Akademie věd České republiky
Ústav teorie informace a automatizace

Czech Academy of Sciences
Institute of Information Theory and Automation

RESEARCH REPORT

Martin BRANDA, Michal ČERVINKA, Alexandra SCHWARTZ

SPARSE ROBUST PORTFOLIO OPTIMIZATION VIA NLP REGULARIZATIONS

No. 2358 December 2016

ÚTIA AV čR, P. O. Box 18, 182 08 Prague, Czech Republic
Fax: (+420)(2)688 4903
E-mail: utia@utia.cas.cz

Sparse robust portfolio optimization via NLP regularizations

Martin BRANDA[1], Michal ČERVINKA[2], Alexandra SCHWARTZ[3]

Abstract

We deal with investment problems where we minimize a risk measure under a condition on the sparsity of the portfolio. Various risk measures are considered including Value-at-Risk and Conditional Value-at-Risk under normal distribution of returns and their robust counterparts are derived under moment conditions, all leading to nonconvex objective functions. We propose four solution approaches: a mixed-integer formulation, a relaxation of an alternative mixed-integer reformulation and two NLP regularizations. In a numerical study, we compare their computational performance on a large number of simulated instances taken from the literature.

1 Introduction

In portfolio optimization, two basic types of decision-making frameworks are adopted: the utility maximization and the return-risk trade-off analysis, see, e.g., Levy [28] for properties and relations between these two approaches. In the latter, it is important to define a risk that the concerned system has. In optimization problems governed by uncertain inputs, typically represented as random variables, the risk is explicitly quantified by a risk measure.

In return-risk analysis, widely used both in theory and practice, an investor faces a trade-off between expected return and associated risk. In his pioneering work in 1952, Markowitz [29] adopted variance as a measure of risk in his mean-variance analysis. Variance measures equally both positive and negative

[1]Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, CZ-186 75 Prague 8, Czech Republic.
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, CZ-182 08 Prague 8 , Czech Republic.
branda@karlin.mff.cuni.cz

[2]Institute of Economic Studies, Faculty of Social Sciences, Charles University, Smetanovo nábřeží 6, CZ-110 01 Prague 1, Czech Republic.
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, CZ-182 08 Prague 8 , Czech Republic.
michal.cervinka@fsv.cuni.cz; cervinka@utia.cas.cz

[3]Graduate School of Computational Engineering, Department of Mathematics, Technische Universität Darmstadt, Dolivostraße 15, D-64293 Darmstadt, Germany.
aschwartz@mathematik.tu-darmstadt.de

fluctuations. In finance, however, attention is mostly given to losses. So already Markowitz himself was aware of this short-coming of variance [30] and proposed a downside risk as an alternative to variance.

Many other alternatives were introduced since then. Nowadays, Value-at-Risk (VaR), which measures the maximum loss that can be expected during a specific time horizon with a probability β (β close to 1), is widely used in the banking and insurance industry as a downside risk measure. Despite its popularity, VaR lacks some important mathematical properties. Artzner at al. [1] presented an axiomatic definition of risk measures and coined a coherent risk measure which has certain reasonable properties. Conditional Value-at-Risk (CVaR), the mean value of losses that exceed the value of VaR, exhibits favorable mathematical properties such as coherence implying convexity. Rockafellar and Uryasev [37, 38] proposed to minimize CVaR for optimizing a portfolio so as to reduce the risk of high losses without prior computation of the corresponding VaR while computing VaR as a by-product. Their CVaR minimization formulation results usually in convex or even linear programs which proved attractive for financial optimization and risk management in practice due to their tractability for larger real life instances. For each of these risk measures, one can formulate corresponding mean-risk portfolio optimization problems.

Regardless of the risk measure used, these models are strongly dependent on the underlying distribution and its parameters, which are typically unknown and have to be estimated, cf. Fabozzi et al. [17]. Investors usually face the so called estimation risk as they rely on a limited amount of data to estimate the input parameters. Portfolios constructed using these estimators perform very poorly in terms of their out-of-sample mean and variance as the resulting portfolio weights fluctuate substantially over time, cf. e.g. Michaud [31] and Chopra and Ziemba [12]. As some reformulations of mean-risk portfolio problems depend on the assumption of normality, poor performance can also be caused by deviations of the empirical distribution of returns from normality. One can thus also consider the distribution ambiguity in the sense that no knowledge of the return distribution for risky assets is assumed while the mean and variance/covariance are assumed to be known. For these reasons, we examine portfolio policies based on robust estimators.

Robust optimization is another approach for optimization under uncertainty. Robust optimization has been rapidly developed since the pioneering work of Ben-Tal and Nemirovski [2]. Robust portfolio selection deals with eliminating the impacts of estimation risk and/or distribution ambiguity. Goldfarb and Iyengar [23] studied the robust portfolio in the mean-variance framework. Instead of the precise information on the mean and the covariance matrix of asset returns, they introduced some types of uncertainty, such as box uncertainty and ellipsoidal uncertainty. They also considered the robust VaR portfolio selection problem by assuming a normal distribution. Chen et al. [11] minimized the worst-case CVaR risk measure over all distributions with fixed first and second moment information. The reader is referred to El Ghaoui et al. [16] and Popescu [36] for other studies on portfolio optimization with distributional robustness. Paç and Pınar [33] extend Chen et al. [11] to the case where a risk-free asset is also included and distributional robustness is complemented with ellipsoidal mean return ambiguity. Other choices of the ambiguity set for VaR and CVaR are considered e.g. by Tütüncü and Koening [42], Pflug and Wozabal [35], Zhu and Fukushima [43], DeMiguel and Nogales [15] and Delage and Ye [13]. For survey of recent approaches to construct robust portfolios, we refer to Kim et al. [27].

Recently, several authors focused on inducing sparsity of the solution which improved the out-of-sample performance of resulting portfolios, cf. e.g. DeMiguel et al. [14]. Reduction of transaction costs can also support the use of sparse portfolios in practice. While some studies impose a penalty on the l_1-norm of the asset weight vector or its alternatives, e.g. Fastrich et al. [18], some consider so called cardinality constraints. The portfolio optimization problem resulting from the latter can be viewed as a mixed-integer problem and it is considered computationally challenging. The examples of solution techniques include exact branch-and-bound methods, e.g., Borchers and Mitchell [6], Bertsimas and Shioda [3]; exact branch-and-cut methods, e.g., Bienstock [5]; heuristic algorithms, e.g., Chang et al. [10]; and relaxation algorithms, e.g., Shaw et al. [41], Murray and Shek [32] and Burdakov et al. [7, 8].

Despite the vast literature on robust portfolio optimization and many works on sparse portfolio optimization, there are only few works that concern both sparse and robust portfolios, cf. e.g. Bertsimas and Takeda [4].

In this paper, we consider the cardinality constrained minimization of VaR and CVaR under normality of asset returns and their robust counterparts under distribution ambiguity. We assume that both first and second order moments are known. Our contribution is to apply the recently proposed NLP-reformulation in Burdakov et al [7, 8], which is further studied in Červinka et al. [9]. We perform a numerical experiment to compare performance of four solution methods: GUROBI to solve a mixed-integer formulation of the problems, SNOPT to solve a relaxed NLP reformulation of the problems, and two regularization methods. A similar numerical study has been reported in Burdakov et al. [7] for a cardinality constrained (and non-robust) mean-variance model, where the objective function was convex quadratic, whereas we investigate the investment problems with VaR and CVaR introduced above which lead to nonconvex problems even after relaxing the conditions on sparsity.

The paper is organized as follows: In Section 2, we introduce the risk measures and investment problems with a condition on portfolio sparsity. Moreover, we discuss the solution approaches. Section 3 provides an extensive numerical study.

A word on notation: By $e \in \mathbb{R}^n$ we denote the vector with all components equal to one. For two vectors $x, y \in \mathbb{R}^n$ the vector $x \circ y \in \mathbb{R}^n$ denotes the componentwise (Hadamard) product of x and y.

2 Minimizing robust VaR and CVaR under distribution ambiguity with cardinality constraints

We consider a market with n risky financial assets. The disposable wealth is to be allocated into a portfolio $x \in \mathbb{R}^n$, such that each component x_i denotes the fraction of disposable wealth to be invested into the i-th asset, $i = 1, \ldots, n$. We do not allow short-sales, i.e. we assume $x \geq 0$. Furthermore, we demand that the whole disposable budget is invested, i.e. $e^\top x = 1$. Thus, for a vector x of allocations to n risky assets and a random vector ξ of return rates for these assets, we consider the following loss function

$$\ell(x, \xi) = -x^\top \xi.$$

Assume that ξ follows a probability distribution π from the ambiguity (uncertainty) set $D = \{\pi \mid E_\pi[\xi] = \mu, Cov_\pi(\xi) = \Gamma \succ 0\}$ of distributions with expected value μ and positive semidefinite covariance matrix Γ.

Markowitz [29] considered variance $\sigma^2(x) = x^\top \Gamma x$ as a risk measure associated with portfolio x. In the 90s, the investment bank J.P. Morgan reinvented the quantile risk measure (quantile premium principle) used by actuaries for investment banking, giving rise to Value-at-Risk (VaR). Associated with a confidence level β,

$$\mathrm{VaR}_\beta(x) = \min\{z \mid P_\pi(\ell(x, \xi) \leq z) \geq \beta\}.$$

Artzner et al. [1] defined coherent risk measure as a measure satisfying monotonicity, translation invariance, subadditivity and positive homogeneity. It is known, that VaR is not a coherent risk measure as it fails subadditivity. On the other hand, the conditional value-at-risk (CVaR) introduced by Rockafellar and Uryasev [37] turns out to be a convex and coherent risk measure. CVaR at level β is defined as the expected value of loss that exceeds $\mathrm{VaR}_\beta(x)$. Alternatively, Rockafellar and Uryasev [37] showed that calculation of CVaR and VaR can be achieved simultaneously by minimizing the auxiliary function with respect to $\alpha \in \mathbb{R}$

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1-\beta} E[(\ell(x, \xi) - \alpha)_+],$$

where $(v)_+ = \max\{0, v\}$. Thus,

$$\mathrm{CVaR}_\beta(x) = \min_\alpha F_\beta(x, \alpha)$$

and $\mathrm{VaR}_\beta(x)$ is the left endpoint of the interval $\mathrm{argmin}_\alpha F_\beta(x, \alpha)$.

Let us assume normality of returns ξ. Denote by ϕ and Φ density and cumulative distribution function of the standard normal distribution, respectively. Following Fabozzi et al. [17], originating in Rockafellar and Uryasev [37], the value-at-risk can be expressed as

$$\mathrm{VaR}_\beta(x) = \zeta_\beta \sqrt{x^\top Q x} - \mu^\top x, \tag{1}$$

where $\zeta_\beta = -\Phi^{-1}(1-\beta)$, and assuming $\beta > 0.5$, the conditional value-at-risk reduces to

$$\mathrm{CVaR}_\beta(x) = \eta_\beta \sqrt{x^\top Q x} - \mu^\top x, \tag{2}$$

where $\eta_\beta = \frac{-\int_{-\infty}^{\Phi^{-1}(1-\beta)} t\phi(t)dt}{1-\beta}$.

Further, we consider the worst case VaR for a fixed x with respect to the ambiguity set D defined as

$$\mathrm{RVaR}_\beta(x) = \sup_{\pi \in D} \mathrm{VaR}_\beta(x).$$

Analogously, we consider the worst case CVaR for a fixed x with respect to set D defined as

$$\mathrm{RCVaR}_\beta(x) = \sup_{\pi \in D} \mathrm{CVaR}_\beta(x) = \sup_{\pi \in D} \min_\alpha F_\beta(x, \alpha).$$

Based on Chen et. al [11, proof of Theorem 2.9], further generalized in Paç and Pınar [33] using Shapiro [40, Theorem 2.4], we have that under distribution ambiguity,

$$\mathrm{RVaR}_\beta(x) = \frac{2\beta - 1}{2\sqrt{\beta(1-\beta)}} \sqrt{x^\top Q x} - \mu^\top x \tag{3}$$

and

$$\mathrm{RCVaR}_\beta(x) = \sqrt{\frac{\beta}{1-\beta}} \sqrt{x^\top Q x} - \mu^\top x. \tag{4}$$

We now formulate cardinality constrained portfolio selection models for each of the risk measures (1)–(4). For the sake of brevity, we replace a particular function from (1)–(4) by a general risk function $r(x)$. Consider the cardinality constrained problem

$$\begin{aligned} \min_x \ & r(x) \\ \text{s.t.} \ & e^\top x = 1, \\ & 0 \le x \le u, \\ & \|x\|_0 \le \kappa, \end{aligned} \tag{5}$$

where $u \in \mathbb{R}^n$ are given upper bounds in the investments in the n risky assets, $\kappa \in \mathbb{N}$ is a natural number and $\|x\|_0$ denotes the cardinality of the support of the vector x, i.e. the number of its nonzero elements. Naturally, we assume that $\kappa < n$.

The problem (5) is difficult to solve due to the cardinality constraint. It can be readily reformulated as a mixed integer problem using binary decision variables

$$\begin{aligned} \min_{x,z} \quad & r(x) \\ \text{s.t.} \quad & e^\top x = 1, \\ & 0 \leq x \leq u \circ z, \\ & z \in \{0,1\}^n, \\ & e^\top z \leq \kappa. \end{aligned} \tag{6}$$

If x_i is positive, then the corresponding z_i must be equal to one and by the reformulated cardinality constraint $e^\top z \leq \kappa$ this can happen at most κ times.

As even for simple instances of cardinality constrained problems Bienstock [5] showed the problem to be NP-complete, solving problems (6) even using specialized global solution techniques can be computationally very time demanding. Thus, we consider the following relaxed NLP reformulation of (6) introduced in Burdakov et al. [8].

$$\begin{aligned} \min_{x,y} \quad & r(x) \\ \text{s.t.} \quad & e^\top x = 1, \\ & 0 \leq x \leq u, \\ & 0 \leq y \leq e, \\ & x \circ y = 0, \\ & e^\top y \geq n - \kappa. \end{aligned} \tag{7}$$

Here, in contrast to the previous reformulation, whenever x_i is positive, the corresponding y_i has to be equal to zero. Due to the reformulated cardinality constraint $e^\top y \geq n - \kappa$ this can again occur at most κ times. Note that this problem can be considered a mathematical program with complementarity constraints (MPCC) due to the complementarity constraints $x \geq 0, y \geq 0, x \circ y = 0$.

In our numerical experiments, we consider the two techniques to regularize the complementarity constraints developed in Scholtes [39] and Kanzow and Schwartz [25] for MPCCs. The latter has already been successfully modified and applied to cardinality-constrained mean-variance portfolio problems in Burdakov et al. [7]. For the regularization based on Scholtes [39], we replace the constraints $x \geq 0, y \geq 0, x \circ y = 0$ in (7) by the inequalities

$$x \geq 0, y \geq 0, x \circ y \leq te \tag{8}$$

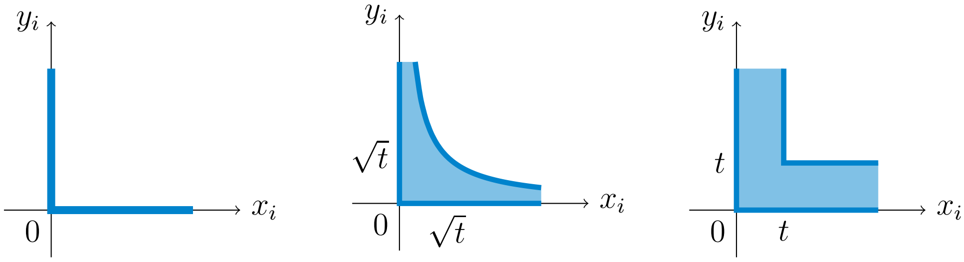
for some small regularization parameter $t > 0$. Analogously, for the regularization by Kanzow and Schwartz [25], we replace the same constraints in (7) by the inequalities

$$\Phi(x, y; t) \leq 0, x \geq 0, y \geq 0, \tag{9}$$

where $\Phi_i(x, y; t) = \varphi(x_i, y_i; t)$ with

$$\varphi(a, b; t) = \begin{cases} (a-t)(b-t) & \text{if } a+b \geq 2t, \\ -\frac{1}{2}\left[(a-t)^2 + (b-t)^2\right] & \text{if } a+b < 2t. \end{cases}$$

It is not difficult to see that for $t \geq 0$ the inequality $\varphi(a, b; t) \leq 0$ is equivalent to $\min\{a, b\} \leq t$. Figure 1 provides an illustration of the respective feasible sets of the complementarity constraint and the two regularizations. In both cases, the idea of the regularization method is to solve a sequence of parameterized nonlinear programs NLP(t_k) for $t_k \searrow 0$. Comparison of theoretical properties and numerical experience with these regularizations for general MPCCs, along with other types of regularizations that we do not consider here, can be found in Kanzow and Schwartz [26]. We refer to Burdakov et al. [7] for details on the properties of NLP(t_k) and the convergence properties for the approach from Kanzow and Schwartz [25] adapted to cardinality constrained programs.



(a) Complementarity constraints

(b) Scholtes regularization

(c) Kanzow-Schwartz regularization

Figure 1: Illustration of the complementarity constraints and the two regularizations

3 Numerical study

In this section, we apply the introduced solution approaches to the investment problems with the VaR and CVaR measures under normality assumption and to the robust VaR and CVaR. We will consider each problem for several levels of β, in particular we select $\beta \in \{0.9, 0.95, 0.99\}$. Table 1 contains the values of the corresponding quantiles and generalized quantiles, which appear in the exact reformulations of the risk measures.

Table 1: Quantiles and generalized quantiles

	β	0.9	0.95	0.99
VaR	ζ_β	1.2816	1.6449	2.3263
CVaR	η_β	1.7550	2.0627	2.6652
RVaR	$\frac{2\beta-1}{2\sqrt{\beta(1-\beta)}}$	1.3333	2.0647	4.9247
RCVaR	$\sqrt{\frac{\beta}{1-\beta}}$	3.0000	4.3589	9.9499

We use 90 simulated instances with mean vectors and variance matrices which were already employed by [19] and are freely available at website [20]. The generation of the data was described by [34]. Various problems with $n = 200, 300$ and 400 assets are included in the dataset.

We will compare the performance of the following solution approaches:

1. `GUROBI_60`: Solve the mixed integer formulation (6) using the commercial mixed-integer solver `GUROBI`, version 6.5, with time limit 60s and start vector $x^0 = 0$, $z^0 = e$.
2. `GUROBI_300_40`: Same as above but with time limit 300s and node limit 40.
3. `Relaxation_01`: Solve the relaxed nonlinear problem (7) using the sparse SQP method `SNOPT`, version 7.5, with start vector $x^0 = 0$, $y^0 = e$.
4. `Relaxation_00`: Same as above but with start vector $x^0 = 0$, $y^0 = 0$.
5. `Scholtes_01`: Solve a sequence of Scholtes regularizations (8) using `SNOPT` with starting point $x^0 = 0$, $y^0 = e$.
6. `Scholtes_00`: Same as above but with start vector $x^0 = 0$, $y^0 = 0$.
7. `KanzowSchwartz_01`: Solve a sequence of Kanzow–Schwartz regularizations (9) using `SNOPT` with starting point $x^0 = 0$, $y^0 = e$.
8. `KanzowSchwartz_00`: Same as above but with start vector $x^0 = 0$, $y^0 = 0$.

All computations were done in `MATLAB R2014a`. A few details on the implementation of the respective solution approaches:

More information on the solver `GUROBI` and its various options can be found at [24]. To be able to solve the mixed-integer problem (6) with `GUROBI`, we had to reformulate it in the following form :

$$\begin{aligned}
\min_{x,z,w,v} \quad & c_\beta v - \mu^\top x \\
\text{s.t.} \quad & e^\top x = 1, \\
& 0 \leq x \leq u \circ z, \\
& z \in \{0,1\}^n, \\
& e^\top z \leq \kappa, \\
& v \geq 0, \\
& w = Q^{\frac{1}{2}} x, \\
& v^2 \geq w^\top w,
\end{aligned}$$

where c_β is the respective constant from Table 1 for the different risk measures. Since we used $x^0 = 0$ as start vector, we also used $w^0 = 0$ and $v^0 = 0$.

Note that `GUROBI` is a global solver, i.e. it tries to verify that a candidate solution is indeed a global minimum. Since the other solution approaches do not provide any guarantee of finding a global solution, we set the option `mipfocus` to one in order to encourage `GUROBI` to try to find good solutions fast. Additionally we set the option `timelimit` to 60s at first. However, we found that `GUROBI` sometimes spent the whole time by looking for a feasible solution without moving to the branch-and-bound tree. Thus we increased `timelimit` to 300s and added the condition on the maximal number of computed nodes `nodelimit` = 40 to obtain results less dependent on slight variations in computation time.

The relaxed problem (7) and the regularized problems (8) and (9) are all solved using the sparse SQP method `SNOPT`, see [22, 21] for more information. We started both regularization methods with an initial parameter $t_0 = 1$ and decreased t_k in each iteration by a factor of 0.01. Both regularization methods where terminated if either $\|x^k \circ y^k\|_\infty \leq 10^{-6}$ or $t_k < 10^{-8}$.

The constraints $e^\top x = 0$ and $0 \leq x \leq u$ were usually satisfied in the solutions x^* found by all methods (except for `GUROBI` which occasionally did not return a feasible solution at all, see below). In order to check whether the cardinality constraint $\|x\|_0 \leq \kappa$ are also satisfied, we count the number of all components $x_i^* > 10^{-6}$.

Table 2 contains results for a particular problem with 400 assets (pard400-e-400). We can see that `GUROBI` running 60s was not able to provide a feasible solution for problem with $\mathrm{RVaR}_{0.99}$. The Scholtes regularization starting from point $x^0 = 0$, $y^0 = e$ was not successful for $\mathrm{RCVaR}_{0.95}$. However, in all other cases the Scholtes regularization starting from $x^0 = 0$, $y^0 = e$ provided the best solution with a runtime around 1s. We also report the relative gap:

$$\textit{relative gap} = (f - f_{best})/f_{best},$$

where f is the objective value obtained by an algorithm and f_{best} denotes the lowest objective value for a problem.

Summary results for all problems are reported in Tables 3, 4, 5. For each problem with a particular risk measure, level β, number of assets and algorithm we report the following descriptive statistics over 30 instances of problems: average relative gap with respect to the minimal objective value, average computation time (in seconds), number of cases when the algorithm found the best solution, number of cases when the result was infeasible with respect to the sparsity or orthogonality. All computations were done on two computers with comparable performance indicators. Nonetheless, the given computation times should only be used for a qualitative comparison of the methods.

It can be observed that the best results were obtained by approach 5: the Scholtes regularization starting from $x^0 = 0$, $y^0 = e$. When the results of this regularization were feasible, they correspond to the best obtained solutions. However, in many cases the portfolios obtained by the regularizations were infeasible. Also the relaxed problems (approach 3) behaved badly showing an average relative gap greater than 100%.

To further investigate the behavior, we changed the starting point of relaxations and regularizations to $x^0 = 0$, $y^0 = 0$. In this case, the obtained optimal values were slightly worse for the regularizations, but we have reduced the problems with infeasible solutions. Moreover, for the starting point $x^0 = 0$, $y^0 = 0$, the behavior of the relaxation approach improved significantly such that it is fully comparable with the regularizations.

Figures 2, 3, 4 present performance plots for each problem size and algorithm. We identified the minimal objective value for each problem found by any of the eight considered algorithms algorithms and then compared it with the remaining objective values using the ratio: actual objective value/minimal objective value. The graphs report the relative number of problems (y-value), where the ratio is lower or equal to the x-value. We would prefer algorithms with the curve close to the upper-left corner, i.e. which produce good and feasible solutions. Since infeasible problems are considered with an infinite objective function value, not all algorithm curves touch the upper bound 1. This is the case for the regularized problems with $x^0 = 0$, $y^0 = e$ for all problem sizes. For the largest problems with 400 assets, even `GUROBI` with 60s limit and Kanzow–Schwartz regularization starting from $x^0 = 0$, $y^0 = 0$ were not able to reach the upper bound 1.

4 Conclusions

We proposed and compared several solution approaches for cardinality constrained portfolio optimization problems. We minimized VaR and CVaR risk

Table 2: Results for a problem with 400 assets (pard400-e-400)

	VaR			CVaR			RVaR			RCVaR		
β	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Alg.	Objective value											
1	62.29	42.79	58.87	49.58	55.28	64.67	64.81	52.08	–	80.31	117.98	272.07
2	32.70	40.55	63.97	49.58	55.28	64.67	36.62	52.73	133.68	79.21	117.98	272.07
3	53.20	68.28	96.57	72.85	85.62	110.64	55.35	85.71	204.43	124.53	180.94	413.04
4	29.76	38.20	54.03	40.76	47.91	61.90	30.97	47.96	114.39	69.68	101.24	231.11
5	25.94	33.30	47.12	35.54	41.76	53.99	27.00	41.80	99.86	60.79	–	201.92
6	29.76	38.20	54.03	40.76	47.91	61.90	30.97	47.95	114.34	69.68	101.21	231.02
7	27.30	35.05	49.58	37.40	43.94	56.80	28.41	44.00	104.92	63.94	92.86	201.45
8	29.76	38.20	54.03	40.76	47.91	61.90	30.97	47.95	114.39	69.68	101.24	231.11
Alg.	Relative gap											
1	1.40	0.29	0.25	0.39	0.32	0.20	1.40	0.25	–	0.32	0.27	0.35
2	0.26	0.22	0.36	0.39	0.32	0.20	0.36	0.26	0.34	0.30	0.27	0.35
3	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	0.95	1.05
4	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.09	0.15
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	0.00
6	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.14	0.15	0.09	0.15
7	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.00	0.00
8	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.09	0.15
Alg.	Computation time (s)											
1	60	61	62	60	61	69	67	73	–	60	68	68
2	300	300	258	300	300	300	300	300	189	288	302	300
3	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03	0.03
4	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.05	0.05	0.05
5	1.03	0.78	1.09	1.07	1.06	1.16	0.97	1.17	1.14	0.71	–	1.04
6	1.53	1.47	1.52	1.64	1.51	1.39	1.50	1.52	1.37	1.41	1.25	1.20
7	0.77	0.78	0.76	0.76	0.72	0.81	0.73	0.83	0.69	0.77	0.71	0.69
8	0.87	0.91	0.80	0.81	0.92	0.80	0.81	0.88	0.83	0.86	0.83	0.83

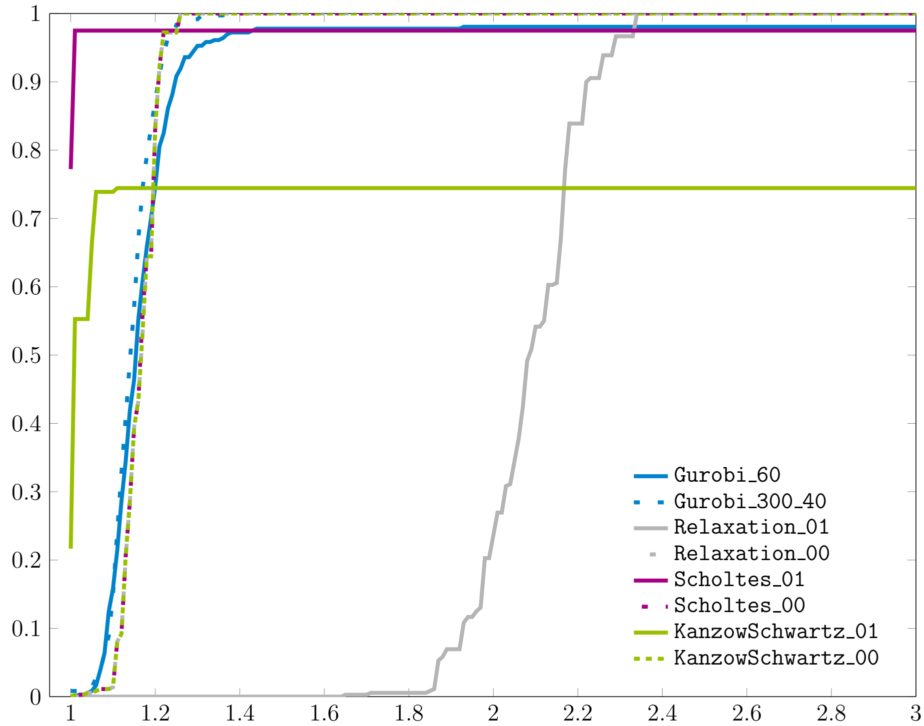


Figure 2: Performance plot of the objective function values for $n = 200$ assets

Table 3: Results for 30 instances with 200 assets

	VaR			CVaR			RVaR			RCVaR		
β	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Alg.	Average relative gap											
1	0.126	0.157	0.143	0.158	0.171	0.180	0.164	0.170	0.154	0.167	0.179	0.188
2	0.158	0.149	0.146	0.133	0.130	0.155	0.123	0.139	0.145	0.143	0.147	0.164
3	1.091	1.087	1.093	1.092	1.088	1.089	1.073	1.085	1.096	1.094	1.092	1.096
4	0.166	0.163	0.167	0.166	0.164	0.165	0.156	0.162	0.169	0.167	0.167	0.169
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.166	0.163	0.167	0.166	0.164	0.165	0.156	0.162	0.169	0.167	0.167	0.169
7	0.004	0.010	0.016	0.013	0.020	0.014	0.008	0.020	0.017	0.014	0.015	0.000
8	0.166	0.163	0.167	0.166	0.164	0.165	0.156	0.162	0.169	0.167	0.167	0.169
Alg.	Average computation time (s)											
1	60	60	60	60	60	60	67	67	67	60	60	60
2	59	72	51	63	63	80	56	59	71	56	66	71
3	0.02	0.02	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02
4	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
5	0.25	0.26	0.27	0.26	0.28	0.25	0.27	0.28	0.26	0.27	0.27	0.26
6	0.31	0.31	0.34	0.31	0.31	0.30	0.31	0.32	0.33	0.31	0.32	0.31
7	0.23	0.20	0.18	0.18	0.19	0.19	0.19	0.20	0.18	0.19	0.17	0.19
8	0.21	0.21	0.22	0.22	0.21	0.20	0.21	0.21	0.20	0.21	0.20	0.20
Alg.	Best solution found (out of 30)											
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	0	1	1	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	29	29	28	28	28	28	26	28	27	28	27	30
6	0	0	0	0	0	0	1	0	0	0	0	0
7	17	20	14	19	16	14	16	15	11	14	11	9
8	0	0	0	0	0	0	0	0	0	0	0	0
Alg.	Solution was infeasible (out of 30)											
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	0	3	1	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0
7	6	4	5	3	5	5	5	5	13	6	14	21
8	0	0	0	0	0	0	0	0	0	0	0	0

Table 4: Results for 30 instances with 300 assets

	VaR			CVaR			RVaR			RCVaR		
β	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Alg.	Average relative gap											
1	0.256	0.245	0.224	0.219	0.232	0.202	0.204	0.196	0.217	0.220	0.248	0.232
2	0.209	0.224	0.216	0.206	0.197	0.187	0.229	0.200	0.215	0.207	0.234	0.230
3	1.093	1.086	1.082	1.083	1.082	1.085	1.093	1.082	1.094	1.090	1.094	1.094
4	0.170	0.166	0.163	0.164	0.164	0.165	0.170	0.164	0.171	0.168	0.171	0.171
5	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
6	0.170	0.166	0.163	0.164	0.164	0.165	0.170	0.164	0.171	0.168	0.169	0.171
7	0.016	0.017	0.011	0.021	0.014	0.014	0.018	0.014	0.018	0.020	0.022	0.008
8	0.170	0.166	0.163	0.164	0.164	0.165	0.170	0.164	0.171	0.168	0.171	0.168
Alg.	Average computation time (s)											
1	60	60	61	60	60	60	67	67	64	60	60	60
2	141	135	132	148	158	127	135	149	139	144	125	156
3	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02
4	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
5	0.49	0.48	0.56	0.53	0.58	0.53	0.49	0.56	0.56	0.59	0.57	0.55
6	0.69	0.68	0.69	0.68	0.67	0.69	0.70	0.70	0.73	0.72	0.69	0.67
7	0.47	0.43	0.40	0.40	0.41	0.38	0.42	0.42	0.39	0.43	0.39	0.41
8	0.47	0.46	0.44	0.44	0.45	0.45	0.46	0.47	0.43	0.44	0.44	0.44
Alg.	Best solution found (out of 30)											
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	29	28	21	28	26	23	29	25	25	25	26	22
6	0	0	0	1	0	0	0	0	0	0	0	0
7	13	10	18	8	9	13	14	8	11	11	10	11
8	0	0	0	0	0	0	0	0	0	0	0	0
Alg.	Solution was infeasible (out of 30)											
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	0	2	2	2	2	2	0	3	0	1	0	0
6	0	0	0	0	0	0	0	0	0	0	1	0
7	1	4	6	4	3	11	1	3	13	12	12	17
8	0	0	0	0	0	0	0	0	0	0	0	1

Table 5: Results for 30 instances with 400 assets

	VaR			CVaR			RVaR			RCVaR		
β	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99	0.9	0.95	0.99
Alg.	Average relative gap											
1	0.266	0.259	0.235	0.226	0.277	0.241	0.311	0.228	0.209	0.253	0.288	0.269
2	0.205	0.238	0.244	0.198	0.215	0.214	0.206	0.199	0.226	0.212	0.258	0.250
3	1.180	1.191	1.195	1.188	1.172	1.201	1.182	1.178	1.201	1.201	1.197	1.200
4	0.171	0.177	0.179	0.175	0.168	0.183	0.173	0.171	0.183	0.183	0.181	0.182
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
6	0.171	0.177	0.179	0.175	0.167	0.183	0.172	0.170	0.183	0.183	0.181	0.183
7	0.016	0.023	0.017	0.016	0.011	0.017	0.015	0.011	0.006	0.013	0.003	0.000
8	0.171	0.177	0.179	0.175	0.167	0.183	0.173	0.171	0.180	0.183	0.181	0.183
Alg.	Average computation time (s)											
1	62	61	67	62	62	64	67	67	65	63	64	63
2	231	243	220	201	206	205	217	221	224	222	183	210
3	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.03
4	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.04	0.04	0.05	0.04
5	0.94	1.03	1.09	1.04	1.02	1.04	0.93	1.04	1.07	1.07	1.06	1.06
6	1.36	1.36	1.38	1.38	1.36	1.36	1.37	1.50	1.32	1.37	1.32	1.30
7	0.85	0.77	0.76	0.75	0.76	0.77	1.00	0.76	0.73	0.73	0.74	0.72
8	0.86	0.86	0.84	0.84	0.83	0.84	0.85	0.85	0.82	0.85	0.83	0.79
Alg.	Best solution found (out of 30)											
1	0	0	0	0	0	0	0	1	0	0	0	0
2	0	0	1	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	24	25	19	22	13	20	24	14	16	19	16	21
6	0	0	0	0	0	0	0	0	0	0	0	0
7	14	8	12	9	16	12	12	15	16	14	16	11
8	0	0	0	0	1	0	0	0	0	0	0	0
Alg.	Solution was infeasible (out of 30)											
1	0	0	0	0	0	0	0	0	20	0	0	1
2	0	0	0	0	0	0	0	0	1	0	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	5	1	2	5	0	0	2	0
6	0	0	0	0	0	0	1	0	0	0	0	0
7	2	1	12	2	5	12	4	5	12	11	13	19
8	1	1	2	1	1	0	1	1	2	0	2	0

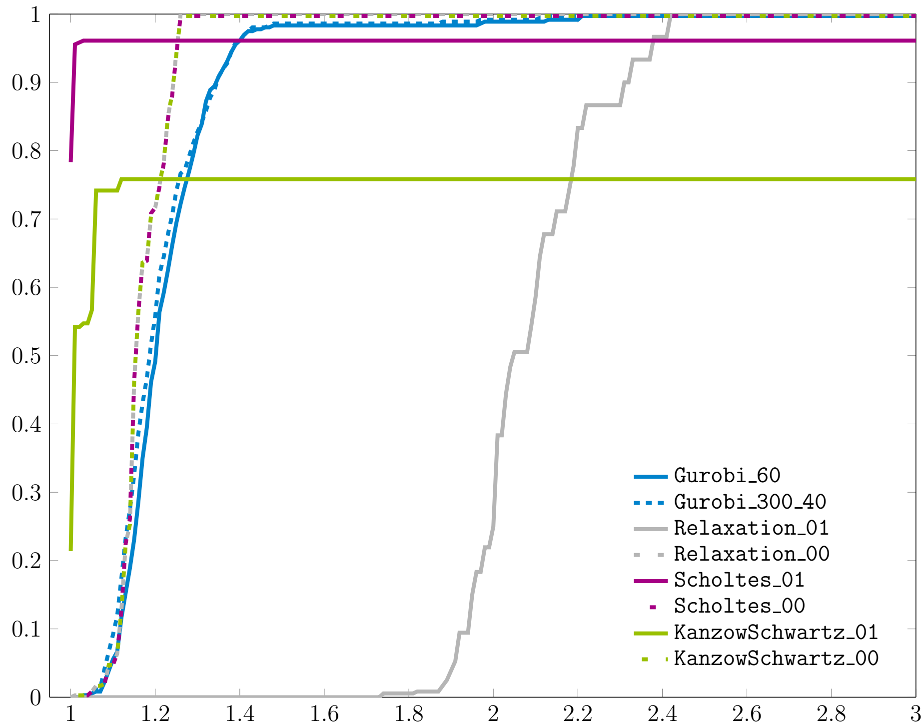


Figure 3: Performance plot of the objective function values for $n = 300$ assets

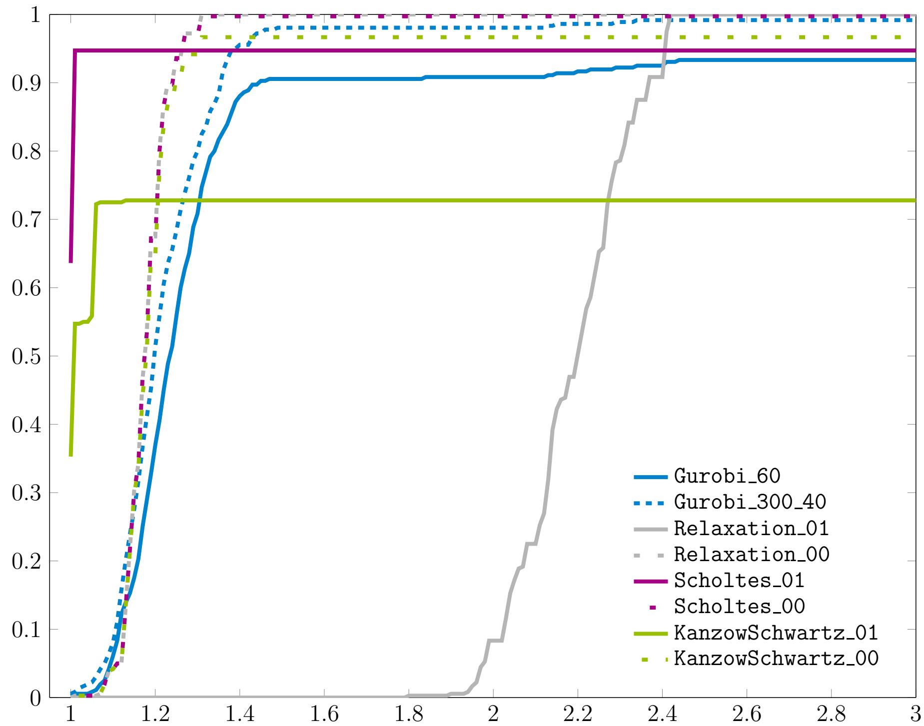


Figure 4: Performance plot of the objective function values for $n = 400$ assets

measures under the assumption of normality and distributional ambiguity. Future research will be devoted to developing a global solution strategy based on several starting points and combinations of the proposed methods.

Acknowledgments

We would like to acknowledge the support of the Czech Science Foundation (GA ČR) under the grants GA13-01930S and GA15-00735S. The work of Alexandra Schwartz is supported by the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt.

References

[1] Artzner, P.; Delbaen, F.; Eber, J.-M.; Heath, D.: Coherent measures of risk. *Mathematical Finance* **9**(3), 203–228, 1999.

[2] Ben-Tal, A,; Nemirovski, A.: Robust convex optimization. *Math. Oper. Res.* **23**(4), 769–805, 1998.

[3] Bertsimas, D.; Shioda, R.: Algorithm for cardinality-constrained quadratic optimization. *Comput. Optim. Appl.* **43**, 1–22, 2009.

[4] Bertsimas, D.; Takeda, A.: Optimizing over coherent risk measures and non-convexities: a robust mixed integer optimization approach. *Comput. Optim. Appl.* **62**, 613–639, 2015.

[5] Bienstock, D.: Computational study of a family of mixed-integer quadratic programming problems. *Math. Program.* **74**, 121–140, 1995.

[6] Borchers, B.; Mitchell, J.E.: An improved branch and bound algorithm for mixed integer nonlinear programs. *Comput. Oper. Res.* **21**, 359–367, 1994.

[7] Burdakov, O.P.; Kanzow, Ch.; Schwartz, A.: On a reformulation of mathematical programs with cardinality constraints. In *Advances in Global Optimization*, D.Y. Gao, N. Ruan, amd W.X. Xing (eds), Proceedings of the 3rd World Congress of Global Optimization (Huangshan, China, 2013), Springer, New York, 121–140, 2015.

[8] Burdakov, O.P.; Kanzow, Ch.; Schwartz, A.: Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method, *SIAM J. Optim.* **26**(1), 397–425, 2016.

[9] Červinka, M.; Kanzow, Ch.; Schwartz, A.: Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Math. Program., Ser. A* **160**(1), 353–377, 2016.

[10] Chang, T.-J.; Meade, N.; Beasley, J.E.; Sharaiha, Y.M.: Heuristics for cardinality constrained portfolio optimization. *Comput. Oper. Res.* **27**, 1271–1302, 2000.

[11] Chen, L.; He, S.; Zhang, S.: Tight Bounds for some risk measures, with applications to robust portfolio selection, *Oper. Res.* **59**(4), 847–865, 2011.

[12] Chopra, V.K.; Ziemba, W.T.: The effect of errors in means, variances and covariances on optimal portfolio choice. *Journal of Portfolio Management* **19**(2), 6–11, 1993.

[13] Delage, E.; Ye, Y.: Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.* **58**(3), 595–612, 2010.

[14] DeMiguel, V.; Garlappi, L.; Nogales, J.; Uppal, R.: A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Manag. Sci.* **55**(5), 798–812, 2009.

[15] DeMiguel, V.; Nogales, F.J.: Portfolio selection with robust estimation. *Oper. Res.* **57** (3), 560-577, 2009.

[16] El Ghaoui, L.; Oks, M; Oustry, F.: Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. *Oper. Res.* **51**(4), 543–556, 2003.

[17] Fabozzi, F.J.; Huang, D.; Zhou, G.: Robust portfolios: contributions from operations research and finance. *Ann. Oper. Res.* **176**(1), 191–220, 2010.

[18] Fastrich, B.; Paterlini, S.; Winkler, P.: Constructing optimal sparse portfolios using regularization methods. *Comput. Manag. Sci.* **12**(3), 417–434, 2015.

[19] Frangioni, A.; Gentile, C.: SDP diagonalizations and perspective cuts for a class of nonseparable MIQP, *Oper. Res. Lett.* **35**, 181–185, 2007.

[20] Frangioni, A.; Gentile, C.: Mean-variance problem with minimum buy-in constraints, data and documentation, online at http://www.di.unipi.it/optimize/Data/MV.html.

[21] Gill, P.E.; W. Murray, W.; Saunders, M.A.; Wong, E.: SNOPT 7.5 User's Manual, *CCoM Technical Report* **15-3**, Center for Computational Mathematics, University of California, San Diego.

[22] Gill, P.E.; W. Murray, W.; Saunders, M.A.: SNOPT: An SQP algorithm for large-scale constrained optimization, *SIAM Review* **47**, 99-131, 2005.

[23] Goldfarb, D.; Iyengar, G.: Robust portfolio selection problems. *Math. Oper. Res.* **28**(1), 1–38, 2003.

[24] Gurobi Optimization, Inc.: Gurobi Optimizer Reference Manual, online at http://www.gurobi.com, 2016.

[25] Kanzow, Ch.; Schwartz, A.: A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. *SIAM J. Optim.* **23**, 770–798, 2013.

[26] Kanzow, Ch.; Schwartz, A.: The Price of Inexactness: Convergence Properties of Relaxation Methods for Mathematical Programs with Complementarity Constraints Revisited, *Math. Oper. Res.* **40**(2), 253–275, 2015.

[27] Kim, J.; Kim, W.; Fabozzi, F.: Recent developments in robust portfolios with a worst-case approach. *Journal of Optimization Theory and Applications* **161** (1), 103–121, 2014.

[28] Levy, H. (Ed.): *Stochastic Dominance: Investment Decision Making under Uncertainty.* Springer, New York, 2006.

[29] Markowitz, H.M.: Portfolio selection. *The Journal of Finance* **7**, 77-91, 1952.

[30] Markowitz, H.M.: *Portfolio selection: Efficient diversification of investments.* John Wiley & Sons, New York, 1959.

[31] Michaud, R.O.: The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts J.* **45**(1), 31–42, 1989.

[32] Murray, W.; Shek, H.: A local relaxation method for the cardinality constrained portfolio optimization problem. *Comput. Optim. Appl.* **53**, 681–709, 2012.

[33] Paç, A.B.; Pınar, M.Ç.: Robust portfolio choice with CVaR and VaR under distribution and mean return ambiguity. *TOP* **22**, 875–891, 2014.

[34] Pardalos, P.M.; Rodgers, G.P.: Computing aspects of a branch and bound algorithm for quadratic zero-one programming. *Computing* **144**, 45–131, 1990.

[35] Pflug, G.; Wozabal, D.: Ambiguity in portfolio selection. *Quantitative Finance*, 435–442, 2007.

[36] Popescu, I.: Robust mean-covariance solutions for stochastic optimization. *Oper. Res.* **55**(1), 98–112, 2005.

[37] Rockafellar, R.T.; Uryasev, S.: Optimization of conditional value-at-risk. *Journal of Risk* **2**, 21–41, 2000.

[38] Rockafellar, R.T.; Uryasev, S.: Conditional Value-at-Risk for General Loss Distributions. *J. Bank. Financ.* **26**, 1443-1471, 2002.

[39] Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.* **11**, 918–936, 2001.

[40] Shapiro, A.: *Topics in stochastic programming.* CORE lecture series. Université catholique de Louvain, Louvain-la-nueve, 2011.

[41] Shaw, D.X.; Liu, S.; Kopman, L.: Lagrangian relaxation procedure for cardinality-constrained porfolio optimization. *Optim. Methods Softw.* **23**, 411–420, 2008.

[42] Tütüncü, R.H.; Koening, M.: Robust asset allocation. *Ann. Oper. Res.* **132**, 157–187, 2004.

[43] Zhu, S.; Fukushima, M.: Worst-case Conditional Value-at-Risk with application to robust portfolio management. *Oper. Res.* **57**(5), 1155-1168, 2009.