Application of a General Risk Management Model to Portfolio Optimization Problems with elliptical distributed returns for risk neutral and risk averse decision makers.

Bahar Kaynar
Risk management Department, Garanti Bank, Levent Nispatiye, Besiktas 34240 Istanbul, Turkey.
baharkaynar@garanti.com.tr

Ş. İlker Birbil
Faculty of Engineering and Natural Sciences, Sabancı University, 34956 Istanbul, Turkey.
sibirbil@sabanciuniv.edu

J.B.G. Frenk
Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands.
frenk@few.eur.nl

Econometric Institute Report E.I. 2007-12

ABSTRACT: We discuss a class of risk measures for portfolio optimization with linear loss functions, where the random returns of financial instruments have multivariate elliptical distribution. Under this setting we pay special attention to two risk measures, Value-at-Risk and Conditional-Value-at-Risk and differentiate between risk neutral and risk averse decision makers. When the so-called disutility function is taken as the identity function, the optimization problem is solved for a risk neutral investor. In this case, the optimal solutions of the two portfolio problems using the Value-at-Risk and Conditional-Value-at-Risk measures are the same as the solution of the classical Markowitz model. We adapt an existing less known finite algorithm to solve the Markowitz model. Its application within finance seems to be new and outperforms the standard quadratic programming procedure `quadprog` within MATLAB. When the disutility function is taken as a convex increasing function, the problem at hand is associated with a risk averse investor. If the Value-at-Risk is the choice of measure we show that the optimal solution does not differ from the risk neutral case. However, when Conditional-Value-at-Risk is preferred for the risk averse decision maker, the corresponding portfolio problem has a different optimal solution. In this case the used objective function can be easily approximated by Monte Carlo simulation. For the actual solution of the Markowitz model, we modify and implement the less known finite step algorithm and explain its core idea. After that we present numerical results to illustrate the effects of two disutility functions as well as to examine the convergence behavior of the Monte Carlo estimation approach.

Keywords: Elliptical distributions; linear loss functions; value-at-risk; conditional value-at-risk; portfolio optimization; disutility

---

1. Introduction. In the world of finance and engineering it is desirable to make decisions that minimize risk. However, quantifying risk differs according to the measure used. There are two central approaches for modelling risk: it can be identified as a function of the deviation from an expected value or as a function of absolute loss. The former approach is the main idea of the Markowitz mean-variance model. The latter approach involves two recent risk measures, namely Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). VaR is the quantile of the distribution function of random loss associated with a given portfolio at a specified probability level. CVaR is the conditional expectation of losses above VaR value with corresponding probability level [12].

In this paper, we analyze a general risk management model applied to portfolio problems when dealing with risk neutral and risk averse decision makers. The used risk measures are VaR and CVaR and the returns of the assets belong to the elliptical world. A similar approach was initiated in [12] for the risk neutral decision maker and the special case of multivariate normally distributed returns on assets. When the term *elliptical world* is used, we actually refer to returns of financial instruments having elliptical distributions. For instance, normal and student-t are two typical elliptical distributions. Within this framework, we assume a linear loss function coupled with a disutility function. The type of the disutility function sets forth the difference between a risk neutral and a risk averse investor. In the risk neutral case, the disutility function is the identity function while for the risk averse case the disutility function is increasing and convex. We first give a short new proof of the known equivalence for the risk neutral decision maker between the use of the VaR and CVaR risk measures and the well-known mean-variance approach of Markowitz. Actually this equivalence holds for the larger class of positive homogeneous and translation invariant risk measures. At the same time we explain in our computational section an adapted version of an algorithm for special quadratic programming problems original proposed by Michelot [11] and use this algorithm to solve these mean-variance Markowitz quadratic programming problems. The number of steps in this algorithm to find the optimal allocation of the assets is finite and equals at most the number of used assets. It boils down to iteratively finding the analytical solutions of projection onto canonical simplices and elementary cones and its main computational burden consists of inverting matrices. In our computational section we show that this algorithm is much faster than the standard quadratic programming solver `quadprog` used in MATLAB. If we deal with a risk averse (risk sensitive) decision maker with a general increasing convex (concave) disutility function and VaR is taken as the decision measure it is shown that these decision makers take the same portfolio decision as a risk neutral decision maker having a linear disutility function. This observation seems to be new and might be used as a criticism for not using this well-known risk measure. However, the same observation does not hold for the CVaR measure, and in this case a separate convex optimization problem has to be solved. Its solution is different from the solution given by the Markowitz mean-variance approach used by a risk neutral decision maker. However, modulo an unknown constant representing the VaR measure for a known univariate spherically distributed random variable the objective function has a nice analytical form. To evaluate this unknown constant we may use Monte Carlo estimation of a simple expectation. This observation implies that within the elliptical world for risk averse decision makers it is possible to avoid the difficult task of generating scenarios outside this world as done by Rockafellar and Uryasev [12]. This focus on the elliptical allows significant reduction in the simulation time. Contrary to Rockafellar and Uryasev [12] we do not deal with applications of financial concepts such as hedging which actually lead to similar models. The important risk measure CVaR is also discussed by Embrechts et. al. under the name *expected shortfall* or *mean excess loss* together with properties of elliptical distributions [5]. This work significantly helped us to classify the elliptical world with respect to risk measures and types of decision makers. The discussion about different decision makers is not given in their work. The theory of coherent risk measures that we use in our study is thoroughly discussed Artzner et.al. [3].

The outline of this paper is as follows. In Section 2 we give a brief introduction to a general risk management model with a special focus on VaR and CVaR measures. Since we assume that the returns have elliptical distribution, an introduction to the elliptical world is presented in Section 3. In the first part of Section 3 we analyze the problem for a risk neutral decision maker. In the second part we consider the portfolio for a risk averse decision maker. We devote Section 4 to our computational study. We start with a discussion about a finite step algorithm, and then we modify the algorithm to solve the Markowitz model. This is followed by a presentation of some numerical results to illustrate the effects of disutility functions when used with VaR and CVaR measures. We conclude the paper in Section 5.

2. A general risk management model. In this section we introduce a general risk management model discussed in [12], and show how to apply this model to portfolio optimization. In portfolio optimization the decision maker tries to allocate capital to n financial instruments in such a way that a given risk measure is minimized. To start with this model, we observe that a decision maker faces loss. This loss is given by some real valued random function $f(\mathbf{x}, \mathbf{Y})$, where $\mathbf{x} \in \mathbb{X} \subseteq \mathbb{R}^n$ is a decision vector taken by the decision maker from a closed convex set, and $\mathbf{Y} \in \mathbb{R}^n$ is a random vector denoting uncertainty. The function $(\mathbf{x}, \mathbf{z}) \mapsto f(\mathbf{x}, \mathbf{z})$ is called the loss function. In particular the vector $\mathbf{x}$ (called a portfolio) represents the allocations of resources to the n different financial instruments, whereas the vector $\mathbf{Y} \in \mathbb{R}^n$ denotes the uncertain returns of these financial instruments. It is assumed that short-selling is not allowed, and hence, $\mathbf{x}$ must be nonnegative. The cumulative distribution function on $\mathbb{R}$ of the random loss $f(\mathbf{x}, \mathbf{Y})$ is then given by

$$\Psi_{\mathbf{x}}(\alpha) := \mathbb{P}\{f(\mathbf{x}, \mathbf{Y}) \leq \alpha\}$$

and this cumulative distribution function is assumed to be continuous. Its corresponding inverse cumulative distribution function, also known as the quantile function, defined on $(0, 1)$ is given by

$$\Psi_{\mathbf{x}}^{\leftarrow}(\beta) := \min\{\alpha \in \mathbb{R} : \Psi_{\mathbf{x}}(\alpha) \geq \beta\}.$$

Within mathematical statistics the value $\Psi_{\mathbf{x}}^{\leftarrow}(\beta)$ for $0 < \beta < 1$ is known as the βth quantile of the cumulative distribution function $\Psi_{\mathbf{x}}$, while within risk management it is called the Value-at-Risk, $\mathrm{VaR}_\beta(f(\mathbf{x}, \mathbf{Y}))$ of the loss $f(\mathbf{x}, \mathbf{Y})$ at probability level β. Moreover, for fixed β the function $\alpha_\beta : \mathbb{X} \to \mathbb{R}$ given by

$$\alpha_\beta(\mathbf{x}) := \Psi_{\mathbf{x}}^{\leftarrow}(\beta) \tag{1}$$

is within risk management a popular risk measure and used as a selection criterion in the following way. Given the loss function $f(\mathbf{x}, \mathbf{Y})$ one tries to solve the optimization problem

$$\min\{\alpha_\beta(\mathbf{x}) : \mathbf{x} \in \mathbb{X}\}. \tag{VP}$$

However, in this framework the function α_β is in general nonconvex even for convex loss functions, and usually difficult to evaluate. Another risk measure is the function $\phi_\beta : \mathbb{X} \to \mathbb{R}$ given by

$$\phi_\beta(\mathbf{x}) := (1-\beta)^{-1}\mathbb{E}\left(f(\mathbf{x}, \mathbf{Y})\mathbf{1}_{\{f(\mathbf{x},\mathbf{Y})\geq \alpha_\beta(\mathbf{x})\}}\right),$$

and this risk measure is called the Conditional-Value-at-Risk, $\mathrm{CVaR}_\beta(f(\mathbf{x}, \mathbf{Y}))$ of the loss function $f(\mathbf{x}, \mathbf{Y})$ at level β [12]. To explain the name of this risk measure we observe by the continuity of $\Psi_\mathbf{x}$ that $\Psi_\mathbf{x}(\Psi_\mathbf{x}^{\leftarrow}(u)) = u$ for every $0 < u < 1$ [15]. This shows by relation (1) that

$$\mathbb{P}\{f(\mathbf{x}, \mathbf{Y}) \geq \alpha_\beta(\mathbf{x})\} = \mathbb{P}\{f(\mathbf{x}, \mathbf{Y}) > \alpha_\beta(\mathbf{x})\} = 1 - \Psi_\mathbf{x}(\Psi_\mathbf{x}^{\leftarrow}(\beta)) = 1 - \beta \tag{2}$$

and so

$$\phi_\beta(\mathbf{x}) = \mathbb{E}\left(f(\mathbf{x}, \mathbf{Y}) \mid f(\mathbf{x}, \mathbf{Y}) \geq \alpha_\beta(\mathbf{x})\right),$$

which justifies the name Conditional-Value-at-Risk. The main reason for introducing this new risk measure is that the function $\phi_\beta : \mathbb{X} \to \mathbb{R}$ is convex on the convex set $\mathbb{X}$ for any loss function f satisfying $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{z})$ is convex for fixed $\mathbf{z}$. Therefore the optimization problem

$$\min\{\phi_\beta(\mathbf{x}) \ : \ \mathbf{x} \in \mathbb{X}\} \tag{CVP}$$

becomes a convex programming problem. To show the convexity of ϕ_β, we introduce for $0 < \beta < 1$ the function $F_\beta : \mathbb{X} \times \mathbb{R} \to \mathbb{R}$ given by

$$F_\beta(\mathbf{x}, \alpha) := \alpha + (1-\beta)^{-1}\mathbb{E}\left(\max\{f(\mathbf{x}, \mathbf{Y}) - \alpha, 0\}\right).$$

We now require the following result shown in [12].

Lemma 2.1 *It follows for every $\mathbf{x} \in X$ that*

$$\phi_\beta(\mathbf{x}) = \min_{\alpha \in \mathbb{R}} F_\beta(\mathbf{x}, \alpha).$$

Moreover, if the closed interval $S_\beta(\mathbf{x})$ denotes the set of optimal solutions of the above optimization problem, then the left point of this interval equals VaR $_\beta(f(\mathbf{x}, \mathbf{Y}))$.

An easy corollary of the above lemma is given by the following result.

Lemma 2.2 *If the function $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{z})$ is convex for every $\mathbf{z}$, then the function $\phi_\beta : \mathbb{X} \to \mathbb{R}$ is convex.*

Proof. Since $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{z})$ is convex for every $\mathbf{z}$ we obtain for every realization $\mathbf{Y}(\omega)$ of the random vector $\mathbf{Y}$ that $(\mathbf{x}, \alpha) \mapsto \max\{f(\mathbf{x}, \mathbf{Y}(\omega)) - \alpha, 0\}$ is convex on $\mathbb{X} \times \mathbb{R}$. This shows that $F_\beta : \mathbb{X} \times \mathbb{R} \to \mathbb{R}$ is convex on $\mathbb{X} \times \mathbb{R}$ and by Lemma 2.1 the function ϕ_β is also convex. $\square$

By selecting a special loss function it is possible to apply the above risk model to a portfolio management problem. Throughout this work, we will take $f(\mathbf{x}, \mathbf{Y}) = u(-\mathbf{x}^\intercal \mathbf{Y})$, where u is either a linear or convex disutility function. For u linear (with slope 1) the resulting loss function represents the disutility of the portfolio $\mathbf{x}$ for a risk neutral decision maker. If u is a convex increasing function, then we obtain the disutility of the same portfolio for a risk averse decision maker [7]. Observe in [12] only the risk neutral decision maker is considered. In both cases it is easy to see that the conditions of Lemma 2.2 are satisfied. In the next section we will focus on a class of distributions associated with the random return $\mathbf{Y}$ of the n financial instruments.

3. An introduction to the elliptical world. To analyze our general risk model for portfolio management we first introduce the following class of multivariate distributions [5, 6]. Recall that a mapping U is called orthogonal, if $U^\intercal U = UU^\intercal = I$. We also adopt the following notation; $\mathbf{X} =_d \mathbf{Z}$ means that the random vector (variable) $\mathbf{X}$ has the same distribution as $\mathbf{Z}$, while $\mathbf{X} \sim F$ means that the random vector $\mathbf{X}$ has the cumulative distribution function F. Moreover, $\mathbb{R}_{++}$ denotes the positive real numbers.

DEFINITION 3.1 *A random vector* $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ *has a spherical distribution if for any orthogonal mapping* $U : \mathbb{R}^n \to \mathbb{R}^n$, *it holds that,*

$$U\mathbf{X} =_d \mathbf{X}.$$

Since for any spherical distributed random vector $\mathbf{X}$ and an orthogonal mapping U it follows that $U\mathbb{E}(\mathbf{X}) = \mathbb{E}(U\mathbf{X}) = \mathbb{E}(\mathbf{X})$ we obtain that its expectation equals $\mathbf{0}$. Also it can be shown [6] that the random vector $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ has a spherical distribution if and only if there exists some real-valued function $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that the characteristic function $\psi(\mathbf{t}) := \mathbb{E}(\exp(i\mathbf{t}^\intercal\mathbf{X}))$ is given by $\psi(\mathbf{t}) = \phi(\|\mathbf{t}\|^2)$. By this result we immediately obtain for every $t \in \mathbb{R}$ and $1 \leq j \leq n$ that

$$\mathbb{E}(\exp(it\mathbf{X}_j)) = \phi(t^2). \tag{3}$$

Using the above characterization of a spherical distribution another useful description can also be derived [6]. For completeness a short proof is listed.

LEMMA 3.1 *The random vector* $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ *has a spherical distribution if and only if* $\mathbf{a}^\intercal\mathbf{X} =_d \|\mathbf{a}\|\mathbf{X}_1$ *for every* $\mathbf{a} \in \mathbb{R}^n$.

PROOF. If the random vector $\mathbf{X}$ has a spherical distribution, then there exists some function $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that $\mathbb{E}(\exp(i\mathbf{a}^\intercal\mathbf{X})) = \phi(\|\mathbf{a}\|^2)$ for every $\mathbf{a} \in \mathbb{R}^n$. Hence for every $t \in \mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^n$ it follows by relation (3) that

$$\mathbb{E}(\exp(it\mathbf{a}^\intercal\mathbf{X})) = \phi(\|t\mathbf{a}\|^2) = \phi(t^2\|\mathbf{a}\|^2) = \mathbb{E}(\exp(it\|\mathbf{a}\|\mathbf{X}_1)).$$

By using the one to one correspondence between a characteristic function and the cumulative distribution function of the associated random variable [10], we obtain $\mathbf{a}^\intercal\mathbf{X} =_d \|\mathbf{a}\|\mathbf{X}_1$. To prove the reverse implication we observe that

$$\mathbb{E}(\exp(i\mathbf{a}^\intercal\mathbf{X})) = \mathbb{E}(\exp(i\|\mathbf{a}\|\mathbf{X}_1)). \tag{4}$$

This implies for every $\mathbf{a} \in \mathbb{R}^n$ that

$$\mathbb{E}(\exp(-i\mathbf{a}^\top\mathbf{X})) = \mathbb{E}(\exp(i\mathbf{a}^\top\mathbf{X}))$$

and so the function $\mathbf{a} \mapsto \mathbb{E}(\exp(i\mathbf{a}^\top\mathbf{X}))$ is real-valued. Hence by relation (4) the function $\mathbf{a} \mapsto \mathbb{E}(\exp(i\|\mathbf{a}\|\mathbf{X}_1))$ is also real-valued and introducing $\phi : \mathbb{R}_+ \to \mathbb{R}$ given by $\phi(t) := \mathbb{E}(\exp(i\sqrt{t}\mathbf{X}_1))$ it follows again by relation (4) that

$$\mathbb{E}(\exp(i\mathbf{a}^\top\mathbf{X})) = \phi(\|\mathbf{a}\|^2).$$

Applying the characteristic function description of a spherical distribution we conclude that $\mathbf{X}$ has a spherical distribution. □

If $N(\mu, \Sigma)$ denotes the multivariate normal distribution with mean μ and covariance matrix Σ it is well known that $U\mathbf{V} =_d \mathbf{V}$ for any random vector $\mathbf{V} \sim N(\mathbf{0}, I)$. This shows with $\mathbf{S}$ a real valued nonnegative random variable independent of $\mathbf{V}$ that the random vector $\mathbf{X} = \mathbf{S}\mathbf{V}$ has a spherical distribution. The distribution of the random vector $\mathbf{S}\mathbf{V}$ is called a scale mixture of standard multivariate normal distributions, and this class is an important subclass of the set of spherical distributions [6]. Next to the the multivariate normal distribution another well known distribution belonging to the above class is the multivariate t-distribution with ν degrees of freedom. A related class of distributions is given by the following [5, 6].

DEFINITION 3.2 *A random vector* $\mathbf{Y} = (\mathbf{Y}_1, \cdots, \mathbf{Y}_n)^\intercal$ *has an elliptical distribution if there exists an affine mapping* $\mathbf{x} \mapsto A\mathbf{x} + \mu$ *and a random vector* $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ *having a spherical distribution such that* $\mathbf{Y} = A\mathbf{X} + \mu$.

For convenience, an elliptical distributed random vector $\mathbf{Y}$ is denoted by $(A, \mu, \mathbf{X})$. In the remainder of this section we will assume that the returns $\mathbf{Y}$ of the n-financial instruments in our portfolio model have an elliptical distribution.

3.1 Risk neutral decision maker. In our portfolio decision model the random loss for a risk neutral decision maker choosing portfolio $\mathbf{x}$ and having returns $\mathbf{Y}$ on the n financial assets is given by $f(\mathbf{x},\mathbf{Y}) = -\mathbf{x}^\intercal\mathbf{Y}$. In case the random return vector $\mathbf{Y}$ has an elliptical distribution, the following representation is an immediate consequence of Lemma 3.1.

LEMMA 3.2 *If $\mathbf{Y}$ has an elliptical distribution with representation $(A,\mu,\mathbf{X})$ with $\mathbf{X} = (\mathbf{X}_1,\cdots,\mathbf{X}_n)^\intercal$, then*

$$-\mathbf{x}^\intercal\mathbf{Y} =_d \|A\mathbf{x}\|\mathbf{X}_1 - \mathbf{x}^\intercal\mu. \tag{5}$$

for every portfolio $\mathbf{x} \in \mathbb{R}^n$. Moreover, the parameters of the spherical (marginal) distribution of the random variable $\mathbf{X}_1$ are independent of $\mathbf{x}$.

PROOF. Since the elliptical distributed random vector $\mathbf{Y}$ has representation $(A,\mu,\mathbf{X})$ and $\mathbf{X}$ is spherical distributed, it follows that

$$-\mathbf{x}^\intercal\mathbf{Y} = -\mathbf{x}^\intercal A\mathbf{X} - \mathbf{x}^\intercal\mu.$$

Applying Lemma 3.1 with $\mathbf{a} = A^\intercal\mathbf{x}$ and using $\|A^\intercal\mathbf{x}\|^2 = \|A\mathbf{x}\|^2$ yields the desired result. □

Applying Lemma 3.2 one can easily compute the value at risk and the conditional value at risk of the loss $-\mathbf{x}^\intercal\mathbf{Y}$ for $\mathbf{Y}$ having an elliptical distribution. To derive these expressions we need some properties of those risk measures. Since these properties are only mentioned in [5] but not proved, we give a short proof. Observe that the value at risk and the conditional value at risk can be seen as real-valued functions defined on the space of real-valued random variables. In particular for any given real-valued random variable $\mathbf{Z}$ the value at risk $\mathrm{VaR}_\beta(\mathbf{Z})$ of $\mathbf{Z}$ at level β is given by $\mathrm{VaR}_\beta(\mathbf{Z}) := F^{\leftarrow}(\beta)$ with F the cumulative distribution function of $\mathbf{Z}$, while the conditional values at risk $\mathrm{CVaR}_\beta(\mathbf{Z})$ of $\mathbf{Z}$ at level β is defined by $\mathrm{CVaR}_\beta(\mathbf{Z}) := \mathbb{E}\left(\mathbf{Z} \mid \mathbf{Z} \geq F^{\leftarrow}(\beta)\right)$.

DEFINITION 3.3 *Let $\mathcal{B}$ be the space of all real-valued random variables $\mathbf{Z}$. A function $\varrho : \mathcal{B} \to (-\infty,\infty]$ is called positive homogeneous if $\rho(\lambda\mathbf{Z}) = \lambda\varrho(\mathbf{Z})$ for every $\lambda > 0$ and $\mathbf{Z} \in \mathcal{B}$. The function $\varrho(\cdot)$ is called translation invariant if $\varrho(\mathbf{Z} + a) = \rho(\mathbf{Z}) + a$ for every $a \in \mathbb{R}$ and $\mathbf{Z} \in \mathcal{B}$.*

It is now possible to show the following result.

LEMMA 3.3 *The risk measures $VaR_\beta : \mathcal{B} \to (-\infty,\infty]$ and $CVaR_\beta : \mathcal{B} \to (-\infty,\infty]$ are positive homogeneous and translation invariant.*

PROOF. Clearly for every $\lambda > 0$ and $\mathbf{Z} \in \mathcal{B}$ with cumulative distribution function F it follows that

$$\mathbb{P}\left\{\lambda\mathbf{Z} \leq x\right\} = \mathbb{P}\left\{\mathbf{Z} \leq \lambda^{-1}x\right\} = F(\lambda^{-1}x)$$

and hence

$$\mathrm{VaR}_\beta(\lambda\mathbf{Z}) = \inf\{x \ : \ F(\lambda^{-1}x) \geq \beta\} = \lambda\inf\{x \ : \ F(x) \geq \beta\} = \lambda\mathrm{VaR}_\beta(\mathbf{Z}).$$

Since $\mathbb{P}\left\{\mathbf{Z} + a \leq x\right\} = F(x-a)$ we also obtain

$$\mathrm{VaR}_\beta(\mathbf{Z} + a) = \inf\{x \ : \ F(x-a) \geq \beta\} = a + \inf\{x \ : \ F(x) \geq \beta\} = a + \mathrm{VaR}_\beta(\mathbf{Z}),$$

and this shows the result for the value at risk measure. To show the desired properties for the conditional value at risk measure it follows using $\mathrm{VaR}_\beta(\lambda\mathbf{Z}) = \lambda\mathrm{VaR}_\beta(\mathbf{Z})$ that

$$\mathrm{CVaR}_\beta(\lambda\mathbf{Z}) = \mathbb{E}\left(\lambda\mathbf{Z} \mid \lambda\mathbf{Z} \geq \mathrm{VaR}_\beta(\lambda\mathbf{Z})\right) = \mathbb{E}\left(\lambda\mathbf{Z} \mid \mathbf{Z} \geq \mathrm{VaR}_\beta(\mathbf{Z})\right)$$

for every $\lambda > 0$ and $\mathbf{Z} \in \mathcal{B}$. This implies

$$\mathrm{CVaR}_\beta(\lambda\mathbf{Z}) = \lambda\mathbb{E}\left(\mathbf{Z} \mid \mathbf{Z} \geq \mathrm{VaR}_\beta(\mathbf{Z})\right) = \lambda\mathrm{CVaR}_\beta(\mathbf{Z}).$$

Moreover, using $\mathrm{VaR}_\beta(\mathbf{Z} + a) = a + \mathrm{VaR}_\beta(\mathbf{Z})$ we also obtain

$$\mathrm{CVaR}_\beta(\mathbf{Z} + a) = \mathbb{E}\left(\mathbf{Z} + a \mid \mathbf{Z} + a \geq \mathrm{VaR}_\beta(\mathbf{Z} + a)\right) = \mathbb{E}\left(\mathbf{Z} + a \mid \mathbf{Z} \geq \mathrm{VaR}_\beta(\mathbf{Z})\right)$$

for every $a \in \mathbb{R}$ and $\mathbf{Z} \in \mathcal{B}$. This implies

$$\mathrm{CVaR}_\beta(\mathbf{Z} + a) = a + \mathbb{E}(\mathbf{Z}|\mathbf{Z} \geq \mathrm{VaR}\ _\beta(\mathbf{Z})) = a + \mathrm{CVaR}_\beta(\mathbf{Z})$$

and the result is proved. □

By Lemma 3.2 and 3.3 it follows immediately for $\mathbf{Y}$ having an elliptical distribution with representation $(A, \mu, \mathbf{X})$ with $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ that

$$\mathrm{VaR}_\beta(-\mathbf{x}^\intercal\mathbf{Y}) = \|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu \tag{6}$$

and

$$\mathrm{CVaR}_\beta(-\mathbf{x}^\intercal\mathbf{Y}) = \|A\mathbf{x}\| \,\mathrm{CVaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu. \tag{7}$$

for every $\mathbf{x} \in \mathbb{R}^n$ (observe the case $A\mathbf{x} = \mathbf{0}$ is easy to check). Since $\beta > 0.5$ and $\mathbf{X}_1$ has a one-dimensional spherical distribution (see Lemma 3.2) and is therefore symmetric around 0 we obtain that $\mathrm{VaR}_\beta(\mathbf{X}_1)$ and $\mathrm{CVaR}_\beta(\mathbf{X}_1)$ are positive. Hence by relations (6) and (7) the optimization problems (VP) and (CVP) defined in Section 2 reduce to

$$\min_{\mathbf{x}\in\mathbb{X}} \|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu \tag{VP}$$

and

$$\min_{\mathbf{x}\in\mathbb{X}} \|A\mathbf{x}\| \,\mathrm{CVaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu, \tag{CVP}$$

respectively. In particular, in these optimization problems the feasible region is given by

$$\mathbb{X} = \{\mathbf{x} \in \mathbb{R}^n \;:\; \mathbf{e}^\intercal\mathbf{x} = 1,\; \mu^\intercal\mathbf{x} = r,\; \mathbf{x} \geq \mathbf{0}\}. \tag{8}$$

This means that among the set of nonnegative portfolios with the same expected return we like to select that portfolio with minimal risk. Since $\mathrm{VaR}_\beta(\mathbf{X}_1) > 0$ and $\mathrm{CVaR}_\beta(\mathbf{X}_1) > 0$ (independent of $\mathbf{x}$) it is clear for the above feasible region $\mathbb{X}$ that the optimization problems (CVP) and (VP) have the same optimal solutions and this optimal solution can be obtained by solving the corresponding well-known Markowitz mean-variance problem

$$\min\left\{\frac{1}{2}\mathbf{x}^\intercal\Sigma\mathbf{x} \;:\; \mathbf{e}^\intercal\mathbf{x} = 1,\; \mu^\intercal\mathbf{x} = r,\; \mathbf{x} \geq \mathbf{0}\right\}, \tag{MP}$$

where $\Sigma := AA^\intercal$ is the covariance matrix (modulo a multiplicative positive constant) of the elliptical distributed random returns $\mathbf{Y}$. To avoid pathological cases we always assume that the matrix A is invertible, and hence, the covariance matrix Σ is strictly positive definite. In the next section we will propose a fast finite step algorithm to solve problem MP. Actually the above equivalence between the mean-variance approach and the (conditional) value at risk measure for $\mathbf{Y}$ having an elliptical distribution holds for a much larger class of risks measures. Although known [5] we list for completeness a short proof.

Lemma 3.4 *If $\mathbf{Y}$ has an elliptical distribution with representation $(A, \mu, \mathbf{X})$ and the considered risk measure $\rho(\cdot)$ is positive homogeneous, translation invariant and $\rho(\mathbf{X}_1) > 0$, then*

$$\rho(-\mathbf{x}_1^\intercal\mathbf{Y}) \leq \rho(\mathbf{x}_2^\intercal\mathbf{Y}) \leftrightarrow \sigma^2(\mathbf{x}_1^\intercal\mathbf{Y}) \leq \sigma^2(\mathbf{x}_2^\intercal\mathbf{Y})$$

for any two nonzero portfolios $\mathbf{x}_1$ and $\mathbf{x}_2$ satisfying $A\mathbf{x}_1, A\mathbf{x}_2$ nonzero and $r = \mathbb{E}\,(\mathbf{x}_1^\intercal\mathbf{Y}) = \mathbb{E}\,(\mathbf{x}_2^\intercal\mathbf{Y})$.

Proof. Since $\rho(\cdot)$ is translation invariant and $r = \mathbb{E}\,(\mathbf{x}_1^\intercal\mathbf{Y}) = \mathbb{E}\,(\mathbf{x}_2^\intercal\mathbf{Y})$ we obtain

$$\rho(-\mathbf{x}_1^\intercal\mathbf{Y}) \leq \rho(-\mathbf{x}_2^\intercal\mathbf{Y}) \leftrightarrow \rho(-\mathbf{x}_1^\intercal\mathbf{Y} + \mathbb{E}\,(\mathbf{x}_1^\intercal\mathbf{Y})) \leq \rho(-\mathbf{x}_2^\intercal\mathbf{Y} + \mathbb{E}\,(\mathbf{x}_2^\intercal\mathbf{Y}))\,. \tag{9}$$

Since $\mathbf{Y}$ has an elliptical distribution, we know by Lemma 3.2 that $-\mathbf{x}^\intercal\mathbf{Y} + \mathbb{E}\,(\mathbf{x}^\intercal\mathbf{Y}) =_d \|A\mathbf{x}\|\mathbf{X}_1$. This implies by relation (9) and $\rho(\cdot)$ positive homogeneous that

$$\rho(-\mathbf{x}_1^\intercal\mathbf{Y}) \leq \rho(-\mathbf{x}_2^\intercal\mathbf{Y}) \leftrightarrow \|A\mathbf{x}_1\|\rho(\mathbf{X}_1) \leq \|A\mathbf{x}_2\|\rho(\mathbf{X}_1).$$

Using now $\sigma^2(\mathbf{x}^\top\mathbf{Y}) = c\|A\mathbf{x}\|^2$ with c some positive constant and $\rho(\mathbf{X}_1) > 0$ the result follows. □

In the next subsection we discuss the behavior of a risk averse decision maker under the (conditional) value at risk measures.

3.2 Risk averse decision maker. In case we are dealing with a risk averse decision maker the loss function under consideration has the form $f(\mathbf{x}, \mathbf{Y}) = u(-\mathbf{x}^\intercal\mathbf{Y})$ with $u : \mathbb{R} \to \mathbb{R}$ an increasing convex disutility function. As an example we might take $u(t) = (\max\{t - \tau, 0\})^2$ (see also Section 4). For this case the loss function is given by $f(\mathbf{x}, \mathbf{Y}) = (\max\{-\mathbf{x}^\intercal\mathbf{Y} - \tau, 0\})^2$. Here $\tau > 0$ denotes a fixed positive number representing the acceptable loss of an investor (see also [5]). By Lemma 3.2 it follows immediately for the given (convex) disutility function u and $\mathbf{Y}$ having an elliptical distribution with representation $(A, \mu, \mathbf{X})$ and $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ that

$$u(-\mathbf{x}^\intercal\mathbf{Y}) =_d u(\|A\mathbf{x}\|\,\mathbf{X}_1 - \mathbf{x}^\intercal\mu). \tag{10}$$

To show that the value at risk measure for any continuous increasing disutility functions yields the same selection of the optimal portfolio, we need the following property of this risk measure. Observe the proof is technical due to the fact that u might have flat pieces and so we need to introduce for the proof two types of inverse functions (see Appendix). In case the disutility functions are assumed to be strictly increasing the proof considerably simplifies. However, as mentioned above there are important convex disutility functions with flat pieces and so we cover the increasing continuous disutility functions.

LEMMA 3.5 *If the disutility function* $u : \mathbb{R} \to \mathbb{R}$ *is increasing and continuous then* $VaR_\beta(u(\mathbf{Z})) = u(VaR_\beta(\mathbf{Z}))$ *for any* $\mathbf{Z} \in \mathcal{B}$.

PROOF. By definition

$$CVaR_\beta(u(\mathbf{Z})) = \inf\{x \in \mathbb{R} : \mathbb{P}\{u(\mathbf{Z}) \le x\} \ge \beta\}.$$

Since the distribution function F of $\mathbf{Z}$ satisfies $F(\infty) = \lim_{x\uparrow\infty} F(x) = 1$ it follows that $\lim_{x\uparrow\infty} \mathbb{P}\{u(\mathbf{Z}) \le x\} = 1$ and this shows using $0 < \beta < 1$ that the set $\{x \in \mathbb{R} : \mathbb{P}\{u(\mathbf{Z}) \le x\} \ge \beta\}$ is nonempty. Also, by the definition of $u^{\hookleftarrow}$ in the Appendix we obtain

$$\begin{aligned}\{x \in \mathbb{R} : \mathbb{P}\{u(\mathbf{Z}) \le x\} \ge \beta\} &= \{x \in \mathbb{R} : \mathbb{P}\{\mathbf{Z} \le u^{\hookleftarrow}(x)\} \ge \beta\} \\ &= \{x \in \mathbb{R} : F(u^{\hookleftarrow}(x)) \ge \beta\}.\end{aligned}$$

Since by Lemma A.1 the function $u^{\hookleftarrow}$ is increasing and right continuous and F is also increasing and right continuous the nonempty set $\{x \in \mathbb{R} : F(u^{\hookleftarrow}(x)) \ge \beta\}$ is closed. Hence the infimum of this set is attained at some $x_\beta \in \mathbb{R}$ and satisfies $F(u^{\hookleftarrow}(x_\beta)) \ge \beta$. If range($u$) is unbounded from above, then obviously $x_\beta \in cl(range(u))$. Also for range(u) bounded from above and so $\sigma := \sup\{u(y) : y \in \mathbb{R}\}$ is finite we may conclude that $x_\beta \in cl(range(u))$. To show this assume that $x_\beta > \sigma$. This implies for every $y \in (\sigma, x_\beta)$ that $F(u^{\hookleftarrow}(y)) = 1 > \beta$ and this contradicts $x_\beta = \min\{x \in \mathbb{R} : F(u^{\hookleftarrow}(x)) \ge \beta\}$. Therefore $x_\beta \le \sigma$ and hence x_β belongs to $cl(range(u))$. By Lemma A.1 we also know using x_β belongs to $cl(range(u))$ that $u(u^{\hookleftarrow}(x_\beta)) = x_\beta$ and combining the above observations yields

$$CVaR_\beta(u(\mathbf{Z})) = \min\{x \in \mathbb{R} : \mathbb{P}\{u(\mathbf{Z}) \le x\} \ge \beta\} = x_\beta = u(u^{\hookleftarrow}(x_\beta)).$$

In the last part of this proof we only need to verify that $u(u^{\hookleftarrow}(x_\beta)) = u(F^{\leftarrow}(\beta))$. Since $F(u^{\hookleftarrow}(x_\beta)) \ge \beta$ it follows by the definition of $F^{\leftarrow}(\beta)$ that $u^{\hookleftarrow}(x_\beta) \ge F^{\leftarrow}(\beta)$. Suppose now by contradiction that $u^{\leftarrow}(x_\beta) > F^{\leftarrow}(\beta)$. By Lemma A.1 we know that $\lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k) = u^{\leftarrow}(x_\beta)$ for any strictly increasing sequence $x_k \uparrow x_\beta$ and so there exist some $x_{k_0} < x_\beta$ satisfying $u^{\hookleftarrow}(x_{k_0}) > F^{\leftarrow}(\beta)$ implying $F(u^{\hookleftarrow}(x_{k_0})) \ge \beta$. By the definition of x_β given by

$$x_\beta = \min\{x \in \mathbb{R} : F(u^{\hookleftarrow}(x)) \ge \beta\}$$

this yields $x_\beta \le x_{k_0}$ and we obtain a contradiction. Thus we have verified that $u^{\hookleftarrow}(x_\beta) \ge F^{\leftarrow}(\beta) \ge u^{\leftarrow}(x_\beta)$ and by the monotonicity of u it follows that $u(u^{\hookleftarrow}(x_\beta)) \ge u(F^{\leftarrow}(\beta)) \ge u(u^{\leftarrow}(x_\beta))$. Applying Lemma A.1 and $x_\beta \in cl(range(u))$ we obtain $u(u^{\hookleftarrow}(x_\beta)) = u(u^{\leftarrow}(x_\beta))$ and this shows that $u(u^{\hookleftarrow}(x_\beta)) = u(F^{\leftarrow}(\beta))$. □

By relation (6) and Lemma 3.5 it follows immediately for u an increasing and continuous disutility function that

$$\mathrm{VaR}_\beta(u(-\mathbf{x}^\intercal\mathbf{Y})) = u(\|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu). \tag{11}$$

Hence for a risk averse decision maker (with a convex increasing disutility function u) minimizing the value at risk measure we need to solve the optimization problem

$$\min_{\mathbf{x}\in\mathbb{X}} u(\|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu)$$

with X given by relation (8). Since u is increasing and continuous it follows that

$$\min_{\mathbf{x}\in\mathbb{X}}\{u(\|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu) \;:\; \mathbf{x} \in \mathbb{X}\} = u(\min_{\mathbf{x}\in\mathbb{X}}\{\|A\mathbf{x}\| \,\mathrm{VaR}_\beta(\mathbf{X}_1) - \mathbf{x}^\intercal\mu\}).$$

This shows that the optimal portfolio is the same for a risk averse (risk neutral) and even risk seeking decision maker for elliptical returns $\mathbf{Y}$ and the value at risk measure. They all boil down to solving a Markowitz mean-variance problem. By this observation it seems that this measure is not a good objective for both risk averse and risk seeking decision makers. This might be an argument (besides the noncoherency (in the sense of Artzner et al) of the value at risk measure [3]) not to use this measure. In case we use the conditional value at risk measure (this is a coherent risk measure [3]) the following result can be shown.

LEMMA 3.6 *If the disutility function $u : \mathbb{R} \to \mathbb{R}$ is increasing and continuous, then*

$$CVaR_\beta(u(\mathbf{Z})) = \mathbb{E}(u(\mathbf{Z})|\mathbf{Z} \geq u^{\leftarrow}(u(VaR_\beta(\mathbf{Z})))).$$

Moreover, if $VaR_\beta(\mathbf{Z})$ is a point of strict increase of u, then

$$CVaR_\beta(u(\mathbf{Z})) = \mathbb{E}(u(\mathbf{Z})|\mathbf{Z} \geq VaR_\beta(\mathbf{Z})).$$

PROOF. By the definition of $u^{\leftarrow}$ (see Appendix) and Lemma 3.5 we obtain

$$\{u(\mathbf{Z}) \geq VaR_\beta(u(\mathbf{Z}))\} = \{u(\mathbf{Z}) \geq u(VaR_\beta(\mathbf{Z}))\} = \mathbf{Z} \geq u^{\leftarrow}(u(VaR_\beta(\mathbf{Z})))\}.$$

If $VaR_\beta(\mathbf{Z})$ is a point of strict increase of the function u then by Lemma A.1 we obtain $u^{\leftarrow}(u(VaR_\beta(\mathbf{Z}))) = VaR_\beta(\mathbf{Z})$. Applying now the definition of $CVaR_\beta(\mathbf{Z})$ and these observations yields the result. □

If in our portfolio model for a risk averse decision maker with convex increasing disutility function u the value at risk $VaR_\beta(-x^\top \mathbf{Y})$ given in relation (11) with $\mathbf{Y}$ having an elliptical distribution with representation $(A, \mu, \mathbf{X}), \mathbf{X} = (\mathbf{X}_1, ..., \mathbf{X}_n)$ is a point of strict increase of the function u, then by Lemma 3.6 we obtain

$$CVaR_\beta(u(-\mathbf{x}^\top \mathbf{Y})) = \mathbb{E}(u(-\mathbf{x}^\top \mathbf{Y})| - \mathbf{x}^\top \mathbf{Y} \geq VaR_\beta(-\mathbf{x}^\top \mathbf{Y})).$$

This implies for $\|A\mathbf{x}\| > 0$ (sufficient condition: $\|A$ is invertible) by relations (10) and (11) that

$$CVaR_\beta(u(-\mathbf{x}^\top \mathbf{Y})) = \mathbb{E}(u(\|A\mathbf{x}\|\mathbf{X}_1 - \mathbf{x}^\top \mu)|\mathbf{X}_1 \geq VaR_\beta(\mathbf{X}_1))$$

Hence for A invertible and the points of increase assumption the portfolio optimization problem for a risk averse decision maker is given by

$$\min_{\mathbf{x} \in \mathbb{X}} \{\mathbb{E}\left(u(\|A\mathbf{x}\|\, \mathbf{X}_1 - \mathbf{x}^\intercal \mu) \mid \mathbf{X}_1 \geq \mathrm{VaR}_\beta(\mathbf{X}_1)\right)\}.$$

This optimization problem gives in general a different optimal solution for risk neutral and risk averse (with increasing convex disutility function) decision makers. We already saw that for risk neutral decision makers this boils down to solving the Markowitz mean-variance problem. Moreover, for risk averse decision makers with an increasing convex disutility function the above problem is a general convex optimization problem. To approximate for such general functions the above objective we first need to compute $\mathrm{Var}_\beta(\mathbf{X}_1)$. If we consider the special case $\mathbf{Y} \sim N(\mu, \Sigma)$ with $\Sigma = AA^\intercal$ (see the computational results section), and hence, $\mathbf{X}_1$ has a standard univariate $N(0, 1)$ distribution the quantile $\mathrm{VaR}_\beta(\mathbf{X}_1)$ can be approximated accurately by a rational function [16, 1] or it is listed in a table in almost any standard textbook on statistics (see for example [8, 2]). For $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^\intercal$ having a scale mixture of standard normal multivariate distributions one may apply Monte Carlo simulation or numerical techniques to calculate $\mathrm{VaR}_\beta(\mathbf{X}_1)$. After having computed this we might use the (unbiased) Monte Carlo estimator of size m given by

$$\widehat{\mathbf{T}}_m(\mathbf{x}) := \frac{1}{m(1-\beta)} \sum_{i=1}^{m} u(\|A\mathbf{x}\|\, \mathbf{Z}_i - r) \mathbf{1}_{\{\mathbf{Z}_i \geq \mathrm{VaR}_\beta(\mathbf{X}_1)\}}, \tag{12}$$

where $\mathbf{Z}_i$, $1 \leq i \leq m$ are independent copies of the real valued random variable $\mathbf{X}_1$. In general m is a realization of a random variable which depends on the desired width of the chosen $100(1-\alpha)$ percent confidence interval [14]. If we consider the important case that the random vector $\mathbf{X}$ has a scaled mixture of standard multivariate normal distributions or $\mathbf{X} =_d \mathbf{S}\mathbf{V}$ with $\mathbf{V} = (\mathbf{V}_1, \cdots, \mathbf{V}_n)^\intercal \sim N(\mathbf{0}, I)$ and the real valued nonnegative random variable $\mathbf{S}$ independent of $\mathbf{V}$ it is obvious that $\mathbf{X}_1 =_d \mathbf{S}\mathbf{V}_1$ with $\mathbf{V}_1 \sim N(0, 1)$. This representation is helpful in generating a sample $(\mathbf{Z}_1(\omega), \cdots, \mathbf{Z}_m(\omega))$ if it is easy to generate independent copies of the random variable $\mathbf{S}$. Also for u increasing convex it is obvious that the random function $\mathbf{x} \mapsto \widehat{\mathbf{T}}_m(\mathbf{x})$ is convex on $\mathbb{X}$ and this shows that the approximating problem

$$\min_{\mathbf{x} \in \mathbb{X}} \left\{ \frac{1}{m(1-\beta)} \sum_{i=1}^{m} u(\|A\mathbf{x}\|\, \mathbf{Z}_i(\omega) - r) \mathbf{1}_{\{\mathbf{Z}_i(\omega) \geq \mathrm{VaR}_\beta(\mathbf{X}_1)\}} \right\} \tag{13}$$

with a sample $(\mathbf{Z}_1(\omega), \cdots, \mathbf{Z}_m(\omega))$ is a convex optimization problem. For arbitrary distributed random vectors $\mathbf{Y}$ and convex loss functions $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{z})$ ($\mathbf{z}$ fixed) the CVaR measure $\mathbf{x} \mapsto \mathrm{CVaR}_\beta(f(\mathbf{x}, \mathbf{Y})$ is estimated in [12] by generating scenarios from multivariate distributions. Due to the special structure of our loss functions and $\mathbf{Y}$ having an elliptical distribution, it is sufficient in our case to simulate from a univariate distribution.

4. Computational study. We start this section with a finite step algorithm to solve Markowitz problem (MP). Then, we present two disutility functions and illustrate their effects on the considered portfolio problems.

4.1 Modified Michelot algorithm. At the end of Section 3.1, we have emphasized that when $\mathbf{Y}$ has a multivariate elliptical distribution and $f(\mathbf{x}, \mathbf{Y})$ is a bilinear, then minimizing VaR and CVaR measures are equivalent to solving the corresponding Markowitz problem (MP) with the predetermined expected return r.

The algorithm introduced by Michelot finds in finite steps the projection of a given vector onto a special polytope [11]. The main idea of this algorithm is to use the analytic solutions of a sequence of projections onto canonical simplices and elementary cones. To apply Michelot's algorithm, we use a transformation $\mathbf{y} = \Sigma^{\frac{1}{2}}\mathbf{x}$. Then problem (MP) becomes

$$\min\left\{\frac{1}{2}\mathbf{y}^\intercal\mathbf{y} \ : \ \mathbf{d}^\intercal\mathbf{y} = 1, \ \eta^\intercal\mathbf{y} = r, \ A\mathbf{y} \geq \mathbf{0}\right\}, \tag{14}$$

where $\mathbf{d}^\intercal = \mathbf{e}^\intercal\Sigma^{-1/2}$, $\eta^\intercal = \mu^\intercal\Sigma^{-1/2}$ and $A = \Sigma^{-1/2}$. Note that the matrix A is the same as the matrix used for elliptical distributions in Section 3. To modify Michelot's algorithm according to our problem, we need to introduce several sets. Let

$$\begin{aligned} \mathcal{V} &= \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{d}^\intercal\mathbf{y} = 1, \ \eta^\intercal\mathbf{y} = r\}, \\ \mathcal{Y}_\mathcal{I} &= \{\mathbf{y} \in \mathbb{R}^n \mid (A\mathbf{y})_i = 0, \ i \in \mathcal{I}\}, \\ \mathcal{V}_\mathcal{I} &= \mathcal{V} \cap \mathcal{Y}_\mathcal{I}, \end{aligned} \tag{15}$$

where $\mathcal{I} \subseteq \{1, 2, \cdots, n\}$ denotes an index set and $(A\mathbf{y})_i$ denotes the ith component of vector $A\mathbf{y}$. Algorithm 4.1.1 gives the steps of the Modified Michelot Algorithm. The algorithm starts with solving the quadratic programming problem $\min\{\mathbf{y}^\intercal\mathbf{y} \ : \ \mathbf{y} \in \mathcal{V}\}$. It is easy to show that the analytic solution for this quadratic program is given by

$$\bar{\mathbf{y}} = \begin{bmatrix} \mathbf{d}^\intercal \\ \eta^\intercal \end{bmatrix}^\intercal \begin{bmatrix} \mathbf{d}^\intercal\mathbf{d} & \mathbf{d}^\intercal\eta \\ \eta^\intercal\mathbf{d} & \eta^\intercal\eta \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ r \end{bmatrix}. \tag{16}$$

Notice that some of the components $(A\mathbf{y})_i$ may be negative. After identifying the most negative component and initializing the index set $\mathcal{I}$, the algorithm iterates between projections of the incumbent solution $\bar{\mathbf{x}}$ onto subspace $\mathcal{Y}_\mathcal{I}$, and then onto subspace $\mathcal{V}_\mathcal{I}$ until none of the components are negative; i.e., the solution is optimal. The first projection is given by

$$\begin{aligned} P_{\mathcal{Y}_\mathcal{I}}(\bar{\mathbf{y}}) &:= \arg\min\left\{\tfrac{1}{2}\|\mathbf{y} - \bar{\mathbf{y}}\|^2 \ : \ \mathbf{y} \in \mathcal{Y}_\mathcal{I}\right\} \\ &= \left(I - [A]_{i\in\mathcal{I}}^\intercal \left([A]_{i\in\mathcal{I}} [A]_{i\in\mathcal{I}}^\intercal\right)^{-1} [A]_{i\in\mathcal{I}}\right)\bar{\mathbf{y}}, \end{aligned} \tag{17}$$

where $[A]_{i\in\mathcal{I}}$ denotes the submatrix formed by the rows $i \in \mathcal{I}$ of A. Similarly, the second projection [4] yields

$$\begin{aligned} P_{\mathcal{V}_\mathcal{I}}(\bar{\mathbf{y}}) &:= \arg\min\left\{\tfrac{1}{2}\|\mathbf{y} - \bar{\mathbf{y}}\|^2 \ : \ \mathbf{y} \in \mathcal{V}_\mathcal{I}\right\}, \\ &= \bar{\mathbf{y}} - \begin{bmatrix} \mathbf{d}^\intercal \\ \eta^\intercal \\ [A]_{i\in\mathcal{I}} \end{bmatrix}^\intercal \left(\begin{bmatrix} \mathbf{d}^\intercal \\ \eta^\intercal \\ [A]_{i\in\mathcal{I}} \end{bmatrix} \begin{bmatrix} \mathbf{d}^\intercal \\ \eta^\intercal \\ [A]_{i\in\mathcal{I}} \end{bmatrix}^\intercal\right)^{-1} \left(\begin{bmatrix} \mathbf{d}^\intercal \\ \eta^\intercal \\ [A]_{i\in\mathcal{I}} \end{bmatrix}\bar{\mathbf{y}} - \begin{bmatrix} 1 \\ r \\ 0 \end{bmatrix}\right). \end{aligned} \tag{18}$$

Since we have a finite number of assets, the dimension of the index set $\mathcal{I}$ is also finite. This shows that the modified algorithm terminates in at most n iterations (see also; [11]).

Step 1: Set $\bar{\mathbf{y}}$ as in (16). If $A\bar{\mathbf{y}} \geq \mathbf{0}$ then stop; $\bar{\mathbf{y}}$ is optimal. Otherwise, select i with most negative $(A\bar{\mathbf{y}})_i$, set $\mathcal{I} \leftarrow \{i\}$, and then go to *Step 2.*

Step 2: Set $\bar{\mathbf{y}} \leftarrow P_{\mathcal{Y}_\mathcal{I}}(\bar{\mathbf{y}})$ as in (17), and then go to *Step 3.*

Step 3: Set $\bar{\mathbf{y}} \leftarrow P_{\mathcal{V}_\mathcal{I}}(\bar{\mathbf{y}})$ as in (18). If $A\bar{\mathbf{y}} \geq \mathbf{0}$ then stop; $\bar{\mathbf{y}}$ is optimal. Otherwise, select i with most negative $(A\bar{\mathbf{y}})_i$, update $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$, and then go to *Step 2.*

Algorithm 4.1: Modified Michelot Algorithm

To analyze the performance of Algorithm 4.1, we have used MATLAB as our testing environment. All the computational experiments are conducted on a Pentium III 600 Mhz personal computer running Linux. First, we have randomly generated a set of test problems for different numbers of assets (n) as follows:

- ⋄ The components of $n \times n$ matrix $A = \Sigma^{-1/2}$ are sampled uniformly from interval $(-2.5, 5)$.
- ⋄ The components of vector μ are sampled uniformly from interval $(0.1, 5)$, and the first two components are sorted in ascending order; i.e., $\mu_1 \leq \mu_2$.
- ⋄ To ensure feasibility, the value r is then sampled uniformly from interval (μ_1, μ_2).
- ⋄ For each value of n, 10 replications are generated.
- ⋄ In all problems, β value is fixed to 0.95.

Clearly, Problem (MP) can be solved by any quadratic programming solver. In MATLAB the procedure that solves these types of problems is called `quadprog`, which is also used in the financial toolbox. Therefore, to compare the proposed algorithm, we also solved the set of problems with `quadprog`. Table 4.1 shows the statistics of the computation times out of 10 replications. The second and third columns in Table 4.1 shows the averages and the standard deviations of the computation times obtained by Algorithm 4.1, respectively. Similarly, columns four and five give the average and the standard deviation of the computation times found by `quadprog`, respectively.

	Algorithm 4.1		quadprog	
n	**Average**	**Std. Dev.**	**Average**	**Std. Dev.**
25	0.0251	0.0152	0.1094	0.2026
50	0.0767	0.0647	0.1565	0.0869
100	0.3018	0.3564	0.5732	0.2604
200	1.6080	1.2740	4.1343	0.9778
400	11.0991	11.0530	42.9040	10.7764

Table 1: Computation time statistics of `quadprog` and Algorithm 4.1 in seconds.

The average computational times in Table 4.1 show that Modified Michelot Algorithm is approximately four times faster than the MATLAB procedure `quadprog`. However, it is important to note that the MATLAB procedure `quadprog` involves many error checks that may also be the cause of higher computation times. Our implementation of Algorithm 4.1 is straightforward. Therefore, the performance can be improved by a better implementation. In fact the higher standard deviation figures obtained by Algorithm 4.1 suggest that matrix inverse calculations should be carried out more efficiently, since these inversions constitute the main computational burden of Algorithm 4.1. Overall, these results allow us to claim that an efficient implementation of Modified Michelot Algorithm is a fast and finite step alternative for solving problem (MP).

4.2 Two disutility functions. In this section we conduct a numerical study to illustrate the effects of two different disutility functions on VaR and CVaR measures. Again we use MATLAB as our programming environment.

Rockafellar and Uryasev gave a portfolio optimization example involving three instruments [12]. The returns on these instruments are coming from a multivariate normal distribution. The mean return vector and the covariance matrix are given as

$$\mu^{\intercal} = (0.01001110, 0.0043532, 0.0137058)$$

and

$$\Sigma = \begin{bmatrix} 0.00324625 & 0.00022983 & 0.00420395 \\ 0.00022983 & 0.00049937 & 0.00019247 \\ 0.00420395 & 0.00019247 & 0.00764097 \end{bmatrix},$$

respectively. The expected return r is equal to 0.011 and $\beta \in \{0.90, 0.95, 0.99\}$. We select two disutility functions from the literature. The first one is using the max operator (see also [5]), and it is given as

$$u(t) = (\max\{t - \tau, 0\})^2,$$

where we set parameter $\tau = (\max_{i=1,2,3}\{\mu_i\} + \min_{i=1,2,3}\{\mu_i\})/2$. As the second distutility function, we use the exponential function

$$u(t) = \exp(t) - 1.$$

As discussed in [9], exponential-type utility is one of the most frequently used functions. We subtract 1 to have zero disutility to reflect the case of zero loss.

Before applying the disutility functions, we solve the risk neutral decision maker problem with Algorithm 4.1. We then plug the optimal solution into equation (6) and (7) to obtain VaR and CVaR values, respectively. These results are given in Table 2 for different values of β. We note that the optimal solution of this problem is already given in [12]. However, we repeat these results here for ease of reference.

Risk Measure	$\beta = 0.90$	$\beta = 0.95$	$\beta = 0.99$
VaR	0.0678	0.0920	0.1321
CVaR	0.0970	0.1160	0.1530

Table 2: VaR and CVaR values for a risk neutral decision maker (see also [12]).

Recall that in case of VaR measure, the optimal solution is the same as the optimal solution of the corresponding (risk neutral) Markowitz problem. Therefore, we only report for VaR measure the change in the objective function values. When CVaR measure is used, however, the objective function may not have an analytical form. Due to the special structure of the example here, it turns out that the objective functions can be obtained for both disutility functions. In Appendix A relations (20) and (21) give the objective functions for max and exponential disutility functions, respectively. Notice that both relations involve evaluations of the cumulative distribution function of a standard normal variable. There are standard methods in simulation for approximating the standard normal distribution very accurately [13]. In our computational study, we use `MATLAB` procedure `normcdf` to evaluate this approximation. To solve the convex optimization problem (**??**), we use the default constrained nonlinear programming procedure in `MATLAB` called `fmincon`. This procedure requires an initial solution, $\mathbf{x}_0$. In all our experiments we set ${\mathbf{x}_0}^\intercal = (0, 0.2893, 0.7107)$.

Table 3 shows for both disutility functions, the changes in VaR and CVaR values for different values of β. As expected both VaR and CVaR values increase as β increases for both disutility functions. It is interesting to observe that when max operator is used, both risk measures yield lower VaR and CVaR values in comparison with the risk neutral case given in Table 2. On the other hand, when exponential function is used, both VaR and CVaR values increase. This is due to fact that the first disutility function with the max operator penalizes only the deviation from an acceptable loss, whereas the exponential disutility function, in general, penalizes the loss more severely than the risk-neutral case.

Disutility Function	**Risk Measure**	$\beta = 0.90$	$\beta = 0.95$	$\beta = 0.99$
max	VaR	0.0040	0.0073	0.0162
	CVaR	0.0092	0.0129	0.0224
exp	VaR	0.0702	0.0944	0.1413
	CVaR	0.1022	0.1232	0.1655

Table 3: VaR and CVaR values for a risk averse decision maker.

With the optimal CVaR values at hand, we next conduct numerical experiments to illustrate the convergence behavior of the Monte Carlo estimator (12) as sample size increases. We solve with both disutility functions the corresponding convex optimization problems (13), and compare our results with the optimal objective function values given in Table 3. Since the Monte Carlo simulation depends on the selected random seed, we conduct our experiments over 10 different seeds. Tables 4 and 5 include the average values over 10 problems for each combination of β and sample size.

Table 4 illustrates the results obtained by solving the estimation problem (13) with max disutility function. The first column of the table gives the sample size. The average percentage deviations of the estimated CVaR values from the optimal values (see Table 3) are reported in columns 2, 4 and 6, respectively for each value of β. The corresponding average computation times in seconds are given

	$\beta = 0.90$		$\beta = 0.95$		$\beta = 0.99$	
m	**(%)**	**(sec.)**	**(%)**	**(sec.)**	**(%)**	**(sec.)**
10^2	26.2933	0.0270	41.5724	0.0267	76.3254	0.0512
10^3	10.0896	0.0475	14.2461	0.0399	24.7877	0.0362
10^4	3.4124	0.2053	4.7921	0.1476	11.2223	0.0919
10^5	0.9649	0.8869	1.2278	0.5673	3.2685	0.3606
10^6	0.3441	7.0481	0.4342	4.4051	1.0562	2.3380

Table 4: Average CVaR deviations and computation times over 10 replications (max disutility).

in columns 3, 5 and 7. Naturally, as the sample size increases the deviation from the optimal value decreases. Notice that the summation in evaluating the Monte Carlo estimator (12) causes the main computational burden. Therefore, the computation times also increase as the sample size becomes larger. The objective function with CVaR measure involves a conditional expectation, and the sample points passing the condition decreases as β increases. The figures in columns 2, 4 and 6 from left to right confirm this observation. However, since a small number of sample points pass the condition, only a few number of function evaluations are carried out in evaluating the Monte Carlo estimator. One might avoid generating too much realizations not passing the condition by the use of tilted univariate densities and hence apply variance reduction of the Monte Carlo estimator (rare event simulation [14]). However, we did not do this. Hence, the computation times slightly decrease in almost all cases, except $(m = 10^2, \beta = 0.99)$ combination, as β increases (see columns 3, 5 and 6 from left to right).

	$\beta = 0.90$		$\beta = 0.95$		$\beta = 0.99$	
m	**(%)**	**(sec.)**	**(%)**	**(sec.)**	**(%)**	**(sec.)**
10^2	21.3703	0.0303	40.4043	0.0299	78.3227	0.0258
10^3	7.4044	0.0480	11.6809	0.0411	23.4547	0.0362
10^4	2.4530	0.2300	4.0388	0.1477	10.9822	0.0926
10^5	0.7486	0.8903	0.9725	0.5650	3.1080	0.3665
10^6	0.2905	7.2095	0.4120	4.4671	1.0655	2.2760

Table 5: Average CVaR deviations and computation times over 10 replications (exponential disutility).

The layout of Table 5 is exactly the same as Table 4. The figures in the table are obtained by solving the estimation problem (13) with the exponential disutility function. Again we observe a similar pattern: On one hand, increasing the sample size improves the accuracy of the approximation at the expense of longer computation time; on the other hand, higher β values requires larger sample sizes to obtain more accurate results.

5. Conclusion. In this paper, we discuss the portfolio optimization problem when random asset returns have elliptical distributions. Concentrating on VaR and CVaR measures, we analyze the cases of both risk-neutral and risk-averse decision makers. Furthermore, a fast and finite step algorithm is given for the risk neutral case. When dealing with the risk-averse decision makers, we show that the simulation from multivariate distributions can be reduced to generating realizations from a univariate distribution. This becomes important especially when the analytic function of the objective function does not exits, and hence, Monte Carlo estimation becomes necessary. Finally, we conducted numerical experiments to support our findings.

There are several directions in which this research can be extended. A natural extension to our work may also involve hedging. A classification, similar to ours, can be studied for the non-elliptical world. Notice that in the non-elliptical world, if CVaR is chosen as the risk measure, the overall problem is still a convex programming problem for loss functions being convex in the decision variable $\mathbf{x}$ for fixed $\mathbf{z}$. However the objective function in general does not have a nice analytical form. If sampling from univariate distributions is not possible, then the notion of copulas can be analyzed to simulate realizations from multivariate distributions [5].

Appendix A. The Risk Averse Objective Functions. To conduct the numerical experiments for a risk averse decision maker, we need to calculate CVaR objective function

$$\mathbb{E}\left(u(\|A\mathbf{x}\|\mathbf{X}_1 - \mu^\intercal \mathbf{x}) \mid \mathbf{X}_1 \geq \text{VaR}_{\ \beta}(\mathbf{X}_1)\right) = \frac{1}{1-\beta}\mathbb{E}\left(u(\|A\mathbf{x}\|\mathbf{X}_1 - \mu^\intercal \mathbf{x})\mathbf{1}_{\{\mathbf{X}_1 \geq \text{VaR}_{\ \beta}(\mathbf{X}_1)\}}\right). \tag{19}$$

To simplify the notation, define $a := \|A\mathbf{x}\|$, $\alpha := \text{VaR}_{\ \beta}(\mathbf{X}_1)$, and recall that $r = \mu^\intercal \mathbf{x}$. In the numerical results section, it is assumed that the returns are coming from a multivariate normal distribution. Therefore, $\mathbf{X}_1$ becomes a standard normal random variable. This implies that

$$\mathbb{E}\left(u(a\mathbf{X}_1 - r)\mathbf{1}_{\{\mathbf{X}_1 \geq \alpha\}}\right) = \frac{1}{\sqrt{2\pi}}\int_\alpha^\infty u(az-r)\exp\left(\frac{-z^2}{2}\right)dz.$$

When the utility function is given as $u(t) = (\max\{t-\tau, 0\})^2$, then

$$\mathbb{E}\left(u(a\mathbf{X}_1 - r)\mathbf{1}_{\{\mathbf{X}_1 \geq \alpha\}}\right) = \frac{1}{\sqrt{2\pi}}\int_\alpha^\infty (\max\{az-r-\tau, 0\})^2\exp\left(\frac{-z^2}{2}\right)dz.$$

Let $\gamma := \max\{(r+\tau)/a, \alpha\}$, then

$$\mathbb{E}\left(u(a\mathbf{X}_1 - r)\mathbf{1}_{\{\mathbf{X}_1 \geq \alpha\}}\right) = \frac{1}{\sqrt{2\pi}}\int_\gamma^\infty (a^2z^2 - 2(r+\tau)az + (r+\tau)^2)\exp\left(\frac{-z^2}{2}\right)dz.$$

Each term of this integration can now be obtained. The first term is given by

$$\begin{aligned}\frac{a^2}{\sqrt{2\pi}}\int_\gamma^\infty z^2\exp\left(\frac{-z^2}{2}\right)dz &= \frac{a^2}{\sqrt{2\pi}}\left(-z\exp\left(\frac{-z^2}{2}\right)|_\gamma^\infty + \int_\gamma^\infty \exp\left(\frac{-z^2}{2}\right)dz\right)\\ &= \frac{a^2\gamma}{\sqrt{2\pi}}\exp\left(\frac{-\gamma^2}{2}\right) + a^2\Phi(-\gamma),\end{aligned}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable. The second and third terms are obtained, respectively, by

$$\frac{-2(r+\tau)a}{\sqrt{2\pi}}\int_\gamma^\infty z\exp\left(\frac{-z^2}{2}\right)dz = \frac{-2(r+\tau)a}{\sqrt{2\pi}}\exp\left(\frac{-\gamma^2}{2}\right)$$

and

$$\frac{(r+\tau)^2}{\sqrt{2\pi}}\int_\gamma^\infty \exp\left(\frac{-z^2}{2}\right)dz = (r+\tau)^2\Phi(-\gamma).$$

Therefore, in case of max form of the utility, the objective function becomes

$$\mathbb{E}\left(u(a\mathbf{X}_1 - r) \mid \mathbf{X}_1 \geq \alpha\right) = \frac{1}{1-\beta}\left(\frac{a^2\gamma - 2a(r+\tau)}{\sqrt{2\pi}}\exp\left(\frac{-\gamma^2}{2}\right) + a^2\Phi(-\gamma) + (r+\tau)^2\Phi(-\gamma)\right). \tag{20}$$

When the utility function is given as $u(t) = \exp(t) - 1$, then

$$\begin{aligned}\mathbb{E}\left(u(a\mathbf{X}_1 - r)\mathbf{1}_{\{\mathbf{X}_1 \geq \alpha\}}\right) &= \frac{1}{\sqrt{2\pi}}\int_\alpha^\infty \exp(az-r)\exp\left(\frac{-z^2}{2}\right)dz - \Phi(-\alpha)\\ &= \exp(-r)\frac{1}{\sqrt{2\pi}}\int_\alpha^\infty \exp\left(\frac{-z^2+2za}{2}\right)dz - \Phi(-\alpha)\\ &= \exp\left(\frac{a^2-2r}{2}\right)\frac{1}{\sqrt{2\pi}}\int_\alpha^\infty \exp\left(-\frac{(z-a)^2}{2}\right)dz - \Phi(-\alpha)\\ &= \exp\left(\frac{a^2-2r}{2}\right)\Phi(a-\alpha) - \Phi(-\alpha).\end{aligned}$$

Therefore, in case of exponential form of the utility, the objective function becomes

$$\mathbb{E}\left(u(a-r) \mid \mathbf{X}_1 \geq \alpha\right) = \frac{1}{1-\beta}\left(\exp\left(\frac{a^2-2r}{2}\right)\Phi(a-\alpha) - \Phi(-\alpha)\right). \tag{21}$$

Appendix A. On some properties of inverse functions In this appendix we discuss a technical lemma needed for the proof of Lemma 3.5. Let $u : \mathbb{R} \to \mathbb{R}$ be an increasing continuous function and denote by $\mathrm{range}(u)$ its range given by

$$\mathrm{range}(u) := \{u(x) : x \in \mathbb{R}\}.$$

Since u is increasing and continuous its $\mathrm{range}(u)$ is a convex interval. Moreover, if we adopt the convention that $\sup\{\emptyset\} = -\infty$ and $\inf\{\emptyset\} = \infty$ we introduce the increasing inverse functions $u^{\hookleftarrow} : \mathbb{R} \to \mathbb{R}$ and $u^{\leftarrow} : \mathbb{R} \to \mathbb{R}$ given by

$$u^{\hookleftarrow}(x) = \sup\{y \in \mathbb{R} : u(y) \leq x\}$$

and

$$u^{\leftarrow}(x) = \inf\{y \in \mathbb{R} : u(y) \geq x\}$$

One can now show the following technical result. Before proving this result we introduce the following notions. If $h : \mathbb{R} \to \mathbb{R}$ is an increasing function the point x is called a point of strict increase if there exists some $\epsilon > 0$ such that h is strictly increasing on $(x-\epsilon, x+\epsilon)$. A function h is called right continuous if $\lim_{x_k \downarrow x} h(x_k) = h(x)$ for every $x \in \mathbb{R}$ and it is called left continuous if $\lim_{x_k \uparrow x} h(x_k) = h(x)$ for every $x \in \mathbb{R}$.

Lemma A.1 *If $u : \mathbb{R} \to \mathbb{R}$ is an increasing continuous function then $u^{\leftarrow}(x) \leq u^{\hookleftarrow}(x)$ for every $x \in \mathbb{R}$. Moreover the function $u^{\hookleftarrow} : R \to \mathbb{R}$ is right continuous and the function $u^{\leftarrow} : \mathbb{R} \to \mathbb{R}$ is left continuous and*

$$\lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k) = u^{\leftarrow}(x).$$

for every strictly increasing sequence $x_k \uparrow x$ with $x \in \mathbb{R}$. Also $u(u^{\leftarrow}(x)) = u(u^{\hookleftarrow}(x)) = x$ for every $x \in cl(range(u))$ and $u^{\leftarrow}(u(x)) = u^{\hookleftarrow}(u(x)) = x$ for every point x of strict increase.

Proof. For every $x \in range(u)$ it follows that these exists some $y \in \mathbb{R}$ satisfying $u(y) = x$ and this yields $u^{\leftarrow}(x) \leq y \leq u^{\hookleftarrow}(x)$. If $x \notin range(u)$ then exactly one of the sets $\{y \in \mathbb{R} : u(y) \leq x\}$ and $\{y \in \mathbb{R} : u(y) \geq x\}$ is empty. If $\{y \in \mathbb{R} : u(y) \leq x\}$ is empty or equivalently $\{y \in \mathbb{R} : u(y) \geq x\} = \mathbb{R}$ we obtain $u^{\leftarrow}(x) = -\infty$ and $u^{\hookleftarrow}(x) = \sup\{\emptyset\} = -\infty$, while for $\{y \in \mathbb{R} : u(y) \geq x\}$ is empty or equivalently $\{y \in \mathbb{R} : u(y) \leq x\} = \mathbb{R}$ it follows $u^{\hookleftarrow}(x) = \infty$ and $u^{\leftarrow}(x) = \inf\{\emptyset\} = \infty$. Hence we have verified that $u^{\leftarrow} \leq u^{\hookleftarrow}$. To show that the function $u^{\hookleftarrow}$ is right continuous consider some $x \in \mathbb{R}$ and a decreasing sequence $x_k \downarrow x$. This implies that the sequence $u^{\hookleftarrow}(x_k)$ is decreasing with lowerbound $u^{\hookleftarrow}(x)$ and so $\alpha = \lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k)$ exists (possibly $-\infty$) and $\alpha \geq u^{\hookleftarrow}(x)$. Without loss of generality we may assume that $u^{\hookleftarrow}(x) < \infty$. If $x \notin range(u)$ and hence $u^{\hookleftarrow}(x) = -\infty$ assume by contradiction that $\alpha > -\infty$. This implies that there exists some finite M satisfying $\alpha \geq M$. Since u is increasing it follows that the lower level sets $L(r) := \{y \in \mathbb{R} : u(y) \leq r\}$ are convex intervals and so we obtain $(-\infty, M) \subseteq L(x_k)$ for every $k \in \mathbb{N}$. Since $x_k \downarrow x$ we also obtain using $L(x_{k+1}) \subseteq L(x_k)$ for every $k \in \mathbb{N}$ that $L(x) = \cap_{k\in\mathbb{N}} L(x_k)$ and this implies $(-\infty, M) \subseteq L(x)$. However, by assumption $x \notin range(u)$ and so $L(x)$ is empty. This yields a contradiction and we have verified for $x \notin range(u)$ that $\lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k) = u^{\hookleftarrow}(x)$. Also for $x \in range(u)$ we assume by contradiction that $\alpha > u^{\hookleftarrow}(x)$. This means that $u(\alpha) > x$ and since $x_k \downarrow x$ one can find some $k_0 \in \mathbb{N}$ satisfying $u(\alpha) > x_{k_0}$. By the continuity of u there exist some $\alpha_0 < \alpha$ satisfying $u(\alpha_0) > x_{k_0}$ and so $u^{\hookleftarrow}(x_{k_0}) \leq \alpha_0 < \alpha$. Since by definition $u^{\hookleftarrow}(x_k) \downarrow \alpha$ it follows that $\alpha < \alpha$ and we have a contradiction. Therefore $\alpha = u^{\hookleftarrow}(x_\infty)$ and we have shown that $u^{\hookleftarrow}$ is right continuous. Similarly one can show that $u^{\leftarrow}$ is left continuous. To show that $\lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k) = u^{\leftarrow}(x)$ for every strictly increasing sequence $x_k \uparrow x$ we first observe that $x_k < x$ for every $k \in \mathbb{N}$. Since u is increasing this shows $u^{\hookleftarrow}(x_k) \leq u^{\leftarrow}(x)$. Moreover, the sequence $u^{\hookleftarrow}(x_k)$ is increasing and so $\alpha := \lim_{k\uparrow\infty} u^{\hookleftarrow}(x_k)$ exists and $\alpha \leq u^{\leftarrow}(x_\infty)$. Without loss of generality we may assume that $u^{\leftarrow}(x) > -\infty$ and $\alpha > -\infty$. Suppose now by contradiction that $\alpha < u^{\leftarrow}(x)$. This shows for both $x \in range(u)$ and $x \notin range(u)$ that $u(\alpha) < x$ and by the continuity of u and $x_k \uparrow x$ one can find as before some $k_0 \in \mathbb{N}$ and $\alpha_0 > \alpha$ satisfying $u(\alpha_0) < x_{k_0}$. This implies $\alpha < \alpha_0 \leq u^{\leftarrow}(x_{k_0})$ and since $u^{\leftarrow}(x_{k_0}) \leq u^{\hookleftarrow}(x_{k_0})$ and $u^{\hookleftarrow}(x_{k_0}) \uparrow \alpha$ we obtain as before a contradiction. Since for every $x \in range(u)$ the set $\{y : u(y) \leq x\}$ is convex and nonempty and u is increasing and continuous we obtain that this set is actually a closed nonempty set. Hence sup is attained and the same argument applies to inf. By this observation we obtain $u(u^{\hookleftarrow}(x)) \leq x$ and $u(u^{\leftarrow}(x)) \geq x$. Since u is increasing and $u^{\leftarrow}(x) \leq u^{\hookleftarrow}(x)$ this implies $x \leq u(u^{\leftarrow}(x)) \leq u(u^{\hookleftarrow}(x)) \leq x$ and so $u(u^{\leftarrow}(x)) = u(u^{\hookleftarrow}(x)) = x$ for every $x \in range(u)$. If $x \in \mathbb{R}$ belongs to $cl(range(u)) \backslash range(u)$ then x does not belong to range (u) and either $x = \sup\{u(y) : y \in \mathbb{R}\}$ finite or $x = \inf\{u(y) : y \in \mathbb{R}\}$ finite. In both cases it is easy to verify that $u(u^{\leftarrow}(x)) = u(u^{\hookleftarrow}(x)) = x$ and we have shown that $u(u^{\leftarrow}(x)) =$

$u(u^{\hookleftarrow}(x)) = x$ for every $x \in cl(range(u))$. To show the last part let x be a point of strict increase. By the monotonicity of u this implies that $\{y \in \mathbb{R} : u(y) \geq u(x)\} = [x, \infty)$ and $\{y \in \mathbb{R} : u(y) \leq x\} = (-\infty, x]$. This shows the last result. □

References

[1] Abramowitz, M. and I.A. Stegun, *Handbook of mathematical functions (ninth printing)*, Dover Publications, 1972.

[2] Arnold, S.F., *Mathematical statistics*, Prentice Hall, Englewood Cliffs, 1990.

[3] Artzner, P., Delbaen, F., Eber, J.M and D. Heath, *Coherent measures of risk*, Mathematical Finance **9** (1999), no. 3, 203–228.

[4] Bertsekas, D.P., *Nonlinear programming (2nd edition)*, Athena Scientific, Belmont, 1999.

[5] Embrechts,P., McNeil, A. and D. Straumann, *Correlation and dependence in risk management: properties and pittfals*, Risk Value Managemant at Risk and Beyond (Dempster, M.A.H., ed.), Cambridge University Press, Cambridge, 2002, pp. 176–223.

[6] Fang, K.T., Kotz, S. and K.W. Ng, *Symmetric multivariate and related distributions*, Chapman and Hall, New York, 1987.

[7] Fishburn, P.C., *Utility theory for decision making*, Wiley, New York, 1970.

[8] Hogg, R.V. and A.T. Craig, *Introduction to mathematical statistics*, MacMillan, New York, 1978.

[9] Luenberger, D. G., *Investment science*, Oxford University Press, New York, 1998.

[10] Lukacs, E., *Characteristics functions (2nd edition)*, Griffin, London, 1970.

[11] Michelot, C., *A finite algorithm for finding the projection of a point onto the canonical simplex of* $\mathbb{R}^n$, Journal of Optimization Theory and Applications **50** (1986), no. 1, 195–200.

[12] Rockafellar, R.T. and S. Uryasev, *Optimization of conditional value at risk*, The journal of risk **2** (2000), no. 3, 21–41.

[13] Ross, S.M., *Simulation (second edition)*, Statistical Modeling and Decision Science, Academic Press, San Diego, 1997.

[14] ______, *Simulation (3rd edition)*, Academic Press, New York, 2002.

[15] Serfling, R.J., *Approximation theorems in statistics*, Wiley, New York, 1980.

[16] Tijms, H.C., *Stochastic models :an algortihmic approach*, Wiley, New York, 1995.