Proceedings of the International Multiconference on
Computer Science and Information Technology pp. 901–908

ISBN 978-83-60810-22-4
ISSN 1896-7094

Efficient Portfolio Optimization with Conditional Value at Risk

Włodzimierz Ogryczak
Warsaw University of Technology
Institute of Control & Comput. Engg.
00–665 Warsaw, Poland
E-mail: ogryczak@ia.pw.edu.pl

Tomasz Sliwinski
Warsaw University of Technology
Institute of Control & Comput. Engg.
00–665 Warsaw, Poland
E-mail: tsliwins@ia.pw.edu.pl

***Abstract*—The portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem where the expected return is maximized and some (scalar) risk measure is minimized. In the original Markowitz model the risk is measured by the variance while several polyhedral risk measures have been introduced leading to Linear Programming (LP) computable portfolio optimization models in the case of discrete random variables represented by their realizations under specified scenarios. Among them, the second order quantile risk measures, recently, become popular in finance and banking. The simplest such measure, now commonly called the Conditional Value at Risk (CVaR) or Tail VaR, represents the mean shortfall at a specified confidence level. Recently, the second order quantile risk measures have been introduced and become popular in finance and banking. The corresponding portfolio optimization models can be solved with general purpose LP solvers. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model with huge number of variables and constraints thus decreasing the computational efficiency of the model since the number of constraints (matrix rows) is usually proportional to the number of scenarios. while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments. We show that the computational efficiency can be then dramatically improved with an alternative model taking advantages of the LP duality. In the introduced models the number of structural constraints (matrix rows) is proportional to the number of instruments thus not affecting seriously the simplex method efficiency by the number of scenarios.**

I. Introduction

FOLLOWING Markowitz [13], the portfolio selection problem is modeled as a mean-risk bicriteria optimization problem where the expected return is maximized and some (scalar) risk measure is minimized. In the original Markowitz model the risk is measured by the variance but several polyhedral risk measures have been introduced leading to Linear Programming (LP) computable portfolio optimization models in the case of discrete random variables represented by their realizations under specified scenarios. The simplest LP computable risk measures are dispersion measures similar to the variance. Konno and Yamazaki [6] presented the portfolio selection model with the mean absolute deviation (MAD) while Young [33] introduced the Minimax model. Yitzhaki [32] introduced the mean-risk model using Gini's mean (absolute) difference as the risk measure. The Gini's mean difference turn out to be a special aggregation technique of the multiple criteria LP model [18] based on the pointwise comparison of the absolute Lorenz curves. The latter leads the quantile shortfall risk measures directly related to the dual theory of choice under risk [26], [28], [31] which are more commonly used and accepted. Recently, the second order quantile risk measures have been introduced in different ways by many authors [2], [4], [16], [17], [27]. The measure, usually called the Conditional Value at Risk (CVaR) or Tail VaR, represents the mean shortfall at a specified confidence level. The CVaR measures maximization is consistent with the second degree stochastic dominance [19]. Several empirical analyses confirm its applicability to various financial optimization problems [1], [11].

This paper is focused on computational efficiency of the portfolio optimization models based on the CVaR or the Minimax measures. We assume that the instruments returns are represented by their realizations under T scenarios. The basic LP model for the CVaR portfolio optimization contains then T auxiliary variables as well as T corresponding linear inequalities. Actually, the number of structural constraints in the LP model (matrix rows) is proportional to the number of scenarios T, while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments $T+n$. Hence, its dimensionality is proportional to the number of scenarios T. It does not cause any computational difficulties for a few hundreds of scenarios as in computational analysis based on historical data. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios [25]. This may lead to the LP model with huge number of auxiliary variables and constraints thus decreasing the computational efficiency of the model. Actually, in the case of fifty thousand scenarios and one hundred instruments the model may require more than an hour computation time with the state-of-art LP solver (CPLEX code). To overcome this difficulty some alternative solution approaches are searched trying to reformulate the optimization problems as two-stage recourse problems [8] or to employ nondifferential optimization techniques [9]. We show that the computational efficiency can be then dramatically improved with an alternative model formulation taking advantages of the LP duality. In the introduced model the number of structural constraints is proportional to the number of instruments n

while only the number of variables is proportional to the number of scenarios T thus not affecting so seriously the simplex method efficiency. Indeed, the computation time is then below a minute.

The paper is organized as follows. In the next section we introduce briefly basics of the mean-risk portfolio optimization with the LP computable risk measures. In Section 3 we develop and test computationally efficient optimization models taking advantages of the LP duality.

II. Portfolio Optimization and Risk Measures

The portfolio optimization problem considered in this paper follows the original Markowitz' formulation and is based on a single period model of investment. At the beginning of a period, an investor allocates the capital among various securities, thus assigning a nonnegative weight (share of the capital) to each security. Let $J = \{1, 2, \ldots, n\}$ denote a set of securities considered for an investment. For each security $j \in J$, its rate of return is represented by a random variable R_j with a given mean $\mu_j = \mathbf{E}\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,2,\ldots,n}$ denote a vector of decision variables x_j expressing the weights defining a portfolio. The weights must satisfy a set of constraints to represent a portfolio. The simplest way of defining a feasible set Q is by a requirement that the weights must sum to one and they are nonnegative (short sales are not allowed), i.e.

$$Q = \{\mathbf{x} \ : \ \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n\} \tag{1}$$

Hereafter, we perform detailed analysis for the set Q given with constraints (1). Nevertheless, the presented results can easily be adapted to a general LP feasible set given as a system of linear equations and inequalities, thus allowing one to include short sales, upper bounds on single shares or portfolio structure restrictions which may be faced by a real-life investor.

Each portfolio $\mathbf{x}$ defines a corresponding random variable $R_{\mathbf{x}} = \sum_{j=1}^{n} R_j x_j$ that represents the portfolio rate of return while the expected value can be computed as $\mu(\mathbf{x}) = \sum_{j=1}^{n} \mu_j x_j$. We consider T scenarios with probabilities p_t (where $t = 1, \ldots, T$). We assume that for each random variable R_j its realization r_{jt} under the scenario t is known. Typically, the realizations are derived from historical data treating T historical periods as equally probable scenarios ($p_t = 1/T$). Although the models we analyze do not take advantages of this simplification. The realizations of the portfolio return $R_{\mathbf{x}}$ are given as $y_t = \sum_{j=1}^{n} r_{jt} x_j$.

The portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem where the mean $\mu(\mathbf{x})$ is maximized and the risk measure $\varrho(\mathbf{x})$ is minimized. In the original Markowitz model, the standard deviation was used as the risk measure. Several other risk measures have been later considered thus creating the entire family of mean-risk models (c.f., [10], [11]). These risk measures, similar to the standard deviation, are not affected by any shift of the outcome scale and are equal to 0 in the case of a risk-free portfolio while taking positive values for any risky portfolio. Unfortunately, such risk measures are not consistent with the stochastic dominance order [15] or other axiomatic models of risk-averse preferences [29] and coherent risk measurement [2].

In stochastic dominance, uncertain returns (modeled as random variables) are compared by pointwise comparison of some performance functions constructed from their distribution functions. The first performance function $F_{\mathbf{x}}^{(1)}$ is defined as the right-continuous cumulative distribution function: $F_{\mathbf{x}}^{(1)}(\eta) = F_{\mathbf{x}}(\eta) = \mathbf{P}\{R_{\mathbf{x}} \leq \eta\}$ and it defines the first degree stochastic dominance (FSD). The second function is derived from the first as $F_{\mathbf{x}}^{(2)}(\eta) = \int_{-\infty}^{\eta} F_{\mathbf{x}}(\xi) \ d\xi$ and it defines the second degree stochastic dominance (SSD). We say that portfolio $\mathbf{x}'$ dominates $\mathbf{x}''$ under the SSD ($R_{\mathbf{x}'} \succ_{_{SSD}} R_{\mathbf{x}''}$), if $F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta)$ for all η, with at least one strict inequality. A feasible portfolio $\mathbf{x}^0 \in Q$ is called SSD efficient if there is no $\mathbf{x} \in Q$ such that $R_{\mathbf{x}} \succ_{_{SSD}} R_{\mathbf{x}^0}$. Stochastic dominance relates the notion of risk to a possible failure of achieving some targets. As shown by Ogryczak and Ruszczyński [19], function $F_{\mathbf{x}}^{(2)}$, used to define the SSD relation, can also be presented as follows: $F_{\mathbf{x}}^{(2)}(\eta) = \mathbf{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\}$ and thereby its values are LP computable for returns represented by their realizations y_t.

An alternative characterization of the SSD relation can be achieved with the so-called Absolute Lorenz Curves (ALC) [16], [30] which represent the second quantile functions defined as $F_{\mathbf{x}}^{(-2)}(0) = 0$ and

$$F_{\mathbf{x}}^{(-2)}(p) = \int_0^p F_{\mathbf{x}}^{(-1)}(\alpha) d\alpha \tag{2}$$

where $F_{\mathbf{x}}^{(-1)}(p) = \inf \ \{\eta \ : \ F_{\mathbf{x}}(\eta) \ \geq \ p\}$ is the left-continuous inverse of the cumulative distribution function $F_{\mathbf{x}}$. The pointwise comparison of ALCs is equivalent to the SSD relation [20] in the sense that $R_{\mathbf{x}'} \succeq_{_{SSD}} R_{\mathbf{x}''}$ if and only if $F_{\mathbf{x}'}^{(-2)}(\beta) \geq F_{\mathbf{x}''}^{(-2)}(\beta)$ for all $0 < \beta \leq 1$. Moreover,

$$\begin{aligned} F_{\mathbf{x}}^{(-2)}(\beta) &= \max_{\eta \in R} \ \left[\beta\eta - F_{\mathbf{x}}^{(2)}(\eta)\right] \\ &= \max_{\eta \in R} \ [\beta\eta - \mathbf{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\}] \end{aligned} \tag{3}$$

where η is a real variable taking the value of β-quantile $Q_\beta(\mathbf{x})$ at the optimum. For a discrete random variable represented by its realizations y_t problem (3) becomes an LP.

For any real tolerance level $0 < \beta \leq 1$, the normalized value of the ALC defined as

$$M_\beta(\mathbf{x}) = F_{\mathbf{x}}^{(-2)}(\beta)/\beta \tag{4}$$

is called the *Conditional Value-at-Risk (CVaR)* or Tail VaR or Average VaR. The CVaR measure is an increasing function of the tolerance level β, with $M_1(\mathbf{x}) = \mu(\mathbf{x})$. For any $0 < \beta < 1$, the CVaR measure is SSD consistent [20] and coherent [24]. Opposite to deviation type risk measures, for coherent measures larger values are preferred and therefore the measures are sometimes called safety measures [11]. Due to (3), for a discrete random variable represented by its realizations y_t the CVaR measures are LP computable. It is important to notice that although the quantile risk measures (VaR and CVaR) were introduced in banking as extreme risk measures for very small tolerance levels (like $\beta = 0.05$), for the portfolio optimization good results have been provided by rather larger tolerance levels [11]. For β approaching 0, the CVaR measure tends to the Minimax measure $M(\mathbf{x}) = \min_{t=1,\ldots,T} y_t = \min_{t=1,\ldots,T} \sum_{j=1}^{n} r_{jt}x_j$ introduced to portfolio optimization by Young [33].

The commonly accepted approach to implementation of the Markowitz-type mean-risk model is based on the use of a specified lower bound μ_0 on expected returns while optimizing the risk measure. This bounding approach provides a clear understanding of investor preferences and a clear definition of optimal solution portfolio to be sought. For coherent and SSD consistent risk measures we consider the approach results in the following maximization problem:

$$\max\{\ \varrho(\mathbf{x}) : \quad \mu(\mathbf{x}) \geq \mu_0, \quad \mathbf{x} \in Q\} \tag{5}$$

where $\varrho(\mathbf{x}) = M_\beta(\mathbf{x})$ or $\varrho(\mathbf{x}) = M(\mathbf{x})$ respectively

We demonstrate that such portfolio optimization models can be effectively solved for large numbers of scenarios while taking advantages of appropriate dual LP formulations.

III. Computational LP Models

Let us consider portfolio optimization problem with security returns given by discrete random variables with realization r_{jt} thus leading to LP portfolio optimization model (5) for the risk measures we consider.

Following (3) and (4), the CVaR portfolio optimization model can be formulated as the following LP problem:

$$\begin{array}{ll} \max & \eta - \dfrac{1}{\beta}\displaystyle\sum_{t=1}^{T} p_t d_t \\ \text{s.t.} & \displaystyle\sum_{j=1}^{n} \mu_j x_j \geq \mu_0 \\ & \displaystyle\sum_{j=1}^{n} x_j = 1 \\ & x_j \geq 0,\ j = 1,\ldots,n \\ & d_t - \eta + \displaystyle\sum_{j=1}^{n} r_{jt}x_j \geq 0,\ d_t \geq 0,\ t = 1,\ldots,T \end{array} \tag{6}$$

where η is unbounded variable. Except from the core portfolio constraints (1) and the expected return bound, the model (6) contains T nonnegative variables d_t plus single η variable and T corresponding linear inequalities. Hence, its dimensionality is proportional to the number of scenarios T. Exactly, the LP model (6) contains $T + n + 1$ variables and $T + 2$ constraints. It does not cause any computational difficulties for a few hundreds of scenarios as in several computational analysis based on historical data [12]. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model (6) with huge number of variables and constraints thus decreasing the computational efficiency of the model. If the core portfolio constraints contain only linear relations, like (1), then the computational efficiency can easily be achieved by taking advantages of an alternative LP formulation.

Note that, due to the finite distribution of returns, the CVaR measure is well defined by the following optimization

$$\begin{aligned} M_\beta(\mathbf{x}) \quad = \quad & \min_{u_t}\{\sum_{t=1}^{T}(\sum_{j=1}^{n} r_{jt}x_j)u_t : \sum_{t=1}^{T} u_t = 1, \\ & 0 \leq u_t \leq \frac{p_t}{\beta}\ t = 1,\ldots,T\} \end{aligned} \tag{7}$$

that implements directly the ALC formula (2). The entire CVaR portfolio optimization problem may be respectively expressed as

$$\max_{\mathbf{x}\in Q} M_\beta(\mathbf{x}) = \max_{\mathbf{x}\in Q}\min_{\mathbf{u}\in U}\sum_{t=1}^{T}(\sum_{j=1}^{n} r_{jt}x_j)u_t \tag{8}$$

where

$$\begin{aligned} U \quad = \quad & \{(u_1,\ldots,u_T) : \sum_{t=1}^{T} u_t = 1, \\ & 0 \leq u_t \leq \frac{p_t}{\beta} \quad t = 1,\ldots,T\}. \end{aligned}$$

The inner optimization problem represents (7). It is an LP for a given vector $\mathbf{x}$. However, within the entire portfolio optimization problem (8) the objective function $\sum_{t=1}^{T}(\sum_{j=1}^{n} r_{jt}x_j)u_t$ becomes nonlinear for $\mathbf{x}$ being a vector of variables. This difficulty can be overcome by an equivalent dual LP formulation of problem (7). Indeed, introducing dual variable η corresponding to the equation $\sum_{t=1}^{T} u_t = 1$ and variables d_t corresponding to upper bounds on u_t one gets the LP dual

$$\begin{aligned} M_\beta(\mathbf{x}) \quad = \quad & \max_{q,d_t}\ \{\eta - \frac{1}{\beta}\sum_{t=1}^{T} p_t d_t : \\ & \sum_{j=1}^{n} r_{jt}x_j \geq \eta - d_t,\ d_t \geq 0\ \forall\ t\}. \end{aligned} \tag{9}$$

This leads us to the standard LP model (6) for the CVaR portfolio optimization.

An alternative CVaR optimization model can be built by taking advantages of the minimax theorem. Since both sets Q and U are convex polyhedra, formula (8) can be rewritten into a dual form

$$\begin{aligned} & \max_{\mathbf{x}\in Q}\min_{\mathbf{u}\in U}\sum_{t=1}^{T}(\sum_{j=1}^{n} r_{jt}x_j)u_t \\ & = \min_{\mathbf{u}\in U}\max_{\mathbf{x}\in Q}\sum_{t=1}^{T}(\sum_{j=1}^{n} r_{jt}x_j)u_t \\ & = \min_{\mathbf{u}\in U} D(\mathbf{u}) \end{aligned} \tag{10}$$

with the inner optimization problem

$$\begin{aligned} D(\mathbf{u}) &= \max_{x_j} \quad \{\sum_{t=1}^{T} u_t \sum_{j=1}^{n} r_{jt} x_j : \\ & \qquad \sum_{j=1}^{n} \mu_j x_j \geq \mu_0,\ \sum_{j=1}^{n} x_j = 1, \\ & \qquad x_j \geq 0,\ j = 1, \ldots, n\} \\ &= \max_{x_j} \quad \{\sum_{j=1}^{n} (\sum_{t=1}^{T} r_{jt} u_t) x_j : \\ & \qquad \sum_{j=1}^{n} \mu_j x_j \geq \mu_0,\ \sum_{j=1}^{n} x_j = 1, \\ & \qquad x_j \geq 0,\ j = 1, \ldots, n\}. \end{aligned} \tag{11}$$

Again, we may take advantages of the LP dual to the inner problem. Indeed, introducing dual variable q corresponding to the equation $\sum_{j=1}^{n} x_j = 1$ and variable u_0 corresponding to the inequality $\sum_{j=1}^{n} \mu_j x_j \geq \mu_0$ we get the LP dual

$$\begin{aligned} D(\mathbf{u}) = \min_{q,u_0} \quad & \{q - \mu_0 u_0 : \\ & q - \mu_j u_0 - \sum_{t=1}^{T} r_{jt} u_t \geq 0 \quad j = 1, \ldots, n\}. \end{aligned}$$

Hence, an alternative model for the CVaR portfolio optimization (10) can be expressed as the following LP:

$$\begin{aligned} \min \quad & q - \mu_0 u_0 \\ \text{s.t.} \quad & q - \mu_j u_0 - \sum_{t=1}^{T} r_{jt} u_t \geq 0,\ j = 1, \ldots, n \\ & \sum_{t=1}^{T} u_t = 1 \\ & 0 \leq u_t \leq \frac{p_t}{\beta},\ t = 1, \ldots, T \end{aligned} \tag{12}$$

LP model (12) contains $T+1$ variables u_t, but the T constraints corresponding to variables d_t from (6) take the form of simple upper bounds on u_t (for $t = 1, \ldots, T$) thus not affecting the problem complexity. The number of constraints in (12) is proportional to the total of portfolio size n, thus it is independent from the number of scenarios. Exactly, there are $T+1$ variables and $n+1$ constraints. This guarantees a high computational efficiency of the model even for very large number of scenarios. Similarly, other portfolio structure requirements are modeled with rather small number of constraints thus generating small number of additional variables in the model. Actually, the model (12) is the LP dual to the model (6), thus similar to that introduced in [23]. Obviously, the optimal portfolio shares x_j are not directly represented within the solution vector of problem (12) but they are easily available as the dual variables (shadow prices) for inequalities $q - \mu_j u_0 - \sum_{t=1}^{T} r_{jt} u_t \geq 0$.

The Minimax portfolio optimization model can be written as the following LP problem:

$$\begin{aligned} \max \quad & \eta \\ \text{s.t.} \quad & \sum_{j=1}^{n} \mu_j x_j \geq \mu_0, \\ & \sum_{j=1}^{n} x_j = 1 \\ & x_j \geq 0,\ j = 1, \ldots, n \\ & -\eta + \sum_{j=1}^{n} r_{jt} x_j \geq 0,\ t = 1, \ldots, T \end{aligned} \tag{13}$$

which is simpler than the standard CVaR optimization model (6). Except from the portfolio weights x_j, the model contains only one additional variable η. Nevertheless, it still contains T linear inequalities in addition to the core constraints. Hence, its dimensionality is $(T+2) \times (n+1)$.

The Minimax portfolio optimization model representing a limiting case of the CVaR model for β tending to 0. Actually, for any $\beta \leq \min_{t=1,\ldots,T} p_t$ we gets $M_\beta(\mathbf{x}) = M(\mathbf{x})$ thus allowing to represent the Minimax portfolio optimization by the CVaR optimization model (6) and to take advantages of its dual form (12). Due to $\beta \leq p_t$ for all $t = 1, \ldots, T$, the upper bounds on variables u_t becomes redundant and we get the following dual form of the Minimax portfolio optimization:

$$\begin{aligned} \min \quad & q - \mu_0 u_0 \\ \text{s.t.} \quad & q - \mu_j u_0 - \sum_{t=1}^{T} r_{jt} u_t \geq 0,\ j = 1, \ldots, n \\ & \sum_{t=1}^{T} u_t = 1 \\ & u_t \ \geq 0,\ t = 1, \ldots, T \end{aligned} \tag{14}$$

The model dimensionality is only $(n+1) \times (T+2)$ thus guaranteeing a high computational efficiency even for very large number of scenarios.

The Mean Absolute Deviation (MAD) risk measure is directly given by the value of the second order cdf $F_{\mathbf{x}}^{(2)}$ at the mean $\bar{\delta}(\mathbf{x}) = \mathbf{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = F_{\mathbf{x}}^{(2)}(\mu(\mathbf{x}))$ [19]. Therefore, its leads to an LP portfolio optimization model very similar to that for the CVaR optimization (6). Indeed, we get:

$$\begin{aligned} \max \quad & -\sum_{t=1}^{T} p_t d_t \\ \text{s.t.} \quad & \sum_{j=1}^{n} \mu_j x_j \geq \mu_0, \quad \sum_{j=1}^{n} x_j = 1 \\ & d_t \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j,\ d_t \geq 0,\ t = 1, \ldots, T \\ & x_j \geq 0,\ j = 1, \ldots, n \end{aligned} \tag{15}$$

with $T+n$ variables and $T+2$ constraints. The LP dual model

TABLE I
COMPUTATIONAL TIMES (IN SECONDS) FOR THE STANDARD MEAN-RISK MODELS

Scenarios (T)	Securities (n)	CVaR (6) $\beta = 0.05$	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.4$	$\beta = 0.5$	Minimax (13)	MAD (15)
5 000	76	2.5	2.6	3.7	5.2	6.7	7.6	0.5	19.0
7 000	76	4.1	4.4	6.9	9.3	11.8	14.8	0.9	9.3
10 000	76	6.5	8.3	13.9	18.4	24.2	29.8	1.3	18.4
50 000	50	3275.4	4876.6	–	–	–	–	7.8	–
50 000	100	–	–	–	–	–	–	24.1	–

TABLE II
COMPUTATIONAL TIMES (IN SECONDS) FOR THE DUAL MEAN-RISK MODELS

Scenarios (T)	Securities (n)	CVaR (12) $\beta = 0.05$	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.4$	$\beta = 0.5$	Minimax (14)	MAD (16)
5 000	76	0.9	0.9	1.1	1.3	1.4	1.6	0.5	2.1
7 000	76	1.2	1.4	1.8	2.0	2.2	2.3	0.7	3.1
10 000	76	2.0	2.3	2.9	3.4	3.9	4.0	1.0	10.8
50 000	50	14.9	19.4	24.0	26.8	28.2	27.9	3.9	25.8
50 000	100	40.0	54.6	68.7	77.7	78.2	78.8	8.2	76.7

takes then the form:

$$\begin{aligned} \min \quad & q - \mu_0 u_0 \\ \text{s.t.} \quad & q \geq \mu_j u_0 - \sum_{t=1}^{T} (\mu_j - r_{jt}) u_t, \; j = 1, \ldots, n \\ & 0 \leq u_t \; \leq p_t, \; t = 1, \ldots, T \end{aligned} \tag{16}$$

with dimensionality $n \times (T + 2)$. Hence, there is again guaranteed the high computational efficiency even for very large number of scenarios.

We have run two groups of computational tests. The medium scale tests of 5 000, 7 000 and 10 000 scenarios and 76 securities were generated following the FTSE 100 related data [5]. The large scale tests instances developed by Lim et al. [9] were generated from a multivariate normal distribution for 50 or 100 securities with the number of scenarios 50 000 just providing an adequate approximation to the underlying unknown continuous price distribution. When applying the lower bound on the required expected return, its value μ_0 was defined as the expected return of the portfolio with equal weights (market value). All computations were performed on a PC with the Pentium 4 2.6GHz processor and 3GB RAM employing the simplex code of the CPLEX 9.1 package.

In Tables I and II there are presented computation times for all the above primal and dual models. All results are presented as the averages of 10 different test instances of the same size. For the medium scale test problems the solution times of the dual CVaR models (12) ranging from 0.9 to 4.0 seconds are not much shorter than those for the primal models ranging from 2.5 to 29.8 seconds, respectively. However, an attempt to solve the primal CVaR model (6) of the large scale test problems was successful only for $\beta = 0.05$ and $\beta = 0.1$ and the times were dramatically longer than those for the dual model. For other values of β the timeout of 6000 seconds occurred (marked with '–'). For 100 securities the primal model was not solvable within the given time limit, while the dual models could be successfully solved in 40.0 to 78.8 seconds.

The Minimax models are computationally very easy. Running the computational tests we were able to solve the medium scale test instances of the dual model (14) in times below 1 second and the large scale test instances in up to 8.2 seconds on average. In fact, even the primal model could be solved in reasonable time between 0.5 and 24 seconds, for medium and large scale test instances, respectively.

The MAD models are computationally similar to the CVaR models. Indeed, only medium scale test instances of the primal model (15) could be solved within the given time limit. Much shorter computing times could be achieved for the dual MAD model (16) – not more than 10.8 seconds for the medium scale and 76.7 seconds for the large scale test instances.

To see how the value of the required expected return affects the solution times we have performed additional tests for the reformulated CVaR (with $\beta = 0.1$) and Minimax models with increased value μ_0. The increased value was set in the middle between the expected return of the market value and the maximum possible return for single security portfolio. The computational times are generally comparable with those for the market value constraints. Actually, for the medium scale problems (10 000 scenarios) the CVaR optimization time has remained unchanged (2.3 seconds) while the Minimax optimization time has only raised from 1.0 to 1.1 seconds and for the MAD model optimization it has raised from 10.8 to 13.9. For the large scale problems (50 000 scenarios) we have noticed even a drop in computation times for the CVaR model (reduction from 54.6 to 49.9 seconds) and similar reduction (from 76.7 to 55.4 seconds) has occured for the MAD model optimization while the Minimax optimization time has increased from 8.2 to 10.5 seconds.

IV. GINI'S MEAN DIFFERENCE AND RELATED MODELS

Yitzhaki [32] introduced the portfolio optimization model using Gini's mean difference (GMD) as risk measure. The GMD is given as $\Gamma(\mathbf{x}) = \frac{1}{2} \int \int |\eta - \xi| dF_{\mathbf{x}}(\eta) dF_{\mathbf{x}}(\xi)$ although several alternative formulae exist. For a discrete random variable represented by its realizations y_t, the measure

$$\Gamma(\mathbf{x}) = \sum_{t'=1}^{T} \sum_{t'' \neq t'-1} \max\{y_{t'} - y_{t''}, 0\} p_{t'} p_{t''}$$

is LP computable (when minimized) leading to the following portfolio optimization model:

$$\begin{array}{ll} \max & -\sum_{t=1}^{T} \sum_{t' \neq t} p_t p_{t'} d_{tt'} \\ \text{s.t.} & \sum_{j=1}^{n} \mu_j x_j \geq \mu_0, \quad \sum_{j=1}^{n} x_j = 1 \\ & d_{tt'} \geq \sum_{j=1}^{n} r_{jt} x_j - \sum_{j=1}^{n} r_{jt'} x_j \\ & d_{tt'} \geq 0, \ t \neq t' = 1, \ldots, T \\ & x_j \geq 0, \ j = 1, \ldots, n \end{array} \tag{17}$$

which contains $T(T-1)$ nonnegative variables $d_{tt'}$ and $T(T-1)$ inequalities to define them. This generates a huge LP problem even for the historical data case where the number of scenarios is 100 or 200. Krzemienowski and Ogryczak [7] have shown with the earlier experiments that the CPU time of 7 seconds on average for $T = 52$ has increased to above 30 sec. with $T = 104$ and even more than 180 sec. for $T = 156$. However, similar to the CVaR models, variables $d_{tt'}$ are associated with the singleton coefficient columns. Hence, while solving the dual instead of the original primal, the corresponding dual constraints take the form of simple upper bounds (SUB) which are handled implicitly outside the LP matrix. For the simplest form of the feasible set (1) the dual GMD model takes the following form:

$$\begin{array}{l} \min q - \mu_0 u_0 \\ q \geq \mu_j u_0 + \sum_{t=1}^{T} \sum_{t' \neq t} (r_{jt} - r_{jt'}) u_{tt'}, \ j = 1, \ldots, n \\ 0 \leq u_{tt'} \ \leq p_t p_{t'}, \ t, t' = 1, \ldots, T; t \neq t' \end{array} \tag{18}$$

where original portfolio variables x_j are dual prices to the inequalities. The dual model contains $T(T-1)$ variables $u_{tt'}$ but the number of constraints (excluding the SUB structure) $n+1$ is proportional to the number of securities. The above dual formulation can be further simplified by introducing variables:

$$\bar{u}_{tt'} = u_{tt'} - u_{t't}, \quad t, t' = 1, \ldots, T; t < t' \tag{19}$$

which allows us to reduce the number of variables to $T(T-1)/2$ by replacing (18) with the following:

$$\begin{array}{l} \min q - \mu_0 u_0 \\ q \geq \mu_j u_0 + \sum_{t=1}^{T} \sum_{t' > t} (r_{jt} - r_{jt'}) \bar{u}_{tt'}, \ j = 1, \ldots, n \\ -p_t p_{t'} \leq \bar{u}_{tt'} \ \leq p_t p_{t'}, \ t < t' = 1, \ldots, T \end{array} \tag{20}$$

Such a dual approach may dramatically improve the LP model efficiency in the case of larger number of scenarios. Actually, as shown with the earlier experiments of [7], the above dual formulations let us to reduce the optimization time below 10 seconds for $T = 104$ and $T = 156$. Nevertheless, the case of really large number of scenarios still may cause computational difficulties, due to huge number of variables ($T(T-1)/2$). This may require some column generation techniques [3] or nondifferentiable optimization algorithms [9].

As shown by Yitzhaki [32] for the SSD consistency of the GMD model one needs to maximize the complementary measure

$$\mu_\Gamma(\mathbf{x}) = \mu(\mathbf{x}) - \Gamma(\mathbf{x}) = \mathbf{E}\{R_\mathbf{x} \wedge R_\mathbf{x}\} \tag{21}$$

where the cumulative distribution function of $R_\mathbf{x} \wedge R_\mathbf{x}$ for any $\eta \in \mathbf{R}$ is given as $F_\mathbf{x}(\eta)(2 - F_\mathbf{x}(\eta))$. Hence, (21) is the expectation of the minimum of two independent identically distributed random variables (i.i.d.r.v.) $R_\mathbf{x}$ thus representing the *mean worse return*. This provides us with another LP model although it is not more compact than that of (17) and its dual (18). Alternatively, the GMD may be expressed with integral of the absolute Lorenz curve as

$$\begin{aligned} \Gamma(\mathbf{x}) &= 2\int_0^1 (\alpha\mu(\mathbf{x}) - F_\mathbf{x}^{(-2)}(\alpha))d\alpha \\ &= 2\int_0^1 \alpha(\mu(\mathbf{x}) - M_\alpha(\mathbf{x}))d\alpha \end{aligned}$$

and respectively

$$\mu_\Gamma(\mathbf{x}) = 2\int_0^1 F_\mathbf{x}^{(-2)}(\alpha)d\alpha = 2\int_0^1 \alpha M_\alpha(\mathbf{x})d\alpha \tag{22}$$

thus combining all the CVaR measures. In order to enrich the modeling capabilities, one may treat differently some more or less extreme events. In order to model downside risk aversion, instead of the Gini's mean difference, the *tail Gini's* measure introduced by Ogryczak and Ruszczyński [21], [20] can be used:

$$\begin{aligned} \mu_{\Gamma_\beta}(\mathbf{x}) &= \mu(\mathbf{x}) - \frac{2}{\beta^2}\int_0^\beta (\mu(\mathbf{x})\alpha - F_\mathbf{x}^{(-2)}(\alpha))d\alpha \\ &= \frac{2}{\beta^2}\int_0^\beta F_\mathbf{x}^{(-2)}(\alpha)d\alpha \end{aligned} \tag{23}$$

In the simplest case of equally probable T scenarios with $p_t = 1/T$ (historical data for T periods), the tail Gini's measure for $\beta = K/T$ may be expressed as the weighted combination of CVaRs $M_{\beta_k}(\mathbf{x})$ with tolerance levels $\beta_k = k/T$ for $k = 1, 2, \ldots, K$ and properly defined weights [21]. In a general case, we may resort to an approximation based on some reasonably chosen grid β_k, $k = 1, \ldots, m$ and weights w_k expressing the corresponding trapezoidal approximation of the integral in the formula (23). Exactly, for any $0 < \beta \leq 1$, while using the grid of m tolerance levels $0 < \beta_1 < \ldots < \beta_k < \ldots < \beta_m = \beta$ one may define weights:

$$\begin{aligned} w_k &= \frac{(\beta_{k+1} - \beta_{k-1})\beta_k}{\beta^2}, \ k = 1, \ldots, m-1 \\ w_m &= \frac{\beta - \beta_{m-1}}{\beta} \end{aligned} \tag{24}$$

where $\beta_0 = 0$. This leads us to the Weighted CVaR (WCVaR) measure [12] defined as

$$\begin{aligned} & M_{\mathbf{w}}^{(m)}(\mathbf{x}) = \sum_{k=1}^{m} w_k M_{\beta_k}(\mathbf{x}) \\ & \sum_{k=1}^{m} w_k = 1, \quad w_k > 0,\ k = 1, \ldots, m \end{aligned} \tag{25}$$

We emphasize that despite being only an approximation to (23), any WCVaR measure itself is a well defined LP computable measure with guaranteed SSD consistency and coherency, as a combination of the CVaR measures. Hence, it needs not to be built on a very dense grid to provide proper modeling of risk averse preferences. While analyzed on the real-life data from the Milan Stock Exchange the weighted CVaR models have usually performed better than the GMD itself, the Minimax or the extremal CVaR models [12].

Here we analyze only computational efficiency of the LP models representing the WCVaR portfolio optimization. For returns represented by their realizations we get the following LP optimization problem:

$$\begin{aligned} \max \quad & \sum_{k=1}^{m} w_k \eta_k - \sum_{k=1}^{m} \frac{w_k}{\beta_k} \sum_{t=1}^{T} p_t d_{tk} \\ \text{s.t.} \quad & \sum_{j=1}^{n} \mu_j x_j \geq \mu_0, \quad \sum_{j=1}^{n} x_j = 1 \\ & x_j \geq 0,\ j = 1, \ldots, n \\ & d_{tk} \geq \eta_k - \sum_{j=1}^{n} r_{jt} x_j \\ & d_{tk} \geq 0,\ t = 1, \ldots, T; k = 1, \ldots, m \end{aligned} \tag{26}$$

where η_k (for $k = 1, \ldots, m$) are unbounded variables taking the values of the corresponding β_k-quantiles (in the optimal solution). The problem dimensionality is proportional to the number of scenarios T and to the number of tolerance levels m. Exactly, the LP model contains $m \times T + n$ variables and $m \times T + 2$ constraints. The LP problem structure is similar to that of repereseting the so-called WOWA optimization in fuzzy approaches [22]. It does not cause any computational difficulties for a few hundreds scenarios and a few tolerance levels, as in a simple computational analysis based on historical data [12]. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands scenarios. This may lead to the LP model (26) with huge number of variables and constraints thus decreasing the computational efficiency of the model. If the core portfolio constraints contain only linear relations, like (1), then the computational efficiency can easily be achieved by taking advantages of the LP dual to model (26). The LP dual model takes the following form:

$$\begin{aligned} \min \quad & q - \mu_0 u_0 \\ \text{s.t.} \quad & q - \mu_j u_0 - \sum_{t=1}^{T} r_{jt} \sum_{k=1}^{m} u_{tk} \geq 0,\ j = 1, \ldots, n \\ & \sum_{t=1}^{T} u_{tk} = w_k \\ & 0 \leq u_{tk} \leq \frac{p_t w_k}{\beta_k},\ t = 1, \ldots, T; k = 1, \ldots, m \end{aligned} \tag{27}$$

that contains $m \times T$ variables u_{tk}, but the $m \times T$ constraints corresponding to variables d_{tk} from (26) take the form of simple upper bounds on u_{tk} thus not affecting the problem complexity. Hence, again the number of constraints in (27) is proportional to the total of portfolio size n and the number of tolerance levels m, thus it is independent from the number of scenarios. Exactly, there are $m \times T + 2$ variables and $m + n$ constraints thus guaranteeing a high computational efficiency of for very large number of scenarios.

TABLE III
COMPUTATIONAL TIMES (IN SECONDS) FOR THE DUAL WCVAR MODELS

Scenarios (T)	Securities (n)	Model (27) $m = 3$	$m = 5$
5 000	76	6.2	13.8
7 000	76	10.0	22.8
10 000	76	16.2	37.7
50 000	50	121.8	281.0
50 000	100	335.3	731.7

We have tested computational efficiency of the dual model (27) using the same randomly generated test instances as for testing of the CVaR and other basic models in Section III. Table III presents average computation times of the dual models for $m = 3$ with tolerance levels $\beta_1 = 0.1$, $\beta_2 = 0.25$, $\beta_3 = 0.5$ and weights $w_1 = 0.1$, $w_2 = 0.4$ and $w_3 = 0.5$, thus representing the parameters leading to good results on real life data [12], as well as for $m = 5$ with uniformly distributed tolerance levels $\beta_1 = 0.1$, $\beta_2 = 0.2$, $\beta_3 = 0.3$, $\beta_4 = 0.4$, $\beta_5 = 0.5$ and weights (24).

V. CONCLUDING REMARKS

The classical Markowitz model uses the variance as the risk measure, thus resulting in a quadratic optimization problem. There were introduced several alternative risk measures which are computationally attractive as (for discrete random variables) they result in solving linear programming (LP) problems. The LP solvability is very important for applications to real-life financial decisions where the constructed portfolios have to meet numerous side constraints and take into account transaction costs. A gamut of LP computable risk measures has been presented in the portfolio optimization literature although most of them are related to the absolute Lorenz curve and thereby the CVaR measures. We have shown that all the risk measures used in the LP solvable portfolio optimization models can be derived from the SSD shortfall criteria. This allows us to guarantee their SSD consistency for any distribution of outcomes.

The corresponding portfolio optimization models can be solved with general purpose LP solvers. However, in the case of more advanced simulation models employed for scenario generation one may get several thousands of scenarios. This may lead to the LP model with huge number of variables and constraints thus decreasing the computational efficiency of the models. For the CVaR model, the number of constraints (matrix rows) is proportional to the number of scenarios. while the number of variables (matrix columns) is proportional to the total of the number of scenarios and the number of instruments. We have shown that the computational efficiency can be then dramatically improved with an alternative model taking advantages of the LP duality. In the introduced model the number of structural constraints (matrix rows) is proportional to the number of instruments thus not affecting seriously the simplex method efficiency by the number of scenarios. In particular, for the case of 50 000 scenarios, it has resulted in computation times below 30 seconds for 50 securities or below a minute for 100 instruments. Similar computational times have also been achieved for the reformulated Minimax model.

Similar reformulation can be developed for other LP computable portfolio optimization models as many of them are related to the Absolute Lorenz Curve [18], [10]. In particular, this applies to the classical LP portfolio optimization models based on the mean absolute deviation as well as to the Gini's mean difference. The standard LP models for the Gini's mean difference [32] and its downside version [7] require T^2 auxiliary constraints which makes them hard already for medium numbers of scenarios, like a few hundred scenarios given by historical data. The models taking advantages of the LP duality allow one to limit the number of structural constraints making it proportional to the number of scenarios T thus increasing dramatically computational performances for medium numbers of scenario although still remaining hard for very large numbers of scenarios.

Acknowledgment

The authors are indebted to Churlzu Lim from the University of North Carolina at Charlotte for providing the test data of 50 000 scenarios portfolio selection problem and to Diana Roman from CARISMA, Brunel University for providing the FTSE 100 related test data of 5 000, 7 000 and 10 000 scenarios we have used in our computational studies.

References

[1] Andersson, F, H Mausser, D Rosen and S Uryasev, (2001), Credit risk optimization with conditional value-at-risk criterion, *Mathematical Programming* 89, 273–291.

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

[3] Desrosiers, J and M Luebbecke (2005), A primer in column generation, in: G Desaulniers, J Desrosier, and M Solomon, eds.: *Column Generation*, Springer, New York, 132.

[4] Embrechts, P, C Klüppelberg and T Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, New York.

[5] Fabian, C I, G Mitra, D Roman and V Zverovich (2010), An enhanced model for portfolio choice with SSD criteria: a constructive approach, *Quantitative Finance*, forthcoming, DOI: 10.1080/14697680903493607.

[6] Konno, H and H Yamazaki (1991), Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market, *Management Science* 37, 519–531.

[7] Krzemienowski, A and W Ogryczak (2005), On extending the LP computable risk measures to account downside risk, *Computational Optimization and Applications* 32, 133–160.

[8] Künzi-Bay, A and J Mayer (2006). Computational aspects of minimizing conditional value-at-risk. *Computational Management Science*, 3, 3–27.

[9] Lim, C, H D Sherali and S Uryasev (2010), Portfolio optimization by minimizing conditional value-at-risk via nondifferentiable optimization, *Computational Optimization and Applications* 46, 391–415

[10] Mansini, R, W Ogryczak and M G Speranza (2003), On LP solvable models for portfolio selection, *Informatica* 14, 37–62.

[11] Mansini, R, W Ogryczak and M G Speranza (2003), LP solvable models for portfolio optimization: A classification and computational comparison, *IMA Journal of Management Mathematics* 14, 187–220.

[12] Mansini, R, W Ogryczak and M G Speranza (2007), Conditional value at risk and related linear programming models for portfolio optimization, *Annals of Operations Research* 152, 227–256.

[13] Markowitz, H M (1952), Portfolio selection, *Journal of Finance* 7, 77–91.

[14] Miller, N and A Ruszczyński (2008). Risk-adjusted probability measures in portfolio optimization with coherent measures of risk. *European Journal of Operational Research*, 191, 193–206.

[15] Müller, A and D Stoyan (2002), *Comparison Methods for Stochastic Models and Risks*, Wiley, Chichester.

[16] Ogryczak, W (1999), Stochastic dominance relation and linear risk measures, in A M J Skulimowski, ed.: *Financial Modelling – Proc. 23rd Meeting EURO WG Financial Modelling, Cracow, 1998*, Progress & Business Publisher, Cracow, 191–212.

[17] Ogryczak, W (2000), Risk measurement: Mean absolute deviation versus Ginis mean difference, in W G Wanka, ed.: *Decision Theory and Optimization in Theory and Practice – Proc. 9th Workshop GOR WG, Chemnitz, 1999*, Shaker Verlag, Aachen, 33–51.

[18] Ogryczak, W (2000), Multiple criteria linear programming model for portfolio selection, *Annals of Operations Research* 97, 143–162.

[19] Ogryczak, W and A Ruszczyński (1999), From stochastic dominance to mean-risk models: semideviations as risk measures, *European Journal of Operational Research* 116, 33–50.

[20] Ogryczak, W and A Ruszczyński (2002), Dual stochastic dominance and related mean-risk models, *SIAM J. Optimization* 13, 60–78.

[21] Ogryczak, W and A Ruszczyński (2002), Dual Stochastic dominance and quantile risk measures, *International Transactions in Operational Research* 9, 661–680.

[22] Ogryczak W and T Śliwiński (2009), On efficient WOWA optimization for decision support under risk, *International Journal of Approximate Reasoning* 50, 915–928.

[23] Ogryczak W and T Śliwiński (2010), On solving the dual for portfolio selection by optimizing Conditional Value at Risk, *Computational Optimization and Applications*, forthcoming, DOI 10.1007/s10589-010-9321-y.

[24] Pflug, G Ch (2000), Some remarks on the value-at-risk and the conditional value-at-risk, in S Uryasev, ed.: *Probabilistic Constrained Optimization: Methodology and Applications*, Kluwer AP, Dordrecht.

[25] Pflug, G Ch (2001), Scenario tree generation for multiperiod financial optimization by optimal discretization, *Mathematical Programming* 89, 251-271.

[26] Quiggin, J (1982), A theory of anticipated utility, *Journal of Economic Behavior and Organization* 3, 323–343.

[27] Rockafellar, R T and S Uryasev (2000), Optimization of conditional value-at-risk, *Journal of Risk* 2, 21–41.

[28] Roell, A (1987), Risk aversion in Quiggin and Yaari's rank-order model of choice under uncertainty, *Economic Journal* 97, 143–159.

[29] Rothschild, M and J E Stiglitz (1969), Increasing risk: I. A definition, *Journal of Economic Theory* 2, 225–243.

[30] Shorrocks, A F (1983), Ranking income distributions, *Economica* 50, 3–17.

[31] Yaari, M E (1987), The dual theory of choice under risk, *Econometrica* 55, 95–115.

[32] Yitzhaki, S (1982), Stochastic dominance, mean variance, and Gini's mean difference, *The American Economic Revue* 72, 178–185.

[33] Young, M R (1998), A minimax portfolio selection rule with linear programming solution, *Management Science* 44, 673–683.