J Glob Optim (2013) 56:1409–1423
DOI 10.1007/s10898-012-9842-2

Convex relaxations and MIQCQP reformulations for a class of cardinality-constrained portfolio selection problems

X. T. Cui · X. J. Zheng · S. S. Zhu · X. L. Sun

Received: 27 December 2011 / Accepted: 3 January 2012 / Published online: 14 January 2012

Abstract In this paper we investigate a class of cardinality-constrained portfolio selection problems. We construct convex relaxations for this class of optimization problems via a new Lagrangian decomposition scheme. We show that the dual problem can be reduced to a second-order cone program problem which is tighter than the continuous relaxation of the standard mixed integer quadratically constrained quadratic program (MIQCQP) reformulation. We then propose a new MIQCQP reformulation which is more efficient than the standard MIQCQP reformulation in terms of the tightness of the continuous relaxations. Computational results are reported to demonstrate the tightness of the SOCP relaxation and the effectiveness of the new MIQCQP reformulation.

Keywords Portfolio selection · Cardinality constraint · Mixed 0–1 QCQP reformulation · Second-order cone program · Semidefinite program

This work was supported by National Natural Science Foundation of China under Grants 10971034, 11101092 and 71071036, and by the Joint NSFC/RGC grants under Grant 71061160506.

X. T. Cui · X. L. Sun
Department of Management Science, School of Management, Fudan University, Shanghai 200433, People's Republic of China
e-mail: xtcui@fudan.edu.cn

X. L. Sun
e-mail: xls@fudan.edu.cn

X. J. Zheng (✉)
School of Economics and Management, Tongji University, Shanghai 200092, People's Republic of China
e-mail: xjzheng@fudan.edu.cn

S. S. Zhu
Department of Finance and Investment, Sun Yat-Sen Business School, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China
e-mail: sszhu@fudan.edu.cn

1 Introduction

Since the pioneering work of Markowitz [24,25], portfolio optimization has become an important modeling and analysis tool in modern investment. In Markowitz's mean-variance model, the mean and variance are used to measure the expected return and risk of the portfolio, respectively. Consider a market with n assets. Let R_i be the random return of the ith asset and $x = (x_1, x_2, \dots, x_n)^T$ is the weight vector of the portfolio. The random return of the portfolio is given by $\sum_{i=1}^n R_i x_i$. So the mean and variance of the portfolio return are $r(x) = \mu^T x$ and $\sigma(x) = x^T \Sigma x$, where $\mu = (\mu_1, \dots, \mu_n)^T$ with $\mu_i = E(R_i)$ and $\Sigma = (\sigma_{ij})$ with $\sigma_{ij} = E[(R_i - \mu_i)(R_j - \mu_j)]$. The classical mean-variance model can be then expressed as:

$$\text{(MV)} \quad \begin{aligned} &\min \ x^T \Sigma x \\ &\text{s.t. } \mu^T x \geq \rho, \\ &\quad\ x \in X, \end{aligned}$$

where ρ is the prescribed return level and X is the set formed by other deterministic constraints on x such as budget constraint, sector constraint, no shorting constraint, lower and upper bounds and index tracking error control constraint.

Motivated by Markowitz's work, different extensions of the mean-variance models have been proposed over the last six decades. Konno and Yamazaki [22] proposed a mean absolute deviation (MAD) model which can be solved by linear program. Young [35] established a minimax portfolio selection model in which the return of the worst-case scenario is maximized. Value-at-Risk (VaR) [12,20] was proposed as a risk measure to evaluate the maximum loss under certain confidence level. The portfolio selection models using VaR as a risk measure, however, are difficult to solve due to the nonconvexity of the VaR as a function of the portfolio weight vector. To overcome this difficulty, Rockafellar and Uryasev [30,31] proposed to use conditional VaR (CVaR), the expected value of the extreme loss beyond VaR, as a risk measure. The resulting CVaR based portfolio model can be solved by linear programming and thus is computationally tractable. Angelelli et al. [2] investigated the CVaR and MAD models with real features, and introduced some heuristic methods to provide efficient and effective solutions for portfolio selection models with integer variables. Alexander and Baptista [1] investigated the mean-variance model with VaR and CVaR constraints and compared the portfolio implications of VaR and CVaR constraints in mean-variance framework. In [6,7], stock price fluctuations and cross-correlations were studied by network approaches. Li and Ng [23] considered optimal solution of multi-period mean-variance portfolio selection using a dynamic programming approach (see also [26,36]). An excellent survey of optimization models for portfolio selection can be found in [27] (see also [28]).

We focus in this paper on portfolio selection models with real-life trading constraints. In particular, we are interested in portfolio selection models with the following cardinality constraint and minimum buy-in threshold:

$$|\text{supp}(x)| \leq K, \ x_i \geq a_i, \ \forall\, i \in \text{supp}(x), \tag{1}$$

where $\text{supp}(x) = \{i \mid x_i \neq 0\}$, $0 < K < n$ and $0 < a_i < 1$. The cardinality constraint $|\text{supp}(x)| \leq K$ is introduced to control the total number of different assets in the optimal portfolio due to the transaction cost and managerial concerns. Bienstock [5] investigated the cardinality-constrained mean-variance model and presented a specified branch-and-cut method. Bertsimas and Shioda [4] proposed a branch-and-bound procedure where the continuous relaxation of subproblems were solved by Lemke's pivoting method. Bonami and Lejeune

[8] investigated portfolio selection problem with probabilistic and integer constraints including cardinality and buy-in threshold constraints. An exact branch-and-bound approach was proposed in [8]. Shaw et al. [33] proposed a branch-and-bound method based on Lagrangian relaxation for cardinality constrained mean-variance model. Subgradient method was used in [33] to solve the Lagrangian dual problems of each subproblem in the branch-and-bound process. Frangioni and Gentile [15–17] investigated perspective cut and relaxation for a class of convex 0–1 mixed integer programs with semi-continuous variables which include cardinality constrained quadratic program as a special case.

We consider in this paper a class of cardinality constrained mean-variance portfolio selection problems where factor models are used to estimate the return of the portfolio. Factor models have been popular and effective tools to grasp the characteristics of asset returns. Sharpe [32] first proposed a single factor model where the single factor is usually the market index. Notable factor models include the three-factor model developed by Fama and French [13,14], the macroeconomic factors proposed by Burmeister [9] and the statistical factor models (see [11]). In addition to the constraints (1), we also include a convex quadratic constraint to represent the risk control or tracking error constraint in the mean-variance model. The resulting optimization model is a quadratically constrained quadratic program (QCQP) with cardinality and minimum threshold constraints. We propose in this paper a new Lagrangian decomposition scheme for this class of cardinality constrained QCQP which splits the problem into a convex quadratic subproblem and a semidefinite program (SDP) subproblem. We show that the dual problem can be reduced to a second-order cone program (SOCP) which is tighter than the continuous relaxation of the standard MIQCQP reformulation. This SOCP relaxation leads to a new MIQCQP reformulation which is more efficient than the standard MIQCQP reformulation in the sense that its continuous relaxation is tighter than that of the standard MIQCQP reformulation. It is well-known that SDP and SOCP problems are polynomial solvable by interior-point methods (see [29]). To test the new MIQCQP reformulation, we use `CPLEX 12.1` to solve different MIQCQP reformulations of cardinality constrained portfolio selection problems using historical data of 10 sector indexes of Standard & Poor's 500 index. Computational results indicate that the new MIQCQP reformulation outperforms the standard MIQCQP reformulation in terms of the computing time, final gap and the number of nodes explored by `CPLEX 12.1`.

The paper is organized as follows. In Sect. 2, we describe the cardinality constrained portfolio selection problem using factor models. We discuss in Sect. 3 how to use Lagrangian decomposition technique to construct tight convex relaxations. In particular, we show that the dual problem can be reduced to an SOCP problem. A new MIQCQP reformulation is then proposed. In Sect. 4, we report our comparison numerical results on the performance of different MIQCQP reformulations. Some concluding remarks are given in Sect. 5.

Notations: Throughout the paper, we denote by $v(\cdot)$ the optimal value of problem $(\cdot)$. We denote by $\Re^n_+$ the nonnegative orthant of $\Re^n$, $\mathcal{S}^n$ the set of all $n \times n$ symmetric matrices. The standard inner product in $\mathcal{S}^n$ is defined as $A \bullet B = \mathrm{trace}(AB) = \sum_{i,j=1}^n a_{ij} b_{ij}$. Notation $A \succeq B$ implies that matrix $A - B$ is positive semidefinite.

2 Cardinality constrained portfolio selection models

In this section, we first describe a general portfolio selection model with cardinality and minimum buy-in threshold constraints. Several special cases of cardinality constrained portfolio selection models using different risk measures and tracking error constraint are then discussed.

The general portfolio selection model considered in this paper can be expressed as follows:

$$\begin{aligned}
\text{(P)} \quad & \min\ f(x) := x^T(Q_1 + D_1)x + c_1^T x, \\
& \text{s.t.}\ \ g(x) := x^T(Q_2 + D_2)x + c_2^T x \le \sigma_0, \\
& \qquad Ax \le b, \\
& \qquad |\mathrm{supp}(x)| \le K, \\
& \qquad 0 \le x \le u,\ x_i \ge a_i, i \in \mathrm{supp}(x),
\end{aligned}$$

where $Q_i \in \mathcal{S}^n$ $(i = 1, 2)$ are positive semidefinite matrices, $D_i \in \mathcal{S}^n$ $(i = 1, 2)$ are nonnegative diagonal matrices, $c_i \in \Re^n$ for $i = 1, 2$, $A \in \Re^{m\times n}$, $b \in \Re^m$, and $u \in \Re^n$ is the upper bound vector of x. It will be seen below that problem (P) is a general form of many cardinality constrained portfolio selection problems where different risk measures based on factor models are used or tracking error against a benchmark index is required to be controlled in the model. The reader is referred to [10,21] for more discussions on portfolio selection models using multiple risk measures.

We assume that the random return R_i is driven by a group of factors:

$$R_i = \alpha_i + \beta_i^T f + \epsilon_i, \tag{2}$$

where $f \in \Re^m$ is the vector of random factors, α_i is the intercept representing the alpha value of the asset and $\beta_i \in \Re^m$ is the factor loading sensitivities, and ϵ_i is a random scalar representing the asset-specific return. Let $\alpha = (\alpha_1, \ldots, \alpha_n)^T$, $B = (\beta_1, \ldots, \beta_n)$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)^T$. We assume that ϵ_i $(i = 1, \ldots, n)$ has zero mean and is uncorrelated to the factors and ϵ_j for $j \ne i$. The covariance matrix of the asset returns can be then expressed as

$$\Sigma = B^T F B + D, \tag{3}$$

where F is the covariance matrix of the factors, and D is a diagonal matrix with the ith element being the variance of ϵ_i. It follows from (3) that the variance of the portfolio x is

$$x^T \Sigma x = x^T B^T F B x + x^T D x. \tag{4}$$

In the risk decomposition (4), the first part $x^T B^T F B x$ is called *systematic risk* and the second part $x^T D x$ is called *nonsystematic risk* or *specific risk*, which is contributed by the individual assets and can be reduced in a well-diversified portfolio.

As different choices of factors in the factor model (2) bring us different risk information, it is reasonable to combine two different factor models in one portfolio selection strategy. One way to realize this is to minimize the risk measure derived from one factor model whilst controlling the risk measure derived from another factor model. Let $\Sigma_1 = B_1^T F_1 B_1 + D_1$ and $\Sigma_2 = B_2^T F_2 B_2 + D_2$ be the covariance matrices of the asset returns driven by the factor model (2) using two different sets of factors. The resulting cardinality constrained portfolio selection model can be expressed as follows:

$$\begin{aligned}
(\mathrm{P}_1) \quad & \min\ x^T B_1^T F_1 B_1 x + x^T D_1 x, \\
& \text{s.t.}\ x^T B_2^T F_2 B_2 x + x^T D_2 x \le \sigma_0, \\
& \qquad \mu^T x \ge \rho,\ e^T x = 1, \\
& \qquad |\mathrm{supp}(x)| \le K, \\
& \qquad 0 \le x \le u,\ x_i \ge a_i,\ \forall i \in \mathrm{supp}(x),
\end{aligned}$$

where $e = (1, \ldots, 1)^T \in \Re^n$ and ρ is the target return level.

Another special case of problem (P) is the portfolio model in which the overall risk is minimized whilst the specific risk is required to be controlled. The motivation of controlling the specific risk is that the risk coming from specific sources of the individual assets is more volatile and uncertain. In some cases, investors are willing to take extra risk to ensure that the specific risk is below certain level. The resulting model can be written as

$$
\begin{aligned}
(\mathrm{P}_2) \qquad & \min\ x^T B^T F B x + x^T D x, \\
& \text{s.t. } x^T D x \le \sigma_0, \\
& \quad \mu^T x \ge \rho,\ e^T x = 1, \\
& \quad |\mathrm{supp}(x)| \le K, \\
& \quad 0 \le x \le u,\ x_i \ge a_i,\ \forall i \in \mathrm{supp}(x).
\end{aligned}
$$

Finally, we consider the cardinality constrained portfolio selection model with tracking error constraint. Nowadays, fund managers are often evaluated by the total return performance relative to a benchmark, usually a broadly diversified index of assets. Tracking error is a widely used measure for such relative performance. Let x_B be the index weight vector. The difference of the returns between the portfolio x and the index x_B is a random variable: $R^T(x - x_B)$, where R is the random return vector of the n assets driven by the factor model (2). The variance of the random return difference $R^T(x - x_B)$ can be then used to measure the tracking error:

$$
\text{tracking error} = (x - x_B)^T (B^T F B + D)(x - x_B).
$$

Minimizing the total portfolio volatility with tracking error control, the cardinality constraint and the minimum buy-in threshold results in the following problem:

$$
\begin{aligned}
(\mathrm{P}_3) \qquad & \min\ x^T B^T F B x + x^T D x, \\
& \text{s.t. } (x - x_B)^T (B^T F B + D)(x - x_B) \le \sigma_0, \\
& \quad \mu^T x \ge \rho,\ e^T x = 1, \\
& \quad |\mathrm{supp}(x)| \le K, \\
& \quad 0 \le x \le u,\ x_i \ge a_i,\ \forall i \in \mathrm{supp}(x).
\end{aligned}
$$

Alternatively, we can construct a portfolio model where the tracking error is minimized whilst the total volatility of the portfolio is controlled.

From the above discussion, we see that problem (P) is a general form of many cardinality constrained portfolio selection models when multiple risk measures are involved or tracking error constraint is imposed in the model.

3 Lagrangian decomposition and MIQCQP reformulation

In this section, we discuss how to use Lagrangian decomposition technique to get tight convex relaxation for problem (P). An SOCP relaxation is then derived from the dual problem via the Lagrangian decomposition scheme. Finally, we give a new MIQCQP reformulation whose continuous relaxation is exactly the SOCP relaxation.

3.1 Lagrangian decomposition

We first note that problem (P) can be reformulated to the following mixed integer quadratically constrained quadratic program (MIQCQP) by introducing a 0–1 variable $y_i \in \{0, 1\}$ to indicate the zero or nonzero status of each decision variable x_i:

$$
\begin{aligned}
(\mathrm{MIQCQP}_0) \qquad & \min\ x^T(Q_1+D_1)x + c_1^T x \\
& \text{s.t. } x^T(Q_2+D_2)x + c_2^T x \le \sigma_0, \\
& \quad Ax \le b,\ e^T y \le K, \\
& \quad a_i y_i \le x_i \le u_i y_i,\ y_i \in \{0,1\},\ i=1,\dots,n.
\end{aligned}
$$

A natural convex relaxation of (MIQCQP_0) and thus (P) is the convex QCQP problem obtained by relaxing the binary set $\{0,1\}^n$ to the box $[0,1]^n$. Let (QCQP_0) denote such a convex relaxation. In the following, we derive a new convex relaxation which is tighter or at least as tight as (QCQP_0).

By introducing a redundant constraint $z = x$, problem (MIQCQP_0) can be rewritten as:

$$
\begin{aligned}
& \min\ x^T Q_1 x + z^T D_1 z + c_1^T x \\
& \text{s.t. } x^T Q_2 x + z^T D_2 z + c_2^T x \le \sigma_0, \\
& \quad z = x,\ 0 \le x \le u \\
& \quad Ax \le b,\ e^T y \le K, \\
& \quad a_i y_i \le z_i \le u_i y_i,\ y_i \in \{0,1\},\ i=1,\dots,n.
\end{aligned}
$$

Associating the constraint $x^T Q_2 x + z^T D_2 z + c_2^T x \le \sigma_0$ with multiplier $\lambda \ge 0$ and the constraint $z = x$ with multiplier vector $\pi \in \Re^n$, we obtain the following Lagrangian relaxation:

$$
d(\lambda,\pi) = -\sigma_0\lambda + d_1(\lambda,\pi) + d_2(\lambda,\pi),
$$

where

$$
\begin{aligned}
d_1(\lambda,\pi) = & \min\ x^T(Q_1+\lambda Q_2)x + (c_1+\lambda c_2 - \pi)^T x && (5)\\
& \text{s.t. } Ax \le b, \\
& \quad 0 \le x \le u \\
d_2(\lambda,\pi) = & \min\ z^T(D_1+\lambda D_2)z + \pi^T z && (6)\\
& \text{s.t. } e^T y \le K, \\
& \quad a_i y_i \le z_i \le u_i y_i,\ i=1,\dots,n, \\
& \quad y \in \{0,1\}^n.
\end{aligned}
$$

The dual problem is

$$
\begin{aligned}
(\mathrm{D}) \qquad & \max\ d(\lambda,\pi) \\
& \text{s.t. } \lambda \ge 0,\ \pi \in \Re^n.
\end{aligned}
$$

By weak duality, we always have $v(\mathrm{D}) \le v(\mathrm{MIQCQP}_0) = v(\mathrm{P})$.

3.2 SOCP relaxation

In the following, we show that the dual problem (D) can be reduced to an SOCP problem. We first note that if $Q_1 + \lambda Q_2 \succeq 0$, then the subproblem (5) is a convex quadratic program so that strong duality between (5) and its dual problem holds. Using Shor's homogeneous technique, we can express the dual problem of (5) as the following SDP problem:

$$\begin{aligned} d_1(\lambda, \pi) &= \max\ \theta \\ \text{s.t.}\ &\begin{pmatrix} Q_1 + \lambda Q_2 & \frac{1}{2}(c_1 + \lambda c_2 - \pi + A^T\gamma + \xi - \eta) \\ \frac{1}{2}(c_1 + \lambda c_2 - \pi + A^T\gamma + \xi - \eta) & -\gamma^T b - \xi^T u - \theta \end{pmatrix} \succeq 0, \\ &\xi, \eta \in \Re^n_+,\ \gamma \in \Re^m_+. \end{aligned}$$

Let $D_k = \text{diag}(d_1^k, \ldots, d_n^k)$ $(k = 1, 2)$. For each $i = 1, \ldots, n$, let

$$q_i = \min\{(d_i^1 + \lambda d_i^2)z_i^2 + \pi_i z_i \mid z_i \in \{0\} \cup [a_i, u_i]\}. \tag{7}$$

Then, $q_i \le 0$ for $i = 1, \ldots, n$. It follows that $d_2(\lambda, \pi)$ is the sum of the K smallest elements of vector $q = (q_1, \ldots, q_n)$. Let $S_K(p)$ denote the sum of the K largest elements of vector $p \in \Re^n$. Then, we have

$$d_2(\pi, \lambda) = \max\{-t \mid S_K(-q) \le t\}. \tag{8}$$

The following lemma is a special case of the LMI representation of the sum of the K largest eigenvalues of a symmetric matrix (see, e.g., [3, p. 147]).

Lemma 1 *For any vector $p \in \Re^n$, the following two sets A and B are identical:*

$$\begin{aligned} A &= \{(p, t) \mid S_K(p) \le t\} \\ B &= \{(p, t) \mid \exists\, (v, s) \in \Re^n \times \Re \text{ satisfying (a), (b) and (c)}\} \end{aligned}$$

where (a) $t - Ks - e^T v \ge 0$, (b) $v - p + se \ge 0$ *and* (c) $v \ge 0$.

It follows from (8) and Lemma 1 that the subproblem (6) can be rewritten as

$$\begin{aligned} d_2(\lambda, \pi) = \max\ & -t \\ \text{s.t. } & t - Ks - e^T v \ge 0,\ v + q + se \ge 0,\ v \ge 0. \end{aligned} \tag{9}$$

On the other hand, by (7), we can express q_i as

$$\begin{aligned} q_i &= \min\{(d_i^1 + \lambda d_i^2)z_i^2 + \pi_i z_i \mid z_i \in \{0\} \cup [a_i, u_i]\} \\ &= \min\{0, \min_{z_i \in [a_i, u_i]} [(d_i^1 + \lambda d_i^2)z_i^2 + \pi_i z_i]\} \\ &= \max\ \beta_i \\ &\quad\ \text{s.t. } \beta_i \le 0, \\ &\quad\quad\ \beta_i \le \min_{z_i \in [a_i, u_i]} [(d_i^1 + \lambda d_i^2)z_i^2 + \pi_i z_i] \\ &= \max\ \beta_i \\ &\quad\ \text{s.t. } \beta_i \le 0,\ \beta_i \le \tau_i, \\ &\quad\quad \begin{pmatrix} d_i^1 + \lambda d_i^2 & \frac{1}{2}(\pi_i + \phi_i - \psi_i) \\ \frac{1}{2}(\pi_i + \phi_i - \psi_i) & a_i \psi_i - u_i \phi_i - \tau_i \end{pmatrix} \succeq 0, \\ &\quad\quad \phi_i \ge 0,\ \psi_i \ge 0, \end{aligned}$$

where the last equality is due to the strong duality for the convex quadratic program $\min_{z_i \in [a_i, u_i]} [(d_i^1 + \lambda d_i^2)z_i^2 + \pi_i z_i]$, and $\phi_i \ge 0$ and $\psi_i \ge 0$ are the multipliers for the box constraint $a_i \leq z_i \leq u_i$. Using (9) and the above expression of q_i $(i = 1, \ldots, n)$, we can rewrite the subproblem (6) as the following SDP problem:

$$\begin{aligned} d_2(\lambda, \pi) = \max \ & -t \\ \text{s.t. } & t - Ks - e^T v \geq 0, \ v + \beta + se \geq 0, \ \tau - \beta \geq 0, \\ & \begin{pmatrix} d_i^1 + \lambda d_i^2 & \frac{1}{2}(\pi_i + \phi_i - \psi_i) \\ \frac{1}{2}(\pi_i + \phi_i - \psi_i) & a_i \psi_i - u_i \phi_i - \tau_i \end{pmatrix} \succeq 0, \ i = 1, \ldots, n, \\ & v, \psi, \phi, -\beta \in \Re^n_+, \ \tau \in \Re^n. \end{aligned}$$

Summarizing the above discussion, we conclude that the dual problem (D) can be reduced to the following SDP problem:

$$\begin{aligned} (\mathrm{D_{sdp}}) \quad & \max \ -\sigma_0 \lambda + \theta - t, \\ & \text{s.t. } t - Ks - e^T v \geq 0, \ v + \beta + se \geq 0, \ \tau - \beta \geq 0, \\ & \begin{pmatrix} Q_1 + \lambda Q_2 & \frac{1}{2}(c_1 + \lambda c_2 - \pi + A^T \gamma + \xi - \eta) \\ \frac{1}{2}(c_1 + \lambda c_2 - \pi + A^T \gamma + \epsilon - \eta) & -\gamma^T b - \xi^T u - \theta \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} d_i^1 + \lambda d_i^2 & \frac{1}{2}(\pi_i + \phi_i - \psi_i) \\ \frac{1}{2}(\pi_i + \phi_i - \psi_i) & a_i \psi_i - u_i \phi_i - \tau_i \end{pmatrix} \succeq 0, \ i = 1, \ldots, n \\ & \xi, \eta, v, \psi, \phi, -\beta \in \Re^n_+, \ \gamma \in \Re^m_+, \ \tau, \pi \in \Re^n, \ \theta, s, t \in \Re, \lambda \in \Re_+. \end{aligned}$$

It can be verified that the conic dual of problem ($\mathrm{D_{sdp}}$) is

$$\begin{aligned} (\mathrm{SDP}) \quad \min \ & Q_1 \bullet X + c_1^T x + \sum_{i=1}^n d_i^1 \delta_i, \\ \text{s.t. } & Q_2 \bullet X + c_2^T x + \sum_{i=1}^n d_i^2 \delta_i \leq \sigma_0, \\ & Ax \leq b, \ e^T y \leq K, \ 0 \leq y \leq 1, \\ & a_i y_i \leq x_i \leq u_i y_i, \ i = 1, \ldots, n, \\ & x_i^2 \leq \delta_i y_i, \ \delta_i \geq 0, \ i = 1, \ldots, n, \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0, \end{aligned}$$

where $X \in \mathcal{S}^n$. By the conic duality theorem (see, e.g., [34]), if (SDP) is strictly feasible, then the strong duality holds, i.e., $v(\mathrm{D_{sdp}}) = v(\mathrm{SDP})$. We assume in the sequel that the Slater condition for ($\mathrm{QCQP_0}$) holds, i.e., there exists a relative interior point in the feasible set of ($\mathrm{QCQP_0}$). It is easy to see that this assumption implies the strict feasibility of (SDP).

Note that $Q_1 \succeq 0$ and $Q_2 \succeq 0$. We can eliminate X in (SDP) by replacing $X \bullet Q_1$ and $X \bullet Q_2$ with $x^T Q_1 x$ and $x^T Q_2 x$, respectively. Then, problem (SDP) can be simplified to the following problem:

$$\begin{aligned} (\mathrm{QCQP_1}) \quad \min \ & f(x, \delta) := x^T Q_1 x + c_1^T x + \sum_{i=1}^n d_i^1 \delta_i, \\ \text{s.t. } & g(x, \delta) := x^T Q_2 x + c_2^T x + \sum_{i=1}^n d_i^2 \delta_i \leq \sigma_0, \\ & Ax \leq b, \ e^T y \leq K, \ 0 \leq y \leq 1, \end{aligned}$$

$$a_i y_i \leq x_i \leq u_i y_i,\ i = 1, \ldots, n,$$
$$x_i^2 \leq \delta_i y_i,\ \delta_i \geq 0,\ i = 1, \ldots, n.$$

Note that $g(x, \delta) \leq \sigma_0$ is a convex quadratic constraint. Also,

$$x_i^2 \leq \delta_i y_i,\ \delta_i \geq 0,\ y_i \geq 0 \Leftrightarrow \left\| \begin{pmatrix} \frac{\delta_i - y_i}{2} \\ x_i \end{pmatrix} \right\|_2 \leq \frac{\delta_i + y_i}{2}.$$

Therefore, (QCQP_1) is a second-order cone program (SOCP).

The following theorem shows that (QCQP_1) is tighter than or at least as tight as (QCQP_0), the continuous relaxation of the standard MIQCQP reformulation of (P).

Theorem 1 *It holds that* $v(\text{QCQP}_0) \leq v(\text{QCQP}_1) \leq v(\text{P})$.

Proof Let (x, y, δ) be any feasible solution to (QCQP_1). By the constraints $x_i^2 \leq \delta_i y_i, \delta_i \geq 0$ and $0 \leq y_i \leq 1$ in (QCQP_1), we deduce that $\delta_i \geq x_i^2$. Note that $D_2 = \text{diag}(d_1^2, \ldots, d_n^2)$ is a nonnegative diagonal matrix. Thus, $g(x, \delta) \geq x^T Q_2 x + c_2^T x + x^T D_2 x = g(x)$. Similarly, we have $f(x, \delta) \geq f(x)$. Therefore, (x, y) is feasible to (QCQP_0) with $f(x, \delta) \geq f(x)$, which implies that $v(\text{QCQP}_1) \geq v(\text{QCQP}_0)$. The inequality $v(\text{QCQP}_1) \leq v(\text{P})$ is due to the relation:

$$v(\text{QCQP}_1) = v(\text{D}_{\text{sdp}}) = v(\text{SDP}) = v(\text{D}) \leq v(\text{P}).$$

This completes the proof. $\square$

The following example shows that the strict inequality may holds in the inequality $v(\text{QCQP}_0) \leq v(\text{QCQP}_1)$ in Theorem 1.

Example 1 Consider the cardinality constrained problem:

$$\begin{aligned} \min\ & 3x_1^2 - 4x_1x_2 + 5x_2^2, \\ \text{s.t.}\ & x_1^2 + x_2^2 \leq 5, \\ & -2x_1 - 3x_2 \leq -6, \\ & |\text{supp}(x)| \leq 1, \\ & x_i \geq 1,\ \forall i \in \text{supp}(x), \\ & 0 \leq x_i \leq 3,\ i = 1, 2. \end{aligned}$$

The example has an optimal solution $x^* = (0, 2)^T$ with optimal value 20. In this example, we can set

$$Q_1 = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix},\ D_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, Q_2 = 0,\ D_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The optimal value of the standard continuous relaxation is $v(\text{QCQP}_0) = 5.5775$ while the optimal value of (QCQP_1) is $v(\text{QCQP}_1) = 12.0510 > v(\text{QCQP}_0)$.

3.3 MIQCQP reformulation

We see that the SOCP relaxation (QCQP_1) is the continuous relaxation of the following MIQCQP problem:

$$
\begin{aligned}
(\mathrm{MIQCQP}_1) \quad & \min \ f(x,\delta) := x^T Q_1 x + c_1^T x + \sum_{i=1}^n d_i^1 \delta_i, \\
& \text{s.t.} \ \ g(x,\delta) := x^T Q_2 x + c_2^T x + \sum_{i=1}^n d_i^2 \delta_i \le \sigma_0, \\
& \qquad Ax \le b, \ e^T y \le K, \ \ y \in \{0,1\}^n, \\
& \qquad a_i y_i \le x_i \le u_i y_i, \ \ i = 1, \ldots, n, \\
& \qquad x_i^2 \le \delta_i y_i, \ \ i = 1, \ldots, n.
\end{aligned}
$$

The following theorem shows that (MIQCQP_1) is equivalent to (MIQCQP_0).

Theorem 2 *A solution* (x, y, δ) *solves* (MIQCQP_1) *if and only if* (x, y) *solves* (MIQCQP_0). *Moreover,* $v(\mathrm{MIQCQP}_1) = v(\mathrm{MIQCQP}_0)$.

Proof For any feasible solution (x, y, δ) of (MIQCQP_1), using the similar argument as in the proof of Theorem 1, we can show that (x, y) is a feasible solution to (MIQCQP_0) with $f(x,\delta) \ge f(x)$. Thus, $v(\mathrm{MIQCQP}_1) \ge v(\mathrm{MIQCQP}_0)$.

On the other hand, suppose (x, y) is a feasible solution to (MIQCQP_0). Let $\delta_i = x_i^2$ for $i = 1, \ldots, n$. Then, $x_i^2 \le \delta_i y_i$ for any $y_i \in \{0, 1\}$, $f(x,\delta) = f(x)$ and $g(x,\delta) = g(x)$. Thus, (x, y, δ) is a feasible solution to (MIQCQP_1) with $f(x,\delta) = f(x)$. Therefore, $v(\mathrm{MIQCQP}_0) \ge v(\mathrm{MIQCQP}_1)$. This proves the theorem. □

From Theorems 1 and 2, we see that (MIQCQP_1) is a more efficient reformulation than the standard reformulation (MIQCQP_0) in the sense that its continuous relaxation is always tighter than or at least as tight as that of the standard reformulation (MIQCQP_0). It is therefore expected that continuous relaxation-based branch-and-bound methods perform better on (MIQCQP_1) than (MIQCQP_0), as witnessed in our numerical experiment in the next section.

We point out that the reformulation (MIQCQP_1) can be also obtained by applying the perspective reformulation proposed in [15] (see also [16–18]) to the standard MIQCQP reformulation (MIQCQP_0). Our derivation of (MIQCQP_1) is different from the perspective approach in [15] where perspective function is used to convexify separable quadratic functions on semi-continuous variables.

4 Numerical experiments

In this section, we report our numerical results on the tightness of the SOCP relaxation and the performance of the new reformulation (MIQCQP_1) when compared with the standard reformulation (MIQCQP_0).

Our test problems are instances of the three cardinality constrained portfolio selection problems (P_1), (P_2) and (P_3) described in Sect. 2. To build the factor models involved in these three types of test problems, we use the weekly return data of 481 stocks from Standard & Poor's 500 index between 2005 and 2010. We use linear regressions on 10 sector indexes of Standard & Poor's 500 index to construct a sector factor model. We also construct a statistic factor model by performing principal component analysis on the weekly return data of 481 stocks. The quadratic objective function and the quadratic constraint in (P_1), (P_2) and (P_3) are explained as follows.

– The objective function of (P_1) is the variance obtained by the sector factor model while the quadratic constraint function is the variance obtained from the statistic factor model.

Table 1 Numerical results for problem (P_1)

n	$MIQCQP_0$		$MIQCQP_1$		$QCQP_0$	$QCQP_1$
	CPU time (final gap)	No. of nodes	CPU time (final gap)	No. of nodes	Gap ratio	Gap ratio
100	19.8	47	8.8	5	1.17	0.09
100	29.3	80	11.4	13	0.83	0.04
100	25.2	107	8.1	7	1.55	0.17
100	16.3	46	8.4	5	0.41	0.03
100	16.8	41	9.1	5	1.50	0.12
200	151.2	155	22.0	15	2.98	0.37
200	96.8	68	15.2	5	2.51	0.01
200	126.6	114	19.4	11	2.46	0.21
200	102.6	61	22.4	11	0.90	0.04
200	154.3	133	16.3	5	2.82	0.18
300	2,323	1,175	48.0	19	3.74	0.56
300	1,347	573	41.7	15	4.06	0.26
300	1,039	464	31.5	9	3.68	0.44
300	1,174	503	39.5	15	3.49	0.45
300	1,239	551	34.8	7	3.90	0.11
400	1,129	161	33.4	5	5.56	0.46
400	3,600 (1.6%)	601	56.2	9	3.55	0.16
400	3,600 (2.7%)	606	93.7	15	4.59	0.23
400	3,600 (2.9%)	520	48.0	7	4.56	0.42
400	3,600 (2.2%)	545	67.6	23	4.32	0.45

- The objective function of (P_2) is the same as that of (P_1) while the quadratic constraint is the constraint of controlling the specific risk $x^T Dx$ in the sector factor model.
- The objective function of (P_3) is the same as that of (P_1) while the quadratic constraint is the tracking error control with Standard & Poor's 500 index as the benchmark index.

We generate 5 instances for each test problem with the same size ($n = 100, 200, 300$ and 400). For each instance, we randomly choose n stocks from the 481 stocks. For each test problem, we set the prescribed weekly return level $\rho = 0.2\%$, $a_i = 0.01$ and $u_i = 0.3$ ($i = 1, \ldots, n$). The cardinality upper bound K is set equal to 10.

The numerical tests were implemented in Matlab 7.9 and run on a PC (2.4 GHz, 3 GB RAM). The mixed integer quadratic reformulations $(MIQCQP_0)$ and $(MIQCQP_1)$ are solved by the mixed integer QCP solver in `CPLEX 12.1` with Matlab interface using continuous relaxation for generating lower bounds [19]. The experiments were conducted by using the default setting of `CPLEX 12.1` with the maximum CPU time set equal to 3,600 s.

To measure the quality of the convex relaxations $(QCQP_i)$ ($i = 0, 1$), we calculate the following gap ratio between the lower bounds generated by $(QCQP_i)$ ($i = 0, 1$) at the root node and the optimal value of (P):

$$\text{gap ratio} = \frac{v(\mathrm{P}) - v(\mathrm{QCQP_i})}{v(\mathrm{P})}(\%),$$

Table 2 Numerical results for problem (P_2)

n	$MIQCQP_0$		$MIQCQP_1$		$QCQP_0$	$QCQP_1$
	CPU time (final gap)	No. of nodes	CPU time (final gap)	No. of nodes	Gap ratio	Gap ratio
100	15.2	135	5.2	5	1.71	0.02
100	24.5	125	5.3	5	1.85	0.19
100	22.5	229	8.2	19	2.57	0.23
100	57.1	784	19.4	81	3.01	0.51
100	248.5	4,131	8.3	18	6.60	0.30
200	607.4	2,017	15.0	23	2.77	0.11
200	53.1	87	14.7	15	1.60	0.12
200	89.0	207	12.0	9	1.96	0.07
200	1,129	3,495	12.8	9	5.50	0.15
200	101.7	233	8.4	5	2.75	0.10
300	882.0	1,328	28.2	21	3.13	0.17
300	931.6	1,118	34.9	33	4.07	0.20
300	3,077	3,444	43.9	49	5.42	0.59
300	3,600 (1.8%)	4,341	37.1	29	5.37	0.38
300	3,600 (3.0%)	4,553	72.5	77	5.58	0.68
400	3,600 (1.0%)	1,854	25.4	11	3.54	0.42
400	3,600 (21.0%)	1,761	42.5	23	7.45	0.40
400	3,600 (5.2%)	1,574	93.1	87	6.57	0.07
400	3,600 (3.0%)	1,529	113.3	133	4.36	0.14
400	3,600 (1.9%)	1,630	78.8	55	4.48	0.45

where $v(P)$ is obtained from solving ($MIQCQP_1$) via `CPLEX 12.1`. If the instance is not solved within the maximum CPU time, we also record the final gap of `CPLEX12.1` which is defined by

$$\text{final gap} = \frac{(\text{upper bound} - \text{lower bound})}{\text{upper bound}} (\%),$$

where "upper bound" is the objective value of the current best feasible solution.

The computing time, the number of nodes explored by `CPLEX 12.1` and the relaxation gap ratio for test problems (P_1), (P_2) and (P_3) are summarized in Tables 1, 2 and 3. Comparing the numerical results in Tables 1, 2 and 3, we see that the new reformulation ($MIQCQP_1$) is much more efficient than the standard reformulation ($MIQCQP_0$) in terms of the computing time used by `CPLEX 12.1`. In particular, `CPLEX 12.1` can solve ($MIQCQP_1$) for each instances of the three types of test problems to global optimality within 3,600 CPU seconds, while `CPLEX 12.1` did not terminate within 3,600 CPU seconds for some large-scale instances of ($MIQCQP_0$).

Figures 1, 2 and 3 illustrate the average gap ratio versus the average CPU time used by `CPLEX 12.1` for test problems (P_1), (P_2) and (P_3), respectively. From Figures 1, 2 and 3, it is clear that the gap ratio of the continuous relaxation of ($MIQCQP_1$) is much less than that of ($MIQCQP_0$). This suggests that the SOCP relaxation ($QCQP_1$) is significantly tighter

Table 3 Numerical results for problem (P_3)

n	$MIQCQP_0$		$MIQCQP_1$		$QCQP_0$	$QCQP_1$
	CPU time (final gap)	No. of nodes	CPU time (final gap)	No. of nodes	Gap ratio	Gap ratio
100	132.9	899	8.6	5	3.00	0.03
100	879.9	6,607	9.8	3	8.82	0.18
100	3,600 (6.3%)	27,877	13.5	49	9.13	0.27
100	3,600 (1.8%)	32,274	22.8	93	6.89	0.54
100	3,600 (0.6%)	32,168	25.7	91	6.02	0.23
200	3,600 (2.4%)	5,011	22.5	17	5.81	0.39
200	3,600 (0.5%)	6,287	38.2	43	5.81	0.77
200	3,600 (5.6%)	5,008	36.2	40	7.68	0.18
200	3,600 (7.7%)	4,807	29.8	31	8.64	0.25
200	3,600 (21.2%)	4,439	644.4	1,676	15.35	0.92
300	3,600 (13.8%)	1,379	85.0	108	10.61	0.35
300	3,600 (6.7%)	1,374	37.7	21	6.57	0.37
300	3,600 (8.8%)	1,309	46.6	38	8.39	0.39
300	3,600 (9.5%)	1,605	113	126	10.61	1.25
300	3,600 (16.2%)	1,636	109.2	151	10.83	0.30
400	3,600 (8.2%)	478	80.2	59	5.69	0.44
400	3,600 (10.1%)	714	86.7	64	9.95	0.38
400	3,600 (11.8%)	544	77.1	57	11.12	0.16
400	3,600 (15.4%)	544	98.6	68	9.29	0.71
400	3,600 (13.9%)	632	183	141	11.84	0.54

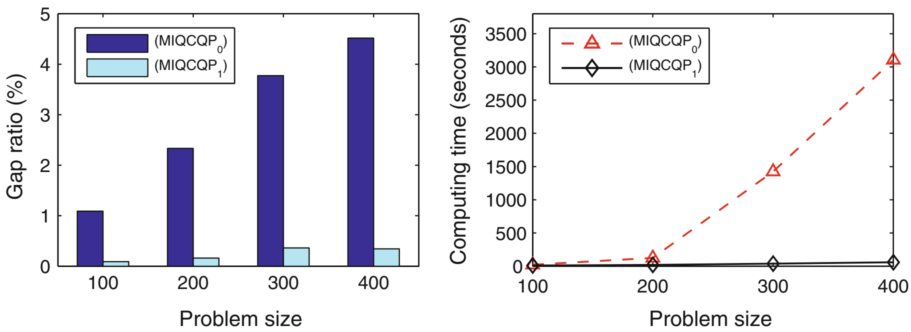

Fig. 1 Average relaxation gap ratio versus computing time for test problem (P_1)

than the continuous relaxation ($QCQP_0$) for our test problems. The evolution of the average computing time in Figures 1, 2 and 3 indicates that `CPLEX 12.1` uses much less time to solve ($MIQCQP_1$) than ($MIQCQP_0$). Again, this is mainly because of the tightness of the continuous relaxation of ($MIQCQP_1$) when compared with ($MIQCQP_0$).

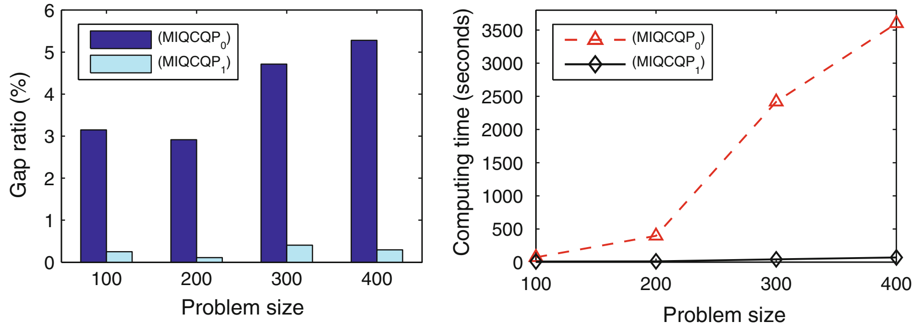

Fig. 2 Average relaxation gap ratio versus computing time for test problem (P_2)

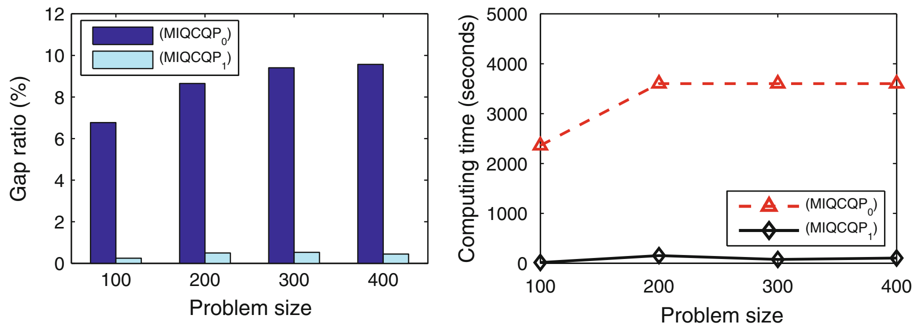

Fig. 3 Average relaxation gap ratio versus computing time for test problem (P_3)

5 Conclusion

We have investigated a class of portfolio optimization problems with cardinality and minimum buy-in threshold constraints. From computational point of view, this class of problems are NP-hard due to the discrete nature of the cardinality constraint. We have showed that the dual problem obtained from a Lagrangian decomposition scheme can be reduced to an SOCP problem. This SOCP problem leads to a new MIQCQP reformulation of the original problem which is more efficient than the standard reformulation in term of the tightness of the continuous relaxations. Our numerical results have demonstrated the outperformance of the new MIQCQP reformulation for three types of cardinality constrained portfolio selection problems using mixed integer QCP solver in `CPLEX 12.1`.

References

1. Alexander, G.J., Baptista, A.M.: A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. Manag. Sci. **50**, 1261–1273 (2004)
2. Angelelli, R., Mansini, R., Speranza, M.G.: A comparison of MAD and CVaR models with real features. J. Banking Finance **32**, 1188–1197 (2007)
3. Ben-Tal, A., Nemirovski, A.S.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. MPS-SIAM Series on Optimization. Society for Industrial Mathematics, Philadelphia, PA (2001)

4. Bertsimas, D., Shioda, R.: Algorithm for cardinality-constrained quadratic optimization. Comput. Optim. Appl. **43**, 1–22 (2009)
5. Bienstock, D.: Computational study of a family of mixed-integer quadratic programming problems. Math. Program. **74**, 121–140 (1996)
6. Boginski, V., Butenko, S., Pardalos, P.M.: Statistical analysis of financial networks. Comput. Stat. Data Anal. **48**, 431–443 (2005)
7. Boginski, V., Butenko, S., Pardalos, P.M.: Mining market data: a network approach. Comput. Oper. Res. **33**, 3171–3184 (2006)
8. Bonami, P., Lejeune, M.A.: An exact solution approach for portfolio optimization problems under stochastic and integer constraints. Oper. Res. **57**, 650–670 (2009)
9. Burmeister, E., Roll, R., Ross, S.A.: Using macroeconomic factors to control portfolio risk. Tech. rep. (2003). Available at: http://www.birr.com/Using_Macroeconomic_Factors.pdf
10. Ceria, S., Margot, F., Renshaw, A., Saxena, A.: Novel approaches to portfolio construction: multiple risk models and multisolution generation. In: Satchell, S. (ed.) Optimizing Optimization: The Next Generation of Optimization Applications & Theory, pp. 23–52. Academic Press and Elsevier, Amsterdam (2010)
11. Connor, G.: The three types of factor models: a comparison of their explanatory power. Financial Analysts J. **51**, 42–46 (1995)
12. Duffie, D., Pan, J.: An overview of value at risk. J. Deriv. **4**, 7–49 (1997)
13. Fama, E.F., French, K.R.: The cross-section of expected stock returns. J. Finance **47**, 427–465 (1992)
14. Fama, E.F., French, K.R.: Size and book-to-market factors in earnings and returns. J. Finance **50**, 131–155 (1995)
15. Frangioni, A., Gentile, C.: Perspective cuts for a class of convex 0–1 mixed integer programs. Math. Program. **106**, 225–236 (2006)
16. Frangioni, A., Gentile, C.: SDP diagonalizations and perspective cuts for a class of nonseparable MIQP. Oper. Res. Lett. **35**, 181–185 (2007)
17. Frangioni, A., Gentile, C.: A computational comparison of reformulations of the perspective relaxation: SOCP vs. cutting planes. Oper. Res. Lett. **37**, 206–210 (2009)
18. Günlük, O., Linderoth, J.: Perspective reformulations of mixed integer nonlinear programs with indicator variables. Math. Program. **124**, 183–205 (2010)
19. IBM ILOG CPLEX: IBM ILOG CPLEX 12.1 User's Manual for CPLEX (2010)
20. Jorion, P.: Value at Risk: The New Benchmark for Controlling Market Risk. McGraw-Hill, Chicago (1997)
21. Kolbert, F., Wormald, L.: Robust portfolio optimization using second-order cone programming. In: Satchell, S. (ed.) Optimizing Optimization: The Next Generation of Optimization Applications & Theory, pp. 3–22. Academic Press and Elsevier, Amsterdam (2010)
22. Konno, H., Yamazaki, H.: Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. Manag. Sci. **37**, 519–531 (1991)
23. Li, D., Ng, W.L.: Optimal dynamic portfolio selection: multiperiod mean-variance formulation. Math. Finance **10**, 387–406 (2000)
24. Markowitz, H.M.: Portfolio selection. J. Finance **7**, 77–91 (1952)
25. Markowitz, H.M.: Portfolio Selection: Efficient Diversification of Investments. Wiley, New York (1959)
26. Moazeni, S.: A note on the dynamic liquidity trading problem with a mean-variance objective. Optim. Lett. **5**, 113–124 (2011)
27. Pardalos, P.M., Sandström, M., Zopounidis, C.: On the use of optimization models for portfolio selection: a review and some computational results. Comput. Econ. **7**, 227–244 (1994)
28. Pardalos, P.M., Tsitsiringos, V. (eds.): Financial Engineering, Supply Chain and E-Commerce. Kluwer, Dordrecht (2002)
29. Pardalos, P.M., Wolkowicz, H. (eds.): Topics in Semidefinite and Interior-Point Methods. Fields Institute Communications Series: vol. 18. American Mathematical Society, Providence (1998)
30. Rockafellar, R.T., Uryasev, S.: Optimization of conditional Value-at-Risk. J. Risk **2**, 21–42 (2000)
31. Rockafellar, R.T., Uryasev, S.: Conditional Value-at-Risk for general loss distributions. J. Banking Finance **26**, 1443–1471 (2002)
32. Sharpe, W.F.: A simplified model for portfolio analysis. Manag. Sci. **9**, 277–293 (1963)
33. Shaw, D., Liu, S., Kopman, L.: Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. Optim. Methods Softw. **23**, 411–420 (2008)
34. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev. **38**, 49–95 (1996)
35. Young, M.R.: A minimax portfolio selection rule with linear programming solution. Manag. Sci. **44**, 673–683 (1998)
36. Zhou, X.Y., Li, D.: Continuous-time mean-variance portfolio selection: a stochastic LQ framework. Appl. Math. Optim. **42**, 19–33 (2000)