

Convex Risk Measures for Portfolio Optimization and Concepts of Flexibility

Hans-Jakob Lüthi¹ Jörg Doege¹

January 26, 2005

Abstract

Due to their axiomatic foundation and their favorable computational properties convex risk measures are becoming a powerful tool in financial risk management. In this paper we will review the fundamental structural concepts of convex risk measures within the framework of convex analysis. Then we will exploit it for deriving strong duality relations in a generic portfolio optimization context. In particular, the duality relationship can be used for designing new, efficient approximation algorithms based on Nesterov's smoothing techniques for non-smooth convex optimization. Furthermore, the presented concepts enable us to formalize the notion of flexibility as the (marginal) risk absorption capacity of a technology or (available) resources.

1 Introduction

In their seminal paper Artzner et.al. (ADEH99) presented an axiomatic foundation of coherent risk measures to quantify and compare uncertain future cash-flows within financial institutions. Recently, Föllmer et.al. (FS02a) extended the notion of coherent risk measures to convex risk measures. It is evident from the axiomatic structure referred above that convex analysis plays a crucial role. Indeed, as we will outline in this paper, convex optimization with its embedded duality is the underlying operational and computational technique for its applications. This was implicitly addressed by Artzner et.al. (ADEH99), Delbaen (De00) and Föllmer et.al. (FS02a) in their proofs of the fundamental representation theorems. But to our knowledge, the fundamental concepts have never been presented explicitly under, at least from a computational point of view, “natural” perspective of convex optimization as we will do in this paper for the restricted but important finite dimensional case. Independently Ruszczyński et.al. (RS04) studied the intimate relation between convex risk functions and duality structure in topological vector spaces of measurable functions. Indeed, the spirit of their work is very similar to ours but they use advanced theory of convex analysis of measurable functions. Restricting the analysis to the

¹ETH Zürich – Swiss Federal Institute of Technology

Dept. of Mathematics – Institute for Operations Research (www.ifor.math.ethz.ch)
Clausiusstrasse 45/47, CH-8092 Zürich – e-Mail: {luethi,doege}@ifor.math.ethz.ch

finite dimensional case, the representation results can be derived by applying elementary concepts of convex analysis based on the impressive work of R. T. Rockafellar. At the same time, this perspective opens up the computational aspects of convex risk measures in portfolio optimization and incidentally, leads to an attractive economical concept, namely the valuation of flexibility.

In section 2 the axiomatic foundation with the associated representation theorems due to Artzner et.al. (ADEH99) and Föllmer et.al. (FS02a) will be reviewed in a self contained manner. From the elementary proofs, which are given in an appendix, one observes that duality theory of convex optimization is the essence of the representation theorems of *any* convex or coherent risk measures. In particular, it enlightens the nontrivial computational task associated with any such risk measure. As prominent examples we present a convex, generalized version of the popular Conditional Value-at-Risk (CVaR) (see Rockafellar et.al. (RU02)) with its dual linear programming representation and discuss the family of entropy-based risk measures.

Having re-established these fundamental representation theorems of risk measures for the finite dimensional case, we will exploit the revealed duality structure to indicate computational effective schemes for portfolio optimization in section 3. Our contributions focuses on a duality analysis of a generic portfolio optimization model. We will establish a strong duality relationship for the model which incidentally shows that the generic model falls into a limited but attractive class of optimization models algorithmically investigated in a recent paper by Nesterov (Ne03a). As an important example, to compute CVaR-efficient portfolios in the sense of Markowitz (Ma52), the strong duality representation can be used to design attractive approximation algorithms which can handle very large, specially structured problem instances.

On the conceptual side, the generic model built upon the axiomatic framework of convex risk measures supports the introduction of a novel notion, namely the value of flexibility, which in this context is defined by the (marginal) risk absorption capacity of a technology or certain types of (available) resources. This economic topic is addressed in section 4 and is the applied contribution of our paper. For a real world example, we refer to electricity production management where this concept was fundamental for establishing the value of flexibility for a system of hydro power plants (see Doege et.al. (DLSU03)).

2 Convex Risk Measures - An Optimization View

The notion of convex risk measures was introduced by Föllmer et.al. (FS02a). This section contains a review of their fundamental work from an optimization point of view. Indeed, as already mentioned, we use elementary techniques from convex analysis to derive the dual representation theorems.

2.1 Convex Risk Measures in Short

Let \mathcal{X} denote the set of random variables $X : \Omega \mapsto \mathbb{R}$ of future (discounted) net worth of a portfolio defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover we define Ω as a set of possible scenarios, where we will assume throughout the paper that $|\Omega|$ is finite¹. One can think of $X(\omega)$ as the value of the portfolio given scenario $\omega \in \Omega$. Eventually, only one of the possible scenarios will take place and hence at the outset one is faced with the uncertainty of the outcome. In order to measure and control its associated risk Artzner et.al. (ADEH99) introduced an axiomatic framework of *coherent* risk measures which recently was "generalized" by Föllmer et.al. (FS02a) to *convex* risk measures as follows:

Definition 2.1 (Convex Risk Measures). *A mapping $\rho : \mathcal{X} \mapsto \mathbb{R}$ is called a convex risk measure, if and only if it is*

- convex: $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$ for $\lambda \in [0, 1]$.
- monotone: $X \leq Y$ then $\rho(X) \geq \rho(Y)$, where $X \leq Y \Leftrightarrow X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$
- translation invariant: if a is constant then $\rho(a\mathbf{1} + X) = -a + \rho(X)$

Adding to these properties positive homogeneity one obtains:

Definition 2.2 (Coherent Risk Measures). *A convex risk measure ρ is called coherent if in addition*

- positive homogeneity: if $\lambda \geq 0$ then $\rho(\lambda \cdot X) = \lambda \cdot \rho(X)$

holds.

Positive homogeneity assumes that the risk grows proportional to the volume of the portfolio X . Indeed, if liquidity cannot be assured in the market, then the exposure might grow faster than linear in the volume, a phenomena which among others motivated the relaxed notion of *convex* risk measures.

Coherent risk measures received much attention in the quantitative risk management community partially due to the fact that the most popular risk measure, the so-called Value-at-Risk (VaR)², is not coherent (nor convex) which has been shown by Delbaen (De00). It is important to note that violating the convexity in this case means that the subadditivity property

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \tag{1}$$

does not hold. Hence, if (1) is violated, then controlling the risk of subsidies by ρ and adding them up might not lead to an upper estimate of the total exposures measured in ρ . (See Acerbi et.al. (AT02a) or Frey et.al. (FM02) for

¹W.L.o.G. we will assume $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$.

²For a definition see equation (9) or JP Morgan's Risk Metrics (JP94). An excellent overview about the different risk measures is given by Szegö (Sz02).

examples of VaR violating subadditivity.)

The property of translation invariance has been questioned and banked by many researcher despite its explanation by Artzner et.al. (ADEH99). Rockafellar et.al. (RUZ02) stipulated an alternative class of risk functionals, called *deviation measures*, which are axiomatically based on $X - \mathbf{E}(X)$, and hence cannot be translation invariant. In the domain of applications of interest to us, we argue that the translation invariance represents a numeraire for risk capital which leads to important economical interpretations and hence translation invariance is a crucial requirement.

Recently, Artzner et.al. (ADEHK04) extended the axiomatic framework from a one-period model to a multi-period model wherein the notion of *time-consistency* is introduced as an additional property to relate risk measures consistently in time (see Cheridito et.al. (CDK04)).

In what follows, we will restrict to one-period measures and we call a position X *acceptable* if it has no risk at all at the end of the time horizon:

Definition 2.3 (Acceptance Set). *The acceptance set for any risk measure $\rho : \mathcal{X} \mapsto \mathbb{R}$ is defined by*

$$\mathcal{A}_\rho := \{X \in \mathcal{X} | \rho(X) \leq 0\}.$$

Obviously (see Föllmer et.al. (FS02a)), the axioms of ρ imply

$$\mathcal{A}_\rho \text{ closed, convex, and non-empty} \tag{2}$$

$$X \in \mathcal{A}_\rho, \quad Y \geq X \quad \text{then} \quad Y \in \mathcal{A}_\rho. \tag{3}$$

Conversely, any convex set $\mathcal{A} \neq \emptyset$ that satisfies properties (2) and (3) induces a convex risk measure, namely

$$\rho_{\mathcal{A}}(X) := \inf\{m \mid m \cdot \mathbf{1} + X \in \mathcal{A}\},$$

where one assumes that $\rho_{\mathcal{A}}(0)$ is bounded from below.

2.2 Representation Theorems and Duality

The famous representation theorem in Artzner et.al. (ADEH99) states that any *coherent* risk measure ρ is a worst case over a non-empty convex set of probability measures \mathcal{Q} , called *generalized scenarios or test probabilities*. Hence, one can compute a coherent risk measure for a position X by

- calculating, under each test probability $\mathbb{Q} \in \mathcal{Q}$, the average of the future net worth $-X$ (loss!) of the position: $\mathbf{E}_{\mathbb{Q}}[-X]$,
- and taking the maximum of all numbers found $\rho(X) = \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbf{E}_{\mathbb{Q}}[-X]$.

In the finite case this representation theorem follows either from a direct application of corollary 12.2.1 in Rockafellar (Ro70) or as a corollary of

Theorem 2.4 (Representation Theorem due to Föllmer et.al. (FS02a)). Denote \mathcal{P} the set of all probability measures over Ω , where we assume $|\Omega|$ to be finite. Any convex risk measure $\rho : \mathcal{X} \mapsto \mathbb{R}$ can be represented in the form

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{P}} (\mathbf{E}_{\mathbb{Q}}[(-X) - \alpha(\mathbb{Q})]) \quad (4)$$

where the “penalty function” $\alpha : \mathcal{P} \mapsto (-\infty, \infty]$ is convex and closed.

The proof, given in Appendix 6, relies on conjugate functions. In particular, from (57)

$$\alpha(\mathbb{Q}) := \rho^*(-\mathbb{Q}) \quad \forall \mathbb{Q} \in \mathcal{P}, \quad (5)$$

where ρ^* denotes the conjugate function of ρ in $\mathbb{R}^{|\Omega|}$

$$\rho^*(u) = \sup_{X \in \mathcal{X}} \{\langle u, X \rangle - \rho(X)\}, \quad (6)$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product³. Recently, Ruszczyński et.al. [see Theorem 2 in (RS04)] proved a very general version of this dual characterization in topological vector spaces of measurable functions using convex analysis results from the space of measurable functions.

From the definition 2.3 of the acceptance set and the dual representation (5) one obtains a characterization of α :

Corollary 2.5 (Penalty Term (FS02b)). Given the acceptance set \mathcal{A}_ρ the corresponding “penalty term” α in the representation theorem 2.4 is given by

$$\alpha(\mathbb{Q}) = \sup_{X \in \mathcal{A}_\rho} \{ \mathbf{E}_{\mathbb{Q}}[-X] \}. \quad (7)$$

Again we refer to Appendix 6 for the details of the proof.

Remark 2.6. From equation (60) we observe that $\alpha(\mathbb{Q})$ is the “normalized” support function of \mathcal{A}_ρ . Hence it can be used to reconstruct \mathcal{A}_ρ using the duality of support and indicator functions (see theorem 13.2 in Rockafellar (Ro70)).

In the special case of coherent measures, α is, due to homogeneity, an indicator function (see (FS02b)):

Corollary 2.7 (Representation of Coherent Risk Measure). ρ is coherent if and only if the “penalty term” in theorem 2.4 is an indicator function

$$\alpha(\mathbb{Q}) = \begin{cases} 0, & \text{if } \mathbb{Q} \in \mathcal{Q} \subset \mathcal{P}, \quad \text{for } \mathcal{Q} \text{ nonempty, closed and convex} \\ \infty, & \text{else.} \end{cases}$$

³Indeed, as a referee pointed out, a general definition of a conjugate in the context of random variables would use $\langle u, X \rangle := \mathbf{E}_{\mathbb{P}}(u^T X)$ instead.

Note that given the dual representation theorem 2.4 the subgradients of ρ are obtained by solving an optimization problem

$$\partial\rho(X) = \mathbf{conv}(-\mathbb{Q}|\mathbb{Q} \in \operatorname{argmax}_{\mathbb{Q} \in \mathcal{P}}\{\mathbf{E}_{\mathbb{Q}}[-X] - \alpha(\mathbb{Q})\}), \quad (8)$$

where \mathbf{conv} denotes the convex hull operator. Certainly, there are instances where due to the simplicity of α the right term in (8) is easy computable as it is the case for the measures discussed below. But in general, to compute a subgradient of ρ one needs to solve a (simple) convex optimization problem. This observation will be taken up again in section 3.

2.3 Examples of Convex Risk Measures

By theorem 2.4 any convex function defined on \mathcal{P} induces a convex risk measure. Hence, to make this concept meaningful an underlying economic interpretation is needed. Furthermore, an analytic expression for the penalty term α will enable efficient computational methods for portfolio optimization as outlined in section 3. Therefore, we will restrict to those examples with an explicit analytic expression for α^4 .

The most popular and widely used coherent risk measure is the Conditional Value-at-Risk (CVaR) as discussed in Rockafellar et.al. (RU02):

Example 2.8 (Conditional Value at Risk). *The Value-at-Risk with confidence level $1 > \beta > 0$ is defined by*

$$VaR_{\beta} = \inf \{m \mid \mathbb{P}[X + m < 0] \leq \beta\}. \quad (9)$$

Note: VaR is the $(1 - \beta)$ -quantile⁵ of the (loss) distribution $(-X)$.

Given the fact that Conditional Value-at-Risk is defined as the expected tail loss and based on equation (9) one would intuitively define CVaR with confidence level $1 > \beta > 0$ as

$$CVaR_{\beta}(X) = \mathbf{E}_{\mathbb{P}}[(-X)|(-X) \geq VaR_{\beta}].$$

This popular definition of CVaR will in general not result in a coherent measure⁶. In order to make CVaR coherent, one needs to care about splitting atoms in discrete distributions. The following definition based on corollary 2.7 yields not only a coherent measure but also relates CVaR directly to its computational ease (see Delbaen (De00))

$$CVaR_{\beta}(X) = \sup_{\mathbb{Q} \in \hat{\mathcal{Q}}} \{\mathbf{E}_{\mathbb{Q}}[-X]\}, \quad (10)$$

⁴Alternatively, Föllmer et.al. (FS02a) derive relevant examples of convex risk measures from the acceptance set using corollary 2.5.

⁵ β has to be small in our case.

⁶A fact that has been noted by many authors (see e.g. Acerbi et.al. (AT02a), Artzner et.al. (ADEH99), Rockafellar et.al. (RU00) (RU02)).

where

$$\hat{\mathcal{Q}} = \{\mathbb{Q} \in \mathcal{P} \mid \mathbb{Q} \leq \frac{\mathbb{P}}{\beta}\}. \quad (11)$$

which is called risk envelope in Rockafellar et.al. (RUZ02).

Hence CVaR is defined by a linear program (10) which is efficiently solvable by ordering the variables according to their efficiencies.

By duality of linear programming we obtain the well known representation of CVaR, namely

$$\text{CVaR}_\beta(X) = \inf_{\eta} [\eta + \frac{1}{\beta} \mathbf{E}_{\mathbb{P}}[(-X - \eta)_+]], \quad (12)$$

where $(a)_+ = \max\{0, a\}$ ⁷. In the sequel denote η_{VaR} the optimal solution to (12).

Due to its economical foundation and its computational ease, CVaR has been applied in many real world applications - even though it fails to differentiate among the severity of risky outcomes. To partially overcome this issue, Acerbi (AT02b) introduced a quantile weighted measure, called spectral measure. Similar in spirit but restricted to the β -quantile we will now introduce a generalization of CVaR as a convex risk measure, which partially can cope with extreme risky outcomes and yet keeps the attractive computational features of CVaR.

Example 2.9 (Generalized CVaR). A slight change of the underlying penalty term in the definition of CVaR to

$$\alpha_L(\mathbb{Q}) = \max_{\omega \in \Omega} \{L[\mathbb{Q}(\omega) - \frac{\mathbb{P}(\omega)}{\beta}]_+\},$$

where $L \geq 0$ is a fixed parameter results in an interesting family of convex risk measures.

First one notes that for $\mathbb{Q} \in \hat{\mathcal{Q}}$, where $\hat{\mathcal{Q}}$ is defined in (11), $\alpha_L(\mathbb{Q}) = 0$. The generalized CVaR is defined as

$$\text{GCVaR}_{\beta,L}(X) = \sup_{\mathbb{Q} \in \mathcal{P}} \{\mathbf{E}_{\mathbb{Q}}[-X] - \alpha_L(\mathbb{Q})\}. \quad (13)$$

Note that (13) is an LP, and by duality of linear programming, the generalized CVaR is alternatively characterized by its dual

$$\text{GCVaR}_{\beta,L}(X) = \inf_{\eta} [\eta + \frac{1}{\beta} \mathbf{E}_{\mathbb{P}}[(-X - \eta)^+]], \quad (14)$$

⁷In Rockafellar et.al. (RU02) the mathematical properties of CVaR have been investigated in a more general setting - resulting in a coherent CVaR definition as in (12) but for general probability measures.

where η satisfies

$$\sum_{\omega \in \Omega} (-X(\omega) - \eta)_+ \leq L. \quad (15)$$

Referring to (12), one immediately observes, that for $L \geq L^* := \sum_{\omega \in \Omega} (-X(\omega) - \eta_{\text{VaR}})_+$:

$$\text{GCVaR}_{\beta,L}(X) = \text{CVaR}_{\beta}(X),$$

where η_{VaR} is determined by (12). On the other hand, for $L = 0$

$$\text{GCVaR}_{\beta,0}(X) = \min_{\omega \in \Omega} [-X(\omega)] = \text{MaxLoss}(X) \quad (16)$$

and in general, for $0 < L < L^*$:

$$\text{GCVaR}_{\beta,L}(X) = \eta(L) + \frac{1}{\beta} \mathbf{E}_{\mathbb{P}}[(-X - \eta(L))_+], \quad (17)$$

where $\eta(L)$ is the unique solution to

$$\sum_{\omega \in \Omega} (-X(\omega) - \eta)_+ = L.$$

Furthermore, $\text{GCVaR}_{\beta,L}(\cdot)$ is a convex, monotone decreasing function in L . In particular:

Let $L = \lambda L^*$ for some $0 \leq \lambda \leq 1$, then

$$\text{CVaR}_{\beta}(X) \leq \text{GCVaR}_{\beta,L}(X) \leq \text{CVaR}_{\beta}(\lambda X) + (1 - \lambda) \text{MaxLoss}(X). \quad (18)$$

λ in (18) can be interpreted as the fraction of the acceptable, accumulated capital at risk in the β -tail of the underlying distribution \mathcal{P} . GCVaR encompasses CVaR , yet eventually, it penalizes large losses in the tail in accordance with L relatively small, i.e. for $\lambda \rightarrow 0$. Certainly, the strength of discrimination depends on the difference between $\text{MaxLoss}(X)$ and $\text{CVaR}_{\beta}(X)$.

The next class of convex risk measures shares with GCVaR the property of controlling the risk tolerance by a parameter while simplifying its computation. Furthermore, it generalizes the exponential risk measure introduced by Föllmer et.al. (Example 12 in (FS02a)).

Example 2.10 (Entropy-based Risk Measures). Let us take for the penalty term the well known entropy function defined on \mathcal{P} by

$$\alpha_{\mu}(\mathbb{Q}) = \sum_{\omega \in \Omega} \mu(\omega) \frac{\mathbb{Q}(\omega)}{\mathbb{P}(\omega)} \ln \left(\frac{\mathbb{Q}(\omega)}{\mathbb{P}(\omega)} \right). \quad (19)$$

Then, depending on the choice of the weights $\mu(\omega) \geq 0$ interesting families of convex risk measures are defined by

$$\rho_{\mu} = \max_{\mathbb{Q} \in \mathcal{P}} [\mathbf{E}_{\mathbb{Q}}(-X) - (\alpha_{\mu}(\mathbb{Q}) - \alpha_{\mu}(\mathbb{Q}^0))], \quad (20)$$

where the normalization term $\alpha_{\mu}(\mathbb{Q}^0)$ is specified by

$$\alpha_{\mu}(\mathbb{Q}^0) = \min_{\mathbb{Q} \in \mathcal{P}} \alpha_{\mu}(\mathbb{Q}). \quad (21)$$

The Karush-Kuhn-Tucker conditions for (20) have the form

$$-\mathbb{P}(\omega)X(\omega) - \mu(\omega)[1 + \ln(\tilde{\mathbb{Q}}(\omega))] = \tau\mathbb{P}(\omega), \quad \text{for } \omega \in \Omega \quad (22)$$

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega)\tilde{\mathbb{Q}}(\omega) = 1, \quad (23)$$

where we have set $\mathbb{Q}(\omega) := \mathbb{P}(\omega)\tilde{\mathbb{Q}}(\omega)$.

For any fixed τ the equalities (22) are satisfied by the following values:

$$\tilde{\mathbb{Q}}(\omega, \tau) = \exp\left[\frac{-\mathbb{P}(\omega)[X(\omega) + \tau]}{\mu(\omega)} - 1\right]. \quad (24)$$

Thus, to satisfy the equation (23) we have to appropriately choose τ in (24): Existence and uniqueness of such a τ are due to the fact that the left hand side in (24) is larger than 1 for τ small enough and decreases monotonically in τ , eventually for $\tau \rightarrow \infty$ it converges to e^{-1} .

We now consider some particular choices of the weights μ :

- **Expected Exponential Loss**

Choosing the weights $\mu(\omega) = \mathbb{P}(\omega)$ results (see Föllmer et.al. (FS02a) in

$$\rho_1^E(X) = \ln \mathbf{E}_{\mathbb{P}}[e^{-X}],$$

as an elementary computation will reveal. Furthermore, multiplying μ by a positive constant γ yields

$$\rho_{\gamma}^E(X) = \gamma \ln \mathbf{E}_{\mathbb{P}}[e^{-\frac{X}{\gamma}}] \geq \mathbf{E}_{\mathbb{P}}[-X].$$

This measure is called “entropic risk measure” by Barrieu et.al. (BK04). In particular, for $\gamma \rightarrow 0$ the risk goes to $\text{MaxLoss}(X)$ and for $\gamma \rightarrow \infty$ to $\mathbf{E}_{\mathbb{P}}[-X]$. Furthermore, $\rho_{\gamma}^E(\cdot)$ is convex in $\gamma \geq 0$ for X fixed and monotone decreasing: For $0 \leq \gamma_1 \leq \gamma_2$

$$\text{MaxLoss}(X) \geq \rho_{\gamma_1}^E(X) \geq \rho_{\gamma_2}^E(X) \geq \mathbf{E}_{\mathbb{P}}[-X].$$

Indeed, the monotonicity in γ is strict unless for the trivial case where $X(\omega) = c$ for all $\omega \in \Omega$.

- **Risk Adapted Probability**

Alternatively, one can choose the weights in (19) constant, i.e. $\mu(\omega) = \gamma$. Then, (24) becomes

$$\tilde{\mathbb{Q}}^*(\omega) = e^{\left[\frac{-\mathbb{P}(\omega)[X(\omega) + \tau^*(\gamma)]}{\gamma} - 1\right]}, \quad (25)$$

where $\tau^*(\gamma)$ is determined by equation (23).

It is easy to see from (22) and (23), that $\tau^*(\gamma)$ is a strictly monotone decreasing function in γ .

Recall that $\mathbb{Q}^*(\omega) = \mathbb{P}(\omega)\tilde{\mathbb{Q}}^*(\omega)$, hence we can see from (25) that depending on the threshold $\tau^*(\gamma)$, the probability of high losses are reinforced whenever

$$\mathbb{P}(\omega)[-X(\omega) - \tau^*(\gamma)] \geq \gamma,$$

whereas the probability of high gains are exponentially decreased, eventually. As a matter of fact, without giving the details, we remark that the risk measure defined in (20), now denoted by ρ_γ , becomes for $\gamma = 0$:

$$\rho_0(X) = \text{MaxLoss}(X)$$

and for $\gamma \rightarrow \infty$

$$\rho_\infty = \mathbf{E}_{\mathbb{Q}^0}[-X]$$

where \mathbb{Q}^0 is given in (21).

As before, convexity and monotonicity in $\gamma \geq 0$ for X fixed still hold.

For general γ we conclude from (22) and (23)

$$\rho_\gamma(X) = \tau^* + \gamma \left[\sum_{\omega \in \Omega} \tilde{\mathbb{Q}}^*(\omega) + \alpha_1(\mathbb{Q}^0) \right]$$

where τ^* and $\tilde{\mathbb{Q}}^*$ are defined in (25) and \mathbb{Q}^0 as above.

For the standard setting where all scenarios are equally likely, i.e. $\mathbb{P}(\omega) = 1/|\Omega|$, we get an explicit expression for ρ :

$$\rho_\gamma(X) = |\Omega|\gamma \ln \left[\sum_{\omega \in \Omega} \frac{1}{|\Omega|} e^{-\frac{X(\omega)}{|\Omega|\gamma}} \right]$$

and $\mathbb{Q}^0 = \mathbb{P}$, i.e. $\rho_\gamma(X) = \rho_{\gamma|\Omega|}^E(X)$.

3 Portfolio Optimization: Some Duality-Relations

In a finite dimensional case α determines the conjugate function of ρ as outlined in (57). If we assume $\alpha(\mathbb{Q})$ to be known, then we can exploit the conjugate property in an obvious (dual) portfolio optimization framework given below. The specific model used will yield a strong duality relation which can be exploited in an algorithmic setting for portfolio optimization. Indeed we argue that for special structured instances Nesterov's smoothing method (see (Ne03a)) can be very powerful in this context. Furthermore, the generic model allows to react specifically on each outcome $\omega \in \Omega$ by an appropriate design of actions. It is in this particular framework where we will model the concept of flexibility in section 4.

3.1 Generic Optimization Models

In the generic portfolio optimization model (26) we assume ρ to be any convex risk measure specified by α as in the representation theorem 2.4. Moreover, we assume that by a set of variables s , representing the decision variables, one can steer and control the underlying risky portfolio X in order to minimize its risk:

$$\begin{aligned} & \inf_{X \in \mathcal{X}, s \in \mathbb{R}^m} \rho(X) & (26) \\ \text{s.t.} & \quad X \leq As + b, \\ & \quad s \in S \\ \text{where} & \quad A \in \mathbb{R}^{|\Omega| \times m}, \quad b \in \mathbb{R}^{|\Omega|} \\ \text{and} & \quad S \subset \mathbb{R}^m \quad \text{nonempty, convex, compact.} \end{aligned}$$

Due to the monotonicity of ρ one can assume equality in (26). Hence $As + b$ models explicitly the impact of decision s on the outcome X in the scenario space Ω and at the same time stipulates the separation (independence) of S from the space of scenarios. But note that s impacts each outcome $\omega \in \Omega$ directly, i.e. one can think of s of being a policy or strategy designed to react contingent on the outcomes.

Furthermore, the formulation is very general and it encompasses the case $X \leq h(s)$, where $h_i(s) : S \mapsto \mathbb{R}$ is concave for $i = 1, \dots, |\Omega|$ by introducing $|\Omega|$ auxiliary variables.

We will derive the Lagrange dual of (26) by making use of the conjugate function ρ^* as in the proof of theorem 2.4. A simple computation reveals that the Lagrange dual function of (26) is defined on \mathcal{P} and given by

$$g(\mathbb{Q}) = \begin{cases} -\alpha(\mathbb{Q}) + \mathbf{E}_{\mathbb{Q}}[-b] + \inf_{s \in S} \{\mathbf{E}_{\mathbb{Q}}[-As]\}, & \mathbb{Q} \in \mathcal{P} \\ -\infty, & \text{otherwise,} \end{cases} \quad (27)$$

where $\alpha(\mathbb{Q})$ is the ‘‘penalty function’’ of the convex risk measure ρ as specified in theorem 2.4.

Using the well established weak duality relations of convex optimization, we conclude

$$\rho(\hat{X}) \geq g(\hat{\mathbb{Q}}),$$

where \hat{X} is a feasible solution to the primal (26) and $\hat{\mathbb{Q}} \in \mathcal{P}$ a feasible solution to the dual (27).

In addition, under the compactness assumptions on S in the generic problem setting (26) strong duality holds:

Theorem 3.1. *Consider the Primal*

$$f^* = \inf_{s \in S} \{ \sup_{\mathbb{Q} \in \mathcal{P}} \{-\alpha(\mathbb{Q}) + \mathbf{E}_{\mathbb{Q}}[-As - b]\} \} \text{ of problem (26)}$$

and its **Dual**

$$d^* = \sup_{\mathbb{Q} \in \mathcal{P}} \{-\alpha(\mathbb{Q}) + \inf_{s \in S} \{\mathbf{E}_{\mathbb{Q}}[-As - b]\}\}.$$

If S is nonempty, compact and convex then strong duality holds, i.e. $f^* = d^*$ (and the optimal value is finite).

Proof. Define

$$f(s) = \sup_{\mathbb{Q} \in \mathcal{P}} \{-\alpha(\mathbb{Q}) + \mathbf{E}_{\mathbb{Q}}[-As - b]\}, \quad \text{for } s \in S, \quad (28)$$

and

$$\phi(\mathbb{Q}) = -\alpha(\mathbb{Q}) + \inf_{s \in S} \{\mathbf{E}_{\mathbb{Q}}[-As - b]\}, \quad \text{for } \mathbb{Q} \in \mathcal{P} \quad (29)$$

then we need to show

$$\inf_{s \in S} f(s) = \sup_{\mathbb{Q} \in \mathcal{P}} \phi(\mathbb{Q}). \quad (30)$$

By construction weak duality holds, furthermore compactness assures that both optimal values are finite and attained.

Now, let $s^* \in S$ be an optimal solution for the primal, then by optimality there exists a subgradient $\xi \in \partial f(s^*)$ such that

$$\xi^T (s - s^*) \geq 0, \quad \text{for all } s \in S. \quad (31)$$

By definition, equation (28) of f using the compactness of \mathcal{P} :

$$\xi \in \partial f(s^*) \Leftrightarrow \xi = -A^T \hat{\mathbb{Q}}, \quad (32)$$

where $\hat{\mathbb{Q}} \in \operatorname{argmax}_{\mathbb{Q} \in \mathcal{P}} \{-\alpha(\mathbb{Q}) + \mathbf{E}_{\mathbb{Q}}[-As^* - b]\}$ (see theorem 23. 2 in (Ro70)).

Using (32) in (31) we get

$$s^* \in \operatorname{argmin}_{s \in S} \{\mathbf{E}_{\hat{\mathbb{Q}}}[-As - b]\}. \quad (33)$$

Hence

$$\begin{aligned} f(s^*) &= -\alpha(\hat{\mathbb{Q}}) + \{\mathbf{E}_{\hat{\mathbb{Q}}}[-As^* - b]\} \\ &= -\alpha(\hat{\mathbb{Q}}) + \min_{s \in S} \{\mathbf{E}_{\hat{\mathbb{Q}}}[-As - b]\}, \quad \text{using (33)} \\ &= \phi(\hat{\mathbb{Q}}). \end{aligned}$$

Therefore, (30) follows now from weak duality. \square

3.2 Optimization Methods

The strong duality in theorem 3.1 is a special instance of Nesterov's adjoint model for applying smooth minimization techniques (see (Ne03a)). In particular, the primal function $f(s)$ defined in (28) is continuously differentiable at any

$s \in S$ whenever $\alpha(\mathbb{Q})$ is continuous and strongly convex on \mathcal{P} and its gradient is given by

$$\nabla f(s) = -A^T \mathbb{Q}^*(s), \quad (34)$$

where $\mathbb{Q}^*(s)$ is the (unique) optimal solution of (28). Furthermore, as shown by Nesterov (theorem 1 in (Ne03a)) the gradient is Lipschitz continuous. In this case, under the reasonable assumptions, that projections onto S can efficiently be computed, the optimal scheme for smooth optimization as described by Nesterov ((Ne03a)) can be applied with an efficiency estimate of the order $\mathcal{O}(\sqrt{\frac{L}{\epsilon}})$, where ϵ is the desired accuracy of the approximate solution and L is the Lipschitz constant for the gradient. For example, these assumptions are satisfied by the entropy-based risk measures discussed in section 2.3.

Furthermore, under the reasonable assumption, that the computation of $f(s)$ in (28) and of its dual $\phi(\mathbb{Q})$ in (29) is simple due to the simple structure of S , and which still remains true by adding a so-called prox-function for S , then we can apply a smooth minimization framework to the generally non-smooth function defined in (29) to efficiently approximate its optimal solution in $\mathcal{O}(\frac{1}{\epsilon})$, where ϵ is the desired accuracy of the approximate solution (see Nesterov ((Ne03a), (Ne03b))). The optimal scheme only needs gradient computations of the form (34) and hence can be carried for very large sample sizes $|\Omega|$.

As an illustrative example, consider the computation of CVaR-efficient portfolios for a simple set S : Assume we are given n financial instruments i , where $L^{(i)}(\omega)$, $i = 1, \dots, n$ represents the *loss* corresponding to instrument i under scenario $\omega \in \Omega$. Denote $L = (L^{(1)}, \dots, L^{(n)}) \in \mathbb{R}^{|\Omega| \times n}$ the corresponding loss-matrix.

Now, consider the following simple, but not trivial, portfolio optimization problem⁸:

$$\min_{x \in S} \text{CVaR}_\beta(Ls) \quad (35)$$

where

$$S := \{s \in \mathbb{R}_+^n : \sum_{i=1}^n s_i = 1, \underline{s} \leq s \leq \bar{s}\},$$

and \underline{s} and \bar{s} are respectively lower and upper bounds for the portfolio vector s .

Using the dual representation of CVaR given in (10) the optimization problem (35) becomes

$$\min_{s \in S} \max_{\mathbb{Q} \in \hat{\mathcal{Q}}} \mathbf{E}_{\mathbb{Q}}[Ls] \quad (36)$$

⁸Whenever a coherent risk measure is specified by an indicator function over a polytope in \mathcal{P} and furthermore the set S is polyhedral as well, then by duality of linear programming the primal (dual) is an instance of a linear program, as shown for the CVaR in (12). In (RU00) a similar setting is used for *approximating* CVaR-efficient portfolios in a general context of risk measures as defined in (ADEH99) using Monte-Carlo sampling.

where $\hat{\mathcal{Q}}$ is defined in (11). Applying theorem 3.1, denote

$$f(s) := \max_{\mathbb{Q} \in \hat{\mathcal{Q}}} \{\mathbf{E}_{\mathbb{Q}}[Ls]\} \quad \text{and} \quad \phi(\mathbb{Q}) := \min_{s \in S} \{\mathbf{E}_{\mathbb{Q}}[Ls],\} \quad (37)$$

then we can write (36) in its primal form, as

$$\min_{s \in S} f(s) \quad \text{or its dual, as} \quad \max_{\mathbb{Q} \in \hat{\mathcal{Q}}} \phi(\mathbb{Q}) \quad (38)$$

Problem (36) has exactly the same structure as the ‘‘adjoint’’ class studied by Nesterov in (Ne03a) and (Ne03b). Note that the functions defined in (37) are both convex but non-smooth in s and \mathbb{Q} , respectively. Thus, we can apply the optimal gradient scheme for smooth optimization as described by Nesterov in (Ne03a) to solve the portfolio optimization problem provided that the simplicity of S allows an efficient computation of the gradient-projection steps. In particular, using the entropy distance for the prox-function (see (Ne03a)) one can establish the following rate of convergence

$$\frac{4\sqrt{\ln n \ln |\Omega|}}{k+1} \cdot \max_{i,\omega} |L_{\omega}^{(i)}| \quad (39)$$

Therefore, an absolute accuracy of 1% of $CVaR_{\beta}$ in terms of the maximal loss can be achieved in no more than

$$400\sqrt{\ln n \ln |\Omega|} \quad (40)$$

gradient iterations.

Currently we are investigating its numerical efficiency in comparison with linear programming techniques. Indeed, for a huge set of scenarios linear programming might fail to work at all. For a detailed description of the method we refer to (El04).

4 Value of Flexibility

As mentioned before, the generic portfolio model (26) is formulated to encompass situations, where the particular choice of policy s has an effect on $X(\omega)$ for each scenario $\omega \in \Omega$. Hence the model encompasses decision situations contingent to (uncertain) outcomes. In other words, (26) models the ability of an economic system to react on unexpected changes in order to limit the associated threat of losses. Such abilities are summarized under the term *system flexibility*. In what follows we argue that the generic portfolio model (26) is an appropriate setting for formalizing and quantifying this elsewhere rather vague notion (see Vollert (Vo03) for a in-depth discussion).

In the framework presented here, the notion of risk is specified in definition 2.1. In this axiomatic foundation, risk reduction is postulated in the translation invariance property of ρ , namely if a is constant then

$$\rho(a\mathbf{1} + X) = -a + \rho(X), \quad (41)$$

i.e. by adding (risk-free capital) a to the cash-flow X , the risk is reduced by exactly the same amount (see definition 2.1).

We will use property (41) of ρ to derive the *value* of system flexibility in the following way:

Consider the model (26) and interpret its feasibility set S (or part of it) as a domain of flexibility to hedge against adverse outcomes $X(\omega)$. In this context we think of s of being a policy or strategy designed to react contingent to the outcomes $\omega \in \Omega$, hence as flexible as possible.

Let $\hat{S} \supset S$, which we will interpret as enhancing the underlying system flexibility. What is this enhancement worth?

Definition 4.1 (Value of Flexibility). *Consider problem (26) and let $\hat{S} \supset S$ be given. Then the reduction in the risk measure ρ will be called the **value of flexibility** of \hat{S} at S :*

$$V(\hat{S}|S) := \hat{\rho}(S) - \hat{\rho}(\hat{S}) \quad (42)$$

where $\hat{\rho}(S)$, $\hat{\rho}(\hat{S})$ denote the “corresponding” optimal values of (26).

Note that in this definition we assume that the (additional) flexibility is optimally exercised, i.e. that a change in strategy might take place.

In accordance with the translation invariance property (41) of convex risk measures, the value of flexibility can directly be linked to capital or cash equivalent, thus, it is comparable to the amount of risk-free capital needed to reduce the risk by the same amount $V(\hat{S}|S)$. Therefore, the value (42) has an attractive economic interpretation which rests on the axiomatic foundation of risk. In particular, if the investment capital needed to enhance S to \hat{S} is less than $V(\hat{S}|S)$ than such an investment is attractive, since it is more profitable than an investment in the risk-free market.

In special settings such as

$$S(y) = \{s \in S^* | w^T s \leq y\}, \quad \text{where } S^* \text{ is closed and convex} \quad (43)$$

the *marginal* value of flexibility is given by the optimal dual variable (shadow price) associated with the single constraint $w^T s \leq y$ in its dual formulation: Without loss of generality, we will assume throughout that there exists $y \in \mathbb{R}$ with $S(y) \neq \emptyset$ and *compact*. Then we say $y \in \mathbb{R}$ is feasible if and only if $S(y) \neq \emptyset$. For y feasible, denote

$$\hat{\rho}(y) = \inf_{s \in S(y)} \{ \sup_{\mathbb{Q} \in \mathcal{P}} \{-\alpha(\mathbb{Q}) + \mathbf{E}_{\mathbb{Q}}[-As - b]\} \} \quad (44)$$

As it is well known, $\hat{\rho}(y)$ is a closed, convex function in y on its domain of feasibility (which is an interval of the form $[a, \infty)$). Let $y_0 \in \mathbb{R}$ be feasible and $y > y_0$ then we consider the directional derivative of $\hat{\rho}$ at y_0 in direction $y - y_0$ denoted by

$$\hat{\rho}'(y_0; y - y_0).$$

As is shown in Rockafellar (see theorem 23. 1 in (Ro70)) , directional derivatives always exist and they are positive homogenous. We denote the normalized directional derivative by

$$\hat{\rho}'(y_0; 1) = \frac{\hat{\rho}'(y_0; y - y_0)}{(y - y_0)},$$

which is independent on the particular choice of $y > y_0$. Note that in this particular setting, the directional derivatives coincide with the right derivative of the convex function $\hat{\rho}(y)$ at y_0 (see Rockafellar (Ro70)). This observation justifies the following definition:

Definition 4.2 (Marginal Value of Flexibility). *Given y_0 with $S(y_0)$ nonempty then the marginal value of flexibility of $S(\cdot)$ at y_0 is defined by*

$$-\hat{\rho}'(y_0; 1) \tag{45}$$

where $\hat{\rho}(\cdot)$ is defined in (44).

Note that definition 4.2 assures that the marginal value of flexibility is always nonnegative for y_0 feasible.

Proposition 4.3 (Marginal Value of Flexibility). *Assume the dual problem (29) given $S(y_0)$ as specified in (43) is feasible and has a unique solution \mathbb{Q}^* , and denote $\hat{\rho}(y_0)$ the corresponding optimal primal value of (44). Consider the problem*

$$\inf_{s \in S(y_0) = \{s \in S^* \mid w^T s \leq y_0\}} \{\mathbf{E}_{\mathbb{Q}^*}[-As]\}. \tag{46}$$

Then there exists a Kuhn-Tucker multiplier $\lambda^*(y_0) \geq 0$ with

$$\inf_{s \in S^*} (\mathbf{E}_{\mathbb{Q}^*}[-As] + \lambda^*(y_0)(w^T s - y_0)) = \inf_{s \in S(y_0)} \{\mathbf{E}_{\mathbb{Q}^*}[-As]\} \tag{47}$$

If $\lambda^*(y_0)$ is unique, then $\lambda^*(y_0)$ corresponds to the marginal value of flexibility of $S(\cdot)$ at y_0 .

Proof. The existence of a Kuhn-Tucker vector for (47) given that y_0 is feasible follows from strong Lagrange-duality for the problem

$$\inf_{s \in S(y_0)} \{\mathbf{E}_{\mathbb{Q}^*}[-As]\} = \sup_{\lambda \geq 0} \{-\lambda y_0 + \inf_{s \in S^*} \{\mathbf{E}_{\mathbb{Q}^*}[-As] + \lambda w^T s\}\}, \tag{48}$$

using the special partial affine structure of $S(y_0)$ and its compactness as specified in (43). Denote $\hat{\rho}(y_0)$ the corresponding optimal primal value of (44) then by theorem 3.1

$$\hat{\rho}(y_0) = \sup_{\mathbb{Q} \in \mathcal{P}, \lambda \geq 0} \{-\alpha(\mathbb{Q}) - \lambda y_0 + \inf_{s \in S^*} \{\mathbf{E}_{\mathbb{Q}}[-As] + \lambda w^T s\}\}. \tag{49}$$

Using theorem 29.1 in Rockafellar (Ro70) and the fact that for y_0 feasible the primal is finite and Kuhn-Tucker multiplier exist,

$$\partial \hat{\rho}(y_0) = \{-\lambda \mid (\mathbb{Q}, \lambda) \in \operatorname{argmax}_{\mathbb{Q} \in \mathcal{P}, \lambda \geq 0} \{-\alpha(\mathbb{Q}) - \lambda y_0 + \inf_{s \in S^*} \{\mathbf{E}_{\mathbb{Q}}[-As] + \lambda w^T s\}\}\}. \tag{50}$$

Hence the second part of the statement follows from definition 4.2 and the assumed uniqueness of the Kuhn-Tucker vector in (50). \square

Remark 4.4. *Note that from (50) the value of flexibility in the degenerate case is given by the smallest value of the Lagrange-multiplier λ specified in (50).*

The reader is referred to Doege et.al. (DLSU03) for a real world application of this concept in the energy market, where the value of flexibility for a (virtual) hydro storage plant was established.

5 Conclusion

In this paper the recent axiomatic concepts of convex risk measures in finite dimensional spaces were discussed in the framework of convex analysis and conjugate functions. Indeed, exploiting the underlying structure for portfolio optimization reveals a strong duality relationship which can be used in numerical methods. Furthermore, we showed that the framework allows an operational definition of the value of flexibility, a hidden issue which is present in many domains of portfolio optimization. In particular, understanding the price of flexibility might support the design and valuation of contracts where flexibility is either bought or sold in order to actively manage risk.

6 Appendix

The appendix contains all proofs of section 2.2 in a self-contained manner using an convex analysis in \mathbb{R}^N . But we note, that most of the results can be stated in more abstract spaces as has been done in (De00), and more ecently in (RS04) using convex analysis in topological vector spaces of measurable functions.

For our purposes, it suffices to concentrate on a finite dimensional scenario space Ω where numerical computation takes place.

Theorem 2.4. The proof relies on conjugate functions.

Given the representation (4) it is straightforward to show that ρ is a convex risk measure.

Now assume ρ to be a convex risk measure. The conjugate function ρ^* is defined as

$$\rho^*(u) = \sup_{X \in \mathcal{X}} \{u^T X - \rho(X)\}. \quad (51)$$

Making use of the translation invariance property of ρ , i. e. adding to a fixed portfolio X a constant term $a \in \mathbb{R}$ leads to

$$\rho^*(u) \geq \sup_{\tilde{X} \in \mathcal{X}} \{u^T \tilde{X} - \rho(\tilde{X})\}, \quad (52)$$

where $\tilde{X} := X + a\mathbf{1}$:

$$\rho^*(u) \geq \sup_a \{u^T(X + a\mathbf{1}) - \rho(X + a\mathbf{1})\} \quad (53)$$

$$= \sup_a \{a(u^T \mathbf{1} + 1)\} + \{u^T X - \rho(X)\}. \quad (54)$$

The first addend is only bounded if $u^T \mathbf{1} = -1$, hence this gives a necessary condition for $\rho^*(u)$ to be finite.

Now, consider an arbitrary but fixed $X \geq 0$, then for any $\lambda \geq 0$ $\lambda X \geq 0$ and the monotonicity property yields $\rho(\lambda X) \leq \rho(0)$.

Therefore

$$\rho^*(u) \geq \sup_{\lambda \geq 0} \{u^T(\lambda X) - \rho(\lambda X)\} \quad (55)$$

$$\geq \sup_{\lambda \geq 0} \{u^T(\lambda X)\} - \rho(0). \quad (56)$$

Again the conjugate $\rho^*(u)$ can only be finite if $u^T X \leq 0, \forall X \geq 0$, therefore $u \leq 0$.

Summing up: the conjugate ρ^* reduces to

$$\rho^*(u) = \begin{cases} \sup_{X \in \mathcal{X}} \{u^T X - \rho(X)\} & \text{for } u^T \mathbf{1} = -1 \text{ and } u \leq 0 \\ \infty & \text{otherwise.} \end{cases}$$

Now define

$$\alpha(\mathbb{Q}) := \rho^*(-\mathbb{Q}) \quad \forall \mathbb{Q} \in \mathcal{P}, \quad (57)$$

and using the fact that the conjugate of the conjugate of a closed convex function is the original function (see theorem 12.2 in (Ro70)) :

$$(\rho^*)^*(X) = \rho(X).$$

One obtains

$$(\rho^*)^*(X) = \sup_u \{X^T u - \rho^*(u)\}. \quad (58)$$

Therefore (57) yields

$$\rho(X) = (\rho^*)^*(X) = \sup_{\mathbb{Q} \in \mathcal{P}} \{\mathbf{E}_{\mathbb{Q}}[-X] - \alpha(\mathbb{Q})\}. \quad (59)$$

□

Corollary 2.5. Using equation (57) for the “penalty term” it remains to be shown that for any $u \leq 0$ with $u^T \mathbf{1} = -1$:

$$\sup_{X \in \mathcal{X}} \{u^T X - \rho(X)\} = \sup_{X \in \mathcal{A}_\rho} \{u^T X\}. \quad (60)$$

By definition 2.3 for all $X \in \mathcal{A}_\rho$: $u^T X - \rho(X) \geq u^T X$ and $\rho(X) \leq 0$. Hence

$$\rho^*(u) \geq \sup_{X \in \mathcal{A}_\rho} \{u^T X - \rho(X)\} \geq \sup_{X \in \mathcal{A}_\rho} \{u^T X\}. \quad (61)$$

To prove (60) we assume by contradiction that $\exists u$ with $u^T \mathbf{1} = -1$ and $u \leq 0$ such that

$$\rho^*(u) > \sup_{X \in \mathcal{A}_\rho} \{u^T X\}. \quad (62)$$

Hence there exists $Y \in \mathcal{X}$ such that

$$u^T Y - \rho(Y) > u^T X \quad \text{for all } X \in \mathcal{A}_\rho. \quad (63)$$

Using the fact that $(Y + \rho(Y)\mathbf{1}) \in \mathcal{A}_\rho$ we obtain

$$\begin{aligned} u^T Y - \rho(Y) &> u^T [Y + \mathbf{1}\rho(Y)] \\ &= u^T Y + \rho(Y) \underbrace{u^T \mathbf{1}}_{=-1} \\ &= u^T Y - \rho(Y) \end{aligned}$$

which yields a contradiction. □

References

- Acerbi, C., and D. Tasche. On the coherence of expected shortfall. In: Szegö, G. (Ed.), “Beyond VaR” (Special Issue). *Journal of Banking & Finance* **26**, (2002).
- Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion. In: Szegö, G. (Ed.), “Beyond VaR” (Special Issue). *Journal of Banking & Finance* **26**, (2002).
- Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath. Coherent Risk Measures. *Mathematical Finance* **9**, (1999) 203–228.
- Artzner, Ph., F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent Multiperiod Risk-Adjusted Values and Bellmann’s Principle. Preprint, ETH Zürich (2004).
- Barrieu, P. and N. El Karoui Optimal Derivatives Design under Dynamic Risk Measures. In: “Mathematics of Finance”. *Contemporary Mathematics* (AMS Proceedings) (2004).
- Boyd, S., and L. Vandenberghe. *Convex Optimization* (Cambridge University Press, 2003)
- Delbaen, F. Coherent Risk Measures. Lecture Notes at Cattedra Galileiana at Scuola Normale di Pisa (2000).
- Cheridito, P., F. Delbaen, and M. Kupper. Dynamic Monetary Risk Measures for Bounded Discrete-Time Processes. Preprint, ETH Zürich (2004).
- Doege, J., H. -J. Lüthi, Ph. Schiltknecht, and G. Unger. Risk Management of Power Portfolios and Valuation of Flexibility. Technical Report, IFOR, ETH Zürich (2003).
- Dos Santos Eleutério, V. Credit Risk Portfolio Optimization Notes for the Approximation Algorithm. Technical Report, IFOR, ETH Zürich (2004).
- Föllmer, H., and A. Schied. Convex Measures of Risk and Trading Constraints. *Finance & Stochastics* **6(4)**, (2002) 429–447.
- Föllmer, H., and A. Schied. Robust Preferences and Convex Measures of Risk. *Finance & Stochastics* **6(4)**, (2002) 429–447.
- Frey, R., and A. McNeil. VaR and expected shortfall in credit portfolios: Conceptual and practical insights. In: Szegö, G. (Ed.), “Beyond VaR” (Special Issue). *Journal of Banking & Finance* **26**, (2002).
- JP Morgan. Risk Metrics. New York (1994).
- Markowitz, H. M. Portfolio Selection. *Journal of Finance* **7**, (1952) 77–91.
- Nesterov, Y. Smooth Minimization of Non-Smooth Functions. CORE DP2003/12 (2003) -to appear in *Mathematical Programming*-.
- Nesterov, Y. Excessive gap technique in non-smooth convex optimization, CORE DP 2003/35.
- Rockafellar, R. T. *Convex Analysis* (Princeton University Press, 1970)

- Rockafellar, R. T. and S. Uryasev. Optimization of Conditional Value-at-Risk. *The Journal of Risk* **2(3)**, (2000) 21–41.
- Rockafellar, R. T. and S. Uryasev. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance* **26**, (2002) 1443–1471.
- Rockafellar, R. T., S. Uryasev and M. Zabarankin. Deviation measures in risk analysis and optimization. Research Report 2002-07. Department of Industrial and Systems Engineering, University of Florida (2002).
- Ruszczynski, A. and A. Shapiro. Optimization of Convex Risk Functions. Economics Working Paper Archive at WUSTL (2004).
- Szegö, G. Measures of risk. *Journal of Banking & Finance* **26**, (2002) 1253–1272.
- Vollert, A. *A Stochastic Control Framework for Real Options in Strategic Valuation* (Birkhäuser Verlag, 2003)