Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 15, No. 1, pp. 55 – 67, 2019

CONVEX MIQP REFORMULATIONS FOR SEMI-CONTINUOUS QUADRATIC PROGRAMMING WITH LOW PRICE*

Xiaojin Zheng, Yutong Pan and Baiyi Wu†

Abstract: In this paper, we propose a new convex mixed-integer quadratic reformulation for quadratic programming with semi-continuous variables, with no price of introducing new variables and new constraints. That is, the new convex mixed-integer quadratic programming (MIQP) has the same size of the original problem. Furthermore, the convex MIQP, whose continuous relaxation is at least as tight as that of perspective reformulation, can be obtained explicitly rather than solving a semidefinite programming problem (SDP) which will limit the application of the method since it is still an intractable task to find the solution of the SDP for practical large-scale problems. The only price of obtaining the convex MIQP is to lift the quadratic term involving x only in the original objective function to a quadratic term of x and y, where y is also a variable of the original problem. We report promising numerical results applying the new convex MIQP reformulation to solve Markowitz mean-variance portfolio selection problems whose number of assets ranges from 400 to 1000.

Key words: *semidefinite programming; semi-continuous variables; mixed integer quadratic programming; perspective cut reformulation; quadratic convex reformulation*

Mathematics Subject Classification: *90C11, 90C20, 90C22*

---

1 Introduction

In many real-world problems, we always encounter semi-continuous variables. A variable $x \in \Re$ is termed semi-continuous if $x \in \{0\} \cup [\alpha, \beta]$ for some $0 < \alpha \leq \beta$. One application of semi-continuous variables is in portfolio selection problems in financial optimization. Because of market frictions in real-life market, such as management and transaction fee, there is often a buy-in threshold or a minimum transaction level. Therefore, an investor can not hold some assets with a very small amount. This situation can be modeled by semi-continuous variables. We can also find many other models with this semi-continuous structure in design problems by [5], portfolio optimization problems by [12], unit commitment problems by [1] and many others by [7].

The models with semi-continuous variables we consider in this paper has the following form

$$(\mathrm{P}) \qquad \min \left\{ f(x,y) = x^T Q x + c^T x + h^T y \mid (x,y) \in \mathcal{F} \right\},$$

*This research was supported by National Natural Science Foundation of China under grants 11671300, 11701106, 11801099 and by the Fundamental Research Funds for the Central Universities under Grant 22120180277.

†Corresponding Author

© 2019 Yokohama Publishers

where Q is an $n \times n$ positive semidefinite symmetric matrix, $c, h \in \Re^n$,

$$\mathcal{F} = \{(x, y) \in \Re^n \times \{0,1\}^n \mid Ax + By \leq d,\ \alpha_i y_i \leq x_i \leq \beta_i y_i,\ i = 1, \ldots, n\},$$

with $A, B \in \Re^{m\times n}$, $d \in \Re^m$ and $0 < \alpha_i \leq \beta_i$, $i = 1, \ldots, n$. Denote by $\overline{\mathcal{F}}$ the relaxation of set $\mathcal{F}$, obtained by replacing $y \in \{0,1\}^n$ with $y \in [0,1]^n$ in $\mathcal{F}$.

The perspective reformulation of (P), which is proposed by [3], has the formulation

$$(\mathrm{PR}(\rho)) \qquad \min \left\{ f_\rho(x, y) = x^T(Q - \mathrm{diag}(\rho))x + c^T x + h^T y + \sum_{i=1}^{n} \rho_i x_i^2 / y_i \mid (x, y) \in \mathcal{F} \right\},$$

where $\rho \in \Omega$ with

$$\Omega = \{\rho \in \Re^n_+ \mid Q - \mathrm{diag}(\rho) \succeq 0\}. \tag{1.1}$$

The continuous relaxation of problem $(\mathrm{PR}(\rho))$, which is denoted by $(\overline{\mathrm{PR}}(\rho))$, is much tighter than that of problem (P). However, efficient solution methods proposed by state-of-the-art solvers can not be applied to solve problem $(\mathrm{PR}(\rho))$ directly due to the fractional terms in the objective function. Therefore necessary cost or price needs to be paid to make sure $(\mathrm{PR}(\rho))$ can be solved by off-the-shelf solver. Two tractable reformulations have been proposed to overcome this issue. The first reformulation, with the price of introducing n second order cone constraints, is a second order cone program (SOCP), which can be found in [4], [7], [12]. The other reformulation is a semi-infinite mixed-integer linear programming, addressed in [2], via representing the value x_i^2/y_i by the supremum of a set of infinite hyperplanes. It has been illustrated by [4] that the semi-infinite reformulation has a computational advantage over the SOCP reformulation for solving Markovitz mean-variance portfolio selection problems and unit commitment problems. The basic reason for this phenomenon is that the different algorithms are used to solve continuous relaxations of the subproblems in the process of a branch-and-bound method. That is, interior point methods for SOCP reformulation is not as efficient as tailed-made, dynamic cutting plane generating algorithms for the semi-infinite reformulation, where active-set (simplex-like) method for quadratic programs is fully utilized. However, the continuous relaxations of the two reformulations of problem $(\mathrm{PR}(\rho))$ are substantially more complex than that of the original problem (P). That costs much more time in solving the continuous relaxation in each branch in the process of a branch-and-bound method, which leads to inefficiency of the overall algorithm as the problem size grows.

By exploring the "inherent piecewise nature" of the perspective function of nonconvex functions corresponding to disjunctions, which has been studied by [10], a piecewise-quadratic programming reformulation is proposed by [5]. However, that approach can only be applied under the assumption that $B \equiv 0$ in problem (P).

The previous discussion motivates us to get a convex MIQP reformulation for problem (P) with no price of introducing new constraints and new variables. The continuous relaxation of the new MIQP should be at least as tight as that of perspective reformulation. And there should be no more extra assumption on problem (P) so that the reformulation can be applied to many classes of problems which has the general formulation of problem (P). The price of obtaining the convex MIQP reformulation should not be expensive. At least it should be cheaper than solving an SDP problem so that our reformulation can be applied to solve large-scale problems. Furthermore, during the process of a branch-and-bound method, the continuous relaxation of our new MIQP reformulation should be at least as tight as that of the perspective reformulation at every child node so that the new MIQP can be solved much more efficiently. In this paper, we are going to accomplish all those missions.

The contribution of this paper can be summarized as follows.

- We propose a convex MIQP reformulation for problem (P) with no price of introducing new variables and new constraints.
- The continuous relaxation of the new MIQP reformulation is better than that of perspective reformulation, or as good as it even in the worst case. The new MIQP reformulation can be obtained explicitly, as will be shown in Theorem 2.2.
- During the process of a branch-and-bound method, the continuous relaxation of our new MIQP reformulation is at least as tight as that of the perspective reformulation at every child node. The continuous relaxation of our new MIQP reformulation is a quadratic programming with the same size of original problem and thus can be solved much more efficiently than that of the SOCP or the semi-infinite reformulation.

The remainder of this paper is organized as follows. In section 2, we propose a new mixed-integer convex quadratic reformulation for problem (P) with low price. In section 3, the application of the new convex MIQP reformulation is investigated for Markowitz mean-variance portfolio selection problems whose number of assets ranges from 400 to 1000. Finally, we conclude in section 4.

Notations: Throughout the paper, we denote val$(\cdot)$ as the optimal value of problem $(\cdot)$, $\Re^n_+$ as the nonnegative orthant of $\Re^n$, e as the all-one vector and I_n as the identity matrix of rank n. For any $a \in \Re^n$, we denote by $\mathrm{diag}(a) = \mathrm{diag}(a_1, \ldots, a_n)$ the diagonal matrix with a_i being the ith diagonal element. Finally, for any $a \in \Re$, we define that $a/0$ is equal to ∞ if $a > 0$, 0 if $a = 0$ and $-\infty$ if $a < 0$.

2 Convex Mixed-Integer Quadratic Programming (MIQP) Reformulation

In this section, we will show that problem (P) can be reformulated as a mixed-integer convex quadratic programming with no price of introducing new constraints and new variables. Furthermore, for some $\rho \in \Omega$, we can get the new convex MIQP explicitly, whose continuous relaxation will be at least as tight as that of problem (PR(ρ)).

Note that for any feasible solution $(x, y) \in \mathcal{F}$, it has

$$x_i y_i = x_i,\ y_i^2 = y_i,\ i = 1, \ldots, n.$$

Then for any $(u, v) \in \Re^n \times \Re^n$, problem (P) is equivalent to the following mixed-integer quadratic programming

$$(\mathrm{P}(u,v)) \qquad \min \quad f_{u,v}(x,y) = x^T Q x + c^T x + h^T y + \sum_{i=1}^n [u_i(x_i y_i - x_i) + v_i(y_i^2 - y_i)]$$
$$\text{s.t.} \quad (x, y) \in \mathcal{F}.$$

As we can see from above, the feasible region of (P(u,v)) is exactly the same with (P) and the objective function of (P(u,v)) only contains variables x and y as well. Hence the new MIQP reformulation dose not involve any new variable or constraint and preserves the same size as the primal problem. $f_{u,v}(x,y)$ may not be a convex function of (x,y) over $\Re^n \times \Re^n$ and then problem (P(u,v)) may not be a convex MIQP. $f_{u,v}(x,y)$ is convex if and only if $(u,v) \in \Lambda$, where

$$\Lambda = \left\{ (u,v) \in \Re^n \times \Re^n \mid \begin{pmatrix} Q & \mathrm{diag}(u)/2 \\ \mathrm{diag}(u)/2 & \mathrm{diag}(v) \end{pmatrix} \succeq 0 \right\}. \tag{2.1}$$

For any $(u,v)\in\Lambda$, problem $(\mathrm{P}(u,v))$ is a convex mixed-integer quadratic reformulation of problem (P). Obviously, (u,v) with $u\equiv 0$, $v\geq 0$ is an element of Λ. Relaxing $y\in\{0,1\}^n$ to $y\in[0,1]^n$, we get continuous relaxation of problem $(\mathrm{P}(u,v))$, denoted by $(\overline{\mathrm{P}}(u,v))$.

As far as we know, the semi-infinite reformulation and the SOCP reformulation, derived from the perspective reformulation $(\overline{\mathrm{PR}}(\rho))$ for some $\rho\in\Omega$, are most efficient in solving problem (P). As proposed by [12], the best perspective reformulation, in the sense that $(\overline{\mathrm{PR}}(\rho))$ provides the tightest continuous relaxation bound among $\rho\in\Omega$, can be obtained via solving a "large" SDP problem, which finally limit the application of the "best" reformulation since it is still an intractable task to solve an SDP for practical large-scale problems.

It is shown in [11] that, for any $\rho\in\Omega$ we can get a convex mixed-integer quadratic reformulation whose continuous relaxation provides a lower bound that is as tight as that of the perspective reformulation $(\mathrm{PR}(\rho))$. The price is to solve the continuous relaxation of problem $(\mathrm{PR}(\rho))$, which can be reformulated as an SOCP problem. That is, the price is to solve an SOCP problem, which is cheaper than solving an SDP. However, the price is not cheap enough for us to solve large-scale problems. Furthermore, when we apply branch-and-bound methods to the new convex MIQP, the continuous relaxations at children nodes is in general not as tight as that of the perspective reformulation. The bound equivalence only occurs at the root node.

For the sake of completeness of the paper, we cite the major result in [11].

Theorem 2.1 (Theorem 5 in [11])**.** *For any $\rho\in\Omega$, let $(\widehat{x},\widehat{y})$ be an optimal solution to problem $(\overline{\mathrm{PR}}(\rho))$. Define $(\widehat{u},\widehat{v})\in\Re^n\times\Re^n$ as following,*

$$\widehat{u}_i=-2\rho_i\frac{\widehat{x}_i}{\widehat{y}_i},\quad \widehat{v}_i=\rho_i\frac{\widehat{x}_i^2}{\widehat{y}_i^2},\quad i=1,\ldots,n, \tag{2.2}$$

Then

(i) $(\widehat{u},\widehat{v})\in\Lambda$,

(ii) $\mathrm{val}(\overline{\mathrm{P}}(\widehat{u},\widehat{v}))=\mathrm{val}(\overline{\mathrm{PR}}(\rho))$.

Our major concern in this paper is the issue that for some $\rho\in\Omega$, how to choose $(u,v)\in\Lambda$ such that the continuous relaxation of problem $(\mathrm{P}(u,v))$ is at least as tight as that of problem $(\mathrm{PR}(\rho))$, without the requirement of solving SDP or SOCP problems. That is, for some $\rho\in\Omega$, how to choose $(u,v)\in\Lambda$ explicitly such that $\mathrm{val}(\overline{\mathrm{P}}(u,v))\geq\mathrm{val}(\overline{\mathrm{PR}}(\rho))$? The following Theorem 2.2 will solve this issue.

Theorem 2.2. *For any $\rho\in\overline{\Omega}$, define $(\overline{u},\overline{v})\in\Re^n\times\Re^n$ as following,*

$$\overline{u}_i=-\rho_i(\alpha_i+\beta_i),\quad \overline{v}_i=\rho_i\alpha_i\beta_i,\quad i=1,\ldots,n, \tag{2.3}$$

where

$$\overline{\Omega}=\left\{\rho\in\Re^n_+ \mid Q-\mathrm{diag}(\omega)\mathrm{diag}(\rho)\succeq 0,\right\}, \tag{2.4}$$

and $\omega_i=(\alpha_i+\beta_i)^2/(4\alpha_i\beta_i)$, $i=1,\ldots,n$. Then

(i) $\rho\in\Omega$ *and* $(\overline{u},\overline{v})\in\Lambda$,

(ii) $\mathrm{val}(\overline{\mathrm{P}}(\overline{u},\overline{v}))\geq\mathrm{val}(\overline{\mathrm{PR}}(\rho))$.

Proof. Based on (1.1) and (2.4), together with the fact that $\omega_i \geq 1$, $i = 1, \ldots, n$, it has $\rho \in \Omega$. In order to prove $(\overline{u}, \overline{v}) \in \Lambda$, it suffices to prove that

$$\overline{Q} = \begin{pmatrix} Q & \mathrm{diag}(\overline{u})/2 \\ \mathrm{diag}(\overline{u})/2 & \mathrm{diag}(\overline{v}) \end{pmatrix} \succeq 0.$$

Denote $D = \mathrm{diag}(d)$ where $d_i = 1$ if $\overline{v}_i = 0$ and $d_i = 1/\overline{v}_i$ if $\overline{v}_i \neq 0$, $i = 1, \ldots, n$. Then we have, $\mathrm{diag}(\overline{u})D\mathrm{diag}(\overline{v}) = \mathrm{diag}(\overline{v})D\mathrm{diag}(\overline{u}) = \mathrm{diag}(\overline{u})$. We denote

$$P = \begin{pmatrix} I_n & 0 \\ -D\mathrm{diag}(\overline{u})/2 & I_n \end{pmatrix}$$

and $\widehat{Q} = P^T\overline{Q}P$. Noticed P is an invertible matrix, so matrices $\widehat{Q}$ and $\overline{Q}$ are called congruent. According to Sylvester's law of inertia([9]), the numbers of positive, negative and zeros eigenvalues of $\widehat{Q}$ and $\overline{Q}$ are equal, so we have:

$$\begin{aligned} \overline{Q} \succeq 0 \quad &\Leftrightarrow \quad \widehat{Q} = \begin{pmatrix} I_n & -\mathrm{diag}(\overline{u})D/2 \\ 0 & I_n \end{pmatrix} \overline{Q} \begin{pmatrix} I_n & 0 \\ -D\mathrm{diag}(\overline{u})/2 & I_n \end{pmatrix} \succeq 0 \\ &\Leftrightarrow \quad \widehat{Q} = \begin{pmatrix} Q - \mathrm{diag}(\overline{u})D\mathrm{diag}(\overline{u})/4 & 0 \\ 0 & \mathrm{diag}(\overline{v}) \end{pmatrix} \succeq 0. \end{aligned}$$

Thus, it suffices to prove $\widehat{Q} \succeq 0$. Since $\overline{v} \geq 0$ according to (2.3), together with (2.3) and (2.4), we have $\widehat{Q} \succeq 0$.

(ii) For any feasible solution $(x, y) \in \overline{\mathcal{F}}$, we have

$$\begin{aligned} \mathrm{val}(\overline{\mathrm{PR}}(\rho)) - \mathrm{val}(\overline{\mathrm{P}}(\overline{\mathrm{u}}, \overline{\mathrm{v}})) &= f_\rho(x, y) - f_{u,v}(x, y) \\ &= \sum_{i=1}^n [-\rho_i x_i^2 + \rho_i x_i^2/y_i - \overline{u}_i(x_i y_i - x_i) - \overline{v}_i(y_i^2 - y_i)] \\ &= \sum_{i=1}^n \rho_i \frac{1 - y_i}{y_i} x_i^2 + \rho_i(\alpha_i + \beta_i)(x_i y_i - x_i) - \rho_i \alpha_i \beta_i (y_i^2 - y_i) \\ &= \sum_{i=1}^n \rho_i \frac{1 - y_i}{y_i} x_i^2 + \rho_i(\alpha_i + \beta_i)(y_i - 1)x_i - \rho_i \alpha_i \beta_i (y_i - 1) y_i \\ &= \sum_{i=1}^n \rho_i \frac{1 - y_i}{y_i} x_i^2 - \rho_i \frac{1 - y_i}{y_i}(\alpha_i + \beta_i) x_i y_i + \rho_i \frac{1 - y_i}{y_i} \alpha_i \beta_i y_i^2 \\ &= \sum_{i=1}^n \rho_i \frac{1 - y_i}{y_i}(x_i^2 - (\alpha_i + \beta_i) x_i y_i + \alpha_i \beta_i y_i^2) \\ &= \sum_{i=1}^n \rho_i \frac{1 - y_i}{y_i}(x_i - \alpha_i y_i)(x_i - \beta_i y_i) \\ &\leq 0, \end{aligned}$$

where the second equality holds due to (2.3) and the last inequality holds due to the fact that $\alpha_i y_i \leq x_i \leq \beta_i y_i$, $i = 1, \ldots, n$. □

Remark 2.3. From Theorem 2.2, for any $\rho \in \overline{\Omega}$, we can get the new convex MIQP $(\mathrm{P}(\overline{u}, \overline{v}))$ explicitly. That is, there is no price of getting the new convex MIQP.

Remark 2.4. Since the parameters $(\overline{u}, \overline{v})$ in the new convex MIQP $(\mathrm{P}(\overline{u}, \overline{v}))$ depend only on the parameters α and β, which will not be changed in children nodes of a branch-and-bound tree, the relationship $\mathrm{val}(\overline{\mathrm{P}}(\overline{u}, \overline{v})) \geq \mathrm{val}(\overline{\mathrm{PR}}(\rho))$ will always hold in each child node of the branch-and-bound tree. That is, the tightness of the bound from the continuous relaxation can be ensured for each subproblem of the branch-and-bound method.

Remark 2.5. Obviously, $\overline{\Omega} \subseteq \Omega$. For $\rho \in \Omega \cap \overline{\Omega}^c$, unfortunately , we may not be able to find $(\overline{u}, \overline{v}) \in \Lambda$ which is independent of the optimal solution of problem $(\overline{\mathrm{PR}}(\rho))$, such that $\mathrm{val}(\overline{\mathrm{P}}(\overline{u}, \overline{v})) \geq \mathrm{val}(\overline{\mathrm{PR}}(\rho))$. However, according to Theorem 2.1, for $\rho \in \Omega \cap \overline{\Omega}^c$, based on the optimal solution of problem $(\overline{\mathrm{PR}}(\rho))$, we can obtain a $(\widehat{u}, \widehat{v}) \in \Lambda$ such that $\mathrm{val}(\overline{\mathrm{P}}(\widehat{u}, \widehat{v})) = \mathrm{val}(\overline{\mathrm{PR}}(\rho))$.

Remark 2.6. There are two obvious ways of finding $\rho \in \overline{\Omega}$.

- Let $\rho_i = \lambda_{\min}/\omega_i$, $i = 1, \ldots, n$ with $\lambda_{\min}$ being the minimum eigenvalue matrix Q. It is easy to check that $\rho \in \overline{\Omega}$ due to the fact that $Q - \mathrm{diag}(\omega)\mathrm{diag}(\rho) = Q - \mathrm{diag}(\lambda_{min} e) \succeq 0$.
- Find $\rho \in \overline{\Omega}$ via solving the following small SDP problem:

$$(\mathrm{SDP}_s) \qquad \max\{e^T \rho \mid Q - \mathrm{diag}(\omega)\mathrm{diag}(\rho) \succeq 0,\ \rho \geq 0\}.$$

Remark 2.7. The inequality $\mathrm{val}(\overline{\mathrm{PR}}(\rho)) \leq \mathrm{val}(\overline{\mathrm{P}}(\overline{\mathrm{u}}, \overline{\mathrm{v}}))$ still holds when $y_i = 0$ because

$$\begin{aligned}
& \lim_{y \to 0} (\mathrm{val}(\overline{\mathrm{PR}}(\rho)) - \mathrm{val}(\overline{\mathrm{P}}(\overline{\mathrm{u}}, \overline{\mathrm{v}}))) \\
= \quad & \lim_{y \to 0} (f_\rho(x, y) - f_{u,v}(x, y)) \\
= \quad & \lim_{y \to 0} \sum_{i=1}^n [-\rho_i x_i^2 + \rho_i x_i^2 / y_i - \overline{u}_i (x_i y_i - x_i) - \overline{v}_i (y_i^2 - y_i)] \\
= \quad & \sum_{i=1}^n \lim_{y \to 0} [-\rho_i x_i^2 + \rho_i x_i^2 / y_i - \overline{u}_i (x_i y_i - x_i) - \overline{v}_i (y_i^2 - y_i)] \\
= \quad & \sum_{i=1}^n [-\rho_i \lim_{y \to 0} x_i^2 + \rho_i \lim_{y \to 0} x_i^2 / y_i - \overline{u}_i \lim_{y \to 0} (x_i y_i - x_i) - \overline{v}_i \lim_{y \to 0} (y_i^2 - y_i)] \\
= \quad & \sum_{i=1}^n \rho_i \lim_{y \to 0} x_i^2 / y_i
\end{aligned}$$

where the last equality holds since x_i is semicontinuous variable, or more explicitly, $x_i = 0$ when $y_i = 0$, as

$$\lim_{y \to 0} x_i^2 = 0, \lim_{y \to 0} (x_i y_i - x_i) = 0, \lim_{y \to 0} (y_i^2 - y_i) = 0$$

As for x_i^2 / y_i, we have

$$\begin{aligned}
& 0 \leq \rho_i \lim_{y \to 0} x_i^2 / y_i \leq \rho_i \lim_{y \to 0} (\beta_i y_i)^2 / y_i = \rho_i \lim_{y \to 0} \beta_i^2 y_i = 0 \\
& \Rightarrow \rho_i \lim_{y \to 0} x_i^2 / y_i = 0 \ \Rightarrow \lim_{y \to 0} (\mathrm{val}(\overline{\mathrm{PR}}(\rho)) - \mathrm{val}(\overline{\mathrm{P}}(\overline{\mathrm{u}}, \overline{\mathrm{v}}))) = 0
\end{aligned}$$

Without loss of generality, we define $x_i^2 / y_i = 0$ when $y_i = 0$, and $\mathrm{val}(\overline{\mathrm{PR}}(\rho)) = \mathrm{val}(\overline{\mathrm{P}}(\overline{\mathrm{u}}, \overline{\mathrm{v}}))$ holds when $y_i = 0$ as well.

We will illustrate by the following Example 2.8 that the continuous relaxation of new MIQP (P(u, v)) is tighter than that of problem (PR(ρ)).

Example 2.8. Consider the following example:

$$\begin{array}{ll} \min & x^T \begin{pmatrix} 135 & 25 & 24 & 71 \\ 25 & 126 & 72 & 51 \\ 24 & 72 & 150 & 63 \\ 71 & 51 & 63 & 112 \end{pmatrix} x \\ \text{s.t.} & 6x_1 + 8x_2 + 4x_3 + 9x_4 \geq 6, \\ & x_1 + x_2 + x_3 + x_4 = 1, \\ & y_1 + y_2 + y_3 + y_4 \leq 2, \\ & 0.1y_i \leq x_i \leq 0.9y_i, y_i \in \{0, 1\},\ i = 1, 2, 3, 4. \end{array}$$

The optimal solution is $(x^*, y^*) = (0.4787, 0.5213, 0, 0, 1, 1, 0, 0)$ and the optimal value val(P) = 77.654. The bound from the continuous relaxation of the problem is val($\overline{\text{P}}$) = 69.4585.

(i) if $\rho_i = \lambda_{min}/w_i,\ i = 1, \ldots, 4$, i.e., $\rho = (14.1668, 14.1668, 14.1668, 14.1668)^T$, val($\overline{\text{PR}}(\rho)$) = 72.5089. Based on (2.3) we have $\overline{u}_1 = \overline{u}_2 = \overline{u}_3 = \overline{u}_4 = -14.1668$, $\overline{v}_1 = \overline{v}_2 = \overline{v}_3 = \overline{v}_4 = 1.2750$ and then val($\overline{\text{P}}(\overline{u}, \overline{v})$) $= 73.7901 > 72.5089 =$ val($\overline{\text{PR}}(\rho)$).

(ii) If ρ is obtained via solving the "small" SDP $\left(\text{SDP}\right)_s$, we get $\rho = (23.8288, 27.0713, 16.4238, 8.4611)^T$ for this example and then val($\overline{\text{PR}}(\rho)$) = 73.3146. Based on (2.3) we have $\overline{u} = (-23.8288, -27.0713, -16.4238, -8.4611)^T$, $\overline{v} = (2.1446, 2.4364, 1.4781, 0.7615)^T$ and val($\overline{\text{P}}(\overline{u}, \overline{v})$) $= 74.6213 > 73.3146 =$ val($\overline{\text{PR}}(\rho)$).

3 Mean-Variance Portfolio Selection Model

The classical Markowitz mean-variance portfolio selection model in financial optimization, proposed by [8], can be described as follows. Suppose that in a financial market, n risky assets with random return vector $R = (R_1, \ldots, R_n)^T$ are available. Denote by μ and Q the expected return vector and the covariance matrix of R, respectively. The mean-variance model solves the following quadratic programming

$$\min\ \{x^T Q x\ \mid\ \mu^T x \geq d,\ \sum_{i=1}^{n} x_i = 1,\ x \geq 0\},$$

where x_i represents the fraction of the total capital invested in asset i and d is a desired return level by the investor. However, in real-life application, many other constraints will be required for the portfolio selection problems. Typically, due to market frictions, such as management and transaction fee, there is minimum and maximum transaction level for each asset i. Therefore, the investors can not hold some assets with a very small amount, or invest too much capital on some assets. Furthermore, the investors will also impose the constraint on the maximum numbers of stock to invest. Hence, in this section, we will stick with the following mixed-integer quadratic programming

$$\min \left\{ x^T Q x \;\middle|\; \begin{array}{l} \mu^T x \geq d, e^T x = 1, \\ \alpha_i y_i \leq x_i \leq \beta_i y_i,\ e^T y \leq K,\ y \in \{0, 1\}^n \end{array} \right\}. \tag{3.1}$$

Note that there are semi-continuous variables x_i in problem (3.1). For any $\rho \in \Omega$, the perspective reformulation of problem (3.1) is obtained as following.

$$(\mathrm{PR}_{\mathrm{MV}}(\rho)) \qquad \min \quad x^T(Q - \mathrm{diag}(\rho))x + \sum_{i=1}^{n} \rho_i x_i^2 / y_i$$

$$\text{s.t.} \quad \mu^T x \ge d, \ e^T x = 1, \ e^T y \le K, \tag{3.2}$$

$$\alpha_i y_i \le x_i \le \beta_i y_i, \ y \in \{0,1\}^n. \tag{3.3}$$

For any $\rho \in \overline{\Omega}$, together with Theorem 2.2, the convex MIQP reformulation of problem (3.1), whose continuous relaxation bound is at least as tight as that of problem $(\mathrm{PR}_{\mathrm{MV}}(\rho))$, has the following formulation

$$(\mathrm{MIQP}(u,v)) \qquad \min \quad x^T Q x + \sum_{i=1}^{n} [u_i x_i y_i - u_i x_i + v_i y_i^2 - v_i y_i]$$

$$\text{s.t.} \quad \mu^T x \ge d, \ e^T x = 1, \ e^T y \le K,$$

$$\alpha_i y_i \le x_i \le \beta_i y_i, \ y \in \{0,1\}^n.$$

where (u_i, v_i) for $i = 1, ..., n$ are defined in (2.3).

3.1 Computational Results

In this section, we illustrate the effectiveness of the new convex mixed-integer quadratic reformulation $(\mathrm{MIQP}(u,v))$ by conducting comparison with the perspective reformulations $(\mathrm{PR}_{\mathrm{MV}}(\rho))$. A key issue in implementing this test is how to choose ρ. According to the technique used by [3], to test our new reformulation, we use the following two choices of ρ:

- $(\mathrm{PC_e})$, $(\mathrm{MIQP_e})$: the reformulations $(\mathrm{PR}_{\mathrm{MV}}(\rho))$ and (MIQP(u, v)), respectively, with $\rho_i = \lambda_{\min}/\omega_i$, $i = 1, \ldots, n$ and $\lambda_{\min}$ being the minimum eigenvalue of covariance matrix Q;
- $(\mathrm{PC_s})$, $(\mathrm{MIQP_s})$: the reformulations $(\mathrm{PR}_{\mathrm{MV}}(\rho))$ and (MIQP(u, v)), respectively, with ρ being the optimal solution to the following simple SDP problem:

  $$(\mathrm{SDP_s}) \qquad \max\{e^T \rho \mid Q - \mathrm{diag}(\omega)\mathrm{diag}(\rho) \succeq 0, \ \rho \ge 0\}, \tag{3.4}$$

  which is called "small" SDP problem in this paper. Here w is defined in Theorem 2.2.

For perspective reformulations, we implemented a branch-and-cut method using CPLEX 12.3 through its C programming interface, in which the perspective cuts can be dynamically generated via *cutcallback* procedures. Our implementation follows [4] and separation is done twice at each node of the branch-and-bound tree. For our convex quadratic reformulation, we also use CPLEX 12.3 to solve it through its C programming interface. In our test, CPLEX default settings are used, including the dual simplex quadratic programming optimizer for the subproblems at each node of the branch-and-bound tree.

The "small" SDP problems $(\mathrm{SDP_s})$ was solved by SeDuMi 1.2 within CVX 1.2 by [6], which is a Matlab-based modeling system for convex optimization. We perform the numerical tests 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 conduct the numerical tests on 90 instances of problem (6). The 90 instances had the following structure:

- there are 30 instances each for $n = 400, 600, 1000$;
- for each n, the 30 instances are further divided into three subsets denoted by n^+, n^0 and n^-. Each subset has a different diagonal dominance in the covariance matrix Q.

The 30 instances for $n = 400$ can be found at: http:www.di.unipi.it/optimize/Data/MV.html. For $n = 600, 1000$, we use the same random generator as in [4] to generate the instances. We tested on all the instances both with the cardinality constraint $e^T y \leq K$ and without the cardinality constraint. Note that because all $\alpha_i \geq 0.075$ in the 90 instances, the number K in the cardinality constraint is at most 13. Thus, we added the cardinality constraint with $K = 6, 10$, respectively. Together with the instances with no cardinality constraint, which we denote as "nc" in table 1, we have 270 instances of problem (3.1) in total.

Table 1: Comparison results of perspective reformulations and our new MIQP reformulations on (MV) data.

Problem	K	time$_s$	(PC$_e$)			(MIQP$_e$)			(PC$_s$)			(MIQP$_s$)		
			gap	time	nodes	gap	time	nodes	gap	time	nodes	gap	time	nodes
400^+	6	87.4	28.41(9)	3594	16933	3.54(7)	2642	21761	18.99(9)	3268	16050	0.01(0)	137	2339
	10	88.3	26.10(9)	3307	19228	20.02(9)	3252	28044	18.26(9)	3253	19006	12.92(8)	3104	45310
	nc	88.4	23.06(9)	3255	28247	21.70(9)	3246	56218	15.29(9)	3244	29812	13.50(8)	3055	78159
400^0	6	92.4	32.09(10)	3600	15345	4.20(8)	3002	25257	24.28(10)	3600	17547	0.45(2)	1705	28776
	10	91.2	29.36(10)	3600	19211	22.87(10)	3600	29878	22.54(10)	3600	20331	16.46(10)	3600	51339
	nc	91.6	26.60(10)	3600	28511	24.96(10)	3600	60412	19.74(10)	3600	29641	17.69(10)	3600	88694
400^-	6	95.1	32.63(10)	3600	15041	5.03(10)	3220	27861	25.74(10)	3600	17143	1.63(6)	2641	42397
	10	94.4	29.64(10)	3600	19633	22.94(10)	3600	32416	23.73(10)	3600	20953	17.47(10)	3600	51953
	nc	93.3	26.06(10)	3600	27598	24.81(10)	3600	57744	19.96(10)	3600	30638	18.25(10)	3600	83084
600^+	6	235.7	65.35(10)	3600	3377	9.97(10)	3600	6883	81.76(10)	3600	4769	0.83(6)	2766	9562
	10	254.3	39.52(10)	3600	4492	28.79(10)	3600	4340	78.59(10)	3600	5859	20.85(10)	3600	10406
	nc	253.5	34.94(10)	3600	6983	32.26(10)	3600	13795	71.56(10)	3600	7156	23.22(10)	3600	18773
600^0	6	264.0	46.12(10)	3600	3806	9.40(10)	3330	6705	67.18(10)	3600	4919	3.29(8)	2942	9537
	10	266.0	40.57(10)	3600	4719	27.39(10)	3600	5159	63.51(10)	3600	6121	20.59(10)	3600	10245
	nc	262.9	32.66(10)	3600	7969	29.89(10)	3600	15982	58.85(10)	3600	8137	22.41(10)	3600	19725
600^-	6	260.3	42.55(10)	3600	3855	10.63(10)	3600	7022	67.53(10)	3600	5088	5.60(10)	3600	10549
	10	255.6	35.56(10)	3600	4822	27.85(10)	3600	4764	63.92(10)	3600	5724	22.11(10)	3600	10687
	nc	254.2	32.67(10)	3600	7435	30.19(10)	3600	15126	57.08(10)	3600	7930	24.39(10)	3600	18989
1000^+	6	1127.5	65.35(10)	3600	3377	9.97(10)	3600	6883	81.76(10)	3600	4769	0.83(9)	2766	9562
	10	1126.7	39.52(10)	3600	4492	28.79(10)	3600	4340	78.59(10)	3600	5859	20.85(10)	3600	10406
	nc	1142.6	34.94(10)	3600	6983	32.26(10)	3600	13795	71.56(10)	3600	7156	23.22(10)	3600	18773
1000^0	6	1000.3	46.12(10)	3600	3806	9.40(10)	3330	6705	67.18(10)	3600	4919	3.29(9)	2942	9537
	10	1056.1	40.57(10)	3600	4719	27.39(10)	3600	5159	63.51(10)	3600	6121	20.59(10)	3600	10245
	nc	1068.4	32.66(10)	3600	7969	29.89(10)	3600	15982	58.85(10)	3600	8137	22.41(10)	3600	19725
1000^-	6	1148.9	42.55(10)	3600	3855	10.63(10)	3600	7022	67.53(10)	3600	5088	5.60(10)	3600	10549
	10	1146.8	35.56(10)	3600	4822	27.85(10)	3600	4764	63.92(10)	3600	5724	22.11(10)	3600	10687
	nc	1141.8	32.67(10)	3600	7435	30.19(10)	3600	15126	57.08(10)	3600	7930	24.39(10)	3600	18989

Table 1 summarizes the average numerical results of the two perspective reformulations and the two convex mixed-integer quadratic reformulations for the 270 instances of problem (3.1). Each line shows the average result of the 10 instances in that group. The notations in Table 1 are explained as follows.

- The columns "Time$_s$" is the computing time (in seconds) for obtaining parameter ρ via solving (SDP$_s$) using CVX.

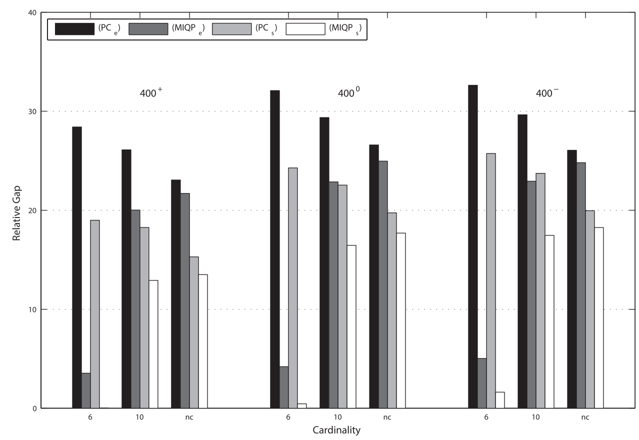

Figure 1: Comparison of relative gap on (MV) data with $n = 400$.

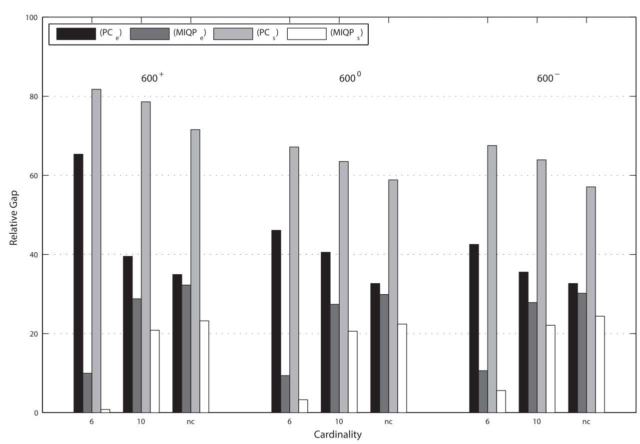

Figure 2: Comparison of relative gap on (MV) data with $n = 600$.

- The column "Gap" is the relative gap (in percentage) of the incumbent solution when CPLEX 12.3 is terminated. The number in parenthesis next to the gap is the number of unsolved instances within 3,600 seconds. The default tolerance of relative gap in CPLEX 12.3 is 0.01%.

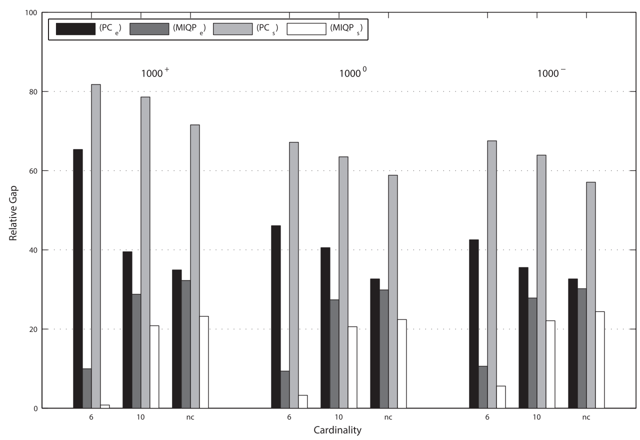

Figure 3: Comparison of relative gap on (MV) data with $n = 1000$.

- The columns "Time" and "Nodes" are the computing time (in seconds) and the number of nodes explored by CPLEX 12.3, respectively.

Figure 1 to 3 display the comparison of relative gap. Together with Table 1, we can see that the average computation time and relative gap of reformulations ($\mathrm{MIQP_e}$) and ($\mathrm{MIQP_s}$) are significantly less than those of ($\mathrm{PC_e}$) and ($\mathrm{PC_s}$), respectively for all instances of types n^+, n^0 and n^-. This is mainly because our new reformulation reduces the effort required at each node while generating a lower bound that is at least as tight as that of the perspective reformulation. The number of nodes of ($\mathrm{MIQP_e}$) and ($\mathrm{MIQP_s}$) are roughly two times that of ($\mathrm{PC_e}$) and ($\mathrm{PC_s}$), respectively. This can be explained from the fact that at each node, when separation is done twice, the perspective algorithm solves two quadratic programming problems, while for our new reformulation, we only need to solve one quadratic programming problem. Then given the same time limit, our new reformulation is able to explore more nodes than the perspective reformulation. Moreover, the continuous relaxation of our new reformulation at each node is at least as tight as that of the continuous relaxation of the perspective reformulation. Given that separation is done only twice, the bound from the perspective algorithm should be worse than the bound from the perspective reformulation. This means that the lower bound of our new reformulation obtained at each node is better than that of the perspective algorithm most of the time.

For instances of all types n^+, n^0 and n^- with small cardinality ($K = 6$), ($\mathrm{MIQP_e}$) and ($\mathrm{MIQP_s}$) appears to be particularly advantageous over ($\mathrm{PC_e}$) and ($\mathrm{PC_s}$), respectively.

It can be noticed from Table 1 that ($\mathrm{MIQP_s}$) performs better than ($\mathrm{MIQP_e}$) in terms of relative gap. However, there is additional price of problem ($\mathrm{MIQP_s}$). That is, we need to solve the small SDP ($\mathrm{SDP_s}$). We can see from Table 1 that the computation time of solving small SDP reaches 1000 seconds for $n = 1000$. This time cost is compensated by smaller relative gaps of ($\mathrm{MIQP_s}$). However, the same relationship between ($\mathrm{PC_e}$) and ($\mathrm{PC_s}$) only holds for $n = 400$. For $n = 600, 1000$, (PC_s) does not perform better than (PC_e), although we take efforts to solve (SDP_s) to get (PC_s).

4 Concluding Remarks

We have presented in this paper a new convex mixed-integer quadratic reformulation (MIQP) for quadratic programming with semi-continuous variables, with no price of introducing new variables and new constraints. That is, our new convex mixed-integer quadratic programming has the same size of the original problem. Furthermore, the convex MIQP, whose continuous relaxation is at least as tight as that of perspective reformulation, can be obtained explicitly rather than solving a large semidefinite programming problem which will limit the application of the method since it is still an intractable task to find the solution of the SDP for practical large-scale problems. The only price of obtaining the convex MIQP is to lift the quadratic term involving x only in the original objective function to a quadratic term of x and y, where y is also a variable of the original problem. When we apply branch-and-bound methods to the new convex MIQP, the property that the continuous relaxation of the new convex MIQP is at least as tight as that of perspective reformulation holds for every child node of the branch-and-bound tree. Together with the fact that our new reformulation significantly reduce the effort required at each node, the performance of our new MIQP reformulation appears to be advantageous over the perspective reformulation. Our preliminary comparison results on Markowitz mean-variance portfolio selection problems whose number of assets ranges from 400 to 1000 indicate that our new convex MIQP can help improve the performance of the MIQP solvers for the problem.

However, there is still an issue unsolved in this paper. It has been shown by [12] that the best $\rho \in \Omega$, in the sense of getting the tightest continuous relaxation bound of the perspective reformulation, can be found via solving an "large" SDP problem. We also have also shown in Remark 2.5 that we can not find the new MIQP corresponding the best ρ which can preserve the "best" continuous relaxation during the process of branch-and-bound method. It is a research topic that we will focus on in the future.

References

[1] A. Borghetti, A. Frangioni, F. Lacalandra and C. A. Nucci, Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment, *IEEE Trans. Power Syst.* 18 (2003) 1–10.

[2] A. Frangioni and C. Gentile, Perspective cuts for a class of convex 0–1 mixed integer programs, *Math. Program.* 106 (2006) 225–236.

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

[4] A. Frangioni and C. Gentile, A computational comparison of reformulations of the perspective relaxation: SOCP vs. cutting planes, *Oper. Res. Lett.* 37 (2009) 206–210.

[5] A. Frangioni, C. Gentile, E. Grande and A. Pacifici, Projected perspective reformulations with applications in design problems, *Oper. Res. Lett.* 59 (2011) 1225–1232.

[6] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 2.1, `http://cvxr.com/cvx`, 2014.

[7] O. Günlük and J. Linderoth, Perspective reformulations of mixed integer nonlinear programs with indicator variables, *Math. Program.* 124 (2010) 183–205.

[8] H. M. Markowitz, Portfolio selection, *J. Finance* 7 (1952) 77–91.

[9] J. J. Sylvester, XIX. A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares, *The London, Edinburgh Dublin Philosophical Magazine J. Science* 4 (1852) 138–142.

[10] M. Tawarmalani and N. V. Sahinidis, Semidefinite relaxations of fractional programs via novel convexification techniques, *J. Global. Optim.* 20 (2001) 133–154.

[11] B. Wu, X. Sun, D. Li and X. Zheng, Quadratic Convex Reformulations for Semicontinuous Quadratic Programming, *SIAM J. Optim.* 27 (2017) 1531–1553.

[12] 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 J. Comput.* 26 (2014) 690–703.

Manuscript received 20 October 2017
revised 10 February 2018
accepted for publication 0 September 2018

Xiaojin Zheng
School of Economics and Management, Tongji University
Shanghai 200092, P.R. China
E-mail address: xjzheng@tongji.eud.cn

Yutong Pan
School of Economics and Management, Tongji University
Shanghai 200092, P.R. China
E-mail address: 11ytpan@tongji.edu.cn

Baiyi Wu
School of Finance, Guangdong University of Foreign Studies
Guangzhou 510420, P.R. China
E-mail address: baiyiwu@outlook.com