

Tight MIQP Reformulations for Semi-Continuous Quadratic Programming: Lift-and-Convexification Approach

Baiyi Wu

Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong,
Shatin, N. T., Hong Kong, bywu@se.cuhk.edu.hk

Xiaoling Sun

Department of Management Science, School of Management, Fudan University, Shanghai 200433, P. R.
China

Duan Li

Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong,
Shatin, N. T., Hong Kong, dli@se.cuhk.edu.hk

Xiaojin Zheng

School of Economics and Management, Tongji University, Shanghai 200092, P. R. China.

Abstract

We consider in this paper a class of semi-continuous quadratic programming problems which arises in many real-world applications such as production planning, portfolio selection and subset selection in regression. We propose a lift-and-convexification approach to derive an equivalent reformulation of the original problem. This lift-and-convexification approach lifts the quadratic term involving x only in the original objective function $f(x, y)$ to a quadratic function of both x and y and convexifies this equivalent objective function. While the continuous relaxation of our new reformulation attains the same tight bound as achieved by the continuous relaxation of the well known perspective reformulation, the new reformulation also retains the linearly constrained quadratic programming structure of the original mix-integer problem. This prominent feature improves the performance of branch-and-bound algorithms by providing the same tightness at the root node as the state-of-the-art perspective reformulation and offering much faster processing time at children nodes. We further combine the lift-and-convexification approach and the quadratic convex reformulation approach in the literature to form an even tighter reformulation. Promising results from our computational tests in both portfolio selection and subset selection problems numerically verify the benefits from these theoretical features of our new reformulations.

1 Introduction

We consider in this paper the following mixed-integer quadratic programming (MIQP) problem:

$$\begin{aligned} \text{(P)} \quad & \min f(x, y) = x^T Q x + c^T x + h^T y \\ & \text{s.t. } Ax + By \leq d, \\ & a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \end{aligned} \tag{1}$$

where Q is an $n \times n$ positive semidefinite symmetric matrix, $c, h \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, and $A, B \in \mathbb{R}^{m \times n}$.

Problem (P) is in general NP-hard (see [7]). Its difficulty arises from the discrete structure induced by the constraint in (1). This constraint is used to model the situation where x_i must rest inside an interval if it is not zero, that is, $x_i \in \{0\} \cup [a_i, b_i]$. These variables x_i are termed *semi-continuous* variables. We assume in our study $a_i < b_i$, $i = 1, \dots, n$. We also assume that the feasible region of problem (P) is nonempty. A recent review on problem (P) and its solution methods can be found in [45].

Semi-continuous variables appear in many real-world optimization problems. For instance, in production planning, the semi-continuous variables are used to describe the state of a production process that is either turned off (inactive), hence nothing is produced, or turned on (active) such that the production level has to lie in certain interval ([22, 26, 27]). Other typical applications of semi-continuous variables include portfolio selection with *minimum buy-in threshold* ([35, 21, 17, 45]) and lot-sizing with minimum order quantity ([3, 41]).

An important instance of (P) involves optimization models with a cardinality constraint:

$$|\text{supp}(x)| \leq K, \quad (2)$$

where $\text{supp}(x) = \{i \mid x_i \neq 0\}$ and K is an integer with $1 \leq K \leq n$. The cardinality constraint is often encountered when the number of nonzero variables has to be limited. The cardinality constraint in (2) can be easily incorporated into problem (P) by introducing an additional linear constraint $\sum_{i=1}^m y_i \leq K$.

A well-known application of semi-continuous variables and cardinality constraint is the *cardinality constrained mean-variance portfolio selection* in financial optimization. The classical mean-variance model of Markowitz is a quadratic programming problem that minimizes the variance subject to linear constraints on expected return and budget availabilities. In real-world applications of portfolio selection models, however, most investors would invest in only a limited number of assets due to market frictions such as management and transaction fees. Moreover, the minimum buy-in threshold is often a mandate trading constraint. Suppose that there are n risky assets in a financial market with a random return vector $R = (R_1, \dots, R_n)^T$. Furthermore, the expected return vector and the covariance matrix of R are assumed to be given as μ and Q , respectively. The portfolio selection model with cardinality and minimum buy-in threshold constraints can be then expressed as:

$$\begin{aligned} \text{(MV)} \quad & \min x^T Q x \\ \text{s.t.} \quad & \sum_{i=1}^n x_i = 1, \\ & \mu^T x \geq \rho, \\ & \sum_{i=1}^n y_i \leq K, \\ & a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \end{aligned}$$

where x_i represents the proportion of the total capital invested in the i th asset, and ρ is a prescribed expected return level set by the investor. Portfolio selection problems with

cardinality and/or minimum threshold constraints have been studied extensively in recent literature. For exact solution methods, please see, e.g., [7, 37, 44, 13, 6, 17, 29]. For inexact solution methods, such as heuristics, local search methods and randomized techniques, please see, e.g., [34, 12, 15, 35, 43, 38, 16, 40, 19, 46].

Another application of (P) with cardinality constraint is the subset selection problem in multivariate linear regression. Given m observed data points (a_i, b_i) with $a_i \in \mathfrak{R}^n$ and $b_i \in \mathfrak{R}$, we need to minimize the least square measure of $\sum_{i=1}^m (a_i^T x - b_i)^2$ with only a subset of the prediction variables in x (see, e.g., [4, 39, 6]). This problem can be formally formulated as:

$$\begin{aligned} \text{(SSP)} \quad & \min \quad \|Ax - b\|^2 \\ & \text{s.t.} \quad |\text{supp}(x)| \leq K, \end{aligned}$$

where $A^T = (a_1, \dots, a_m)$, $b = (b_1, \dots, b_m)^T$, and K is an integer with $1 \leq K \leq n$. When we convert this problem to problem (P), lower bounds and upper bounds on x , i.e. $L \leq x_i \leq U$, can be imposed for a sufficiently large positive number U and a sufficiently small negative number L .

Cardinality constrained linear-quadratic optimal control was investigated in [28]. Furthermore, a polynomially solvable case of the cardinality-constrained quadratic optimization problem was identified in [30].

We focus in this paper on exact solution methods for problem (P). Standard MIQP solvers that are based on branch-and-bound frameworks can be applied to (P) directly. However, the lower bound generated from the continuous relaxation of (P) by relaxing $y_i \in \{0, 1\}$ to $y_i \in [0, 1]$ is often quite loose. Equivalent reformulations with tighter continuous relaxation, i.e., a larger lower bound, have been proposed in the literature [21, 23, 24, 47]. These reformulations are more efficient when solved in MIQP solvers. We propose a lift-and-convexification approach to construct a tight reformulation for problem (P). This lift-and-convexification approach lifts the quadratic term involving x only in the original objective function $f(x, y)$ to a quadratic function of both x and y and convexifies this equivalent objective function in a quadratic form of (x, y) . The new reformulation retains the linearly constrained structure of the MIQP form so that its continuous relaxations can be solved efficiently. At the same time, the lower bound achieved by the continuous relaxation of this newly proposed reformulation is the same as the lower bound obtained from the state-of-the-art perspective reformulation. Thus it improves the performance of branch-and-bound algorithms by providing the same tightness at the root node as the state-of-the-art perspective reformulation and much faster processing time at children nodes. We then further combine our lift-and-convexification approach and the quadratic convex reformulation (QCR) [10, 11] approach in the literature to form an even tighter reformulation. The QCR approach has been applied to zero-one quadratic programs [11] and integer quadratic programs [9]. While the QCR approach cannot be directly applied to problem (P), it can be successfully applied on top of our new lift-and-convexification reformulation. This further reduces the duality gap as we will show in our numerical tests.

The paper is organized as follows: In §2, we review the current state-of-the-art reformulation and exact solution methods for problem (P). In §3, we propose a lift-and-convexification approach to obtain a tight reformulation. We show that this new reformulation is as tight as the state-of-the-art reformulation in terms of the lower bound from its continuous relaxation.

As the continuous relaxation of the new reformulation is a quadratic program, it can be thus solved efficiently. In §4, we conduct numerical experiments to demonstrate the effectiveness of our new lift-and-convexification reformulation. In §5, we review the QCR approach in the literature for the binary and integer quadratic programs. We then combine lift-and-convexification approach and the QCR approach to form an even tighter reformulation. We conclude our paper in §6.

Notation: Throughout this paper, we denote by $v(\cdot)$ the optimal value of problem (\cdot) , and \mathfrak{R}_+^n the nonnegative orthant of \mathfrak{R}^n . For any $a \in \mathfrak{R}^n$, we denote by $\text{diag}(a) = \text{diag}(a_1, \dots, a_n)$ the diagonal matrix with a_i being its i th diagonal element. We denote by e the all-one vector.

2 Literature review and related work

One efficient solution method for (P) is the perspective reformulation proposed by [21, 23], in which problem (P) is transformed into the following equivalent form:

$$\begin{aligned} (\text{PR}(\rho)) \quad \min \quad & f_\rho(x, y) = x^T(Q - \text{diag}(\rho))x + c^T x + h^T y + \sum_{i=1}^n [\rho_i(x_i^2/y_i)] \\ \text{s.t.} \quad & Ax + By \leq d, \\ & a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \end{aligned}$$

where $\rho \in \mathfrak{R}^n$ is chosen such that

$$\rho \geq 0 \text{ and } Q - \text{diag}(\rho) \succeq 0,$$

with an assumption $0/0 = 0$.

The perspective reformulation is very tight, i.e., the lower bound generated from the continuous relaxation of this reformulation is usually much higher than the lower bound generated directly from the continuous relaxation of (P). To deal with the fractional terms in the objective function of (PR(ρ)), two tractable reformulations of (PR(ρ)) were proposed in the literature.

The first reformulation is a second-order cone programming (SOCP) reformulation [2, 32]. For each i , introducing an additional variable $\phi_i = x_i^2/y_i$ and then rewriting the constraint $\phi_i \geq x_i^2/y_i$ as an SOCP constraint yields the following SOCP reformulation:

$$\begin{aligned} (\text{SOCP}(\rho)) \quad \min \quad & x^T(Q - \text{diag}(\rho))x + c^T x + h^T y + \rho^T \phi \\ \text{s.t.} \quad & Ax + By \leq d, \\ & a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \\ & \left\| \begin{array}{c} x_i \\ \frac{\phi_i - y_i}{2} \end{array} \right\| \leq \frac{\phi_i + y_i}{2}, \quad i = 1, \dots, n. \end{aligned}$$

However, as the problem size grows, the time needed to solve the above SOCP relaxation becomes a critical factor. When interior point methods are used, the corresponding branch-and-bound algorithm may converge very slowly.

The second reformulation is the perspective cut (PC) reformulation [21, 23]. Representing the value of x_i^2/y_i by the supremum of a set of infinitely many hyperplanes, which are called perspective cuts, gives rise to the following PC reformulation:

$$\begin{aligned}
(\text{PC}(\rho)) \quad & \min \quad x^T(Q - \text{diag}(\rho))x + c^T x + h^T y + \rho^T \phi \\
& \text{s.t.} \quad Ax + By \leq d, \\
& \quad a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \\
& \quad \phi_i \geq 2\bar{x}_i x_i - \bar{x}_i^2 y_i, \quad \forall \bar{x}_i \in [a_i, b_i], \quad i = 1, \dots, n.
\end{aligned} \tag{3}$$

The perspective cuts in (3) can be added dynamically when $(\text{PC}(\rho))$ is solved in a branch-and-cut framework (see [24]). With the help of warm start and dual methods, quadratic programming relaxations in the perspective cut algorithm can be solved efficiently.

Let $(\overline{\text{PR}}(\rho))$, $(\overline{\text{SOCP}}(\rho))$ and $(\overline{\text{PC}}(\rho))$ denote the continuous relaxations of $(\text{PR}(\rho))$, $(\text{SOCP}(\rho))$ and $(\text{PC}(\rho))$, respectively, by relaxing $y_i \in \{0, 1\}$ to $y_i \in [0, 1]$. It is easy to see that the objective values of these continuous relaxations form the same lower bound for (P). A key issue is how to choose the vector ρ such that this lower bound is as large as possible. One natural way is to set every component of ρ to be the smallest eigenvalue of Q . Frangioni and Gentile [23] proposed a better heuristic and set ρ to be the optimal solution to the following SDP problem:

$$\max\{e^T \rho \mid \rho \geq 0, Q - \text{diag}(\rho) \succeq 0\}. \tag{4}$$

Ideally, the best parameter ρ that maximizes the lower bound $v(\overline{\text{PR}}(\rho))$ can be found by solving the following problem:

$$(\text{MAX}_\rho) \quad \max\{v(\overline{\text{PR}}(\rho)) \mid \rho \geq 0, Q - \text{diag}(\rho) \succeq 0\}. \tag{5}$$

Recently, Zheng et al. [47] established the following interesting result.

Theorem 1 *Problem (MAX_ρ) is equivalent to the following semi-definite programming (SDP) problem:*

$$\begin{aligned}
(\text{SDP}_l) \quad & \max \quad \tau \\
& \text{s.t.} \quad \begin{pmatrix} \rho_i + \mu_i & \frac{1}{2}(c_i - \lambda_i - (a_i + b_i)\mu_i) \\ \frac{1}{2}(c_i - \lambda_i - (a_i + b_i)\mu_i)^T & h_i - \pi_i + (B^T \eta)_i + \mu_i b_i \end{pmatrix} \succeq 0, \\
& \quad i = 1, \dots, n, \\
& \quad \begin{pmatrix} Q - \text{diag}(\rho) & \frac{1}{2}(\lambda + A^T \eta) \\ \frac{1}{2}(\lambda + A^T \eta)^T & -\eta^T d - e^T \pi - \tau \end{pmatrix} \succeq 0, \\
& \quad (\eta, \mu, \pi, \rho) \in \mathfrak{R}_+^m \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n, \\
& \quad (\lambda, \tau) \in \mathfrak{R}^n \times \mathfrak{R}.
\end{aligned}$$

Zheng et al. [47] showed that the perspective cut approach for $(\text{PC}(\rho^*))$ with ρ^* obtained from (SDP_l) is most efficient for solving problem (P) to its optimality.

When $B \equiv 0$ in the constraint $Ax + By \leq d$, Frangioni et al. [25] developed an equivalent MIQP reformulation of $(\text{PR}(\rho))$, whose continuous relaxation becomes a quadratic programming problem. Frangioni et al. [20] also proposed an MIQP reformulation of the original problem (P). But the continuous relaxation of this MIQP reformulation is in general not as tight as that of the perspective reformulation.

3 Lift-and-convexification approach

In this section, we derive a tight

MIQP reformulation of (P) by proposing a lift-and-convexification approach. This approach lifts the quadratic term involving x only in the original objective function $f(x, y)$ to an equivalent quadratic function of both x and y and convexifies this equivalent objective function in a quadratic form of (x, y) .

Contrast to the the perspective reformulation which involves fractional terms, our new reformulation is a quadratic programming problem whose continuous relaxations can be solved efficiently. At the same time, the lower bound achieved by the continuous relaxation of this new reformulation can be proved to achieve the same lower bound obtained from $(\overline{\text{SOCP}}(\rho^*))$ or $(\overline{\text{PC}}(\rho^*))$ with ρ^* calculated from (SDP_l) . To construct the new reformulation, we only need to solve an additional SOCP problem, given the solution for (SDP_l) .

Let us determine first what kind of quadratic functions in the (x, y) -space we need to add to achieve the above mentioned goals.

Theorem 2 *Let $q(x, y)$ be a quadratic function of x and y . If $q(x, y) = 0$ for all $(x, y) \in \{(x, y) \mid a_i y_i \leq x_i \leq b_i y_i, y_i \in \{0, 1\}, i = 1, \dots, n\}$, where $a_i < b_i, i = 1, \dots, n$, then $q(x, y)$ must take the following form:*

$$q(x, y) = \sum_{i=1}^n q_i(x_i, y_i), \quad (6)$$

where

$$q_i(x_i, y_i) = u_i x_i y_i + v_i y_i^2 - u_i x_i - v_i y_i \quad (7)$$

is a quadratic function of (x_i, y_i) parameterized by (u_i, v_i) .

Proof. If $(x, y) \in \{(x, y) \mid a_i y_i \leq x_i \leq b_i y_i, y_i \in \{0, 1\}, i = 1, \dots, n\}$, then for any $i = 1, \dots, n$, if $y_i = 0$, then x_i must be 0.

Let $q(x, y)$ be of the following general form:

$$q(x, y) = (x^T, y^T) \begin{pmatrix} P & F \\ F^T & G \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + p^T x + g^T y,$$

parameterized by $(P, F, G, p, g) \in \mathbb{S}^n \times \mathfrak{R}^{n \times n} \times \mathbb{S}^n \times \mathfrak{R}^n \times \mathfrak{R}^n$. Let x^i be a vector with a non-zero component only in its i th position and y^i be a vector with a non-zero component (which is set at one) only in its i th position. It is clear that $(x^i, y^i) \in \{(x, y) \mid a_i y_i \leq x_i \leq b_i y_i, y_i \in \{0, 1\}, i = 1, \dots, n\}$. If $q(x, y) = 0$ for all $(x, y) \in \{(x, y) \mid a_i y_i \leq x_i \leq b_i y_i, y_i \in \{0, 1\}, i = 1, \dots, n\}$, then

$$q(x^i, y^i) = 0, \quad (8)$$

$$q(x^i + x^j, y^i + y^j) = 0, \quad (9)$$

for any $i, j = 1, \dots, n$. (8) implies

$$P_{ii}x_i^2 + 2F_{ii}x_i + G_{ii} + p_ix_i + g_i = 0. \quad (10)$$

Because (10) must hold for any $x_i \in [a_i, b_i]$ and $a_i < b_i$, we have

$$P_{ii} = 0, \quad (11)$$

$$2F_{ii} = -p_i, \quad (12)$$

$$G_{ii} = -g_i. \quad (13)$$

Furthermore, (9) leads to

$$\begin{aligned} P_{ii}x_i^2 + 2P_{ij}x_ix_j + P_{jj}x_j^2 + 2((F_{ii} + F_{ij})x_i + (F_{jj} + F_{ji})x_j) \\ + G_{ii} + 2G_{ij} + G_{jj} + p_ix_i + p_jx_j + g_i + g_j = 0, \end{aligned}$$

which can be simplified to the following equality by using (11)-(13),

$$P_{ij}x_ix_j + F_{ij}x_i + F_{ji}x_j + G_{ij} = 0.$$

As the above equality holds for any $(x_i, x_j) \in [a_i, b_i] \times [a_j, b_j]$ with $a_i < b_i$ and $a_j < b_j$, we must have $0 = P_{ij} = F_{ij} = F_{ji} = G_{ij}$. Combining the above equality with (11)-(13) yields the following form of $q(x, y)$,

$$q(x, y) = \sum_{i=1}^n 2F_{ii}x_iy_i + G_{ii}y_i^2 - 2F_{ii}x_i - G_{ii}y_i,$$

which is of the same form as (7). □

We propose the following reformulation of (P):

$$\begin{aligned} (\text{P}(u, v)) \quad \min \quad & f_{u,v}(x, y) = f(x, y) + \sum_{i=1}^n q_i(x_i, y_i) \\ \text{s.t.} \quad & Ax + By \leq d, \\ & a_iy_i \leq x_i \leq b_iy_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n, \end{aligned}$$

where $q_i(x_i, y_i)$ is defined in (7). It is easy to see that problem $(\text{P}(u, v))$ is equivalent to (P) and the continuous relaxation of $(\text{P}(u, v))$ is a quadratic program.

The difference among equivalent formulations (P), $(\text{PR}(\rho))$, and $(\text{P}(u, v))$ lies in their objective functions. The following example shows the relative relationship among these three objective functions.

Example 1 Consider a univariate function $f(x) = x^2 - 4x$, where $(x, y) \in \Omega = \{(x, y) \mid y \leq x \leq 3y, y \in [0, 1]\}$. Let $q(x, y) = -xy + y^2 + x - y$. Then, $q(x, y)$ is zero at the region $\{(x, y) \mid y \leq x \leq 3y, y \in \{0, 1\}\}$. Figure 1 illustrates the original function $f(x)$, the lifted quadratic function $f_{u,v}(x, y) := f(x) + q(x, y)$ and the perspective function $f_p(x, y) = \frac{x^2}{y} - 4x$. While the three functions has the same value when $y \in \{0, 1\}$, we can see that $f_{u,v}(x, y)$ always lies between $f(x)$ and $f_p(x, y)$ in the region Ω . This can be numerically verified since $f_{u,v}(x, y) - f(x) = (x - y)(1 - y) \geq 0$ for all $(x, y) \in \{(x, y) \mid y \leq x \leq 3y, y \in [0, 1]\}$. We can also verify that $f_p(x, y) - f_{u,v}(x, y) = (1 - y)y(x^2 + y^2 - xy) \geq 0$ for all $(x, y) \in \{(x, y) \mid y \leq x \leq 3y, y \in (0, 1]\}$.

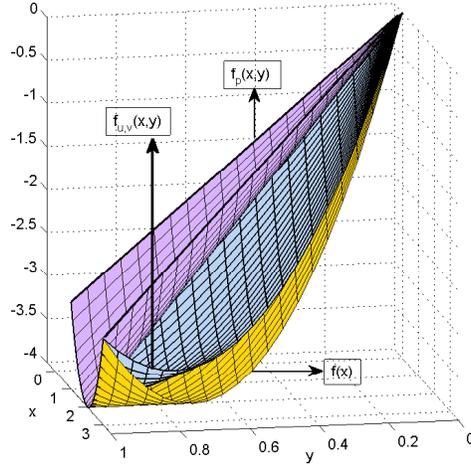


Figure 1: Illustration of the lifting function

Example 1 shows that the objective function of our new reformulation could lie below that of the perspective reformulation. This indicates that our new reformulation may not attain a lower bound that is tighter than the perspective reformulation. However, we will show that our new reformulation can achieve the same lower bound as the perspective reformulation.

Let $u = (u_1, \dots, u_n)^T$ and $v = (v_1, \dots, v_n)^T$. Now, one critical question is “What is the best parameter vector of (u, v) ?” Let $(\bar{P}(u, v))$ denote the continuous relaxation of $(P(u, v))$ by relaxing $y_i \in \{0, 1\}$ to $y_i \in [0, 1]$. It is desirable to choose (u, v) such that the continuous relaxation of $(P(u, v))$ is as tight as possible. This clear goal motivates us to consider the following problem:

$$(\text{MAX}_{uv}) \quad \max\{v(\bar{P}(u, v)) \mid u, v \in \mathfrak{R}^n, f_{u,v}(x, y) \text{ is convex}\}. \quad (14)$$

Theorem 3 *Problem (MAX_{uv}) is equivalent to the following SDP problem:*

$$(\text{SDP}_q) \quad \max \quad \tau$$

$$\text{s.t.} \quad \begin{pmatrix} Q & \frac{1}{2}\text{diag}(u) & \frac{1}{2}\alpha(u, \eta, \mu, \sigma) \\ \frac{1}{2}\text{diag}(u) & \text{diag}(v) & \frac{1}{2}\beta(v, \eta, \mu, \sigma, \lambda, \pi) \\ \frac{1}{2}\alpha(u, \eta, \mu, \sigma)^T & \frac{1}{2}\beta(v, \eta, \mu, \sigma, \lambda, \pi)^T & -\eta^T d - e^T \pi - \tau \end{pmatrix} \succeq 0, \quad (15)$$

$$(\eta, \mu, \sigma, \lambda, \pi) \in \mathfrak{R}_+^m \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n, \quad (16)$$

$$(u, v, \tau) \in \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R},$$

where

$$\alpha(u, \eta, \mu, \sigma) = c - u + A^T \eta - \mu + \sigma, \quad (17)$$

$$\beta(v, \eta, \mu, \sigma, \lambda, \pi) = h - v + B^T \eta + \text{diag}(a)\mu - \text{diag}(b)\sigma - \lambda + \pi. \quad (18)$$

Proof. We first express $(\bar{P}(u, v))$ by its dual form. Associate the following multipliers to the constraints in $(\bar{P}(u, v))$:

- $\eta \in \mathfrak{R}_+^m$ for $Ax + By \leq d$;
- μ_i and $\sigma_i \in \mathfrak{R}_+$ for $a_i y_i \leq x_i$ and $x_i \leq b_i y_i$, respectively, $i = 1, \dots, n$; and
- λ_i and $\pi_i \in \mathfrak{R}_+$ for $y_i \geq 0$ and $y_i \leq 1$, respectively, $i = 1, \dots, n$.

Let $\mu = (\mu_1, \dots, \mu_n)^T$, $\sigma = (\sigma_1, \dots, \sigma_n)^T$, $\lambda = (\lambda_1, \dots, \lambda_n)^T$ and $\pi = (\pi_1, \dots, \pi_n)^T$. Let ω denote the vector formed by $\eta, \mu, \sigma, \lambda$ and π . The Lagrangian function of $(\bar{\mathbf{P}}(u, v))$ is then given by

$$\begin{aligned}
L(x, y, \omega) &= x^T Q x + c^T x + h^T y + \sum_{i=1}^n (u_i x_i y_i + v_i y_i^2 - u_i x_i - v_i y_i) \\
&\quad + \eta^T (Ax + By - d) + \sum_{i=1}^n \mu_i (a_i y_i - x_i) + \sum_{i=1}^n \sigma_i (x_i - b_i y_i) \\
&\quad + \sum_{i=1}^n \lambda_i (-y_i) + \sum_{i=1}^n \pi_i (y_i - 1) \\
&= (x^T, y^T) \begin{pmatrix} Q & \frac{1}{2} \text{diag}(u) \\ \frac{1}{2} \text{diag}(u) & \text{diag}(v) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + (c - u + A^T \eta - \mu + \sigma)^T x \\
&\quad + (h - v + B^T \eta + \text{diag}(a)\mu - \text{diag}(b)\sigma - \lambda + \pi)^T y - \eta^T d - e^T \pi.
\end{aligned}$$

Furthermore, the Lagrangian dual problem of $(\bar{\mathbf{P}}(u, v))$ can be expressed as

$$\max \left\{ \min_{(x, y) \in \mathfrak{R}^n \times \mathfrak{R}^n} L(x, y, \omega) \mid \omega \geq 0 \right\}. \quad (19)$$

Introducing an additional variable τ , we can rewrite (19) as

$$\max \tau \quad (20)$$

$$\text{s.t.} \quad \min_{(x, y) \in \mathfrak{R}^n \times \mathfrak{R}^n} L(x, y, \omega) \geq \tau, \quad (21)$$

$$\omega \geq 0. \quad (22)$$

We see that the constraint in (21) is equivalent to $L(x, y, \omega) - \tau \geq 0$ for all x, y , which is further equivalent to

$$L(x/t, y/t, \omega) - \tau \geq 0, \quad \forall (x, y) \in \mathfrak{R}^n \times \mathfrak{R}^n, \forall t \in \mathfrak{R}, t \neq 0. \quad (23)$$

Multiplying both sides of (23) by t^2 yields a homogeneous quadratic form of (x, y, t) in the left-hand side of (23). Thus, the constraint in (21) is equivalent to the following semidefinite constraint:

$$\begin{pmatrix} Q & \frac{1}{2} \text{diag}(u) & \frac{1}{2} \alpha(u, \eta, \mu, \sigma) \\ \frac{1}{2} \text{diag}(u) & \text{diag}(v) & \frac{1}{2} \beta(v, \eta, \mu, \sigma, \lambda, \pi) \\ \frac{1}{2} \alpha(u, \eta, \mu, \sigma)^T & \frac{1}{2} \beta(v, \eta, \mu, \sigma, \lambda, \pi)^T & -\eta^T d - e^T \pi - \tau \end{pmatrix} \succeq 0, \quad (24)$$

where $\alpha(u, \eta, \mu, \sigma)$ and $\beta(v, \eta, \mu, \sigma, \lambda, \pi)$ are defined by (17) and (18), respectively. Consequently, the problem in (20)-(22) can be expressed as

$$\max \{ \tau \mid \text{constraint (24)}, \omega \geq 0 \}. \quad (25)$$

If $f_{u,v}(x, y)$ is convex, by the strong duality of convex quadratic programming (see, e.g., Proposition 6.5.6 in [5]), the optimal values of $(\overline{P}(u, v))$ and (25) are equal. Thus, we have shown that problem $(\text{MAX}uv)$ is equivalent to an SDP problem in the form of (SDP_q) . \square

Let (u^*, v^*) be the optimal parameters for our new reformulation $(P(u, v))$ by solving the SDP program (SDP_q) . Let ρ^* be the optimal parameters for $(\text{PR}(\rho))$ by solving the SDP program (SDP_l) . It is necessary to compare the tightness of $v(\overline{P}(u^*, v^*))$ and $v(\overline{\text{PR}}(\rho^*))$, i.e., the bounds from the continuous relaxation of the “best” reformulation proposed in this paper and the “best” perspective reformulation. We will show in the following that these two bounds are the same using constructive proofs.

Theorem 4 Define $\overline{\rho} \in \Re^n$ as

$$\overline{\rho}_i = \begin{cases} 0 & \text{if } v_i^* = 0, \\ \frac{u_i^{*2}}{4v_i^*} & \text{otherwise,} \end{cases} \quad i = 1, \dots, n. \quad (26)$$

Then,

- (a) $\overline{\rho}$ is feasible for problem $(\text{MAX}\rho)$,
- (b) $v(\overline{P}(u^*, v^*)) \leq v(\overline{\text{PR}}(\overline{\rho}))$.

Proof. (a) From the proof of Theorem 3, we know that there exists $(\tau^*, \eta^*, \mu^*, \sigma^*, \lambda^*, \pi^*)$ such that $(u^*, v^*, \tau^*, \eta^*, \mu^*, \sigma^*, \lambda^*, \pi^*)$ is optimal to (SDP_q) . The constraint in (15) implies that

$$\begin{pmatrix} Q & \frac{1}{2}\text{diag}(u^*) \\ \frac{1}{2}\text{diag}(u^*) & \text{diag}(v^*) \end{pmatrix} \succeq 0.$$

For any $g \in \Re^n$, we define $p \in \Re^n$ with

$$p_i = \begin{cases} 0 & \text{if } v_i^* = 0, \\ -\frac{u_i^*}{2v_i^*}g_i & \text{otherwise,} \end{cases} \quad i = 1, \dots, n.$$

We then have

$$\begin{aligned} g^T(Q - \overline{\rho})g &= g^T Q g + \sum_{i=1}^n [g_i^2(-\overline{\rho}_i)] \\ &= g^T Q g + \sum_{i=1}^n [u_i^* g_i p_i + v_i^* p_i^2] \\ &= g^T Q g + g^T \text{diag}(u^*) p + p^T \text{diag}(v^*) p \\ &= (g^T, p^T) \begin{pmatrix} Q & \frac{1}{2}\text{diag}(u^*) \\ \frac{1}{2}\text{diag}(u^*) & \text{diag}(v^*) \end{pmatrix} \begin{pmatrix} g \\ p \end{pmatrix} \\ &\geq 0. \end{aligned}$$

Hence $(Q - \bar{\rho}) \succeq 0$ and $\bar{\rho}$ is feasible to problem (MAX ρ).

(b) As $(\overline{\text{PR}}(\bar{\rho}))$ and $(\overline{\text{P}}(u^*, v^*))$ have the same feasible region, for any feasible (x, y) , we can compare their objective values as follows,

$$\begin{aligned}
& f_\rho(x, y; \bar{\rho}) - f_{u,v}(x, y; u^*, v^*) \tag{27} \\
&= \sum_{i=1}^n [\bar{\rho}_i \left(\frac{x_i^2}{y_i}\right) - \bar{\rho}_i x_i^2 - (u_i^* x_i y_i + v_i^* y_i^2 - u_i^* x_i - v_i^* y_i)] \\
&= \sum_{i=1}^n \left[\frac{1}{4v_i^*} (u_i^{*2} x_i^2 \frac{1-y_i}{y_i} + 4u_i^* v_i^* x_i (1-y_i) + 4v_i^{*2} y_i (1-y_i)) \right] \\
&= \sum_{i=1}^n \left[\frac{y_i(1-y_i)}{4v_i^*} \left(\left(\frac{u_i^* x_i}{y_i} \right)^2 + 4u_i^* v_i^* \frac{x_i}{y_i} + 4v_i^{*2} \right) \right] \\
&= \sum_{i=1}^n \left[\frac{y_i(1-y_i)}{4v_i^*} \left(u_i^* \frac{x_i}{y_i} + 2v_i^* \right)^2 \right] \\
&\geq 0.
\end{aligned}$$

The above deduction is valid because if $v_i^* = 0$, then $u_i^* = 0$ due to (15). Thus $v(\overline{\text{P}}(u^*, v^*)) \leq v(\overline{\text{PR}}(\bar{\rho}))$. \square

The following corollary is a direct result of Theorem 4.

Corollary 1 $v(\overline{\text{P}}(u^*, v^*)) \leq v(\overline{\text{PR}}(\rho^*))$.

Next we show the other way around.

Theorem 5 Suppose that (x^*, y^*) is optimal to problem $(\overline{\text{PR}}(\rho^*))$, the continuous relaxation of $(\text{PR}(\rho^*))$. Define $\bar{u}, \bar{v} \in \Re^n$ with

$$\bar{u}_i = \begin{cases} 0 & \text{if } y_i^* = 0, \\ -2\rho_i^* \frac{x_i^*}{y_i^*} & \text{otherwise,} \end{cases} \quad i = 1, \dots, n, \tag{28}$$

$$\bar{v}_i = \begin{cases} 0 & \text{if } y_i^* = 0, \\ \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} & \text{otherwise,} \end{cases} \quad i = 1, \dots, n. \tag{29}$$

Then,

(a) (\bar{u}, \bar{v}) is feasible to problem (MAX u),

(b) $v(\overline{\text{P}}(\bar{u}, \bar{v})) = v(\overline{\text{PR}}(\rho^*))$.

Proof. (a) For any $g, p \in \Re^n$, we have

$$\begin{aligned}
& (g^T, p^T) \begin{pmatrix} Q & \frac{1}{2}\text{diag}(\bar{u}) \\ \frac{1}{2}\text{diag}(\bar{u}) & \text{diag}(\bar{v}) \end{pmatrix} \begin{pmatrix} g \\ p \end{pmatrix} \\
&= g^T Q g + g^T \text{diag}(\bar{u}) p + p^T \text{diag}(\bar{v}) p \\
&= g^T Q g + \sum_{i=1}^n [\bar{u}_i g_i p_i + \bar{v}_i p_i^2] \\
&= g^T (Q - \rho^*) g + \sum_{i=1}^n [\rho_i^* g_i^2 - 2\rho_i^* \frac{x_i^*}{y_i^*} g_i p_i + \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} p_i^2] \\
&= g^T (Q - \rho^*) g + \sum_{i=1}^n [\rho_i^* (g_i - \frac{x_i^*}{y_i^*} p_i)^2] \\
&\geq 0.
\end{aligned}$$

Hence (\bar{u}, \bar{v}) is feasible to problem (MAX_{uv}) .

(b) We first show that (x^*, y^*) is also an optimal solution for $(\bar{\text{P}}(\bar{u}, \bar{v}))$, we compare the gradients of $f_\rho(x, y; \rho^*)$ and $f_{u,v}(x, y; \bar{u}, \bar{v})$ at the point (x^*, y^*) . For $i = 1, \dots, n$, we have

$$\begin{aligned}
& (\nabla f_{u,v}(x^*, y^*; \bar{u}, \bar{v}))_i \\
&= 2Q_i^T x^* + c_i + \bar{u}_i y_i^* - \bar{u}_i \\
&= 2Q_i^T x^* + c_i - 2\rho_i^* \frac{x_i^*}{y_i^*} * y_i^* + 2\rho_i^* \frac{x_i^*}{y_i^*} \\
&= 2Q_i^T x^* + c_i - 2\rho_i^* x_i^* + 2\rho_i^* \frac{x_i^*}{y_i^*} \\
&= (\nabla f_\rho(x^*, y^*; \rho^*))_i
\end{aligned}$$

and

$$\begin{aligned}
& (\nabla f_{u,v}(x^*, y^*; \bar{u}, \bar{v}))_{n+i} \\
&= h_i + \bar{u}_i x_i^* + 2\bar{v}_i y_i^* - \bar{v}_i \\
&= h_i - 2\rho_i^* \frac{x_i^*}{y_i^*} * x_i^* + 2\rho_i^* \frac{x_i^{*2}}{y_i^{*2}} y_i^* - \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} \\
&= h_i - \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} \\
&= (\nabla f_\rho(x^*, y^*; \rho^*))_{n+i}.
\end{aligned}$$

So $\nabla f_{u,v}(x^*, y^*; \bar{u}, \bar{v}) = \nabla f_\rho(x^*, y^*; \rho^*)$. As (x^*, y^*) is assumed to be optimal to problem $(\bar{\text{PR}}(\rho^*))$, the directional derivative of $f_\rho(x, y; \rho^*)$ at (x^*, y^*) along any feasible direction should be non-negative. Since the feasible regions of $\bar{\text{P}}(\bar{u}, \bar{v})$ and $\bar{\text{PR}}(\rho^*)$ are the same, the directional derivative of $f_{u,v}(x, y; \bar{u}, \bar{v})$ at (x^*, y^*) along any feasible direction is also non-negative. So (x^*, y^*) must also be optimal for $(\bar{\text{P}}(\bar{u}, \bar{v}))$ because of the convexity of $f_{u,v}(x, y; \bar{u}, \bar{v})$. (See e.g., Chapter 2.1 of [14].)

Finally, we show $v(\overline{\text{P}}(\overline{u}, \overline{v})) = v(\overline{\text{PR}}(\rho^*))$. Similar to (27), we have

$$\begin{aligned}
& f_\rho(x^*, y^*; \rho^*) - f_{u,v}(x^*, y^*; \overline{u}, \overline{v}) \\
&= \sum_{i=1}^n \left[\rho_i^* \left(\frac{x_i^{*2}}{y_i^*} \right) - \rho_i^* x_i^{*2} - (\overline{u}_i x_i^* y_i^* + \overline{v}_i y_i^{*2} - \overline{u}_i x_i^* - \overline{v}_i y_i^*) \right] \\
&= \sum_{i=1}^n \left[\rho_i^* \left(\frac{x_i^{*2}}{y_i^*} \right) - \rho_i^* x_i^{*2} - \left(-2\rho_i^* \frac{x_i^*}{y_i^*} x_i^* y_i^* + \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} y_i^{*2} + 2\rho_i^* \frac{x_i^*}{y_i^*} x_i^* - \rho_i^* \frac{x_i^{*2}}{y_i^{*2}} y_i^* \right) \right] \\
&= \sum_{i=1}^n \left[\rho_i^* \left(\frac{x_i^{*2}}{y_i^*} \right) - \rho_i^* x_i^{*2} - \left(-2\rho_i^* x_i^{*2} + \rho_i^* x_i^{*2} + 2\rho_i^* \frac{x_i^{*2}}{y_i^*} - \rho_i^* \frac{x_i^{*2}}{y_i^*} \right) \right] \\
&= 0.
\end{aligned}$$

This completes the proof. □

The following corollary is a direct result of Theorem 5.

Corollary 2 $v(\overline{\text{P}}(u^*, v^*)) \geq v(\overline{\text{PR}}(\rho^*))$.

Combining Corollaries 1 and 2 yields the following result.

Theorem 6 $v(\overline{\text{P}}(u^*, v^*)) = v(\overline{\text{PR}}(\rho^*))$.

Thus the bound from our new reformulation is as good as the bound from the perspective reformulation. However, to find (u^*, v^*) , we need to solve (SDP_q) which is an SDP program that has a much larger size than (SDP_l) . Our numerical tests show that (SDP_q) could consume ten times of the computation time of (SDP_l) . Fortunately, based on Theorems 4 and 5, the following corollary becomes evident which reveals the nonnecessity in using (SDP_q) in the calculation.

Corollary 3 (a) Define $\overline{\rho}$ as in (26). Then $\overline{\rho}$ is optimal for problem $(\text{MAX}\rho)$,

(b) Define $(\overline{u}, \overline{v})$ as in (28) and (29). Then $(\overline{u}, \overline{v})$ is optimal to problem $(\text{MAX}uv)$.

Thus, in order to construct our new reformulation $(\text{P}(u^*, v^*))$, we only need to solve (SDP_l) first to get the optimal solution ρ^* for $(\text{MAX}\rho)$ and then solve the SOCP problem $(\overline{\text{SOCP}}(\rho^*))$ and construct (u^*, v^*) according to (28) and (29). Note that $(\overline{\text{PR}}(\rho^*))$ is equivalent to $(\overline{\text{SOCP}}(\rho^*))$.

4 Computational results

In this section, we conduct computational experiments to compare the performance of the perspective cut reformulation $(\text{PC}(\rho))$ and our new reformulation $(\text{P}(u, v))$. To be specific, we compare the performance of standard MIQP solvers between solving the following two reformulations of problem (P):

- (PC): the perspective reformulation (PC(ρ)) with $\rho = \rho^*$, where ρ^* is computed by solving (SDP_l).
- (LCR): our lift-and-convexification reformulation (P(u, v)) with $(u, v) = (u^*, v^*)$, where (u^*, v^*) is obtained by first solving $\overline{\text{SOCP}}(\rho^*)$ and then configuring (u^*, v^*) according to (28) and (29) in Theorem 5.

Although the continuous relaxation of (LCR) is as tight as that of (PC) at the root node of the branch-and-bound tree, the relaxations in (LCR) are in general looser than those in (PC) at children nodes. The advantage of (LCR) is that its continuous relaxations are quadratic programs and thus can be solved much faster than the continuous relaxations of (PC). We would like to test if this advantage of (LCR) would dominate (at least verifying itself as a competitive and useful reformulation).

The time difference between finding ρ^* and (u^*, v^*) is the time needed to solve one SOCP programming problem ($\overline{\text{SOCP}}(\rho^*)$). We will count this amount of time into the computational time for (LCR) in the comparison, although this amount of time is quite small in general.

The two reformulations and $\overline{\text{SOCP}}(\rho^*)$ are all solved in 64-bit IBM ILOG CPLEX Optimization Studio 12.3 (Hereinafter referred to as CPLEX) through its C interface. The perspective cut reformulation (PC) is implemented by means of `user cut callbacks` and `lazy constraint callbacks` in CPLEX. Although [21] suggested to apply the separation procedure only once at each node, we do not limit the times of separation because we find that in our numerical tests, if we allow CPLEX to actively generate mixed integer cuts, the computation would be much faster if the times of separation are unlimited at each node. (SDP_l) is solved using `sedumi` interfaced by `CVX 1.21` ([18, 31]) on `Matlab R2012b`.

All the computation is conducted on a Linux machine (64-bit CentOS Release 5.5) with 48 GB of RAM. All the tests are confined on one single thread (2.99 GHz).

We consider two types of test problems from the cardinality constrained mean-variance portfolio selection (MV) and the subset selection problem (SSP) in our computational experiments.

4.1 Cardinality constrained portfolio selection problem

In this subsection, we compare (PC) and (LCR) for the cardinality constrained mean-variance portfolio selection problem (MV) introduced in the introduction section.

Frangioni and Gentile [23] tested 90 instances of (MV) in their paper, 30 instances each for $n = 200, 300$ and 400 . The 30 instances for each n are divided further into three subsets denoted by n^- , n^0 and n^+ , 10 in each subset, with different diagonal dominance in the matrix Q . We use these 90 instances created in [23] in our test. While Frangioni and Gentile [23] did not consider the cardinality constraint in their models, we add the cardinality constraint to these instances in our numerical experiments. Testing each instance without the cardinality constraint and with $K = 6, 8, 10$, and 12 , we have 450 instances of (MV). The data files of these instances can be downloaded at: <http://www.di.unipi.it/optimize/Data/MV.html>.

Table 1 summarizes the numerical results for the 450 instances of (MV) when the time limit is set at 10000 seconds. Each line reports the average results for the 10 instances in a subset. The notations in the table are given as follows: The column “(time_l)” is the

Table 1: Numerical results of reformulations for (MV)

(MV)	K	(time _{l})	(time _{s})	(PC)		(LCR)	
				time	nodes	time	nodes
200 ⁺	6	27.92	2.30	19.68	65	4.10	26
	8	27.02	2.13	11.20	55	2.29	19
	10	26.19	2.06	6.18	50	2.86	42
	12	26.29	2.05	10.40	111	4.28	95
	nonK	32.60	2.17	13.31	148	8.66	147
200 ⁰	6	26.66	2.54	18.03	86	7.07	65
	8	27.20	2.05	16.02	96	7.59	73
	10	23.83	2.21	9.38	126	4.16	120
	12	24.88	2.10	30.73	256	24.44	217
	nonK	30.64	2.00	33.76	291	27.42	281
200 ⁻	6	26.00	2.39	26.16	204	18.49	248
	8	25.53	2.16	22.49	235	19.96	287
	10	24.67	2.03	19.88	328	11.00	350
	12	25.23	2.11	287.60	3306	329.02	1964
	nonK	28.47	1.85	152.45	1935	219.22	1380
300 ⁺	6	62.05	5.91	82.11	127	14.25	26
	8	65.26	5.54	48.48	133	8.43	25
	10	55.55	5.47	16.85	76	4.20	30
	12	60.31	6.89	18.43	128	10.27	119
	nonK	87.59	16.23	108.44	446	26.60	241
300 ⁰	6	60.45	5.43	43.53	118	28.42	105
	8	54.63	5.14	51.06	190	32.20	123
	10	57.15	5.04	22.00	148	11.96	129
	12	61.08	5.14	75.97	238	80.78	249
	nonK	63.26	4.80	101.63	371	95.79	323
300 ⁻	6	63.27	5.10	55.05	236	48.85	237
	8	62.77	5.72	99.03	399	62.47	328
	10	60.52	6.40	47.20	471	40.47	609
	12	62.08	4.92	35.73	312	48.94	341
	nonK	68.22	4.56	137.15	493	117.96	506
400 ⁺	6	126.97	22.32	70.89	186	22.73	39
	8	122.96	14.19	196.76	587	22.25	38
	10	122.70	12.02	28.43	95	14.23	47
	12	139.76	27.78	38.70	181	14.37	105
	nonK	100.40	25.47	562.95	849	364.24	613
400 ⁰	6	104.11	11.74	105.06	236	88.21	197
	8	124.41	10.54	149.88	435	82.51	170
	10	119.51	10.47	54.55	276	49.06	287
	12	128.22	10.73	71.52	376	40.67	287
	nonK	125.80	13.20	542.91	1132	227.31	846
400 ⁻	6	104.30	11.77	115.01	393	216.02	566
	8	116.33	10.08	341.40	1053	239.26	599
	10	110.30	10.08	83.77	515	95.99	564
	12	118.64	10.60	149.93	495	29.41	343
	nonK	123.05	10.14	750.27	1448	715.51	1373

computation time for solving (SDP_l) and the column “(time_s)” is the computation time for solving $(SOCP(\rho^*))$. The termination threshold of the relative gap (in percentage) between the objective value of the incumbent solution and the best lower bound is set to be 0.01%. (The exact value of the relative gap when CPLEX terminates could range between 0.00% and 0.01%. Rounding this number would make it 0.00% or 0.01%). Because all our instances terminated with a relative gap smaller than 0.01%, the relative gap is not reported here. The columns “time” and “nodes” are the computing time (in seconds) and the number of nodes explored by CPLEX respectively. The “nonK” refers to the instances with no cardinality constraints.

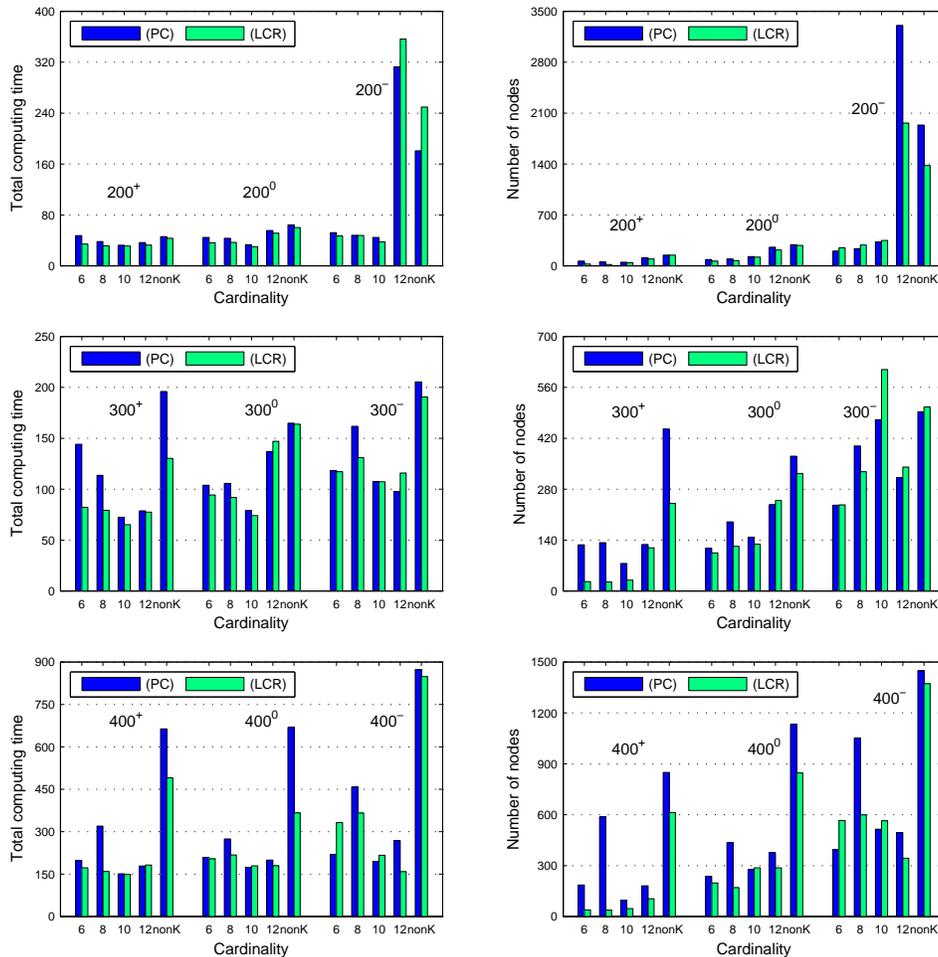


Figure 2: Comparison of total computing time and nodes for (MV).

Figure 2 displays the total computing time and nodes of the two reformulations for (MV). The total computing time for (PC) is the sum of “(time_l)” and the “time” for (PC), and the total computing time for (LCR) is the sum of “(time_l)”, “(time_s)” and the “time” for (LCR).

From Figure 2, we can see that, in terms of the total computing time, (LCR) performs better than (PC) for 25 out of the total 45 cases. If we omit the time for solving the SDP and

Table 2: Numerical results of reformulations for (SSP)

n	K	(time _l)	(time _s)	(PC)		(LCR)	
				time	nodes	time	nodes
50	5	3.86	0.14	2.90	61	0.35	115
	10	3.58	0.21	9.17	205	1.29	445
	15	3.24	0.20	6.82	119	0.71	218
	20	3.44	0.25	9.20	136	0.77	276
100	5	14.39	0.48	8.69	99	3.35	295
	10	14.90	0.37	34.65	334	14.11	993
	15	14.95	0.49	89.46	739	41.38	3767
	20	13.36	0.38	613.27	3737	140.76	11789

SOC_P and only compare the time of the MIQP solver CPLEX, (LCR) performs better than (PC) for 41 out of the total 45 cases. As the perspective cut approach represents the state-of-the-art, the test result for this (MV) data set confirms that using (LCR) reformulation to solve (P) is also efficient. We need to emphasize that, we have tried our best in our numerical tests to optimize the implementation of the perspective cut approach, as the efficiency of the perspective cut approach depends heavily on its implementation details and also on selected parameters of the MIQP solvers.

For the number of nodes explored in (LCR) and (PC), they closely match each other. We might think that because the relaxations in (LCR) are in general looser than the ones in (PC) at children nodes, the number of nodes explored by (LCR) should be larger than (PC) all the time. However, this might not always be the case because the branching schemes, feasible solution heuristics and the branch-and-bound tree inside CPLEX could be quite different for (PC) and (LCR) and the number of nodes explored could demonstrate a more random pattern.

4.2 Subset selection problem

In this subsection, we compare (PC) and (LCR) for the subset selection problem (SSP) introduced in the introduction section.

We use the 40 instances of the subset selection problem from [47] with $n = 50, 100$ and $K = 5, 10, 15, 20$, 5 instances for each (n, K) pair. In those instances, we set $m = 2n$. The elements of A are generated from the standard normal distribution $N(0, 1)$ and $b = A\beta + \epsilon$ where the elements of ϵ are generated from the standard normal distribution $N(0, 1)$ and the elements of β are generated uniformly from $[-1, 1]$. The lower and upper bounds for the solution x_i , $i = 1, \dots, n$, are set, respectively, at -100 and 100 , which are sufficiently large for those instances.

Table 2 summarizes the numerical results for the 40 instances of (SSP). Each line reports the average results for 5 instances for each (n, K) pair. The notations in the table are the same as those in Table 1.

Figure 3 displays the total computing time and nodes of the two reformulations for (SSP). The total computing time for (PC) is the sum of “(time_l)” and the “time” for (PC), and the total computing time for (LCR) is the sum of “(time_l)”, “(time_s)” and the “time” for

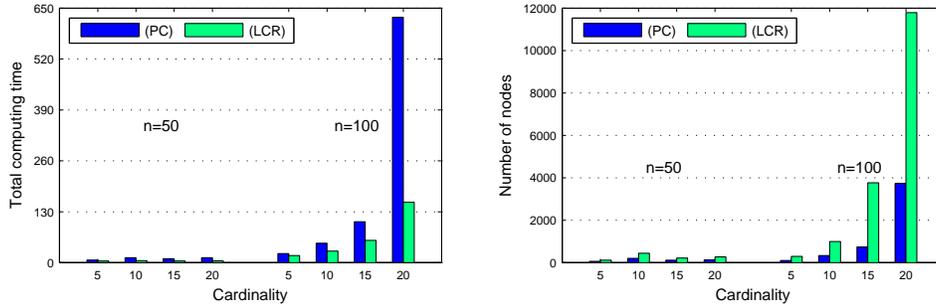


Figure 3: Comparison of total computing time and nodes for (SSP).

(LCR).

From Figure 3, we can see that although (LCR) explores more nodes than (PC), the total computing time for (LCR) is smaller than that of (PC) in all cases and this advantage of (LCR) over (PC) becomes more apparent as the problem size grows and/or as the cardinality increases. Here for the (SSP) data set, trading off tighter children-node bounds with faster processing time indeed has a good payoff. We also remark again that the efficiency of the perspective cut approach depends heavily on its implementation details and also on the selected parameters of the MIQP solvers. We, however, believe that our new reformulation derived from the lift-and-convexification approach provides a good supplement to the current state-of-the-art approaches.

5 Combination of lift-and-convexification and QCR

In this section, we further combine our lift-and-convexification approach and the QCR approach to derive an even tighter reformulation for (P).

Hammer and Rubin [33] pioneered the QCR approach in the following binary quadratic programs:

$$\begin{aligned}
 \text{(BQP)} \quad & \min x^T Q x + c^T x \\
 & \text{s.t. } Ax = d, \\
 & x_i \in \{0, 1\}, i = 1, \dots, n,
 \end{aligned}$$

where Q is indefinite. In the proposed QCR, they added to the objective function a term $\sum_i u(x_i^2 - x_i)$, where u is a scalar and is chosen to be the negative value of the smallest eigenvalue of Q . Billionnet and Elloumi [8] improved this method by adding the term $\sum_i u_i(x_i^2 - x_i)$ with u_i being the optimal dual variables of a certain semi-definite program (SDP). Plateau[42] and Billionnet et al.[10, 11] also utilized the equality $Ax = d$ in QCR and added the term $\sum_i u_i(x_i^2 - x_i) + (Ax - d)^T \text{diag}(w)(Ax - d)$ to the objective, where u and w are chosen to be the dual variables of an enlarged SDP program. Ahlatçioğlu et al. [1] proposed to combine QCR and the convex hull relaxation to solve problem (BQP). The geometric investigation in Li et al. [36] for binary quadratic programs provides some

theoretical support for QCR from another angle. Billionnet et al.[9] extended the QCR approach to general mixed-integer programs by using binary decomposition.

To make our discussion more general, we add equality constraints to (P) and consider the following variant of (P):

$$\begin{aligned}
(P') \quad & \min f(x, y) := x^T Q x + c^T x + h^T y \\
& \text{s.t. } Ax + By \leq d, \\
& \quad Ex + Fy = g, \\
& \quad a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n,
\end{aligned}$$

where $g \in \mathfrak{R}^M$, and $E, F \in \mathfrak{R}^{M \times n}$.

QCR would become beneficial when being applied to the equality and inequality constraints in (P') on a top of our lift-and-convexification reformulation. Let us consider now the following equivalent reformulation of (P'):

$$\begin{aligned}
(P(u, v, w, t)) \quad & \min f(x, y) + \sum_{i=1}^n q_i(x_i, y_i) + (Ex + Fy - g)^T \text{diag}(w)(Ex + Fy - g) \\
& \quad + (Ax + By + s - d)^T \text{diag}(t)(Ax + By + s - d) \\
& \text{s.t. } Ax + By + s = d, \quad s \geq 0, \\
& \quad Ex + Fy = g, \\
& \quad a_i y_i \leq x_i \leq b_i y_i, \quad y_i \in \{0, 1\}, \quad i = 1, \dots, n,
\end{aligned}$$

where $q_i(x_i, y_i)$ is defined in (7).

The best parameter set (u, v, w, t) can be found by solving the following problem:

$$\begin{aligned}
(\text{MAX}_{uvw t}) \\
\max \{v(\bar{P}(u, v, w, t)) \mid \text{The objective function of } (P(u, v, w, t)) \text{ is convex}\},
\end{aligned}$$

where $(\bar{P}(u, v, w, t))$ is the continuous relaxation of $(P(u, v, w, t))$.

Using the same technique in the proof for Theorem 3, we can convert the problem $(\text{MAX}_{uvw t})$ to an SDP problem.

Theorem 7 *The problem $(\text{MAX}_{uvw t})$ is equivalent to the following SDP problem:*

$$\begin{aligned}
(\text{SDP}_a) \quad & \max \tau \\
& \text{s.t. } \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{12}^T & P_{22} & P_{23} & P_{24} \\ P_{13}^T & P_{23}^T & P_{33} & P_{34} \\ P_{14}^T & P_{24}^T & P_{34}^T & P_{44} \end{pmatrix} \succeq 0, \\
& (\delta, \mu, \sigma, \lambda, \pi) \in \mathfrak{R}_+^m \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n \times \mathfrak{R}_+^n, \\
& (\eta, \zeta, u, v, \tau) \in \mathfrak{R}^m \times \mathfrak{R}^M \times \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R},
\end{aligned} \tag{30}$$

where

$$\begin{aligned}
P_{11} &= Q + E^T \text{diag}(w)E + A^T \text{diag}(t)A \\
P_{12} &= \frac{1}{2} \text{diag}(u) + E^T \text{diag}(w)F + A^T \text{diag}(t)B \\
P_{13} &= A^T \text{diag}(t) \\
P_{14} &= \frac{1}{2}(c - u + A^T \eta + E^T \zeta - 2E^T \text{diag}(w)g + 2A^T \text{diag}(t)(-d) - \mu + \sigma) \\
P_{22} &= \text{diag}(v) + F^T \text{diag}(w)F + B^T \text{diag}(t)B \\
P_{23} &= B^T \text{diag}(t) \\
P_{24} &= \frac{1}{2}(h - v + B^T \eta + F^T \zeta - 2F^T \text{diag}(w)g + 2B^T \text{diag}(t)(-d) \\
&\quad + \text{diag}(a)\mu - \text{diag}(b)\sigma - \lambda + \pi) \\
P_{33} &= \text{diag}(t) \\
P_{34} &= \frac{1}{2}(\eta - \delta - 2\text{diag}(t)d) \\
P_{44} &= -\eta^T d - e^T \pi - \zeta^T g + g^T \text{diag}(w)g + d^T \text{diag}(t)d
\end{aligned}$$

Proof. We first express $(\bar{\mathbb{P}}(u, v, w, t))$ by its dual form. Associate the following multipliers to the constraints in $(\bar{\mathbb{P}}(u, v, w, t))$:

- $\eta \in \mathfrak{R}^m$ for $Ax + By + s = d$;
- $\delta \in \mathfrak{R}_+^m$ for $s \geq 0$;
- $\zeta \in \mathfrak{R}^M$ for $Ex + Fy = g$;
- μ_i and $\sigma_i \in \mathfrak{R}_+$ for $a_i y_i \leq x_i$ and $x_i \leq b_i y_i$, respectively, $i = 1, \dots, n$;
- λ_i and $\pi_i \in \mathfrak{R}_+$ for $y_i \geq 0$ and $y_i \leq 1$, respectively, $i = 1, \dots, n$.

Let $\mu = (\mu_1, \dots, \mu_n)^T$, $\sigma = (\sigma_1, \dots, \sigma_n)^T$, $\lambda = (\lambda_1, \dots, \lambda_n)^T$ and $\pi = (\pi_1, \dots, \pi_n)^T$. Let ω denote the vector formed by $\delta, \mu, \sigma, \lambda$ and π . The Lagrangian function of $(\bar{\mathbb{P}}(u, v, w, t))$ is then given by

$$\begin{aligned}
& L(x, y, \omega, \eta, \zeta) \\
&= x^T Q x + c^T x + h^T y + \sum_{i=1}^n (u_i x_i y_i + v_i y_i^2 - u_i x_i - v_i y_i) \\
&\quad + (E x + F y - g)^T \text{diag}(w) (E x + F y - g) \\
&\quad + (A x + B y + s - d)^T \text{diag}(t) (A x + B y + s - d) \\
&\quad + \eta^T (A x + B y + s - d) + \delta^T (-s) + \zeta^T (E x + F y - g) \\
&\quad + \sum_{i=1}^n \mu_i (a_i y_i - x_i) + \sum_{i=1}^n \sigma_i (x_i - b_i y_i) + \sum_{i=1}^n \lambda_i (-y_i) + \sum_{i=1}^n \pi_i (y_i - 1) \\
&= (x, y, s)^T \begin{pmatrix} P_{11} & P_{12} & A^T \text{diag}(t) \\ P_{12}^T & P_{22} & B^T \text{diag}(t) \\ \text{diag}(t)A & \text{diag}(t)B & \text{diag}(t) \end{pmatrix} \begin{pmatrix} x \\ y \\ s \end{pmatrix} \\
&\quad + (c - u + A^T \eta + E^T \zeta - 2E^T \text{diag}(w)g + 2A^T \text{diag}(t)(-d) - \mu + \sigma)^T x \\
&\quad + (h - v + B^T \eta + F^T \zeta - 2F^T \text{diag}(w)g + 2B^T \text{diag}(t)(-d) \\
&\quad + \text{diag}(a)\mu - \text{diag}(b)\sigma - \lambda + \pi)^T y + (\eta - \delta - 2\text{diag}(t)d)^T s \\
&\quad - \eta^T d - e^T \pi - \zeta^T g + g^T \text{diag}(w)g + d^T \text{diag}(t)d.
\end{aligned}$$

Furthermore, the Lagrangian dual problem of $(\bar{P}(u, v, w, t))$ can be expressed as

$$\max \left\{ \min_{(x, y, s) \in \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R}^m} L(x, y, \omega, \eta, \zeta) \mid \omega \geq 0 \right\}. \quad (31)$$

Introducing an additional variable τ , we can rewrite (31) as

$$\max \tau \quad (32)$$

$$\text{s.t.} \quad \min_{(x, y, s) \in \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R}^m} L(x, y, \omega, \eta, \zeta) \geq \tau, \quad (33)$$

$$\omega \geq 0. \quad (34)$$

We see that the constraint in (33) is equivalent to $L(x, y, \omega, \eta, \zeta) - \tau \geq 0$ for all x, y, s , which is further equivalent to

$$L(x/k, y/k, \omega, \eta, \zeta) - \tau \geq 0, \quad \forall (x, y, s) \in \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R}^m, \quad \forall k \in \mathfrak{R}, \quad k \neq 0. \quad (35)$$

Multiplying both sides of (35) by k^2 yields a homogeneous quadratic form of (x, y, s, t) in the left-hand side of (35). Thus, the constraint in (33) is equivalent to the semidefinite constraint (30). Consequently, the problem in (32)-(34) can be expressed as

$$\max \{ \tau \mid \text{constraint (30)}, \omega \geq 0 \}. \quad (36)$$

If the objective function of $(P(u, v, w, t))$ is convex, by the strong duality of convex quadratic programming (see, e.g., Proposition 6.5.6 in [5]), the optimal values of $(\bar{P}(u, v, w, t))$ and (36) are equal. Thus, we have shown that problem $(\text{MAX}uvw t)$ is equivalent to an SDP problem in the form of (SDP_a) . \square

Table 3: Numerical results of bound improvement from combining lift-and-convexification and QCR

n	inst.	impr.	n	inst.	impr.	n	inst.	impr.
200 ⁻	1	17.0%	300 ⁻	1	47.1%	400 ⁻	1	15.1%
200 ⁻	2	28.0%	300 ⁻	2	12.7%	400 ⁻	2	21.8%
200 ⁻	3	31.8%	300 ⁻	3	25.9%	400 ⁻	3	6.5%
200 ⁻	4	18.1%	300 ⁻	4	17.6%	400 ⁻	4	10.1%
200 ⁻	5	12.7%	300 ⁻	5	26.1%	400 ⁻	5	16.5%
200 ⁻	6	17.7%	300 ⁻	6	22.0%	400 ⁻	6	17.9%
200 ⁻	7	31.1%	300 ⁻	7	22.8%	400 ⁻	7	42.1%
200 ⁻	8	29.2%	300 ⁻	8	15.6%	400 ⁻	8	22.5%
200 ⁻	9	30.4%	300 ⁻	9	2.5%	400 ⁻	9	11.5%
200 ⁻	10	33.0%	300 ⁻	10	30.2%	400 ⁻	10	20.1%
average		24.9%			22.2%			18.4%

Although this SDP problem (SDP_a) is very large and takes time to solve, it only needs to be solved once to get the new reformulation. When the original problem (P') is very difficult to solve, solving this SDP to get a better reformulation can gain overall computational advantage.

To test the effectiveness of the new reformulation, we compare the bounds of (SDP_q) and (SDP_a) on the portfolio selection problem data set introduced in §4. We use the 30 instances with the least diagonal dominance. For each instance, we impose additional equality constraints by dividing the stocks into 10 sections and demanding only one stock to be invested from each section. Such an additional constraint is a very practical one, as in real life applications portfolios are often constructed by choosing investment opportunities from different industries and sections. We use the following measure for bound improvement,

$$\text{impr.} = \frac{v(\text{SDP}_a) - v(\text{SDP}_q)}{\text{optimal_objective_value} - v(\text{SDP}_q)}$$

Table 3 shows the bound improvement for the 30 instances, which ranges from 2.5% to 47.1%, resulting an average bound improvement around 20%. This numerical experiment confirms that combining the lift-and-convexification approach and QCR generates a much tighter reformulation on average.

6 Concluding remarks

We have developed in this paper the lift-and-convexification approach to construct a parameterized set of MIQP reformulations for convex quadratic programs with semi-continuous variables. The primary idea behind this approach is to lift the quadratic term in the objective function from the x -space to the (x, y) -space and to convexify the resulting quadratic function of (x, y) . We have proposed an SDP formulation to identify the best MIQP reformulation from among this parameterized set and have proved that the identified best reformulation has a continuous relaxation that is as tight as the continuous relaxation of the well known perspective reformulation. By revealing the relationship between our new reformulation and the perspective reformulation, we further

reduce the computational effort required to construct our new reformulation and show that we only need little extra effort to solve an additional SOCP problem when compared to the perspective reformulation. Most importantly, our new reformulation retains the linearly constrained quadratic programming structure of the original mix-integer problem, which facilitates more effective utilization of commercial mixed integer programming solvers and ensures much faster computational time at children nodes in the branch-and-bound searching process. Our preliminary comparison results indicate that the performance of our new reformulation solved in general MIQP solvers is, at least, competitive to the state-of-the-art perspective cut approach in many cases and provides a good supplement to the state-of-the-art approaches. We further combine our lift-and-convexification approach and the quadratic convex reformulation approach in the literature to obtain an even tighter reformulation. In a broader picture, the lift-and-convexification approach offers an efficient solution framework of tight MIQP reformulation which improves the existing literature on the trade-off between the bound quality and computational complexity.

References

- [1] A. AHLATÇIOĞLU, M. BUSSIECK, M. ESEN, M. GUIGNARD, J.H. JAGLA, AND A. MEERAUS, *Combining QCR and CHR for convex quadratic pure 0–1 programming problems with linear constraints*, Annals of operations research, 199 (2012), pp. 33–49.
- [2] M.S. AKTÜRK, A. ATAMTÜRK, AND S. GÜREL, *A strong conic quadratic reformulation for machine-job assignment with controllable processing times*, Operations Research Letters, 37 (2009), pp. 187–191.
- [3] E.J. ANDERSON AND B.S. CHEAH, *Capacitated lot-sizing with minimum batch sizes and setup times*, International Journal of Production Economics, 30 (1993), pp. 137–152.
- [4] T.S. ARTHANARI AND Y. DODGE, *Mathematical Programming in Statistics*, John Wiley & Sons, New York, 1993.
- [5] D.P. BERTSEKAS, A. NEDIĆ, AND A.E. OZDAGLAR, *Convex Analysis and Optimization*, Athena Scientific Belmont, Mass, 2003.
- [6] D. BERTSIMAS AND R. SHIODA, *Algorithm for cardinality-constrained quadratic optimization*, Computational Optimization and Applications, 43 (2009), pp. 1–22.
- [7] D. BIENSTOCK, *Computational study of a family of mixed-integer quadratic programming problems*, Mathematical Programming, 74 (1996), pp. 121–140.
- [8] A. BILLIONNET AND S. ELLOUMI, *Using a mixed integer quadratic programming solver for the unconstrained quadratic 0-1 problem*, Mathematical Programming, 109 (2007), pp. 55–68.
- [9] A. BILLIONNET, S. ELLOUMI, AND A. LAMBERT, *Extending the QCR method to general mixed-integer programs*, Mathematical programming, 131 (2012), pp. 381–401.
- [10] A. BILLIONNET, S. ELLOUMI, AND M.C. PLATEAU, *Quadratic 0–1 programming: tightening linear or quadratic convex reformulation by use of relaxations*, RAIRO-Operations Research, 42 (2008), pp. 103–121.
- [11] ———, *Improving the performance of standard solvers for quadratic 0-1 programs by a tight convex reformulation: The QCR method*, Discrete Applied Mathematics, 157 (2009), pp. 1185–1197.

- [12] B. BLOG, G. VAN DER HOEK, A.H.G. RINNOOY KAN, AND G.T. TIMMER, *The optimal selection of small portfolios*, Management Science, 29 (1983), pp. 792–798.
- [13] P. BONAMI AND M.A. LEJEUNE, *An exact solution approach for portfolio optimization problems under stochastic and integer constraints*, Operations Research, 57 (2009), pp. 650–670.
- [14] J. BORWEIN AND A. LEWIS, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer, 2006.
- [15] T.J. CHANG, N. MEADE, J.E. BEASLEY, AND Y.M. SHARAIHA, *Heuristics for cardinality constrained portfolio optimisation*, Computer & Operations Research, 27 (2000), pp. 1271–1302.
- [16] Y. CRAMA AND M. SCHYNS, *Simulated annealing for complex portfolio selection problems*, European Journal of Operational Research, 150 (2003), pp. 546–571.
- [17] X.T. CUI, X.J. ZHENG, S.S. ZHU, AND X.L. SUN, *Convex relaxations and MIQCQP reformulations for a class of cardinality-constrained portfolio selection problems*, Journal of Global Optimization, 56 (2013), pp. 1409–1423.
- [18] INC. CVX RESEARCH, *CVX: Matlab software for disciplined convex programming, version 2.0*. <http://cvxr.com/cvx>, Aug. 2012.
- [19] A. FERNÁNDEZ AND S. GÓMEZ, *Portfolio selection using neural networks*, Computers & Operations Research, 34 (2007), pp. 1177–1191.
- [20] A. FRANGIONI, F. FURINI, AND C. GENTILE, *Approximated perspective relaxations: A project and lift approach*, tech. report, 2013. TR-13-04. Available at: <http://compass2.di.unipi.it/TR/Files/TR-13-04.pdf.gz>.
- [21] A. FRANGIONI AND C. GENTILE, *Perspective cuts for a class of convex 0–1 mixed integer programs*, Mathematical Programming, 106 (2006), pp. 225–236.
- [22] —, *Solving nonlinear single-unit commitment problems with ramping constraints*, Operations Research, 54 (2006), pp. 767–775.
- [23] —, *SDP diagonalizations and perspective cuts for a class of nonseparable MIQP*, Operations Research Letters, 35 (2007), pp. 181–185.
- [24] —, *A computational comparison of reformulations of the perspective relaxation: SOCP vs. cutting planes*, Operations Research Letters, 37 (2009), pp. 206–210.
- [25] A. FRANGIONI, C. GENTILE, E. GRANDE, AND A. PACIFICI, *Projected perspective reformulations with applications in design problems*, Operations Research, 59 (2011), pp. 1225–1232.
- [26] A. FRANGIONI, C. GENTILE, AND F. LACALANDRA, *Solving unit commitment problems with general ramp constraints*, International Journal of Electrical Power & Energy Systems, 30 (2008), pp. 316–326.
- [27] —, *Tighter approximated MILP formulations for unit commitment problems*, IEEE Transactions on Power Systems, 24 (2009), pp. 105–113.
- [28] J.J. GAO AND D. LI, *Cardinality constrained linear-quadratic optimal control*, IEEE Transactions on Automatical Control, 56 (2011), pp. 1936–1941.
- [29] —, *Optimal cardinality constrained portfolio selection*, Operations Research, 61 (2013), pp. 745–761.
- [30] —, *A polynomial case of the cardinality-constrained quadratic optimization problem*, Journal of Global Optimization, 56 (2013), pp. 1441–1455.
- [31] M. GRANT AND S. BOYD, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lec-

- ture Notes in Control and Information Sciences, Springer-Verlag Limited, 2008, pp. 95–110. http://stanford.edu/~boyd/graph_dcp.html.
- [32] O. GÜNLÜK AND J. LINDEROTH, *Perspective reformulations of mixed integer nonlinear programs with indicator variables*, Mathematical programming, 124 (2010), pp. 183–205.
- [33] P.L. HAMMER AND A.A. RUBIN, *Some remarks on quadratic programming with 0-1 variables*, RAIRO-Operations Research-Recherche Opérationnelle, 4 (1970), pp. 67–79.
- [34] N.L. JACOB, *A limited-diversification portfolio selection model for the small investor*, Journal of Finance, 29 (1974), pp. 847–856.
- [35] N.J. JOBST, M.D. HORNIMAN, C.A. LUCAS, AND G. MITRA, *Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints*, Quantitative Finance, 1 (2001), pp. 489–501.
- [36] D. LI, X.L. SUN, AND C.L. LIU, *An exact solution method for unconstrained quadratic 0–1 programming: a geometric approach*, Journal of Global Optimization, 52 (2012), pp. 797–829.
- [37] D. LI, X.L. SUN, AND J. WANG, *Optimal lot solution to cardinality constrained mean-variance formulation for portfolio selection*, Mathematical Finance, 16 (2006), pp. 83–101.
- [38] D. MARINGER AND H. KELLERER, *Optimization of cardinality constrained portfolios with a hybrid local search algorithm*, OR Spectrum, 25 (2003), pp. 481–495.
- [39] A.J. MILLER, *Subset Selection in Regression*, Chapman and Hall, 2002. Second edition.
- [40] G. MITRA, F. ELLISON, AND A. SCOWCROFT, *Quadratic programming for portfolio planning: Insights into algorithmic and computational issues. part ii: Processing of portfolio planning models with discrete constraints*, Journal of Asset Management, 8 (2007), pp. 249–258.
- [41] Y.W. PARK AND D. KLABJAN, *Lot sizing with minimum order quantity*, Discrete Applied Mathematics, 181 (2015), pp. 235–254.
- [42] M.C. PLATEAU, *Reformulations quadratiques convexes pour la programmation quadratique en variables 0-1*, PhD thesis, Ph. D. thesis, Conservatoire National d’Arts et Métiers, 2006.
- [43] A. SCHAERF, *Local search techniques for constrained portfolio selection problems*, Computational Economics, 20 (2002), pp. 177–190.
- [44] D.X. SHAW, S. LIU, AND L. KOPMAN, *Lagrangian relaxation procedure for cardinality-constrained portfolio optimization*, Optimization Methods and Software, 23 (2008), pp. 411–420.
- [45] X.L. SUN, X.J. ZHENG, AND D. LI, *Recent advances in mathematical programming with semi-continuous variables and cardinality constraint*, Journal of the Operations Research Society of China, 1 (2013), pp. 55–77.
- [46] J. XIE, S. HE, AND S. ZHANG, *Randomized portfolio selection, with constraints*, Pacific Journal of Optimization, 4 (2008), pp. 89–112.
- [47] X.J. ZHENG, X.L. SUN, AND D. LI, *Improving the performance of MIQP solvers for quadratic programs with cardinality and minimum threshold constraints: A semidefinite program approach*, INFORMS Journal on Computing, 26 (2014), pp. 690–703.