}$$

together with the terminal boundary condition $v(T, y) = \widetilde{L}(y)$. By similar arguments as in the proof of Proposition 8.1 (using convexity of the function $\widetilde{J}$ in y), one can prove that the optimal control of (8.6) is the constant process $r^* \equiv 1$ corresponding to the probability measure $Q^*(y) = P$. Therefore, by Theorem (6.1), the solution of $(\mathbb{S}(x))$ is $X^*(x) \equiv x$ and so the solution to the shortfall risk minimization problem is the trivial reinsurance strategy $\theta^*(x) \equiv 0$. The insurance company optimally reinsures all of its incoming claims.

APPENDIX

A.1. Proof of Theorem 6.1. We recall that the function $\widetilde{L}$ is given by

$$\widetilde{L}(y, \omega) = \max_{0 \leq x \leq B} [-l(B - x) - xy] \tag{A.1}$$

$$= -l\big(B \wedge I(y)\big) - y\big(B - I(y)\big)_+, \qquad y \geq 0, \tag{A.2}$$

and the maximum in (A.1) is attained for

$$\chi(y, \omega) = \big(B - I(y)\big)_+, \qquad y \geq 0. \tag{A.3}$$

The function $\widetilde{L}(\cdot, \omega)$ is convex, differentiable on $(0, \infty)$ with derivative

$$\widetilde{L}'(y, \omega) = -\chi(y, \omega), \qquad y \geq 0. \tag{A.4}$$

(i) Let $y > 0$. Then for all $\delta > 0$, we have

$$\begin{aligned} \frac{\widetilde{J}(y + \delta) - \widetilde{J}(y)}{\delta} \leq \frac{1}{\delta} E\left[\widetilde{L}\left((y + \delta)\frac{dQ^*(y)}{dP}, \omega\right) - \widetilde{L}\left(y \frac{dQ^*(y)}{dP}, \omega\right)\right] \\ + E^{Q^*(y)}\big[A_T\big(Q^*(y)\big) + H_T^0\big] \end{aligned}$$

$$\leq -E^{Q^*(y)}\left[\chi\left((y+\delta)\frac{dQ^*(y)}{dP}, \omega\right)\right] + E^{Q^*(y)}\left[A_T\left(Q^*(y)\right) + H_T^0\right],$$

where we used (A.4) and convexity of $\widetilde{L}(\cdot, \omega)$. By Fatou's lemma, we deduce that

$$\limsup_{\delta \searrow 0^+} \frac{\widetilde{J}(y+\delta) - \widetilde{J}(y)}{\delta} \leq -E^{Q^*(y)}\left[\chi\left(y\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T\left(Q^*(y)\right)\right]. \tag{A.5}$$

Similarly, for all $\delta < 0,\ y+\delta > 0$, we have

$$\frac{\widetilde{J}(y+\delta) - \widetilde{J}(y)}{\delta} \geq -E^{Q^*(y)}\left[\chi\left((y+\delta)\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T\left(Q^*(y)\right)\right].$$

From the expression (A.3) of χ and since $-H_T^0$ is bounded from below, there exists a nonnegative constant cte (independent of δ) such that

$$\mathrm{cte} - A_T\left(Q^*(y)\right) \leq \chi\left((y+\delta)\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T\left(Q^*(y)\right) \tag{A.6}$$

$$\leq B - H_T^0 - A_T\left(Q^*(y)\right). \tag{A.7}$$

Therefore, under the assumption that $v_0(B) < \infty$ and (4.1), one can apply the dominated convergence theorem to deduce that

$$\liminf_{\delta \nearrow 0^-} \frac{\widetilde{J}(y+\delta) - \widetilde{J}(y)}{\delta} \geq -E^{Q^*(y)}\left[\chi\left(y\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T\left(Q^*(y)\right)\right]. \tag{A.8}$$

Relations (A.5)–(A.8) and convexity of the function $\widetilde{J}$ imply the differentiability of $\widetilde{J}$ and provide the expression (6.1) of $\overline{J}'$.

(ii) The function $y \mapsto f_x(y) = \widetilde{J}(y) + xy$ is convex on $(0, \infty)$. Let us check that

$$\lim_{y \to \infty} f_x(y) = \infty \qquad \forall x > v_0(0). \tag{A.9}$$

Indeed, by noting that $\widetilde{L}(y, \omega) \geq -l(B)$, we have

$$\begin{aligned}
\widetilde{J}(y) &\geq \inf_{Q \in \mathcal{P}} E\left[-l(B) - y\frac{dQ}{dP}\left(-H_T^0 - A_T(Q)\right)\right] \\
&= -E[l(B)] - y \sup_{Q \in \mathcal{P}} E^Q[-H_T^0 - A_T(Q)] \\
&= -E[l(B)] - yv_0(0).
\end{aligned}$$

We deduce that $f_x(y) \geq -E[l(B)] + y(x - v_0(0))$, which proves (A.9). We now check that for all $v_0(0) < x < v_0(B)$, there exists $y_0 > 0$ such that $f_x(y_0) < 0$. On the contrary, we should have

$$E\left[\widetilde{L}\left(y\frac{dQ}{dP}, \omega\right) + y\frac{dQ}{dP}\left(A_T(Q) + H_T^0\right)\right] + xy > 0 \qquad \forall y > 0,\ \forall Q \in \mathcal{P}$$

and then

$$x > E\left[-\frac{1}{y}\widetilde{L}\left(y\frac{dQ}{dP}, \omega\right) - \frac{dQ}{dP}\left(H_T^0 + A_T(Q)\right)\right] \qquad \forall y > 0,\ \forall Q \in \mathcal{P}.$$

Since $-\widetilde{L}(y dQ/dP, \omega)/y \geq 0$ and $-\widetilde{L}(y dQ/dP, \omega)/y$ converges to $B dQ/dP$ as y goes to infinity, we deduce by Fatou's lemma that

$$x \geq E^Q\left[B - H_T^0 - A_T(Q)\right] \qquad \forall Q \in \mathcal{P},$$

and then $x \geq v_0(B)$, a contradiction. Since $f_x(0) = 0$, we deduce that for all $v_0(0) < x < v_0(B)$, the function $f_x(\cdot)$ attains an infimum for $\hat{y} > 0$ and since $\widetilde{J}$, and so f_x, is differentiable on $(0, \infty)$, we have $f_x'(\hat{y}) = 0$; that is, $\widetilde{J}'(\hat{y}) = -x$.

Fix some $y > 0$ and let Q be an arbitrary element of $\mathcal{P}$. Denote

$$Q^\varepsilon = (1 - \varepsilon)Q^*(y) + \varepsilon Q, \qquad \varepsilon \in (0, 1).$$

By Lemma 6.1, $Q_\varepsilon \in \mathcal{P}$ so that by definition of $\widetilde{J}$, we have

$$\begin{aligned}
0 &\leq \frac{1}{\varepsilon}E\left[\frac{\widetilde{L}\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - \overline{L}\left(y\frac{dQ^*(y)}{dP}, \omega\right)}{y} + \left(\frac{dQ^\varepsilon}{dP} - \frac{dQ^*(y)}{dP}\right)H_T^0\right] \\
&\quad + \frac{1}{\varepsilon}E\left[\frac{dQ^\varepsilon}{dP}A_T(Q^\varepsilon) - \frac{dQ^*(y)}{dP}A_T\left(Q^*(y)\right)\right] \\
&\leq \frac{1}{\varepsilon}E\left[\frac{\widetilde{L}\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - \widetilde{L}\left(y\frac{dQ^*(y)}{dP}, \omega\right)}{y} + \left(\frac{dQ^\varepsilon}{dP} - \frac{dQ^*(y)}{dP}\right)H_T^0\right] \\
&\quad + E\left[\frac{dQ}{dP}A_T(Q) - \frac{dQ^*(y)}{dP}A_T\left(Q^*(y)\right)\right] \\
&\leq E\left[-\left(\frac{dQ}{dP} - \frac{dQ^*(y)}{dP}\right)\left(\chi\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - H_T^0\right)\right] \\
&\quad + E\left[\frac{dQ}{dP}A_T(Q) - \frac{dQ^*(y)}{dP}A_T\left(Q^*(y)\right)\right],
\end{aligned}$$

where the second inequality follows from the convexity of the function $Q \in \mathcal{P} \mapsto E^Q[A_T(Q)]$ (see Lemma 6.1) and the third inequality from (A.4) and the convexity of $\widetilde{L}$. We obtain then

$$\begin{aligned}
&E^Q\left[\chi\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - H_T^0 - A_T(Q)\right] \\
&\qquad \leq E^{Q^*(y)}\left[\chi\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - H_T^0 - A_T\left(Q^*(y)\right)\right].
\end{aligned} \tag{A.10}$$

Similarly as in (A.7), there exists a nonnegative constant cte independent of ε such that

$$\begin{aligned}\text{cte} - A_T\big(Q^*(y)\big) &\le \chi\left(y\frac{dQ^\varepsilon}{dP}, \omega\right) - H_T^0 - A_T(Q^\varepsilon)\\ &\le B - H_T^0 - A_T\big(Q^*(y)\big).\end{aligned}$$

Therefore, by the dominated convergence theorem and Fatou's lemma applied, respectively, to the right-hand side and left-hand side of (A.10), we have

$$\begin{aligned}&E^Q\left[\chi\left(y\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T(Q)\right]\\ &\qquad \le E^{Q^*(y)}\left[\chi\left(y\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T\big(Q^*(y)\big)\right].\end{aligned}$$

From (6.1) and (A.3), this can be written also as

$$\sup_{Q\in\mathcal{P}} E^Q\left[\chi\left(y\frac{dQ^*(y)}{dP}, \omega\right) - H_T^0 - A_T(Q)\right] \le -\widetilde{J}'(y),$$

for all $y > 0$. By choosing $y = \hat{y}$ defined above, we get

$$\sup_{Q\in\mathcal{P}} E^Q\left[\chi\left(\hat{y}\frac{dQ^*(\hat{y})}{dP}, \omega\right) - H_T^0 - A_T(Q)\right] \le x, \tag{A.11}$$

which proves that $X^*(x) = \chi(\hat{y}\frac{dQ^*(\hat{y})}{dP}, \omega)$ lies in $\mathcal{C}(x)$.

Moreover, by definition (A.1) of $\widetilde{L}$ and by definition of $X^*(x)$, we have for all $X \in \mathcal{C}(x)$,

$$\begin{aligned}&\widetilde{L}\left(\hat{y}\frac{dQ^*(\hat{y})}{dP}, \omega\right) + \hat{y}\frac{dQ^*(\hat{y})}{dP}\big(A_T\big(Q^*(\hat{y})\big) + H_T^0\big)\\ &\qquad = -l\big(B - X^*(x)\big) - \hat{y}\frac{dQ^*(\hat{y})}{dP}\big(X^*(x) - H_T^0 - A_T\big(Q^*(\hat{y})\big)\big)\end{aligned} \tag{A.12}$$

$$\ge -l(B - X) - \hat{y}\frac{dQ^*(\hat{y})}{dP}\big(X - H_T^0 - A_T\big(Q^*(\hat{y})\big)\big). \tag{A.13}$$

Taking expectation under P in (A.12) and (A.13) and using the facts that

$$E^{Q^*(\hat{y})}\big[X^*(x) - H_T^0 - A_T\big(Q^*(\hat{y})\big)\big] = -\widetilde{J}'(\hat{y}) = x, \tag{A.14}$$

$$E^{Q^*(\hat{y})}\big[X - H_T^0 - A_T\big(Q^*(\hat{y})\big)\big] \le x, \tag{A.15}$$

we obtain that

$$E\big[l\big(B - X^*(x)\big)\big] \le E[l(B - X)],$$

which proves that $X^*(x)$ is a solution to problem $(\mathbb{S}(x))$. Relations (A.11) and (A.14) show that $Q^*(\hat{y})$ attains the supremum in $\sup_{Q\in\mathcal{P}} E^Q[X^*(x) - H_T^0 -$

$A_T(Q)]$, and by Theorem 4.1, this proves that there exists $\theta^*(x) \in \mathcal{A}(x)$ such that $X^*(x) = X_T^{x,\theta^*(x)}$, a.s., $\theta^*(x)$ is a solution of the dynamic problem ($\mathbb{P}(x)$) and the associated controlled process is given as in assertion (ii) of Theorem 6.1.

(iii) By definition (A.1) of the function $\widetilde{L}$, we have for all $x \in \mathbb{R}$, $y \geq 0$, $X \in \mathcal{C}(x)$, $Q \in \mathcal{P}$,

$$-l(B - X) - y\frac{dQ}{dP}(X - H_T^0 - A_T(Q)) \leq \widetilde{L}\left(y\frac{dQ}{dP}, \omega\right) + y\frac{dQ}{dP}(A_T(Q) + H_T^0),$$

hence by taking expectation under P,

$$-E[l(B - X)] - yx \leq E\left[\widetilde{L}\left(y\frac{dQ}{dP}, \omega\right) + y\frac{dQ}{dP}(A_T(Q) + H_T^0)\right]$$

and therefore,

$$\sup_{y \geq 0}[-\widetilde{J}(y) - xy] \leq J(x) \qquad \forall x \in \mathbb{R}. \tag{A.16}$$

For $x \geq v_0(B)$, we have $J(x) = 0 = -\widetilde{J}(0)$. Fix now $v_0(0) < x < v_0(B)$. From relations (A.12) and (A.14), we have

$$\begin{aligned}\widetilde{J}(\hat{y}) &= -E\big[l\big(B - X^*(x)\big)\big] - \hat{y}E\left[\frac{dQ^*(\hat{y})}{dP}\big(X^*(x) - H_T^0 - A_T\big(Q^*(\hat{y})\big)\big)\right] \\ &= -J(x) - x\hat{y},\end{aligned}$$

which proves that $J(x) = -\widetilde{J}(\hat{y}) - x\hat{y}$. The proof is complete. $\square$

A.2. Proof of Theorem 7.1. (i) The convexity property of the set Θ follows from the convexity of the set K and the condition (7.5). For any predictable process ζ valued in [0, 1] and for all $\theta^1, \theta^2 \in \Theta$, we have

$$\begin{aligned}&H^{\zeta\theta^1 + (1-\zeta)\theta^2} - \int \zeta\, dH^{\theta^1} - \int (1-\zeta)\, dH^{\theta^2} \\ &\quad = \int \big(h\big(t, \zeta_t\theta_t^1 + (1-\zeta_t)\theta_t^2\big) - \zeta_t h(t, \theta_t^1) - (1-\zeta_t)h(t, \theta_t^2)\big)\, dt,\end{aligned}$$

which lies in $\mathcal{I}$ by the concave property of $h(t, \cdot)$. Hence, property (Hc) is satisfied.

(ii) To prove the closure property, we shall use the following properties of the semimartingale topology stated in Mémin (1980).

(P1) $(V^n)_n$ is a sequence of semimartingales converging to V in the semimartingale topology iff there exists a sequence [also denoted $(V^n)_n$] and a probability measure $Q \sim P$ with bounded density dQ/dP such that $(V^n)_n$ is a Cauchy sequence in $\mathcal{M}^2(Q) \oplus \mathcal{A}(Q)$ where $\mathcal{M}^2(Q)$ is the Banach space of Q-square integrable martingales and $\mathcal{A}(Q)$ is the Banach space of predictable processes with finite Q-integrable variation.

(P2) The semimartingale topology is invariant under a change of equivalent probability measure.

STEP 1. Let

$$U^n = X^{0,\theta^n} - C^n = \int \theta^n \, dS + \int h(t, \theta_t^v) \, dt - C^n, \qquad n \in \mathbb{N},$$

be a sequence converging to U in the semimartingale topology. By (P1) there exists a subsequence [also denoted $(U^n)_n$] and a probability measure $Q \sim P$ with bounded density dQ/dP such that $(U^n)_n$ is a Cauchy sequence in $\mathcal{M}^2(Q) \oplus \mathcal{A}(Q)$. Denote by $S_0 + M^Q + A^Q$ the canonical decomposition of S in the space $\mathcal{M}^2(Q) \oplus \mathcal{A}(Q)$. Then, the canonical decomposition of U^n in $\mathcal{M}^2(Q) \oplus \mathcal{A}(Q)$ is

$$U^n = \int \theta^n \, dM^Q + D^{Q,n},$$

where

$$D^{Q,n} = \int \theta^n \, dA^Q + \int h(t, \theta_t^n) \, dt - C^n.$$

Indeed, fix n and let $N + D$ be the canonical decomposition of U^n in $\mathcal{M}^2(Q) \oplus \mathcal{A}(Q)$. Since θ^n is integrable with respect to the Q-semimartingale S, there exists a decomposition $S = S_0 + M' + A'$ (under Q) such that $\int \theta^n \, dS = \int \theta^n \, dM' + \int \theta^n \, dA'$. Hence U^n can be decomposed into

$$U^n = \int \theta^n \, dM' + D'^n,$$

where

$$D'^n = \int \theta^n \, dA' + \int h(t, \theta_t^n) \, dt - C^n.$$

Since S (resp. U^n) is a special semimartingale, it follows that A' (resp. D'^n) has a locally integrable variation and its predictable compensator is A^Q (resp. D); see Proposition 2.14 in Jacod (1979). Therefore, the predictable compensator of $\int \theta^n \, dA'$ is $\int \theta^n \, dA^Q$ and since $\int h(t, \theta_t^n) \, dt - C^n$ is a predictable process, we deduce by uniqueness of the canonical decomposition that $D = \int \theta^n \, dA^Q + \int h(t, \theta_t^n) \, dt - C^n$ and so $N = \int \theta^n \, dM^Q$. We deduce that $\int \theta^n \, dM^Q$ converges in $\mathcal{M}^2(Q)$ to $\int \theta^n \, dM^Q$ for some predictable process θ, and $D^{Q,n}$ converges in $\mathcal{A}(Q)$.

STEP 2. Let us now prove that $\int \theta^n \, dS$ converges in the semimartingale topology to $\int \theta \, dS$. By (P2), it suffices to prove that $\int \theta^n \, dA^Q$ converges to $\int \theta \, dA^Q$ in $\mathcal{A}(R)$ for some probability measure $R \sim P$ defined later. Denote by $Z = (Z_t)_{0 \le t \le T}$ the density process of the probability measure Q. Since $Z_T =$

dQ/dP is bounded, it follows that $\langle Z, M\rangle$ exists and by Girsanov's theorem, we have

(A.17) $$M_t^Q = M_t - \int_0^t \frac{1}{Z_{u^-}}\, d\langle Z, M\rangle_u, \qquad 0 \le t \le T.$$

Since M is a continuous martingale, we also know that $\langle M^Q\rangle = \langle M\rangle = \int \sigma\, dt$. Since $S = S_0 + M + A = S_0 + M^Q + A^Q$, we deduce from (A.17) that

$$A_t^Q = A_t + \overline{A}_t, \qquad 0 \le t \le T,$$

where

$$\overline{A}_t = \int_0^t \frac{1}{Z_{u^-}}\, d\langle Z, M\rangle_u, \qquad 0 \le t \le T.$$

Now, by writing that $\int \theta^n\, dM^Q$ converges in $\mathcal{M}^2(Q)$ to $\int \theta\, dM^Q$ and that $\langle M^Q\rangle = \int \sigma\, dt$, we have

(A.18) $$E^Q\left[\int_0^T (\theta_t^n - \theta_t)'\sigma_t(\theta_t^n - \theta_t)\, dt\right] \to 0,$$

as n goes to infinity. In the following, we denote by $V(\cdot)$ the variation process of a finite variation process. From (7.2), the variation process of $\int (\theta^n - \theta)'\, dA$ satisfies

$$\begin{aligned}
V\left(\int (\theta^n - \theta)'\, dA\right)_T &= \int_0^T |(\theta_t^n - \theta_t)'\sigma_t\lambda_t|\, dt \\
&\le \int_0^T \sqrt{(\theta_t^n - \theta_t)'\sigma_t(\theta_t^n - \theta_t)}\sqrt{\lambda_t'\sigma_t\lambda_t}\, dt \\
&\le \left(\int_0^T \lambda_t'\sigma_t\lambda_t\, dt\right)^{1/2}\left(\int_0^T (\theta_t^n - \theta_t)'\sigma_t(\theta_t^n - \theta_t)\, dt\right)^{1/2}.
\end{aligned}$$

By condition (7.3) and relation (A.18), this shows that

(A.19) $$\int \theta^{n'}\, dA \to \int \theta'\, dA \quad \text{in } \mathcal{A}(Q),$$

as n goes to infinity.

Let us define the continuous predictable process with finite variation

$$G_t = \exp\left(-\int_0^t \frac{1}{Z_u^2}\, d\langle Z\rangle_u\right), \qquad 0 \le t \le T,$$

and denote $L_t = Z_t^2 G_t$, $0 \le t \le T$. By Itô's formula, we have

$$L_t = 1 + 2\int_0^t Z_u G_u\, dZ_u + \int_0^t G_u\, d([Z] - \langle Z\rangle)_u, \qquad 0 \le t \le T,$$

where $[Z]$ is the quadratic variation process of Z and $\langle Z\rangle$ its predictable compensator. Since G and Z are bounded, it follows that L is a strictly positive

P-martingale with $L_0 = 1$. Hence, it defines a probability measure $R \sim P$ with density process L,

$$\frac{dR}{dP}\bigg|_{\mathcal{F}_t} = L_t, \qquad 0 \le t \le T.$$

By the Kunita–Watanabe inequality and condition (7.1), we have

$$\begin{aligned}
&E^R\biggl[V\biggl(\int(\theta^n-\theta)'\,d\overline{A}\biggr)_T\biggr]\\
&\quad\le \biggl(E^R\biggl[\int_0^T(\theta_t^n-\theta_t)'\sigma_t(\theta_t^n-\theta_t)\,dt\biggr]\biggr)^{1/2}\biggl(E^R\biggl[\int_0^T\frac{1}{Z_t^2}\,d\langle Z\rangle_t\biggr]\biggr)^{1/2}\\
&\quad\le (\mathrm{cte})\biggl(E^Q\biggl[\int_0^T(\theta_t^n-\theta_t)'\sigma_t(\theta_t^n-\theta_t)\,dt\biggr]\biggr)^{1/2}\biggl(E\biggl[\int_0^T\frac{L_t}{Z_t^2}\,d\langle Z\rangle_t\biggr]\biggr)^{1/2}\\
&\quad\le \mathrm{cte}\biggl(E^Q\biggl[\int_0^T(\theta_t^n-\theta_t)'\sigma_t(\theta_t^n-\theta_t)\,dt\biggr]\biggr)^{1/2}(E\langle Z\rangle_T)^{1/2}\\
&\quad\le \mathrm{cte}\biggl(E^Q\biggl[\int_0^T(\theta_t^n-\theta_t)'\sigma_t(\theta_t^n-\theta_t)\,dt\biggr]\biggr)^{1/2},
\end{aligned}\tag{A.20}$$

where we used the fact that $dR/dQ = Z_T G_T$ is bounded, Bayes formula and the properties that $L_t/Z_t^2 = G_t \le 1$ and $E\langle Z\rangle_T = E(Z_T^2) < \infty$. By using (A.18), the last inequality (A.20) shows that

$$\int \theta^{n\prime}\,d\overline{A} \to \int \theta'\,d\overline{A} \quad \text{in } \mathcal{A}(R), \tag{A.21}$$

as n goes to infinity. Therefore, by (A.19), (A.21) and since dR/dQ is bounded, we deduce that

$$\int \theta^{n\prime}\,dA^Q \to \int \theta'\,dA^Q \quad \text{in } \mathcal{A}(R), \tag{A.22}$$

as n goes to infinity.

STEP 3. We now show that under one of the conditions (i) or (ii), $\int h(t, \theta_t^n)\times dt - C^n$ converges to $\int h(t,\theta_t)\,dt - C$, for some $C \in \mathcal{I}$, in the semimartingale topology. By (A.18), we have (possibly along a subsequence)

$$(\theta_t^n - \theta_t)'\sigma_t(\theta_t^n - \theta_t) \to 0 \quad \text{a.s. for all } t \in [0, T], \tag{A.23}$$

as n goes to infinity. Therefore, by condition (7.4), we deduce that

$$\theta_t^n \to \theta_t \quad \text{a.s. for all } t \in [0, T], \tag{A.24}$$

and so $\theta_t \in K$ a.s., for all t. Suppose first that condition (i): K is bounded, is satisfied. Then by (A.24) and the dominated convergence theorem, we have

$$E^Q\biggl[\int_0^T |h(t,\theta_t^n) - h(t,\theta_t)|\,dt\biggr] \to 0, \tag{A.25}$$

as n goes to infinity. Suppose now that condition (ii): σ_t is uniformly elliptic a.s. for all t, is satisfied. Then by (A.18), we deduce also that (A.25) holds. Therefore, this shows that $\theta \in \Theta$ and that $\int h(t, \theta_t^n)\,dt$ converges to $\int h(t, \theta_t)\,dt$ in $\mathcal{A}(Q)$ and in particular for the semimartingale topology. Moreover, since $D^{Q,n} = \int \theta^n\,dA^Q + \int h(t, \theta_t^n)\,dt - C^n$ converges in $\mathcal{A}(Q)$ and in particular in $\mathcal{A}(R)$, it follows from (A.22) that C^n also converges in $\mathcal{A}(R)$ to some predictable process C. Since $|C_t^n - C_t| \le V(C^n - C)_T$, we deduce that C_t^n converges to C_t a.s. (possibly along a subsequence) for all $t \in [0, T]$ and so C inherits the nondecreasing property of C^n, hence $C \in \mathcal{I}$. We have then proved that

$$\begin{aligned} U &= \int \theta\,dS + \int h(t, \theta_t)\,dt - C \\ &= X^{0,\theta} - C, \end{aligned}$$

which ends the proof. □

Acknowledgments. I am grateful to Yuri Kabanov and Christophe Stricker for comments and suggestions which helped to improve the first version of this paper. I also thank Peter Leukert for useful remarks. I am grateful to Nicole El Karoui, Monique Jeanblanc and Florin Avram for pointing out a mistake in an earlier version.

REFERENCES

BRÉMAUD, P. (1981). *Point Processes and Queues.* Springer, New York.

CUOCO, D. (1997). Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. *J. Econom. Theory* **72** 33–73.

CUOCO, D. and CVITANIĆ, J. (1998). Optimal consumption choice for a large investor. *J. Econom. Dynam. Control* **22** 401–436.

CVITANIĆ, J. (2000). Minimizing expected loss of hedging in incomplete and constrained markets. *SIAM J. Control Optim.* **38** 1050–1066.

CVITANIĆ, J. and KARATZAS, I. (1999). On dynamic measures of risk. *Finance Stoch.* **3** 451–482.

CVITANIĆ, J., SCHACHERMAYER, W. and WANG, H. (2000). Utility maximization in incomplete markets with random endowment. *Finance Stoch.* To appear.

DELBAEN, F. and SCHACHERMAYER, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520.

EL KAROUI, N. and JEANBLANC, M. (1998). Optimization of consumption with labor income. *Finance Stoch.* **4** 409–440.

EL KAROUI, N. and QUENEZ, M. C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.* **33** 29–66.

FLEMING, W. and SONER, M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Springer, New York.

FÖLLMER, H. and KRAMKOV, D. (1997). Optional decomposition under constraints. *Probab. Theory Related Fields* **109** 1–25.

FÖLLMER, H. and LEUKERT, P. (1999). Quantile hedging. *Finance Stoch.* **3** 251–273.

FÖLLMER, H. and LEUKERT, P. (2000). Efficient hedging: cost versus shortfall risk. *Finance Stoch.* **4** 117–146.

FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications* **1**. Academic Press, New York.

HEATH, D. (1995). A continuous-time version of Kulldorf's result. Preprint.

JACOD, J. (1979). Calcul stochastique et problémes de martingales. *Lecture Notes in Math.* **714**. Springer, Berlin.

KARATZAS, I. and SHREVE, S. (1998). *Methods of Mathematical Finance*. Springer, New York.

KRAMKOV, D. (1996). Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probab. Theory Related Fields* **105** 459–479.

KRAMKOV, D. and SCHACHERMAYER, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.* **9** 904–950.

KULLDORF, M. (1993). Optimal control of favourable games with a time-limit. *SIAM J. Control Optim.* **31** 52–69.

MÉMIN, J. (1980). Espaces de semimartingales et changement de probabilité. *Z. Wahrsch. Verw. Gebiete* **52** 9–39.

PHAM, H. (2000). Dynamic L^p-hedging in discrete-time under cone constraints. *SIAM J. Control Optim.* **38** 665–682.

SCHWEIZER, M. (1994). Approximating random variables by stochastic integrals. *Ann. Probab.* **22** 1536–1575.

SPIVAK, G. and CVITANIĆ, J. (1999). Maximizing the probability of a perfect hedge. *Ann. Appl. Probab.* **9** 1303–1328.

TOUZI, N. (2000). Optimal insurance demand under marked point processes shocks. *Ann. Appl. Probab.* **10** 283–312.

LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÉATOIRES
CNRS, UMR 7599
UFR MATHÉMATIQUES, CASE 7012
UNIVERSITÉ PARIS 7
2 PLACE JUSSIEU
75251 PARIS CEDEX 05
FRANCE
E-MAIL: pham@gauss.math.jussieu.fr