Minimizing Coherent Risk Measures Of Shortfall In Discrete-Time Models With Cone Constraints

Yumiharu Nakano

Series #547. February 2002

HOKKAIDO UNIVERSITY
PREPRINT SERIES IN MATHEMATICS

#520 A. Inoue, On the worst conditional expectation, 10 pages. 2001.

#521 Yumiharu Nakano, Efficient hedging with coherent risk measure, 10 pages. 2001.

#522 T. Nakazi, Toeplitz operators and weighted norm inequalities on the bidisc, 15 pages. 2001.

#523 T. Mikami, Covariance kernel and the central limit theorem in the total variation distance, 80 pages. 2001.

#524 K. Yamaguchi and T. Yatsui, Geometry of higher order differential equations of finite type associated with symmetric spaces, 43 pages. 2001.

#525 T. Suwa, Residues of Chern classes, 20 pages. 2001.

#526 V. Anh and A. Inoue, Dynamic models of asset prices with long memory, 21 pages. 2001.

#527 T. Izawa and T. Suwa, Multiplicity of functions on singular varieties, 21 pages. 2001.

#528 T. Nakazi and T. Yamamoto, Two dimensional commutative Banach algebras and von Neumann inequality, 18 pages. 2001.

#529 Y. Giga, N. Ishimura and Y. Kohsaka, Spiral solutions for a weakly anisotropic curvature flow equation, 16 pages. 2001.

#530 Y. Giga and P. Rybka, Quasi-static evolution of 3-D crystals grown from supersaturated vapor, 16 pages. 2001.

#531 Y. Tonegawa, Remarks on convergence of the Allen-Cahn equation, 18 pages. 2001.

#532 T. Suwa, Characteristic classes of singular varieties, 26 pages. 2001.

#533 J. Escher, Y. Giga and K. Ito, On a limiting motion and self-intersections for the intermediate surface diffusion flow, 20 pages. 2001.

#534 Y.-H. R. Tsai, Y. Giga and S. Osher, A level set approach for computing discontinuous solutions of a class of Hamilton-Jacobi equations, 30 pages. 2001.

#535 A. Yamagami, On Gouvêa's conjecture in the unobstructed case, 19 pages. 2001.

#536 A. Inoue, What does the partial autocorrelation function look like for large lags, 27 pages. 2001.

#537 T. Nakazi and T. Yamamoto, Norm of a linear combination of two operators of a Hilbert space, 16 pages. 2001.

#538 Y. Giga, On the two-dimensional nonstationary vorticity equations, 12 pages. 2001.

#539 M. Jinzenji, Gauss-Manin system and the virtual structure constants, 25 pages. 2001.

#540 H. Ishii and T.Mikami, Motion of a graph by R-curvature, 28 pages. 2001.

#541 M. Jinzenji and T. Sasaki, $\mathcal{N} = 4$ supersymmetric Yang-Mills theory on orbifold-$T^4/\mathbf{Z}_2$: higher rank case, 17 pages. 2001.

#542 T. Nakazi, The Nevanlinna counting functions for Rudin's orthogonal functions, 7 pages. 2001.

#543 K. Sugano, On H-separable extensions of QF-3 rings, 7pages. 2001.

#544 A. Arai, Non-relativistic limit of a Dirac-Maxwell operator in relativistic quantum electrodynamics, 27pages. 2001.

#545 O. Sawada, On time-local solvability of the Navier-Stokes equations in Besov spaces, 30 pages. 2001.

#546 C. M. Elliott, Y. Giga, and S. Goto, Dynamic boundary conditions for Hamilton-Jacobi equations, 27 pages. 2001.

MINIMIZING COHERENT RISK MEASURES OF SHORTFALL IN DISCRETE-TIME MODELS WITH CONE CONSTRAINTS

Yumiharu Nakano
Department of Mathematics, Faculty of Science
Hokkaido University, Sapporo 060-0810, Japan

January 22, 2002

Abstract

This paper studies the problem of minimizing coherent risk measures of shortfall for general discrete-time financial models with cone-constrained trading strategies, as developed by Pham and Touzi (1999) and Pham (1999). We show that the optimal strategy is obtained by super-hedging a contingent claim, which is represented as a Neyman-Pearson-type random variable.

KEY WORDS: coherent risk measure, shortfall risk, constrained strategy, super-hedging, convex duality.

AMS 2000 Subject Classification: Primary, 91B28; Secondary, 93E20, 49J52

1 Introduction

It is well-known that a frictionless, discrete-time and finite-horizon model of financial market is arbitrage-free if and only if there exists an equivalent martingale measure (the first fundamental theorem of asset pricing). Moreover, the arbitrage-free market is complete if and only if the equivalent martingale measure is unique (the second fundamental theorem of asset pricing). See, e.g., Shiryaev [23, Chapter V]. In arbitrage-free and complete models, any contingent claim H is attainable, that is, there exists a trading strategy ξ^H such that the self-financed wealth process $V(x_0, \xi^H)$ is worth H at the maturity date T. The cost x_0 of replication is given by the expectation of H under the unique martingale measure.

Under the market incompleteness or in the presence of frictions, the standard no-arbitrage arguments are no longer available, and some contingent claims may not be attainable. However, even so, we can still *super-hedge* such claims: starting with an enough initial wealth x_0, an agent can find a trading strategy ξ^H such that

$$V_T(x_0, \xi^H) \geq H, \quad \text{almost surely.}$$

The strategy ξ^H is called a *super-hedging strategy for H*. Then, the super-replication cost x_0 is given by the supremum of expectations of H over a suitable set of probability measures (see, e.g., El Karoui and Quenez [8], Föllmer and Kabanov [9], Karatzas [13], Schäl [21] and the references cited there). We define the *shortfall risk* by the following net profit of hedging loss:

$$-(H - V_T(x, \xi))_+ ,$$

where $(a)_+ = \max(a, 0)$. For a super-hedging strategy ξ^H, the shortfall risk is equal to zero, that is,

$$-\left(H - V_T(x_0, \xi^H)\right)_+ = 0.$$

However, if an investor has only an initial wealth x less than x_0, then he or she cannot necessarily accomplish the super-hedging, that is, the shortfall risk may not be equal to zero.

If we are in a position of such an investor, then we have to accept the possibility of shortfall. As a result, we wish to measure the shortfall risk by some risk measure ρ, and try to find an optimal strategy that solves the minimization problem

$$\inf_{\xi} \rho\left(-(H - V_T(x, \xi))_+\right).$$

See, e.g., Cvitanić [4], Föllmer and Leukert [10], [11], and Pham [18]. These references adopted the $\rho(X) = E^P[l(-X)]$ as risk measures, where $l(\cdot)$ is a loss function and P is the objective probability. They studied the minimization problem

$$\inf_{\xi} E^P\left[l\left((H - V_T(x, \xi))_+\right)\right]$$

in complete or incomplete market models.

In this paper, we work in the general discrete-time models with convex cone-constrained trading strategies. Such models were studied in [18], and Pham and Touzi [19]. Under an assumption, the market model is arbitrage-free if and only if the set $\mathcal{M}^e$ of all equivalent probability measures under which the discounted price process satisfies a generalized martingale property is not empty (see Theorem 2.3 below). Moreover, in such a market, the super-hedging as stated above is possible. Theorem 2.4 below, due to Pham [18], characterizes the super-replication cost of H by the supremum of expectations of H over $\mathcal{M}^e$. In particular, in the case of short-selling prohibition, the set $\mathcal{M}^e$ is given by the set of all "equivalent supermartingale measures".

As risk measures, we use the coherent risk measures introduced by Artzner, Delbaen, Eber, and Heath [2]. They are real-valued functions on a suitable space of random variables satisfying four desirable properties, that is, monotonicity, subadditivity, positive homogeneity, and translation invariance. We recall the precise definition in Section 3. Given a coherent risk measure ρ and a contingent claim H, we study the stochastic control problem

$$\inf_{\xi} \rho\left(-\left(H - V_T(x,\xi)\right)_+\right). \tag{1.1}$$

See Nakano [17], where a similar optimization problem for general frictionless continuous-time semimartingale models is studied.

It is known that a coherent risk measure ρ of a random variable X arises as the supremum of the expected negative of X over a set of "real-world" probability measures or "scenarios", that is,

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E^Q[-X]$$

for some set of probability measures $\mathcal{Q}$ (see [2, Proposition 4.1], Delbaen [7, Theorem 2.3]; see also [17, Theorem 1.2] and Inoue [12, Theorem 1.1]). Thus the problem (1.1) is equivalent to the minimization problem

$$\inf_{\xi} \sup_{Q \in \mathcal{Q}} E^Q\left[\left(H - V_T(x,\xi)\right)_+\right].$$

For a special scenario set $\mathcal{Q}$, problems of this type are studied in Cvitanić and Karatzas [6], and Sekine [20] for frictionless, essentially complete, continuous-time models. The problem (1.1) that consider in this paper is different from those in these references because of the presence of constraints on trading strategies (and the discrete-time setting).

In our approach to the problem (1.1), we use the methods of convex duality and super-hedging. The supermartingale property of $V(x,\xi)$ and the super-hedging method enable us to reduce the dynamic problem (1.1) to the following static one:

$$\inf_{X \in \mathcal{X}(x)} \rho\left(-(H - X)\right), \tag{1.2}$$

where the infimum is taken over a suitable set $\mathcal{X}(x)$ of random variables. To solve the static problem (1.2), we use the convex duality method (cf. [4], Cvitanić and Karatzas [5]). After we enlarge $\mathcal{Q}$ and $\mathcal{M}^e$ to suitable classes $\mathcal{Z}$ and $\mathcal{G}$, respectively, we define the auxiliary dual problem to (1.2) by

$$\sup\left\{E^P\left[H\left(Z \wedge yG\right)\right] - xy : Z \in \mathcal{Z},\ G \in \mathcal{G},\ y \geq 0\right\}, \tag{1.3}$$

where $a \wedge b = \min(a,b)$. Following the manner in [5], we prove the existence of a solution $(\hat{Z}, \hat{G}, \hat{y})$ to the dual problem (1.3). Using the triple, we show that there exists a solution to the problem (1.2) of the form:

$$\hat{X} = H\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + HB\mathbf{1}_{\{\hat{y}\hat{G} = \hat{Z}\}}, \tag{1.4}$$

where B is some random variable taking values in $[0,1]$. This result is similar to [5, Theorem 4.1].

In conclusion, our approach may be summarized as follows: First, we solve the auxiliary dual problem (1.3) to obtain a triple $(\hat{Z}, \hat{G}, \hat{y})$. Next, we find a random variable B such that $\hat{X}$ of the form (1.4) is a solution to the static problem (1.2). Finally, we construct a super-hedging strategy $\hat{\xi}$ for $\hat{X}$. Then, the resulting strategy $\hat{\xi}$ solves (1.1).

This paper is organized as follows: In Section 2, we present the general framework and the basic results for the discrete-time models with cone-constrained trading strategies. We prepare coherent risk measures to be used in Section 3. In Section 4, we formulate the minimization problem (1.1), and reduce the dynamic problem (1.1) to the static one (1.2). Then we present the semi-closed form solution as in (1.4) to the problem (1.2) via the convex duality method. We also give a simple illustrating example. Section 5 is devoted to the proof of a key to the main theorem.

2 Discrete-time models with constraints

Let $T \in \mathbf{N}$. The discounted price process of d stocks is described as an $\mathbf{R}^d$-valued adapted stochastic process $S = \{S_t,\ t = 0, \ldots, T\}$ on some filtered probability space $((\Omega, \mathcal{F}, P), (\mathcal{F}_t)_{t=0,\ldots,T})$. We shall assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and that $\mathcal{F}_T = \mathcal{F}$. For a probability measure Q on $(\Omega, \mathcal{F})$, we denote by E^Q the expectation with respect to Q. We write $L^1(P)$ for $L^1(\Omega, \mathcal{F}, P)$. Here we consider only real-valued function. We assume that $S_t \in L^1(P)$ for $t = 0, \ldots, T$.

For $\mathbf{R}^d$-valued process $Y = (Y_t)_{t=0,\ldots,T}$, we define $\Delta Y_t := Y_t - Y_{t-1}$ for $t = 1, \ldots, T$. We write $a \cdot b$ for the inner product of $a, b \in \mathbf{R}^d$: $a \cdot b := \sum_{k=1}^d a_k b_k$ for $a = (a_k)$ and $b = (b_k)$.

A trading strategy is an $\mathbf{R}^d$-valued predictable process $\xi = (\xi_t)_{t=1,\ldots,T}$. Here, for $i = 1, \ldots, d$, ${\xi_t}^i$ represents the number of shares of the stock S^i held by the investor during $(t-1, t]$ for $t = 1, \ldots, T$. We denote by Ξ the set of all trading strategies. For $x \in \mathbf{R}$ and $\xi \in \Xi$, we define the discounted (self-financed) wealth process $V_t(x, \xi)$ by

$$\begin{aligned} V_t(x, \xi) &= x + \sum_{k=1}^{t} \xi_k \cdot \Delta S_k, \quad t = 1, \ldots, T, \\ V_0(x, \xi) &= x. \end{aligned}$$

Let C be a nonempty closed convex cone in $\mathbf{R}^d$. A constrained trading strategy ξ is a trading strategy such that

$$\xi_t \in C, \quad t = 1, \ldots, T, \quad P \text{ a.s.}$$

We denote by $\Xi(C)$ the set of all constrained trading strategies. Since C is a cone, the (negative) polar cone C° of C is given by

$$C^\circ = \left\{ b \in \mathbf{R}^d : a \cdot b \leq 0 \ \ (\forall a \in C) \right\}$$

(see, e.g., [3, Chapter 1, Section 5]).

Example 2.1. We have the following examples of constraint sets:

(1) *Unconstrained case*: $C = \mathbf{R}^d$. Then $C^\circ = \{0\}$.

(2) *Prohibition of short-selling of some stocks*:

$$C = \left\{a \in \mathbf{R}^d : a^i \geq 0 \ \ (\forall i \in I)\right\},$$

where I is a subset of $\{1, \ldots, d\}$. Then C° is given by

$$C^\circ = \left\{b \in \mathbf{R}^d : b^i \leq 0 \ \ (i \in I), \ b^i = 0 \ \ (i \in \{1, \ldots, d\} \setminus I)\right\}.$$

We define the following convex cone associated with $\Xi(C)$:

$$K = \{V_T(0, \xi) : \xi \in \Xi(C)\}.$$

Denote by $L^0_+(P)$ the space of all nonnegative random variables. We define the *no arbitrage condition* as follows:

Definition 2.2. We say that there is *no arbitrage opportunity* if

$$K \cap L^0_+(P) = \{0\}. \tag{NA}$$

We define the subset $\Xi(x)$ of $\Xi(C)$ by

$$\Xi(x) = \{\xi \in \Xi(C) : V_T(x, \xi) \geq 0, \ P\text{–a.s.}\}.$$

Thus we consider the admissibility condition that imposes the nonnegativity constraint only on the terminal wealth value. Denote $L^\infty(\Omega, \mathcal{F}, P)$ by $L^\infty(P)$, and let $\mathcal{P}$ be the set of all probability measures on $(\Omega, \mathcal{F})$ absolutely continuous with respect to P. As in [18], we introduce the following set of "martingale measures":

$$\mathcal{M}(P) = \left\{ Q \in \mathcal{P} : \begin{array}{l} dQ/dP \in L^\infty(P) \ \text{ and} \\ E^Q[\Delta S_t | \mathcal{F}_{t-1}] \in C^\circ, \ t = 1, \ldots, T, \ \ Q\text{–a.s.} \end{array} \right\},$$
$$\mathcal{M}^e(P) = \{Q \in \mathcal{M}(P) : Q \sim P\}.$$

Writing $V_t(x, \xi) = V_{t-1}(x, \xi) + \xi_t \cdot \Delta S_t$, it is easily seen that for any $\xi \in \Xi(x)$, the process $V(x, \xi)$ is a supermartingale under any $Q \in \mathcal{M}(P)$. In the unconstrained case $C = \mathbf{R}^d$, the set $\mathcal{M}^e(P)$ is actually the set of equivalent probability measures with density in $L^\infty(P)$ under which S is a martingale. On the other hand, in the no short-selling constraints case $C = [0, \infty)^d$, the set $\mathcal{M}^e(P)$ is the set of equivalent probability measures with density in $L^\infty(P)$ under which S is a supermartingale.

Following [18] and [19], we shall make a nondegeneracy assumption on the price process. Set, for $t = 1, \ldots, T$,

$$N(t-1) = \left\{\eta \in L^{0,d}(\mathcal{F}_{t-1}, P) : \eta \cdot \Delta S_t = 0, \ \ P\text{–a.s.}\right\},$$

where $L^{0,d}(\mathcal{F}_t, P)$ denotes the space of all $\mathbf{R}^d$-valued $\mathcal{F}_t$-measurable random variables. We assume that

$$N(t-1) = \{0\} \quad (t = 1, \ldots, T). \tag{B}$$

The condition (B) is satisfied by most standard financial models. For example, the Cox-Ross-Rubinstein model satisfies (B) (see [18, Section 3 and 5]).

Under the condition (B), the no arbitrage condition (NA) implies the "no free lunch" condition (see [18] and [19]). Further, we have the following extended version of first fundamental theorem of asset pricing:

Theorem 2.3 (Pham and Touzi [19]). *Assume* (B). *Then* (NA) *holds if and only if* $\mathcal{M}^e(P) \neq \emptyset$.

We define the super-replication cost x_0 of a nonnegative contingent claim $H \in L^1(P)$ by

$$x_0 = \inf \left\{ x \in \mathbf{R} : V_T(x, \xi) \geq H, \ P\text{—a.s., for some } \xi \in \Xi(x) \right\}.$$

In what follows, we write $\mathcal{M} = \mathcal{M}(P)$ and $\mathcal{M}^e = \mathcal{M}^e(P)$ for simplicity. The next theorem, due to Pham [18], provides a duality result between the initial wealth and the expectation of the contingent claim under probability measures $\mathcal{M}^e$, within a general discrete-time framework with cone constraints.

Theorem 2.4 (Pham [18]). *Assume* (NA) *and* (B). *Then the super-replication cost of a nonnegative contingent claim* $H \in L^1(P)$ *is given by*

$$x_0 = \sup \left\{ E^Q[H] : Q \in \mathcal{M}^e \right\}. \tag{2.1}$$

In (2.1), *we may replace* $\mathcal{M}^e$ *by* $\mathcal{M}$. *Moreover, if* $\sup_{Q \in \mathcal{M}^e} E^Q[H] < \infty$, *then there exists* $\xi^H \in \Xi(x_0)$ *such that* $V_T(x_0, \xi^H) \geq H$, P—a.s. *The strategy* ξ^H *is called a super-hedging strategy for* H.

Remark 2.5. For a super-hedging strategy ξ^H in Theorem 2.4 and $x \geq x_0$, we easily see that $V_T(x, \xi^H) \geq H$.

3 Coherent risk measures

Let us consider a nonnegative contingent claim $H \in L^1(P)$. Following a trading strategy ξ and starting with an initial wealth x, an agent loses the resulting shortfall $(H - V_T(x, \xi))_+$ at time $t = T$. Thus his or her shortfall risk is $-(H - V_T(x, \xi))_+$. We are concerned with the problem of minimizing a coherent risk measure of this shortfall risk.

Definition 3.1 ([2] and [7]). We say that a functional $\rho : L^1(P) \to \mathbf{R}$ is a *coherent risk measure* if the following are satisfied:

(i) For $X \in L^1(P)$ with $X \geq 0$, we have $\rho(X) \leq 0$.

(ii) For X and $Y \in L^1(P)$, we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

(iii) If $X \in L^1(P)$ and $\lambda \in (0, \infty)$, then $\rho(\lambda X) = \lambda \rho(X)$.

(iv) If $X \in L^1(P)$ and $c \in \mathbf{R}$, then $\rho(X + c) = \rho(X) - c$.

Remark 3.2. We refer to [2] for interpretations of the above properties. In [2], they restrict themselves to finite probability spaces. Subsequently, in [7], the definition of coherent risk measures is extended to general probability spaces. In [7], as the space of random variables, the space $L^\infty(P)$ or the space $L^0(P)$ of all random variables is adopted. As in [17], we use the intermediate space $L^1(P)$ here. The space $L^1(P)$ is large enough to be used in our hedging problem since the payoff of, e.g., a European call option belongs to $L^1(P)$ since we have assumed that $S_T \in L^1(P)$.

Let $\rho : L^1(P) \to \mathbf{R}$ be a coherent risk measure that is continuous in the L^1-norm. Recall $\mathcal{P}$ from Section 2. Then, we have the following representation:

$$\rho(X) = \sup_{Q\in\mathcal{Q}} E^Q[-X] \qquad (X \in L^1(P)), \tag{3.1}$$

where

$$\mathcal{Q} = \left\{Q \in \mathcal{P} : dQ/dP \in L^\infty(P),\ E^Q[-X] \le \rho(X) \ \ (\forall X \in L^1(P))\right\}. \tag{3.2}$$

See [17, Theorem 1.2]; see also [2, Proposition 4.1], [7, Theorem 2.3], and [12, Theorem 1.1]. So, every continuous coherent risk measure arises as the supremum of expected negatives of a random variable over a set $\mathcal{Q}$ of "real-world" probability measures or "scenarios". Then, from the uniform boundedness theorem, we easily see that

$$\sup_{Q\in\mathcal{Q}} \|dQ/dP\|_\infty < \infty. \tag{3.3}$$

Remark 3.3. In [17, Theorem 1.2], it is proved that the representation (3.1) also holds for *lower semi-continuous* coherent risk measures ρ. However, even for such ρ, (3.3) still holds by the uniform boundedness theorem. This implies that every lower semi-continuous coherent risk measure turns out to be continuous (see [12, Lemma 2.1]).

We need the next property of coherent risk measures.

Definition 3.4 ([2] and [7]). We say that a coherent risk measure $\rho : L^1(P) \to \mathbf{R}$ is *relevant* if, for every $A \in \mathcal{F}$ with $P(A) > 0$, we have $\rho(-\mathbf{1}_A) > 0$.

The next proposition, which is an analogue of [7, Theorem 3.5], characterizes relevant coherent risk measures.

Proposition 3.5. *Let $\rho : L^1(P) \to \mathbf{R}$ be a continuous coherent risk measure, and let $\mathcal{Q}$ be as in* (3.2). *Then the following are equivalent*:

(i) *ρ is relevant.*

(ii) *The set $\{Q \in \mathcal{Q} : Q \sim P\}$ is not empty.*

Proof. It is easy to prove the implication (ii) $\Rightarrow$ (i). To prove the converse one (i) $\Rightarrow$ (ii), we follow the line of the proof of [7, Theorem 3.5]. Set $\mathcal{Y} = \{dQ/dP : Q \in \mathcal{Q}\}$ and take a sequence $\{Y_n\}_{n=1}^{\infty}$ from $\mathcal{Y}$ so that

$$\lim_{n\to\infty} P(Y_n > 0) = \sup_{Y\in\mathcal{Y}} P(Y > 0).$$

We define the probability measure $\hat{Q} = \hat{Y}P$ by

$$\hat{Y} = \sum_{n=1}^{\infty} 2^{-n} Y_n.$$

Then, since $\{\hat{Y} > 0\} = \cup_{n=1}^{\infty}\{Y_n > 0\}$, we have

$$P(\hat{Y} > 0) \geq \sup_{Y\in\mathcal{Y}} P(Y > 0).$$

However, by (3.3), $\hat{Y} \in L^{\infty}(P)$, whence $\hat{Q} \in \mathcal{Q}$. Thus

$$P(\hat{Y} > 0) = \sup_{Y\in\mathcal{Y}} P(Y > 0). \tag{3.4}$$

Now suppose that $P(\hat{Y} = 0) > 0$. Since ρ is relevant, there exists $Y_0 \in \mathcal{Y}$ such that $E^P[Y_0 \mathbf{1}_{\{\hat{Y}=0\}}] > 0$. This implies

$$P(Y_0 > 0,\ \hat{Y} = 0) > 0.$$

However, if we put $Y' := (\hat{Y} + Y_0)/2 \in \mathcal{Y}$, then we find from (3.4) that

$$\begin{aligned} P(Y_0 > 0,\ \hat{Y} = 0) &= P(\hat{Y} = 0) - P(\hat{Y} = 0,\ Y_0 = 0) \\ &= P(\hat{Y} = 0) + P\Big(\{\hat{Y} > 0\} \cup \{Y_0 > 0\}\Big) - 1 \\ &= P(\hat{Y} = 0) + P(Y' > 0) - 1 \\ &= P(\hat{Y} = 0) + P(\hat{Y} > 0) - 1 = 0, \end{aligned}$$

which is a contradiction. Thus $\hat{Q}$ and P are equivalent. □

Remark 3.6. The *worst conditional expectation* is a typical example of coherent risk measures. Given $\alpha \in (0, 1)$, this measure is defined by, for $X \in L^1(P)$,

$$\mathrm{WCE}_{\alpha}(X) = \sup\big\{E^Q[-X] : Q(\cdot) = P(\cdot|A),\ P(A) > \alpha,\ A \in \mathcal{F}\big\}$$

(see [2, Section 5] and [7, Section 4]). WCE_{α} is the law-invariant, smallest coherent risk measure dominating the value at risk VaR_{α}, which is a popular risk measure, but not a coherent one. Now we define another coherent risk measure $\rho_{\alpha}(\cdot)$ on $L^1(P)$ by

$$\rho_{\alpha}(X) = \sup_{Q\in\mathcal{Q}} E^Q[-X] \qquad (X \in L^1(P)), \tag{3.5}$$

where

$$\mathcal{Q} = \{Q \in \mathcal{P} : dQ/dP \leq 1/\alpha\} .$$

As mentioned in [7, Section 4], if the underlying probability space $(\Omega, \mathcal{F}, P)$ is nonatomic, then ρ_α coincides with WCE_α on $L^1(P)$ since the extreme points of $\mathcal{Q}$ are of the form $\mathbf{1}_A/P(A)$ with $P(A) = \alpha$ (see Lindenstrauss [16]). We easily see that ρ_α is relevant and continuous in the L^1-norm. For related work, we refer to Acerbi and Tasche [1], [12], Kusuoka [15], and [20]. In particular, [20] studies the minimization problem of $\rho_\alpha\left(-\left(H - V_T(x,\xi)\right)_+\right)$ in complete continuous-time market models. In many situations, it seems more convenient to use ρ_α than to use WCE_α itself.

4 The minimization

Let $\rho : L^1(P) \to \mathbf{R}$ be a relevant, continuous coherent risk measure. Then ρ has a representation of the form (3.1). We consider a nonnegative contingent claim $H \in L^1(P)$ that satisfies

$$x_0 := \sup_{Q \in \mathcal{M}} E^Q[H] < +\infty.$$

In this section, we assume the no arbitrage condition (NA) and the nondegeneracy condition (B). Then, it follows from Theorems 2.3 and 2.4 that $\mathcal{M}^e \neq \emptyset$ and that the super-replication cost of H is given by $\sup_{Q \in \mathcal{M}} E^Q[H]$.

Now let $x > 0$ and $\xi_1, \xi_2 \in \Xi(x)$. If

$$\rho\left(-\left(H - V_T(x,\xi_1)\right)_+\right) \leq \rho\left(-\left(H - V_T(x,\xi_2)\right)_+\right)$$

or, equivalently,

$$\sup_{Q \in \mathcal{Q}} E^Q\left[\left(H - V_T(x,\xi_1)\right)_+\right] \leq \sup_{Q \in \mathcal{Q}} E^Q\left[\left(H - V_T(x,\xi_2)\right)_+\right],$$

then we may regard the strategy ξ_1 as preferable to ξ_2. Therefore, an agent, who uses the coherent risk measure ρ as measure of shortfall risk, wishes to minimize $\rho(-(H - V_T(x,\xi))_+)$ for a given initial wealth x. Thus we consider the following optimization problem: for $x > 0$,

$$R(x) := \inf_{\xi \in \Xi(x)} \rho\left(-\left(H - V_T(x,\xi)\right)_+\right) . \tag{4.1}$$

Remark 4.1. For $x \geq x_0$, the super-hedging strategy ξ^H is obviously a solution to the problem (4.1) since $\left(H - V_T(x,\xi^H)\right)_+ = 0$.

We define the set $\mathcal{X}(x)$ by

$$\mathcal{X}(x) = \left\{X \in L^1(P) : 0 \leq X \leq H,\ P\text{–a.s.},\ E^P[GX] \leq x,\ (\forall G \in \mathcal{M})\right\} . \tag{4.2}$$

Here we identify each probability measure $Q \in \mathcal{M}$ with its Radon-Nikodým density $G = dQ/dP$. As in [11, Proposition 3.1], [17, Proposition 1.3], and [18, Proposition 4.1], we reduce the dynamic problem (4.1) to a static one by the next proposition.

Proposition 4.2. *Suppose that $\hat{X}$ is a solution to the static problem*

$$\inf_{X\in\mathcal{X}(x)} \rho(-(H-X)). \tag{4.3}$$

Then there exists a super-hedging strategy $\hat{\xi}\in\Xi(x)$ for $\hat{X}$ that solves the dynamic problem (4.1)*. Moreover, we have*

$$R(x) = \inf_{X\in\mathcal{X}(x)} \rho(-(H-X)).$$

Proof. The proof is similar to that of [18, Proposition 4.1]. Let $\xi\in\Xi(x)$ and $X := H-(H-V_T(x,\xi))_+$. Then $X\in L^1(P)$ and $X\le V_T(x,\xi)$. Now, under every $Q\in\mathcal{M}$, the process $V(x,\xi)$ is a supermartingale, so that $X\in\mathcal{X}(x)$. So we obtain

$$\rho\big(-\big(H-V_T(x,\xi)\big)_+\big) = \rho(-(H-X)) \ge \inf_{X'\in\mathcal{X}(x)} \rho(-(H-X')).$$

Since ξ is an arbitrary element of $\Xi(x)$, we find that

$$R(x) \ge \inf_{X\in\mathcal{X}(x)} \rho(-(H-X)). \tag{4.4}$$

Conversely, suppose that $\hat{X}\in\mathcal{X}(x)$ solves the problem (4.3). We put

$$\hat{x} := \sup_{G\in\mathcal{M}} E^P[G\hat{X}].$$

Then $\hat{x}\le x<+\infty$, and so Theorem 2.4 implies that there exists a super-hedging strategy $\hat{\xi}\in\Xi(x)$ for $\hat{X}$ such that

$$V_T(x,\hat{\xi}) \ge \hat{X}.$$

Since $-(H-\hat{X}) \le -\big(H-V_T(x,\hat{\xi})\big)_+$, we have, from the monotonicity of ρ, that

$$R(x) \le \rho\big(-\big(H-V_T(x,\hat{\xi})\big)_+\big) \le \rho(-(H-\hat{X})). \tag{4.5}$$

The proposition follows from (4.4) and (4.5). $\square$

In what follows, we assume

$$0<x<x_0,$$

and study the static problem (4.3) for the initial wealth x. We follow the convex duality method as in [4] and [5].

We define

$$\mathcal{Z} = \left\{ Z\in L^\infty(P) : \begin{array}{l} Z\ge 0\ (P\text{–a.s.}),\quad E^P[Z]\le 1, \\ E^P[ZX]\le \rho(-X) \quad (\forall X\in L^1_+(P)) \end{array} \right\},$$

where $L^1_+(P)$ denotes the space of all nonnegative random variables in $L^1(P)$. Since $\mathcal{Q}\subset\mathcal{Z}$, it holds that

$$\rho(-X) = \sup_{Z\in\mathcal{Z}} E^P[ZX] \qquad (X\in L^1_+(P)). \tag{4.6}$$

Remark 4.3. In the case of ρ_α in (3.5) with (4.6), we can show that

$$\mathcal{Z} = \{Z \in L^\infty(P) : 0 \le Z \le 1/\alpha,\ E^P[Z] \le 1\}.$$

The next proposition is needed to prove Lemmas 4.7 and 5.1.

Proposition 4.4. *Let $\mathcal{Z}$ be as above.*

(i) *The set $\mathcal{Z}$ is convex and closed under $P-$a.s. convergence.*

(ii) *If a sequence $\{Z_n\}_{n=1}^\infty$ from $\mathcal{Z}$ converges to a random variable Z, $P-$a.s., on the set $\{H > 0\}$ as $n \to \infty$, and if $Z = 0$ on the set $\{H = 0\}$, then we have $Z \in \mathcal{Z}$.*

Proof. Suppose that a sequence $\{Z_n\}$ from $\mathcal{Z}$ converges to a random variable Z, $P-$a.s. Then, by Fatou's lemma, we have

$$E^P[ZX] \le \liminf_{n\to\infty} E^P[Z_n X] \le \rho(-X) < \infty \quad (X \in L^1_+(P)).$$

From this, it easily follows that $Z \in L^\infty(P)$. Similarly, by Fatou's lemma, we have $E^P[Z] \le 1$. Thus $Z \in \mathcal{Z}$, and so (i) follows.

Let $\{Z_n\}$ and Z be as in (ii). By Fatou's lemma, we see that, for $X \in L^1_+(P)$,

$$\begin{aligned} E^P[ZX] &= E^P\left[ZX\mathbf{1}_{\{H>0\}}\right] \le \liminf_{n\to\infty} E^P\left[Z_n X \mathbf{1}_{\{H>0\}}\right] \\ &\le \liminf_{n\to\infty} E^P[Z_n X] \le \rho(-X) < \infty. \end{aligned}$$

In particular, $Z \in L^\infty(P)$. Similarly, we have that $E^P[Z] \le 1$. Thus $Z \in \mathcal{Z}$. This proves (ii). □

As in [4, Section 3] and [5, Section 3], we introduce the set

$$\mathcal{G} := \left\{ G \in L^1_+(P) : \begin{array}{l} E^P[G] \le 1,\ \ E^P[GH] \le x_0, \\ E^P[GX] \le x \ \ (\forall X \in \mathcal{X}(x)) \end{array} \right\}$$

(recall that we have assumed that $x \in (0, x_0)$).

Proposition 4.5. *We have the following:*

(i) *The set $\mathcal{G}$ is convex, closed under $P-$a.s. convergence, and bounded in $L^1(P)$, and it includes the convex hull of $\mathcal{M}$, that is,*

$$\mathrm{conv}\left\{\frac{dQ}{dP}\right\}_{Q\in\mathcal{M}} \subset \mathcal{G}.$$

(ii) *If a sequence $\{G_n\}_{n=1}^\infty$ from $\mathcal{G}$ converges to some random variable G, $P-$a.s., on the set $\{H > 0\}$, and if $G = 0$ on the set $\{H = 0\}$, then we have $G \in \mathcal{G}$.*

Proof. As in the proof of Proposition 4.4, we obtain Proposition 4.5 using Fatou's lemma several times. $\square$

Remark 4.6. The properties (i) and (ii) in Proposition 4.5 are needed to prove Lemma 4.7 and Lemma 5.1, respectively. Therefore, for this purpose, we may replace $\mathcal{G}$ by another set $\mathcal{G}'$ if it satisfies (i) and (ii) in Proposition 4.5. For example, consider the one-period models with no-short-selling constraints. Then we know that

$$\mathcal{M} = \left\{Q \in \mathcal{P} : E^Q[S_1] \leq S_0\right\}. \tag{4.7}$$

We easily find that the set $\mathcal{G}'$ defined by

$$\mathcal{G}' := \left\{G' \in L^1_+(P) : E^P[G'S_1] \leq S_0,\ E^P[G'] \leq 1\right\}$$

satisfies the properties (i) and (ii) in Proposition 4.5 for every H, whence we may replace $\mathcal{G}$ by $\mathcal{G}'$.

Now, as in [5, Section 3], we have the following important observation: for $Z \in \mathcal{Z}$, $G \in \mathcal{G}$, $y \geq 0$, and $X \in \mathcal{X}(x)$,

$$\begin{aligned} E^P\left[Z(H-X)\right] &= E^P[ZH] - E^P\left[X(Z-yG)\right] - yE^P[GX] \\ &\geq E^P[ZH] - E^P\left[H(Z-yG)_+\right] - xy \\ &= E^P\left[H(Z \wedge yG)\right] - xy. \end{aligned} \tag{4.8}$$

We define

$$f(y) = \sup_{Z \in \mathcal{Z}, G \in \mathcal{G}} E^P\left[H(Z \wedge yG)\right] \qquad (y \geq 0); \tag{4.9}$$

$$g(x) = \sup_{y \geq 0}\left(f(y) - xy\right). \tag{4.10}$$

Then we have

$$g(x) \leq \inf_{X \in \mathcal{X}(x)} \sup_{Z \in \mathcal{Z}} E^P\left[Z(H-X)\right] = \inf_{X \in \mathcal{X}(x)} \rho\left(-(H-X)\right). \tag{4.11}$$

Lemma 4.7. *For each $y \geq 0$, there exists a pair $(Z_y, G_y) \in \mathcal{Z} \times \mathcal{G}$ that attains the supremum in* (4.9).

Proof. Let $\{(Z_n, G_n)\}_{n=1}^\infty$ be a sequence from $\mathcal{Z} \times \mathcal{G}$ such that

$$\lim_{n \to \infty} E^P\left[H(Z_n \wedge yG_n)\right] = f(y).$$

Since the set $\mathcal{Z} \times \mathcal{G}$ is bounded in $L^1(P) \times L^1(P)$, the theorem of Komlós [14] (see also Schwartz [22]) implies that there exists a pair $(Z_y, G_y) \in L^1(P) \times L^1(P)$ and a relabeled subsequence $\left\{(Z'_j, G'_j)\right\}_{j=1}^\infty$ of $\{(Z_n, G_n)\}_{n=1}^\infty$ such that

$$\left(\frac{1}{k}\sum_{j=1}^k Z'_j, \frac{1}{k}\sum_{j=1}^k G'_j\right) \to (Z_y, G_y) \quad (k \to \infty), \quad P\text{-a.s.}$$

By the $P-$a.s. closedness of $\mathcal{Z}$ and $\mathcal{G}$, we have $(Z_y, G_y) \in \mathcal{Z} \times \mathcal{G}$. Since

$$|E^P[ZX]| \leq \rho(-|X|) < +\infty \quad (X \in L^1(P), Z \in \mathcal{Z}),$$

it follows from the uniform boundedness theorem that

$$\sup_{Z \in \mathcal{Z}} \|Z\|_\infty < \infty. \tag{4.12}$$

From this, as well as Lebesgue's convergence theorem and the concavity of the function $(s,t) \mapsto s \wedge t$, we obtain

$$\begin{aligned} E^P\left[H(Z_y \wedge yG_y)\right] &= \lim_{k\to\infty} E^P\left[H\left\{\left(\frac{1}{k}\sum_{j=1}^k Z'_j\right) \wedge y\left(\frac{1}{k}\sum_{j=1}^k G'_j\right)\right\}\right] \\ &\geq \lim_{k\to\infty} \frac{1}{k}\sum_{j=1}^k E^P\left[H(Z'_j \wedge yG'_j)\right] \\ &= \lim_{j\to\infty} E^P\left[H(Z'_j \wedge yG'_j)\right] = f(y). \end{aligned}$$

Thus (Z_y, G_y) attains the supremum, as desired. □

Lemma 4.8. *The function $f(\cdot)$ is concave.*

Proof. Let $y_1, y_2 \in [0,\infty)$ and $t \in (0,1)$. If $y_1 = y_2 = 0$ then the concavity immediately follows. So we assume that $y_1 > 0$ or $y_2 > 0$. By Lemma 4.7, there exist $(Z_{y_j}, G_{y_j}) \in \mathcal{Z} \times \mathcal{G}$, $j = 1,2$, such that

$$E^P\left[H(Z_{y_j} \wedge y_j G_{y_j})\right] = f(y_j) \quad (j = 1,2).$$

We put $G := (ty_1 G_{y_1} + (1-t)y_2 G_{y_2}) \,/\, (ty_1 + (1-t)y_2) \in \mathcal{G}$. Then it follows that

$$\begin{aligned} & tf(y_1) + (1-t)f(y_2) \\ &= E^P\left[H\left\{t(Z_{y_1} \wedge y_1 G_{y_1}) + (1-t)(Z_{y_2} \wedge y_2 G_{y_2})\right\}\right] \\ &\leq E^P\left[H\left\{(tZ_{y_1} + (1-t)Z_{y_2}) \wedge (ty_1 G_{y_1} + (1-t)y_2 G_{y_2})\right\}\right] \\ &= E^P\left[H\left\{(tZ_{y_1} + (1-t)Z_{y_2}) \wedge (ty_1 + (1-t)y_2)G\right\}\right] \\ &= f\left(ty_1 + (1-t)y_2\right). \end{aligned}$$

Thus we obtain the lemma. □

Lemma 4.9. *There exists $\hat{y} \equiv \hat{y}_x > 0$ that attains the supremum in* (4.10).

Proof. By Lemma 4.8, the function

$$h(y) := f(y) - xy \quad (y \geq 0)$$

is concave. From (4.12), we see that $h(+\infty) = -\infty$. Clearly, we have

$$h(0) = 0, \quad h(y) \geq -xy \ \ (y > 0).$$

We claim that there exists $y_0 > 0$ such that $h(y_0) > 0$. Suppose otherwise. Then $h(y) \leq 0$ for $y > 0$. Since ρ is relevant, by Proposition 3.5 there exists $Z \in \mathcal{Z}$ such that $Z > 0$, $P-$a.s. We see that, for every $G \in \mathcal{G}$,

$$xy \geq f(y) \geq E^P\left[H(Z \wedge yG)\right] \quad (\forall y > 0).$$

Dividing by y and then letting $y \downarrow 0$, we have that, for any $G \in \mathcal{G}$,

$$E^P[HG] \leq x.$$

However, this contradicts the assumption $x < x_0$ (recall that $x_0 = \sup_{G \in \mathcal{G}} E^P[GH]$). Thus the claim is proved. The lemma now follows from the concavity of $h(\cdot)$. $\square$

Remark 4.10. As mentioned in Remark 4.6, the necessity of considering $\mathcal{Z}$ and $\mathcal{G}$ rather than $\mathcal{Q}$ and $\mathcal{M}$, respectively, is to ensure the existence of a solution to the dual problem $g(x)$.

Now, here is our main theorem.

Theorem 4.11. *Let $\hat{y} > 0$ be as in Lemma 4.9, and let $(\hat{Z}, \hat{G}) \equiv (Z_{\hat{y}}, G_{\hat{y}})$ be an optimal pair for the problem* (4.9) *with $y = \hat{y}$.*

(i) *There exists a $[0,1]$-valued random variable B such that the random variable*

$$\hat{X} := H\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + HB\mathbf{1}_{\{\hat{y}\hat{G} = \hat{Z}\}}$$

is a solution to the static problem (4.3). *Moreover, there is no "duality gap" in* (4.11), *that is, $g(x) = R(x)$.*

(ii) *If $\hat{\xi}$ is a super-hedging strategy for $\hat{X}$, then $\hat{\xi}$ is a solution to* (4.1).

The following proposition is a key to the proof of Theorem 4.11.

Proposition 4.12. *There exists a $[0,1]$-valued random variable B such that $\hat{X}$ of the form in Theorem 4.11 satisfies $\hat{X} \in \mathcal{X}(x)$ as well as the conditions*

$$E^P\left[\hat{G}\hat{X}\right] = x, \tag{4.13}$$

$$\sup_{Z \in \mathcal{Z}} E^P\left[Z(H - \hat{X})\right] = E^P\left[\hat{Z}(H - \hat{X})\right]. \tag{4.14}$$

We prove this proposition in Section 5, following the method of [4] and [5]. In this method, some results from the non-smooth convex analysis (Aubin and Ekeland [3]) are used.

Proof of Theorem 4.11. Using (4.13) and the fact $\hat{X}(\hat{Z} - \hat{y}\hat{G}) = H(\hat{Z} - \hat{y}\hat{G})_+$ in (4.8) with $Z = \hat{Z}$, $X = \hat{X}$, $G = \hat{G}$, and $y = \hat{y}$, we see that

$$\begin{aligned} E^P\left[\hat{Z}(\hat{H} - \hat{X})\right] &= E^P\left[H(\hat{Z} \wedge \hat{y}\hat{G})\right] - x\hat{y} \\ &= f(\hat{y}) - x\hat{y} = g(x), \end{aligned}$$

whence, by (4.14),

$$g(x) = \sup_{Z \in \mathcal{Z}} E^P\left[Z(H - \hat{X})\right] = \rho\left(-(H - \hat{X})\right).$$

Therefore, from (4.11), it follows from that

$$\rho\left(-(H - \hat{X})\right) \leq \inf_{X \in \mathcal{X}(x)} \rho\left(-(H - X)\right).$$

However, $\hat{X}$ is in $\mathcal{X}(x)$, so that $\hat{X}$ attains the infimum above. In particular, by Proposition 4.2, we find that $g(x) = R(x)$ and that (ii) holds. Thus the theorem follows. □

We illustrate Theorem 4.11 by using one-period binomial models.

Example 4.13. Let $0 < d < 1 < u$ and $0 < \pi < 1$. We consider the probability space $\Omega = \{u, d\}$, $\mathcal{F} = 2^\Omega$, $P(\{u\}) = \pi = 1 - P(\{d\})$. Let $\{S_0, S_1\}$ be the discounted price process described as

$$S_0 = 1, \quad S_1 = \begin{cases} u & \text{if } \omega = u, \\ d & \text{if } \omega = d. \end{cases}$$

We consider the no-short-selling constraints. From (4.7), we see that $Q \in \mathcal{M}$ if and only if $Q(u) \leq \hat{\pi} := (1 - d)/(u - d)$.

Let $\alpha \in (0, 1)$. We take ρ_α in (3.5) as the measure of risk here. We consider the European call $H = (S_1 - K)_+$ with $d < K < u$. Then it follows from Remark 4.3 that

$$\mathcal{Z} = \left\{Z : 0 \leq Z \leq \alpha^{-1},\ \pi Z(u) + (1 - \pi)Z(d) \leq 1\right\}.$$

We put

$$\mathcal{G}' = \{G : 0 \leq G(u) \leq \hat{\pi}/\pi,\ \pi G(u) + (1 - \pi)G(d) \leq 1\}.$$

Then $\mathcal{G}'$ satisfies the properties (i) and (ii) in Proposition 4.5. Here, as mentioned in Remark 4.6, we consider $\mathcal{G}'$ rather than $\mathcal{G}$. The function $f(\cdot)$ in (4.9) with $\mathcal{G}$ replaced by $\mathcal{G}'$ is given by

$$f(y) = \sup_{Z \in \mathcal{Z}, G \in \mathcal{G}'} \pi(u - K)\left(Z(u) \wedge yG(u)\right).$$

We easily find that, for $y > 0$, this supremum is attained by, e.g.,

$$\hat{Z}(\omega) = \left(\frac{1}{\alpha} \wedge \frac{1}{\pi}\right) \mathbf{1}_{\{u\}}(\omega), \quad \hat{G}(\omega) = \hat{\pi}\mathbf{1}_{\{u\}}(\omega).$$

Then the dual problem (4.10) is written as

$$g(x) = \sup_{y \geq 0} \left[\pi(u-K)\left\{\left(\alpha^{-1} \wedge \pi^{-1}\right) \wedge y\hat{\pi}\right\} - xy\right].$$

Since $x < x_0 = \hat{\pi}(u-K)$, $\hat{y} = \pi(\alpha^{-1} \wedge \pi^{-1})/\hat{\pi}$ attains this supremum. Therefore, from Theorem 4.11, we deduce that the random variable

$$\begin{aligned} \hat{X} &:= H\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + HB\mathbf{1}_{\{\hat{y}\hat{G} = \hat{Z}\}} \\ &= (u-K)B\mathbf{1}_{\{u\}} \end{aligned}$$

is optimal for some $[0,1]$-valued random variable B. However, from the condition (4.13), we obtain

$$B(u) = \frac{x}{\hat{\pi}(u-K)},$$

whence

$$\hat{X} = \frac{x}{\pi}\mathbf{1}_{\{u\}} = \frac{x}{x_0}H.$$

5 Proof of Proposition 4.12

This section is devoted to the proof of Proposition 4.12. Following [4] and [5], we introduce the Banach space

$$\mathcal{K} := L^1(P) \times L^1(P) \times \mathbf{R}$$

with norm

$$\|(U, V, y)\|_{\mathcal{K}} := E^P\left[|U| + |V|\right] + |y|$$

and its subset

$$\mathcal{L} := \left\{(HZ, yHG, y) \in \mathcal{K} : Z \in \mathcal{Z},\ G \in \mathcal{G},\ y \geq 0\right\}.$$

Lemma 5.1. *The set $\mathcal{L}$ is convex and closed in $\mathcal{K}$.*

Proof. Let $t \in (0,1)$ and $(HZ_j, y_jHG_j, y_j) \in \mathcal{L}$ for $j = 1, 2$. If $y_1 = y_2 = 0$, the convexity immediately follows. Assume that $y_1 > 0$ or $y_2 > 0$. Then we have

$$G := \frac{ty_1G_1 + (1-t)y_2G_2}{ty_1 + (1-t)y_2} \in \mathcal{G},$$

so that $(H(tZ_1 + (1-t)Z_2), H(ty_1G_1 + (1-t)y_2G_2), ty_1 + (1-t)y_2) \in \mathcal{L}$. Thus the convexity follows.

Next, let $(HZ_n, y_nHG_n, y_n) \in \mathcal{L}$ be a sequence that converges to some (U, V, y) in $\mathcal{K}$. Then $y_n \to y$, and $HZ_n \to U$, $P-$a.s. (possibly along a subsequence). Put

$$Z := \begin{cases} U/H & \text{on } \{H > 0\}, \\ 0 & \text{on } \{H = 0\}. \end{cases}$$

Then we have that $U = HZ$ on $\{H > 0\}$ and that $Z_n \to Z$ on $\{H > 0\}$, as $n \to \infty$, $P-$a.s. The property (ii) in Proposition 4.4 implies $Z \in \mathcal{Z}$.

On the other hand, we have

$$E^P[H|y_nG_n - yG_n|] \leq |y_n - y|x_0 \to 0 \quad (n \to \infty),$$

hence

$$E^P[H|yG_n - W|] \to 0 \quad (n \to \infty),$$

where, we set $W = V/H$ on $\{H > 0\}$, and $= 0$ on $\{H = 0\}$. If $y = 0$, then $W = 0 = yHG_1$. If $y > 0$ and if we put $G := W/y$, then we have that $G_n \to G = W/y$, $P-$a.s., on the set $\{H > 0\}$ (possibly along a subsequence) and that $G = 0$ on the set $\{H = 0\}$. The property (ii) in Proposition 4.5 implies $G \in \mathcal{G}$. Thus the closedness of $\mathcal{L}$ follows. □

We now define a functional $\Phi : \mathcal{K} \to \mathbf{R}$ by

$$\Phi(U, V, y) = xy - E^P[U \wedge V].$$

Lemma 5.2. *Let $\hat{Z}$, $\hat{G}$, and $\hat{y}$ be as in Section 4. The functional Φ is proper, convex and lower semi-continuous on $\mathcal{K}$ and attains its infimum over $\mathcal{L}$ at the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$.*

Proof. Since the proof of the properness, convexity, and lower semi-continuity are simple, we omit them. By Lemmas 4.7 and 4.9, we have

$$\begin{aligned} \Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) &= x\hat{y} - E^P\left[H(\hat{Z} \wedge \hat{y}\hat{G})\right] = x\hat{y} - f(\hat{y}) \\ &\leq xy - f(y) \leq xy - E^P[H(Z \wedge yG)] \\ &= \Phi(HZ, yHG, y), \qquad \forall (HZ, yHG, y) \in \mathcal{L}. \end{aligned}$$

Thus the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ is optimal. □

Proof of Proposition 4.8. The methods of the proof is almost same as in [4, Section 3] and [5, Section 4 and 6]. We now consider the dual space

$$\mathcal{K}^* := L^\infty(P) \times L^\infty(P) \times \mathbf{R}$$

of $\mathcal{K}$, and the normal cone (see [3, Definition 4.1.3 and Proposition 4.1.4])

$$\mathcal{N}(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) := \left\{ (T, W, \zeta) \in \mathcal{K}^* : \begin{array}{l} E^P[H\hat{Z}T + \hat{y}H\hat{G}W] + \hat{y}\zeta \geq \\ E^P[UT + VW] + y\zeta, \quad (\forall (U, V, y) \in \mathcal{L}) \end{array} \right\}$$

to the set $\mathcal{L}$ at the point $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) \in \mathcal{L}$. We also consider the subdifferential at this point

$$\partial\Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) := \left\{ (T, W, \zeta) \in \mathcal{K}^* : \begin{array}{r} \Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) - \Phi(U, V, y) \leq \\ E^P \left[T(H\hat{Z} - U) + W(\hat{y}H\hat{G} - V) \right] \\ + \zeta(\hat{y} - y), \quad (\forall (U, V, y) \in \mathcal{K}) \end{array} \right\}.$$

Then, from Lemma 5.2, it follows that the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ solves the problem

$$\inf_{(U,V,y) \in \mathcal{M}} \Phi(U, V, y),$$

so that, by [3, Corollary 4.6.3], there exists a triple $(\hat{T}, \hat{W}, \hat{\zeta}) \in \mathcal{K}^*$ such that $(\hat{T}, \hat{W}, \hat{\zeta}) \in \mathcal{N}(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ and $(-\hat{T}, -\hat{W}, -\hat{\zeta}) \in \partial\Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$. These are equivalent to the following, respectively:

$$E^P \left[H\hat{T}(\hat{Z} - Z) + H\hat{W}(\hat{y}\hat{G} - yG) \right] + \hat{\zeta}(\hat{y} - y) \geq 0, \qquad \forall (Z, G, y) \in \mathcal{Z} \times \mathcal{G} \times [0, +\infty), \tag{5.1}$$

$$E^P \left[\hat{T}(U - H\hat{Z}) + \hat{W}(V - \hat{y}H\hat{G}) + H(\hat{Z} \wedge \hat{y}\hat{G}) - U \wedge V \right] \geq (x + \hat{\zeta})(\hat{y} - y), \qquad \forall (U, V, y) \in \mathcal{K}. \tag{5.2}$$

By letting $y \to \pm\infty$, we see that (5.2) holds only if

$$\hat{\zeta} = -x. \tag{5.3}$$

From (5.1) with $\hat{\zeta} = -x$, $Z = \hat{Z}$, $G = \hat{G}$, and $y = \hat{y} \pm \delta$ $(\delta > 0)$, we have

$$E^P[\hat{G}H\hat{W}] = x. \tag{5.4}$$

On the other hand, (5.1) with $Z = \hat{Z}$ and $y = \hat{y}$ implies

$$E^P[GH\hat{W}] \leq E^P[\hat{G}H\hat{W}], \quad \forall G \in \mathcal{G}, \tag{5.5}$$

and (5.1) with $G = \hat{G}$ and $y = \hat{y}$ implies

$$E^P[ZH\hat{T}] \leq E^P[\hat{Z}H\hat{T}], \quad \forall Z \in \mathcal{Z}. \tag{5.6}$$

Reading (5.2) with

$$U = H\hat{Z} - \mathbf{1}_{\{\hat{W}+\hat{T}>1\}}, \quad V = \hat{y}H\hat{G} - \mathbf{1}_{\{\hat{W}+\hat{T}>1\}}$$

and using (5.3), we have

$$E^P[(\hat{W}+\hat{T}-1)_+] \leq 0.$$

Similarly we obtain $E^P\left[(1-\hat{W}-\hat{T})_+\right] \leq 0$, whence

$$\hat{W}+\hat{T} = 1. \tag{5.7}$$

Hence, the conditions (5.1) and (5.2) can be written as

$$E^P\left[H\hat{W}(\hat{y}\hat{G} - yG + Z - \hat{Z}) + H(Z-\hat{Z})\right] \geq x(\hat{y}-y),$$
$$\forall (Z,G,y) \in \mathcal{Z}\times\mathcal{G}\times[0,+\infty),$$

$$E^P\left[\hat{W}(V-U+H\hat{Z}-\hat{y}H\hat{G}) + U - H\hat{Z} + H(\hat{Z}\wedge\hat{y}\hat{G}) - U\wedge V\right] \geq 0, \tag{5.8}$$
$$\forall (U,V) \in L^1(P)\times L^1(P).$$

Considering (5.8) for $U = H\hat{Z}$, $V = \hat{y}H\hat{G} + \mathbf{1}_A$ with arbitrary $A \in \mathcal{F}$, we see that

$$0 \leq E^P\left[\hat{W}\mathbf{1}_A + H(\hat{Z}\wedge\hat{y}\hat{G}) - H\hat{Z}\wedge(\hat{y}H\hat{G}+\mathbf{1}_A)\right] \leq E^P[\hat{W}\mathbf{1}_A],$$

so that $\hat{W} \geq 0$, $P-$a.s. Similarly we get

$$0 \leq E^P\left[(1-\hat{W})\mathbf{1}_A\right],$$

whence $\hat{W} \leq 1$, $P-$a.s. Therefore

$$0 \leq \hat{W} \leq 1 \qquad P\text{-a.s.} \tag{5.9}$$

The conditions (5.5) and (5.9) imply $H\hat{W} \in \mathcal{X}(x)$. The condition (5.8) for $U = V$ implies

$$E^P\left[H\hat{W}(\hat{Z}-\hat{y}\hat{G})\right] \geq E^P\left[H(\hat{Z}-\hat{y}\hat{G})_+\right]. \tag{5.10}$$

Hence, (5.9) and (5.10) lead to $H\hat{W}(\hat{Z}-\hat{y}\hat{G}) = H(\hat{Z}-\hat{y}\hat{G})_+$, $P-$a.s., so that

$$\hat{W} = 1 \quad \text{on} \quad \{\hat{y}\hat{G} < \hat{Z}\}\cap\{H>0\}, \quad = 0 \quad \text{on} \quad \{\hat{y}\hat{G} > \hat{Z}\}\cap\{H>0\}.$$

Thus, there exist $[0,1]$-valued random variables B and J such that

$$\hat{W} = \mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z},H>0\}} + B\mathbf{1}_{\{\hat{y}\hat{G}\geq\hat{Z},H>0\}} + J\mathbf{1}_{\{H=0\}}. \tag{5.11}$$

Reading (5.8) with $\hat{W}$ in (5.11), $U = HU'$, and $V = HV'$, we have

$$E^P \Big[\Big(\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + B\mathbf{1}_{\{\hat{y}\hat{G} \geq \hat{Z}, H>0\}} \Big) H(U - V + \hat{Z} - \hat{y}\hat{G}) \\ + H(U - \hat{Z} + \hat{Z} \wedge \hat{y}\hat{G} - U \wedge V) \Big] \geq 0, \quad \forall U, V \in L^\infty(P).$$

In particular, reading $U = \hat{Z}$, we have that, for $V \in L^\infty(P)$,

$$E^P \left[H \left(\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + B\mathbf{1}_{\{\hat{y}\hat{G} \geq \hat{Z}, H>0\}} \right) (V - \hat{y}\hat{G}) + H(\hat{Z} \wedge \hat{y}\hat{G} - \hat{Z} \wedge V) \right] \geq 0.$$

So, for $V \in L^\infty(P)$, we obtain

$$\begin{aligned} & E^P \left[BH\mathbf{1}_{\{\hat{y}\hat{G} \geq \hat{Z}\}} (\hat{y}\hat{G} - V) \right] \\ & \leq E^P \left[H(V - \hat{y}\hat{G})\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + H(\hat{Z} \wedge \hat{y}\hat{G} - \hat{Z} \wedge V) \right] \\ & = E^P \Big[H(V - \hat{y}\hat{G})\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + H(\hat{y}\hat{G} - V)\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}, V \leq \hat{Z}\}} \\ & \quad + H(\hat{y}\hat{G} - \hat{Z})\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z} < V\}} + H(\hat{Z} - V)\mathbf{1}_{\{\hat{y}\hat{G} \geq \hat{Z} \geq V\}} \Big] \\ & = E^P \left[H(V - \hat{Z})\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z} < V\}} + H(\hat{Z} - V)\mathbf{1}_{\{\hat{y}\hat{G} \geq \hat{Z} \geq V\}} \right]. \end{aligned} \tag{5.12}$$

From (5.12) with

$$V = \begin{cases} \hat{Z} - \varepsilon & \text{on } \{\hat{y}\hat{G} < \hat{Z}\}, \\ \hat{Z} & \text{on } \{\hat{y}\hat{G} \geq \hat{Z}\} \end{cases}$$

for some $\varepsilon > 0$, we have

$$\{V < \hat{Z}\} = \{\hat{y}\hat{G} < \hat{Z}\},$$

and so

$$E^P \left[BH(\hat{y}\hat{G} - \hat{Z})_+ \right] \leq E^P \left[H(-\varepsilon)\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} \right] \leq 0.$$

This implies $B = 0$ on $\{\hat{y}\hat{G} > \hat{Z}, H > 0\}$, $P-$a.s. Therefore we deduce that

$$H\hat{W} = H(\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + B\mathbf{1}_{\{\hat{y}\hat{G} = \hat{Z}\}}) = \hat{X} \qquad P\text{-a.s.}$$

On the other hand, the conditions (5.4), (5.6) and (5.7) imply that $\hat{X}$ satisfies (4.13) and (4.14). Thus the proposition follows. □

Acknowledgements

I should like to thank Professor Akihiko Inoue for helpful discussions.

References

[1] C. Acerbi and D. Tasche, Expected shortfall: a natural coherent alternative to Value at Risk, to appear in *Wilmott.* `http://www.gloriamundi.org/var/wps.html`

[2] Ph. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, *Math. Finance,* **9** (1999), 203–228.

[3] J. Aubin and J. Ekeland, *Applied nonlinear analysis*, John Wiley & Sons, New York, 1984.

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

[5] J. Cvitanić and I. Karatzas, Generalized Neyman-Pearson lemma via convex duality, *Bernoulli,* **7** (2001), 79–97.

[6] J. Cvitanić and I. Karatzas, On dynamic measures of risk, *Finance Stochat.*, **3** (1999), 904–950.

[7] F. Delbaen, Coherent risk measures on general probability spaces, preprint, ETH Zürich. `http://www.math.ethz.ch/~delbaen/`

[8] N. El Karoui and M.-C. Quenez, Dynamic programming and pricing of contingent claims in an incomplete market, *SIAM J. Control Optim.*, **33** (1995), 29–66.

[9] H. Föllmer and Yu, M. Kabanov, Optional decompositions and Lagrange multipliers, *Finance Stochast.*, **2** (1998), 69–81.

[10] H. Föllmer and P. Leukert, Quantile Hedging, *Finance Stochast.,* **3** (1999), 251–273.

[11] H. Föllmer and P. Leukert, Efficient hedging: cost versus shortfall risk, *Finance Stochast.*, **4** (2000), 117–146.

[12] A. Inoue, On the worst conditional expectation, preprint, Hokkaido University.

[13] I. Karatzas, *Lectures on the Mathematics of Finance*, CRM Monographs series, American Mathematical Society, Providence, 1997.

[14] J. Komlós, A generalization of a problem of Steinhaus, *Acta Math. Acad. Sci. Hung.*, **18** (1967), 217–229.

[15] S. Kusuoka, On law invariant coherent risk measures, in *Advances in Mathematical Economic* (S. Kusuoka and T. Maruyama eds.), **3**, 83–95, Springer-Verlag, Tokyo, 2001.

[16] J. Lindenstrauss, A short proof of Lyapounoff's theorem, *J. Math. Mech.*, **4** (1966), 971–972.

[17] Y. Nakano, Efficient hedging with coherent risk measure, preprint, Hokkaido University.

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

[19] H. Pham and N. Touzi, The fundamental theorem of asset pricing with cone constraints, *J. Math. Econom.*, **31** (1999), 265–280.

[20] J. Sekine, Dynamic minimization of the worst conditional expectation of shortfall under partial information, preprint, Osaka University. `http://www.sigmath.es.osaka-u.ac.jp/~sekine/papers.html`

[21] M. Schäl, Martingale measures and hedging for discrete time financial markets, *Math. Oper. Res.*, **24** (1999), 509–528.

[22] M. Schwartz, New proofs of a theorem of Komlós, *Acta. Math. Hung.*, **47** (1986), 181–185.

[23] A. N. Shiryaev, *Essentials of Stochastic Finance, Facts, Models, Theory*, World Scientific, Singapore, 1999.