BCOL RESEARCH REPORT 16.02

Industrial Engineering & Operations Research
University of California, Berkeley, CA 94720–1777

POLYMATROID INEQUALITIES FOR P-ORDER CONIC MIXED $0-1$ OPTIMIZATION

ALPER ATAMTÜRK AND ANDRÉS GÓMEZ

Abstract. We describe new convex valid inequalities for p-order conic mixed-integer optimization, which includes the important second order conic mixed-integer optimization as a special case. The inequalities are based on the polymatroid inequalities over binary variables for the diagonal case. We prove that the proposed inequalities completely describe the convex hull of a single conic constraint over binary variables and unbounded continuous variables. We then generalize and strengthen the inequalities using other constraints of the optimization problem. Computational experiments for second order conic mixed-integer optimization indicate that the new inequalities strengthen the convex relaxations substantially for the diagonal case as well as the general (non-diagonal) case and lead to significant performance improvements.

Keywords: Polymatroids, submodularity, second order cone, nonlinear cuts

October 2016

A. Atamtürk: Department of Industrial Engineering & Operations Research, University of California, Berkeley, CA 94720. atamturk@berkeley.edu
A. Gómez: Department of Industrial Engineering & Operations Research, University of California, Berkeley, CA 94720. a.gomez@ieor.berkeley.edu .

1. Introduction

Second order conic mixed-integer optimization is a problem of the form

$$\text{(SOCMIO)} \quad \begin{aligned} \min\ & f'x + g'z \\ \text{s.t.}\ & \sqrt{x'Q_i x} \le z_i, && i = 1, \ldots, \ell \\ & (x,z) \in X \subseteq \mathbb{Z}_+^n \times \mathbb{R}_+^m \times \mathbb{R}_+^\ell \end{aligned} \tag{1}$$

where $Q_i \succeq 0$ for $i = 1, \ldots, \ell$. We refer to inequalities (1) as the second order conic constraints. Many design and estimation optimization problems are modeled as SOCMIO (Lobo et al. 1998, Alizadeh and Goldfarb 2003, Atamtürk et al. 2012). In particular, second order conic constraints are frequently used to model probabilistic optimization with Gaussian distributions (Birge and Louveaux 2011) and robust optimization problems with ellipsoidal uncertainty sets (Ben-Tal and Nemirovski 1998, 1999, Ben-Tal et al. 2009).

Linear mixed-integer optimization (LMIO) is a special case of SOCMIO. Strong formulations have proven to be one of the critical components in solving LMIO, and state-of-the-art solvers for LMIO employ a variety of valid inequalities as cutting planes. However, relatively few classes of strong valid inequalities are known to strengthen the convex relaxations of SOCMIO and, more generally, nonlinear mixed-integer optimization.

General valid inequalities for convex nonlinear and/or conic mixed-integer optimization include intersection cuts, disjunctive cuts, and lift-and-project cuts (Ceria and Soares 1999, Stubbs and Mehrotra 1999). Çezik and Iyengar (2005) discuss Gomory cuts for general conic optimization problems. Atamtürk and Narayanan (2010) give conic MIR cuts for conic mixed-integer optimization and Atamtürk and Narayanan (2011) study lifting for conic mixed-integer optimization. Dadush et al. (2011) investigate the split closure of a convex set. Belotti et al. (2015) study the intersection of a convex set and a linear disjunction. Kılınç et al. (2010) and Bonami (2011) discuss the separation of split cuts using outer approximations and nonlinear programming, respectively. Kılınç-Karzan and Yıldız (2015) study disjunctions on the second order cone.

Another stream of research involves generating strong cuts by exploiting structured sets as it is common for the linear integer case. Although the applicability of such cuts is restricted to certain classes of sets, they tend to be far more effective than the general cuts that ignore any special structure. Aktürk et al. (2009, 2010) give second-order representable perspective cuts for a nonlinear scheduling problem with variable upper bounds, which are generalized further by Günlük and Linderoth (2010). Ahmed and Atamtürk (2011) give strong lifted inequalities for maximizing a submodular concave utility function. Atamtürk and Narayanan (2009), Atamtürk and Bhardwaj (2015) study binary knapsack sets defined by a single second-order conic constraint. Modaresi et al. (2016) derive closed form intersection cuts for a number of structured sets.

To goal of the current paper is to contribute to the understanding of convex hull of simple conic mixed-integer sets that form the building blocks of more general constraint sets as relaxations. In a related paper, Atamtürk and Narayanan (2008) give *extended polymatroid inequalities* for second order conic constraints (1) with diagonal Q_i matrices on binary variables, and show that these inequalities describe the convex hull in that case. In this paper, we first extend their results to the mixed-binary case and show that a nonlinear generalization of the polymatroid inequalities is sufficient to describe the convex hull for the mixed-binary case with unbounded continuous variables. We then show how additional constraints, in particular, the upper bounds on the continuous variables, can be used to further generalize and strengthen the first class of inequalities. Interestingly, although the inequalities are derived for the diagonal case, they can be applied to the non-diagonal case as well through a suitable relaxation. Computational experiments indicate that the derived inequalities are quite effective for the diagonal and as well as the non-diagonal cases.

We should note that utilizing the diagonal entries of matrices is standard for constructing convex relaxations in quadratic optimization (e.g. Poljak and Wolkowicz 1995, Anstreicher 2012). In particular, for $x \in \{0,1\}^n$, we have

$$x'Qx \leq z \iff x'(Q - \mathrm{diag}(a))x + a'x \leq z$$

with $a \in \mathbb{R}^n$ such that $Q - \mathrm{diag}(a) \succeq 0$. This transformation is based on the ideal (convex hull) representation of the separable quadratic term $x'\mathrm{diag}(a)x$ as $a'x$ for $x \in \{0,1\}^n$.

A similar approach is also available for convex quadratic optimization with indicator constraints. For $x \in \{0,1\}^n$ and $y \in \mathbb{R}^n$ s.t. $\ell \circ x \leq y \leq u \circ x$, we have

$$y'Qy \leq z \iff y'(Q - \mathrm{diag}(a))y + a't \leq z, \ y_i^2 \leq x_i t_i$$

with $t \in \mathbb{R}^n_+$ (e.g. Aktürk et al. 2009, Günlük and Linderoth 2010). This transformation is based on the ideal representation of each quadratic term $a_i y_i^2$ subject to indicator constraints as a linear term $a_i t_i$ along with a rotated cone constraint $y_i^2 \leq x_i t_i$.

Since in the conic quadratic constraint (1), the terms are not separable even for the diagonal case, simple transformations as in the quadratic cases above are not sufficient to arrive at an ideal formulation. We show that it is necessary to exploit the submodularity of the underlying set function to arrive at the ideal representations.

The rest of the paper is organized as follows. In Section 2 we introduce the notation used throughout the paper and review the results of Atamtürk and Narayanan (2008). In Section 3 we give the complete convex hull description of a single mixed-binary conic constraint. In Section 4 we study mixed-binary conic constraints with upper bounds on the continuous variables. In Section 5 we show how to include additional constraints to generalize and strengthen the inequalities. In Section 6 we report on a computational study done to test the effectiveness of the proposed inequalities for solving SOCMIO, including instances with non-diagonal matrices.

2. Preliminaries

2.1. **Notation.** Let x denote an n-dimensional vector of binary variables, y denote an m-dimension vector of continuous variables, and c and d be *nonnegative* vectors of dimension n and m, respectively, $\sigma \geq 0$ be a constant. Define $N = \{1, \ldots, n\}$ and $M = \{1, \ldots, m\}$. Let $\operatorname{conv}(X)$ denote the convex hull of X.

For $p > 1$ and $y \geq 0$, we study a p-order conic constraint of the form

$$\sqrt[p]{\sum_{i \in N} c_i x_i^p + \sum_{i \in M} d_i y_i^p} \leq z, \tag{2}$$

where the second order conic constraint corresponds to the case $p = 2$. Throughout, instead of the convex inequality (2) we will use

$$\sqrt[p]{\sum_{i \in N} c_i x_i + \sum_{i \in M} d_i y_i^p} \leq z. \tag{3}$$

Constraint (3) is equivalent to (2) over binary x, but it is stronger over the continuous relaxation of x since $x_i^p = x_i$ for $x_i \in \{0,1\}$, but $x_i^p < x_i$ for $x_i \in (0,1)$. Inequality (3) is concave in x, but convex in y. We will exploit both the concavity on x and the convexity on y.

2.2. **Previous work.** In this section we state, without proof, the main results of Atamtürk and Narayanan (2008) for the set

$$K_\sigma = \left\{ (x,z) \in \{0,1\}^n \times \mathbb{R}_+ : \sqrt[p]{\sigma + \sum_{i \in N} c_i x_i} \leq z \right\}.$$

For a given a permutation $((1), (2), \ldots, (n))$ of N, let

$$\begin{aligned} \sigma_{(k)} &= \sigma + \sum_{i=1}^{k-1} c_{(i)}, \text{ and} \\ \pi_{(k)} &= \sqrt[p]{c_{(k)} + \sigma_{(k)}} - \sqrt[p]{\sigma_{(k)}}. \end{aligned} \tag{4}$$

and define the *extended polymatroid inequality* as

$$\sqrt[p]{\sigma} + \sum_{i=1}^{n} \pi_{(i)} x_{(i)} \leq z. \tag{5}$$

Let Π_σ be the set of such coefficient vectors π for *all* permutations of N.

Proposition 1 (Convex hull of K_σ)**.**

$$conv(K_\sigma) = \left\{ (x,z) \in [0,1]^n \times \mathbb{R}_+ : \sqrt[p]{\sigma} + \pi' x \leq z, \ \ \forall \pi \in \Pi_\sigma \right\}.$$

The set function defining K_σ is submodular; therefore, Π_σ form the extreme points of an extended polymatroid. Since the maximization of a linear function over an extended polymatroid can be solved by the greedy algorithm (Edmonds 1970), a point $\bar{x} \in \mathbb{R}^n_+$ can be separated from conv(K_σ) via the greedy algorithm by sorting $\bar{x}_i$ in non-increasing order in $O(n \log n)$.

Proposition 2 (Separation)**.** *A point $\bar{x} \notin conv(K_\sigma)$ such that $\bar{x}_{(1)} \geq \bar{x}_{(2)} \geq \ldots \geq \bar{x}_{(n)}$ is separated from $conv(K_\sigma)$ by inequality* (5)*.*

Atamtürk and Narayanan (2008) also consider a mixed-integer extension and give valid inequalities for the mixed-integer set

$$L_\sigma = \left\{ (x,y,z) \in \{0,1\}^n \times [0,1]^m \times \mathbb{R}_+ : \sqrt[p]{\sigma + \sum_{i \in N} c_i x_i + \sum_{i \in M} d_i y_i^p} \leq z \right\}.$$

Without loss of generality, the upper bounds of the continuous variables in L_σ are set to one by scaling. For $T \subseteq M$, define $d(T) := \sum_{i \in T} d_i$.

Proposition 3 (Valid inequalities for L_σ)**.** *For $T \subseteq M$ inequalities*

$$\sqrt[p]{\sigma + \sum_{i \in T} d_i y_i^p} + \pi' x \leq z, \quad \pi \in \Pi_{\sigma + d(T)} \tag{6}$$

are valid for L_σ.

Inequalities (6) are obtained by setting the subset T of the continuous variables to their upper bounds and relaxing the rest and they dominate any inequality of the form

$$\sqrt[p]{\sigma + \sum_{i \in T} d_i y_i^p} + \xi' x \leq z$$

with $\xi \in \mathbb{R}^n$.

Finally, note that the optimization of a linear function over L_σ:

$$\min\{a'x + b'y + z \ : \ (x,y,z) \in L_\sigma\} \tag{7}$$

is solvable in polynomial time: For a fixed value of x, problem (7) reduces to a (convex) conic quadratic optimization problem in y that can be solved easily.

On the other hand, for fixed a value of y problem (7) reduces to a submodular minimization problem that can be solved by the greedy algorithm (Shen et al. 2003). Without loss of generality, assume that $c_i > 0$ for all i, as otherwise x_i can be set to either 0 or 1, depending on the sign of a_i. Index the binary variables so that $\frac{a_1}{c_1} \leq \ldots \leq \frac{a_n}{c_n}$ (breaking ties arbitrarily) and let $S_i = \{1, \ldots, i\}$ for $i = 1, 2, \ldots, n$. There exists an optimal solution (x^*, y^*) to (7) such that $x_k^* = 1$ if $k \in S_i$ for some $i = 1, \ldots, n$, and $x_k^* = 0$ otherwise. Thus, problem (7) can be solved by fixing the binary variables according to sets S_i one at a time and then solving the remaining conic quadratic optimization problem in polynomial time.

3. Conic constraint with unbounded continuous variables

In this section we consider the mixed-integer set

$$H_\sigma = \left\{ (x, y, z) \in \{0,1\}^n \times \mathbb{R}_+^{m+1} : \sqrt[p]{\sigma + \sum_{i \in N} c_i x_i + \sum_{i \in M} d_i y_i^p} \leq z \right\}.$$

Note that H_σ is the relaxation of L_σ by dropping the upper bounds on the continuous variables y. Thus, the only class of valid inequalities of type (6) are the extended polymatroid inequalities

$$\sqrt[p]{\sigma} + \pi' x \leq z, \quad \forall \pi \in \Pi_\sigma$$

from the "binary-only" relaxation by letting $T = \emptyset$. Here, we define a new class of *nonlinear* valid inequalities for H_σ and prove that they are sufficient to define its convex hull.

Consider the inequalities

$$\sqrt[p]{\left(\sqrt[p]{\sigma} + \pi' x\right)^p + \sum_{i \in M} d_i y_i^p} \leq z, \qquad \pi \in \Pi_\sigma. \tag{8}$$

Proposition 4. *Inequalities* (8) *are valid for* H_σ.

Proof. Consider the extended formulation of H_σ given by

$$\widehat{H}_\sigma = \left\{ (x, y, z, s) \in \{0,1\}^n \times \mathbb{R}_+^{m+2} : \sqrt[p]{s^p + \sum_{i \in M} d_i y_i^p} \leq z, \sqrt[p]{\sigma + \sum_{i \in N} c_i x_i} \leq s \right\}.$$

The validity of inequalities (8) for H_σ follows directly from the validity of the extended polymatroid inequality $\sqrt[p]{\sigma} + \pi' x \leq s$, $\pi \in \Pi_\sigma$ (Proposition 1) for $\widehat{H}_\sigma$. □

Remark 1. For $M = \emptyset$ inequalities (8) reduce to the extended polymatroid inequalities (5).

Remark 2. Although inequalities (8) are nonlinear in the original space of variables, they can be represented as linear inequalities in the extended formulation $\widehat{H}_\sigma$. Such a representation is desirable when they are used as cutting planes in branch-and-cut algorithms.

Remark 3. Since inequalities (8) correspond to extended polymatroid inequalities in an extended formulation, the separation for them is the same as in the binary case and can be done by sorting in $O(n \log n)$ (Proposition 2).

Proposition 5. *Inequalities* (8) *and the bound constraints describe* $conv(H_\sigma)$.

Proof. Consider the optimization of an arbitrary linear function over the convex relaxation of $\widehat{H}_\sigma$:

$$\text{(P1)} \qquad \begin{aligned} \min \ & -a'x - b'y + rz && (9) \\ \text{s.t.} \ & \sqrt[p]{s^p + \sum_{i \in M} d_i y_i^p} \le z && (10) \\ & \sqrt[p]{\sigma} + \pi' x \le s, \quad \forall \pi \in \Pi_\sigma && (11) \\ & x \in [0,1]^n, y \in \mathbb{R}^m_+, z \ge 0, s \ge 0. && (12) \end{aligned}$$

Note that the constraint $\sqrt[p]{\sigma + \sum_{i \in N} c_i x_i} \le s$ in $\widehat{H}_\sigma$ is implied by inequalities (11). We prove that for any linear objective (P1) is either unbounded or has an optimal solution that is integer in x.

Without loss of generality, we can assume that $r > 0$ (if $r < 0$ then the problem is unbounded, and if $r = 0$ then (P1) reduces to a linear program over an integral polyhedron), $r = 1$ (by scaling), $a_i, b_i > 0$ (otherwise $x_i = 0$ or $y_i = 0$ in any optimal solution), and $d_i = 1$ for all $i \in M$ (by scaling y_i). Eliminating the variable z from (P1) we rewrite the problem as

$$\text{(P2)} \qquad \begin{aligned} \min \ & -a'x - b'y + \sqrt[p]{s^p + \sum_{i \in M} y_i^p} \\ \text{s.t.} \ & \sqrt[p]{\sigma} + \pi' x \le s, \quad \forall \pi \in \Pi_\sigma \\ & x \in [0,1]^n, y \in \mathbb{R}^m_+, s \ge 0. \end{aligned}$$

Let $\mu \in \mathbb{R}^m_+$ be the dual variables for constraints $y \ge 0$. From the KKT conditions of (P2) with respect to y, we see that

$$-\mu_k = b_k - \left(s^p + \sum_{i \in M} y_i^p \right)^{\frac{1-p}{p}} y_k^{p-1}, \qquad \forall k \in M.$$

However, the complementary slackness conditions $y_k \mu_k = 0$ imply that $\mu_k = 0$ for all k, as otherwise $-\mu_k = b_k$ contradicts with the assumption that $b_k > 0$. Therefore, it holds that

$$y_k = \sqrt[p-1]{b_k} \cdot \sqrt[p]{s^p + \sum_{i \in M} y_i^p}, \quad \forall k \in M.$$

Defining $\beta = \sum_{i=1}^m b_i^{\frac{p}{p-1}}$, we have

$$\sum_{i \in M} b_i y_i = \beta \sqrt[p]{s^p + \sum_{i \in M} y_i^p}$$

and

$$\sum_{i \in M} y_i^p = \beta \left(s^p + \sum_{i \in M} y_i^p \right). \tag{13}$$

Observe that if $\beta > 1$, equality (13) cannot be satisfied, and the feasible (P2) is dual infeasible, therefore, unbounded. Moreover, if $\beta = 1$ then either the problem is unbounded or $s = 0$ in any optimal solution, which implies that $x = 0$ and all optimal solutions are integral in x. Finally, if $\beta < 1$, we deduce from (13) that

$$\sum_{i \in M} y_i^p = \frac{\beta}{1-\beta} s^p.$$

Replacing the summands in the objective, we rewrite (P2) as

$$\begin{aligned} \text{(P3)} \qquad & \min \; -a'x + (1-\beta)^{\frac{p-1}{p}} s \\ & \text{s.t. } \sqrt[p]{\sigma} + \pi' x \leq s, \quad \forall \pi \in \Pi_\sigma \\ & x \in [0,1]^n, s \geq 0. \end{aligned}$$

As $\beta < 1$, (P3) has an optimal solution and, by Proposition 1, it is integral in x. □

4. Conic constraint with bounded continuous variables

In this section we study the set L_σ, the generalization of K_σ with upper bounded continuous variables. As Example 1 illustrates, $\text{conv}(L_\sigma)$ is significantly more difficult to describe than $\text{conv}(H_\sigma)$.

Example 1. Consider the three-dimensional set given by

$$L_\sigma^2 = \left\{ (x,y,z) \in \{0,1\} \times [0,1] \times \mathbb{R}_+ : \sqrt{\sigma + cx + dy^2} \leq z \right\}.$$

We show in Appendix A that

$$\text{conv}(L_\sigma^2) = \{(x,y,z) \in [0,1] \times [0,1] \times \mathbb{R}_+ : g(x,y) \leq z\},$$

where

$$g(x,y) = \begin{cases} g_1(x,y) = \sqrt{\left(\sqrt{\sigma} + x(\sqrt{c+\sigma} - \sqrt{\sigma})\right)^2 + dy^2} & \text{if } y \leq x + (1-x)\sqrt{\frac{\sigma}{\sigma+c}} \\ g_2(x,y) = \sqrt{\sigma(1-x)^2 + d(y-x)^2} + x\sqrt{\sigma+c+d} & \text{otherwise.} \end{cases}$$

Observe that the inequality $g_1(x,y) \leq z$ is a particular case of (8). The difficulties arise with the function g_2:

(a) The discrete and continuous variables are tied together in the term $\sqrt{\sigma(1-x)^2 + d(y-x)^2}$.

(b) The inequality $g_2(x,y) \le z$ is not valid. In particular, it cuts off the feasible point $(x,y,z) = (1,0,\sqrt{\sigma+c})$. Moreover, the inequality $g_2(x,y) \le z$ cuts off portions of $\text{conv}(L^2_\sigma)$ whenever $y \le x + (1-x)\frac{\sqrt{\sigma}}{\sqrt{\sigma+c}}$.

(c) The condition $y \le x + (1-x)\frac{\sqrt{\sigma}}{\sqrt{\sigma+c}}$ depends both on x and y.

Figure 1 shows functions g_1 and g_2 for a fixed value of x, and illustrates point (b) above. We see that the function g_2 is always "above" the function g_1, and cuts the convex hull of L^2_σ (the shaded region) whenever $y \le x + (1-x)\frac{\sqrt{\sigma}}{\sqrt{\sigma+c}}$.

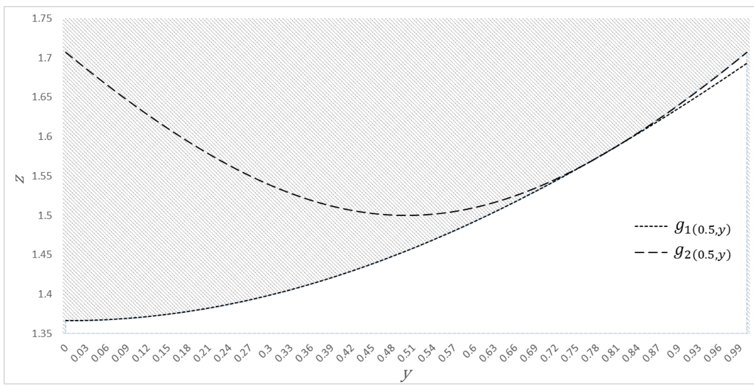

FIGURE 1. Funcs. g_1, g_2 with $\sigma = d = 1$, $c = 2$, restricted to $x = 0.5$.

We now give valid inequalities for $\text{conv}(L_\sigma)$. For $T \subseteq M$, consider the inequalities

$$\sqrt[p]{\left(\sqrt[p]{\sigma + \sum_{i\in T} d_i y_i^p} + \pi' x\right)^p + \sum_{i\in M\setminus T} d_i y_i^p} \le z, \quad \pi \in \Pi_{\sigma+d(T)}. \tag{14}$$

Proposition 6. *Inequalities* (14) *are valid for* L_σ.

Proof. For $T \subseteq M$, let

$$L_\sigma(T) = \left\{(x,y) \in \{0,1\}^n \times [0,1]^m, s \ge 0 : \sqrt[p]{\sigma + \sum_{i\in N} c_i x_i + \sum_{i\in T} d_i y_i^2} \le s\right\},$$

and consider the extended formulation of L_σ given by

$$\widehat{L}_\sigma = \left\{(x,y,s) \in L_\sigma(T), z \ge 0 : \sqrt[p]{s^p + \sum_{i\in M\setminus T} d_i y_i^p} \le z\right\}.$$

The validity of inequalities (14) for L_σ follows from the validity of

$$\sqrt[p]{\sigma + \sum_{i\in T} d_i y_i^p} + \pi' x \le s, \quad \pi \in \Pi_{\sigma+d(T)} \tag{15}$$

for $L_\sigma(T)$ (Proposition 6). □

Remark 4. If $T=\emptyset$, then inequalities (14) coincide with inequalities (8). If $T=M$, then inequalities (14) coincide with inequalities (6). If $T\subset M$, then inequalities (14) dominate inequalities (6).

Remark 5. Inequalities (14) are convex, since they correspond to the projection of convex inequalities (15) in an extended formulation.

Example 2. Consider the set

$$L_0^6=\left\{(x,y,z)\in\{0,1\}^4\times[0,1]^2\times\mathbb{R}_+ : \sqrt{\sum_{i=1}^4 x_i+y_1^2+y_2^2}\leq z\right\}.$$

For the permutation (1,2,3,4) inequalities (14) are

$$\begin{aligned}
&T=\emptyset: && \sqrt{(x_1+0.41x_2+0.32x_3+0.27x_4)^2+y_1^2+y_2^2}\leq z,\\
&T=\{1\}: && \sqrt{(0.41x_1+0.32x_2+0.27x_3+0.24x_4+y_1)^2+y_2^2}\leq z,\\
&T=\{1,2\}: && 0.32x_1+0.27x_2+0.24x_3+0.21x_4+\sqrt{y_1^2+y_2^2}\leq z.
\end{aligned}$$

Observe that for $T=\emptyset$ and $T=\{1\}$, the resulting inequalities dominate the corresponding inequalities obtained from (6), given by $x_1+0.41x_2+0.32x_3+0.27x_4\leq z$ and $0.41x_1+0.32x_2+0.27x_3+0.24x_4+y_1\leq z$, respectively.

Example 1 (Continued). We obtain from (14) the valid inequality

$$g_3(x,y)=\sqrt{\sigma+dy^2}+x\left(\sqrt{\sigma+c+d}-\sqrt{\sigma+d}\right)\leq z$$

for L_σ^2. Observe that if $\sigma=0$, then $g_1(x,y)\leq z$, $g_3(x,y)\leq z$ and the bound constraints give a complete description of $\text{conv}(L_\sigma^2)$ since

$$g_3(x,y)=\sqrt{d}y+x\left(\sqrt{c+d}-\sqrt{d}\right)=\sqrt{d}\,(|y-x|)+x\sqrt{\sigma+c+d}=g_2(x,y)$$

whenever $y\geq x+(1-x)\sqrt{\frac{\sigma}{\sigma+c}}=x$. If $\sigma>0$, then $g_3(x,y)\leq z$ is valid and provides an approximation of $\text{conv}(L_\sigma^2)$ (Figure 2).

5. Strengthened polymatroid inequalities

The polymatroid inequalities of Sections 3 and 4 use the conic constraint and the bounds of the variables. In this section we show how to strengthen the polymatroid inequalities using additional constraints. In particular, given *any* mixed-integer set $X\subseteq\{0,1\}^n\times\mathbb{R}_+^m$, we consider the generalization

$$G_\sigma=\left\{(x,y)\in X, z\geq 0 : \sqrt[p]{\sigma+\sum_{i\in N}c_ix_i+\sum_{i\in M}d_iy_i^p}\leq z\right\}.$$

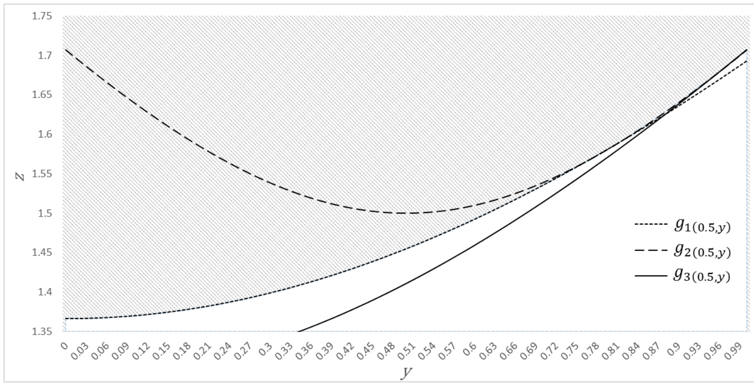

FIGURE 2. Functions g_1, g_2, g_3 with $\sigma = d = 1$, $c = 2$, restricted to $x = 0.5$.

First, in Section 5.1 we describe a lifting procedure for obtaining valid inequalities for G_σ, where computing each coefficient requires solving an integer optimization problem. Then, in Section 5.2 we discuss how the strengthened polymatroid inequalities can be efficiently implemented in practice.

5.1. **Valid inequalities for G_σ.** For a given a permutation $((1),(2),\dots,(n))$ of N and $T \subseteq M$, let

$$h_k(x,y) = \sigma + \sum_{i=1}^{k-1} c_{(i)}x_{(i)} + \sum_{i\in T} d_i y_i^p$$

$$\bar\sigma_{(k)} = \max\left\{h_k(x,y) : (x,y) \in X, x_k = 1\right\}, \text{ and} \tag{16}$$

$$\rho_{(k)} = \begin{cases} \sqrt[p]{c_{(k)} + \bar\sigma_{(k)}} - \sqrt[p]{\bar\sigma_{(k)}} & \text{if } \bar\sigma_{(k)} < \infty \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Consider the inequality

$$\sqrt[p]{\left(\sqrt[p]{\sigma + \sum_{i\in T} d_i y_i^p} + \sum_{i=1}^{n} \rho_{(i)}x_{(i)}\right)^p + \sum_{i\in M\setminus T} d_i y_i^p} \le z. \tag{18}$$

Proposition 7. *Inequalities* (18) *are valid for G_σ.*

Proof. Let

$$G_\sigma(T) = \left\{(x,y) \in X, s \ge 0 : \sqrt[p]{\sigma + \sum_{i\in N} c_i x_i + \sum_{i\in T} d_i y_i^p} \le s\right\},$$

and consider the extended formulation of G_σ given by

$$\hat{G_\sigma} = \left\{ (x,y,s) \in G_\sigma(T), z \geq 0 : \sqrt[p]{s^p + \sum_{i \in M \setminus T} d_i y_i^p} \leq z \right\}.$$

To prove the validity of (18) for G_σ, it is sufficient to show that

$$\sqrt[p]{\sigma + \sum_{i \in T} d_i y_i^p} + \sum_{i=1}^{n} \rho_{(i)} x_{(i)} \leq s \tag{19}$$

is valid for $G_\sigma(T)$. In particular, we prove by induction that

$$\sqrt[p]{\sigma + \sum_{i \in T} d_i y_i^p} + \sum_{i=1}^{k} \rho_{(i)} x_{(i)} \leq \sqrt[p]{\sigma + \sum_{i=1}^{k} c_{(i)} x_{(i)} + \sum_{i \in T} d_i y_i^p} \tag{20}$$

for all $(x,y) \in X$ and $k = 0, \ldots, n$.

Base case: $k = 0$. Inequality (20) holds trivially.

Inductive step. Let $(\bar{x}, \bar{y}) \in X$, and suppose inequality (20) holds for $k-1$. Observe that if $\bar{x}_{(k)} = 0$ or $\rho_{(k)} = 0$, then inequality (20) clearly holds for k. Therefore, assume that $\bar{x}_{(k)} = 1$ and $\bar{\sigma}_{(k)} < \infty$. We have

$$\begin{aligned}
\sqrt[p]{\sigma + \sum_{i=1}^{k} c_{(i)} \bar{x}_{(i)} + \sum_{i \in T} d_i \bar{y}_i^p} &= \sqrt[p]{h_k(\bar{x},\bar{y}) + c_{(k)}} \\
&= \sqrt[p]{h_k(\bar{x},\bar{y})} + \left(\sqrt[p]{h_k(\bar{x},\bar{y}) + c_{(k)}} - \sqrt[p]{h_k(\bar{x},\bar{y})} \right) \\
&\geq \sqrt[p]{h_k(\bar{x},\bar{y})} + \left(\sqrt[p]{\bar{\sigma}_{(k)} + c_{(k)}} - \sqrt[p]{\bar{\sigma}_{(k)}} \right) \qquad (21) \\
&\geq \sqrt[p]{\sigma + \sum_{i \in T} d_i \bar{y}_i^p} + \sum_{i=1}^{k} \rho_{(i)} \bar{x}_{(i)}, \qquad (22)
\end{aligned}$$

where (21) follows from $\bar{\sigma}_{(k)} \geq h_k(\bar{x},\bar{y})$ (by definition of $\bar{\sigma}_{(k)}$) and from the concavity of the root function, and (22) follows from $\sqrt[p]{h_k(\bar{x},\bar{y})} \geq \sqrt[p]{\sigma + \sum_{i \in T} d_i \bar{y}_i^p} + \sum_{i=1}^{k-1} \rho_{(i)} \bar{x}_{(i)}$ (induction hypothesis) and from the definition of $\rho_{(k)}$. □

Example 2 (Continued). Let $X^6 = \left\{ (x,y) \in \{0,1\}^4 \times [0,1]^2 : \sum_{i=1}^{4} x_i + y_1 + y_2 \leq 3 \right\}$ and consider the set $G_0^6 = L_0^6 \cap X^6$. For the permutation (1,2,3,4) inequalities (18) are

$$\begin{aligned}
&T = \emptyset: && \sqrt{(x_1 + 0.41x_2 + 0.32x_3 + 0.32x_4)^2 + y_1^2 + y_2^2} \leq z, \\
&T = \{1\}: && \sqrt{(0.41x_1 + 0.32x_2 + 0.32x_3 + 0.32x_4 + y_1)^2 + y_2^2} \leq z, \\
&T = \{1,2\}: && 0.32x_1 + 0.32x_2 + 0.32x_3 + 0.32x_4 + \sqrt{y_1^2 + y_2^2} \leq z.
\end{aligned}$$

Observe that, in all cases, the resulting inequalities dominate the corresponding inequalities obtained from (14).

Remark 6. If $T = \emptyset$ and $X = \{0,1\}^n \times \mathbb{R}_+^m$, then inequalities (18) reduce to inequalities (8). If $T = \emptyset$ and $X \subset \{0,1\}^n \times \mathbb{R}_+^m$, then inequalities (18) dominate inequalities (8).

Remark 7. If $X = \{0,1\}^n \times [0,1]^m$, then inequalities (18) reduce to inequalities (14). If $X \subset \{0,1\}^n \times [0,1]^m$, then inequalities (18) dominate inequalities (14).

5.2. **Computational efficiency.** Note that computing each coefficient of inequality (18) requires solving the integer optimization problem (16), which may not be practical in most cases. However, observe from Remarks 6 and 7 that solving the optimization problem over *any* relaxation of X results in valid inequalities at least as strong as the ones resulting from using only the bounds constraints.

In particular, assume in problem (16) that for $i \in T$ there exists $u_i \geq 0$ such that $y_i \leq u_i$ (otherwise the problem is unbounded and $\rho_i = 0$) and $u_i = 1$ (by scaling). Moreover let X_P be a polytope such that $X \subseteq X_P$. Convex constraints can also be included in X_P by using a suitable linear outer approximation (Ben-Tal and Nemirovski 2001, Tawarmalani and Sahinidis 2005, Hijazi et al. 2013, Vielma et al. 2015, Lubin et al. 2016).

Given X_P, the approximate coefficients

$$\begin{aligned} &\hat{\rho}_{(k)} = \sqrt[p]{c_{(k)} + \hat{\sigma}_{(k)}} - \sqrt[p]{\hat{\sigma}_{(k)}}, \text{ with} \\ &\hat{\sigma}_{(k)} = \sigma + \max\left\{ \sum_{i=1}^{k-1} c_{(i)} x_{(i)} + \sum_{i \in T} d_i y_i : (x,y) \in X_P, x_k = 1 \right\} \end{aligned} \tag{23}$$

can be computed efficiently by solving a linear program. Moreover, the linear program required to compute $\hat{\sigma}_{(k)}$ differs from the one required for $\hat{\sigma}_{(k-1)}$ in two bound constraints, corresponding to $x_{(k-1)}$ and $x_{(k)}$, and one objective coefficient, corresponding to $x_{(k-1)}$. Therefore, using the simplex method with warm starts, each $\hat{\sigma}_{(k)}$ can be computed efficiently, using only a small number of simplex pivots.

6. Computational experiments

In this section we report computational experiments performed to test the effectiveness of the polymatroid inequalities in solving SOCMIO problems with a branch-and-cut algorithm. In Section 6.1 we test the inequalities introduced in Sections 3 and 4 on problems with bounded continuous variables, in Section 6.2 we test the strengthened polymatroid inequalities introduced in Section 5 for problems with cardinality constraints, and finally in Section 6.3 we combine the ideas presented in the paper to solve instances with nondiagonal quadratic terms. All experiments are done using CPLEX 12.6.2 solver on a workstation with a 2.93GHz Intel®Core$^{\text{TM}}$ i7 CPU and 8 GB main memory and with a single thread. The time limit is set to two hours and CPLEX' default settings are used unless specified otherwise. The inequalities are added only at the root node using callback functions.

6.1. **Instances with bounded continuous variables.** In this section we test the effectiveness of the polymatroid inequalities (8) and (14) in solving optimization problems of the form

$$\min\{-a'x - b'y + \Omega z : (x, y, z) \in L_\sigma\} \tag{24}$$

with $\sigma = 0$ and compare them with default CPLEX with no user cuts. For two numbers $\ell < u$, let $U[\ell, u]$ denote the continuous uniform distribution between ℓ and u. The data for the model is generated as follows: $a_i \sim U[0, 1]$, $\sqrt{c_i} \sim U[0.85a_i, 1.15a_i]$ for $i \in N$, $b_j \sim U[0, 1]$, $\sqrt{d_j} \sim U[0.85b_j, 1.15b_j]$ for $j \in M$, and Ω is the solution[1] of

$$-a(N) - b(M) + \Omega\sqrt{c(N) + d(M)} = 0.$$

Inequalities (8) are added as linear cuts in an extended formulation, as described in Remark 2. For $p = 2$, inequalities (14) are of the form $f(x, y) \leq z$, where

$$f(x, y) = \sqrt{\left(\sqrt{\sigma + \sum_{i\in T} d_i y_i^2} + \pi' x\right)^2 + \sum_{i\in M\setminus T} d_i y_i^2}.$$

As only linear inequalities can be added through callbacks in CPLEX (as of version 12.6.2), we utilize the gradient inequalities for (14). Thus, given a fractional solution $(\bar{x}, \bar{y})$, we add the linear underestimator $g(x, y) \leq z$, where

$$g(x, y) = f(\bar{x}, \bar{y}) + \nabla_x f(\bar{x})'(x - \bar{x}) + \nabla_y f(\bar{y})'(y - \bar{y}).$$

In particular, we have that

$$g(x, y) = \psi + \frac{1}{\psi}\left(\eta\pi'(x - \bar{x}) + \zeta\sum_{i\in T} d_i\bar{y}_i(y_i - \bar{y}_i) + \sum_{i\in M\setminus T} d_i\bar{y}_i(y_i - \bar{y}_i)\right),$$

where

$$\eta = \sqrt{\sigma + \sum_{i\in T} d_i\bar{y}_i^2} + \pi'\bar{x},$$

$$\zeta = \frac{\eta}{\sqrt{\sigma + \sum_{i\in T} d_i\bar{y}_i^2}},$$

$$\psi = \sqrt{\eta^2 + \sum_{i\in M\setminus T} d_i\bar{y}_i^2}.$$

[1] This choice of Ω ensures that the linear and nonlinear components are well-balanced, resulting in challenging instances with large integrality gap.

A greedy heuristic is used to choose $T \subseteq M$ for inequalities (14): if $\bar{y}$ satisfies $\bar{y}_{(1)} \geq \bar{y}_{(2)} \geq \ldots \geq \bar{y}_{(m)}$, then we check for violation inequalities for each T_i of the form $T_i = \{(1), (2), \ldots, (i)\}$ for $i = 0, \ldots, m$. When adding the gradient inequalities corresponding to (14), CPLEX' barrier algorithm is found to be more effective than using the default setting to solve the subproblems of the branch-and-bound tree. Therefore, we report the results for inequalities (14) with the barrier algorithm.

Table 1 presents the results for $n = 100$. Each row represents the average over five instances generated with the same parameters and shows the number of continuous variables (m), the initial gap (`igap`), the root gap improvement (`rimp`), the number of nodes explored (`nodes`), the time elapsed in seconds (`time`), and the end gap (`egap`)[in brackets, the number of instances solved to optimality (#)]. The initial gap is computed as $\texttt{igap} = \frac{t_{\text{opt}} - t_{\text{relax}}}{|t_{\text{opt}}|} \times 100$, where t_{opt} is the objective value of the best feasible solution at termination and t_{relax} is the objective value of the continuous relaxation. The end gap is computed as $\texttt{egap} = \frac{t_{\text{opt}} - t_{\text{bb}}}{|t_{\text{opt}}|} \times 100$, where t_{bb} is the objective value of the best lower bound at termination. The root improvement is computed as $\texttt{rimp} = \frac{t_{\text{root}} - t_{\text{relax}}}{t_{\text{opt}} - t_{\text{relax}}} \times 100$, where t_{root} is the value of the continuous relaxation after adding the valid inequalities to the formulation.

Table 1. Experiments with bounded continuous variables.

m	igap	cpx				inequality (8)				inequality (14) (barrier)			
		rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]
20	1,554.7	0.0	283,747	420	0.0[5]	90.4	19,976	628	0.0[5]	99.5	316	25	0.0[5]
50	724.6	0.0	1,887,926	2,223	0.0[5]	79.4	1,206,283	5,770	65.4[1]	98.8	1,635	857	0.0[5]
100	267.8	0.0	982,945	5,343	16.1[2]	70.1	615,494	7,200	54.6[0]	98.7	1,506	2,959	2.0[3]
Average		**0.0**	**1,051,539**	**2,662**	**5.4[12]**	**80.0**	**613,918**	**4,533**	**40.0[6]**	**99.0**	**1,152**	**1,280**	**0.7[13]**

We observe in Table 1 that the use inequalities (8), which do not exploit the upper bounds of the continuous variables, close 80.0% of the initial gap on average, but the gap improvement does not translate to better solution times or end gaps. On the other hand, inequalities (14), which exploit the upper bounds of the continuous variables, close 99% of the initial gap on average. This improves the performance of the algorithm substantially, reducing the average solution time by half and the end gap from from 5.4% to 0.7%.

6.2. **Instances with a cardinality constraint.** In this section we test the value of strengthening the polymatroid inequalities utilizing additional problem constraints. To do so, we solve optimization problems with a cardinality constraint:

$$\min_{x \in \{0,1\}^n} \left\{ -a'x + \Omega\sqrt{c'x} : \sum_{i=1}^{n} x_i \leq k \right\}, \tag{25}$$

where a and c are generated as in Section 6.1 and $\Omega = \Phi^{-1}(\alpha)$, where Φ is the cumulative distribution function of the normal distribution and $\alpha \in \{0.95, 0.975, 0.99\}$.

We set $n = 200$, and set k to be 15%, 20% and 25% of the total number of variables. Inequalities (5) and (18) are compared with default CPLEX. The coefficients of inequalities (18) are computed using linear programming with warm starts as outlined in Section 5.2—observe that, in this case, the coefficients (23) coincide with (17) since the feasible region is an integral polytope.

Table 2. Experiments with cardinality constraints.

k	α	igap	cpx				inequality (5)				inequality (18)			
			rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]
	0.95	4.4	23.7	7,150,715	2,528	0.3[4]	36.6	3,754,826	2,073	0.4[4]	48.9	2,614,446	1,510	0.2[4]
30	0.975	7.2	7.2	13,632,197	6,120	1.8[1]	23.7	9,573,199	5,945	1.8[1]	39.8	9,235,158	5,797	1.0[1]
	0.99	11.9	4.0	16,867,459	7,200	5.0[0]	14.7	10,899,169	7,200	5.7[0]	31.6	13,328,370	7,200	4.1[0]
Average			**11.6**	**12,550,124**	**5,283**	**2.4[5]**	**25.0**	**8,075,731**	**5,073**	**2.6[5]**	**40.1**	**8,392,658**	**4,836**	**1.8[5]**
	0.95	1.9	20.7	6,235,270	1,674	0.1[4]	70.5	620,389	261	0.0[5]	75.0	90,179	62	0.0[5]
40	0.975	3.3	9.6	13,961,488	4,360	0.4[3]	49.2	3,268,824	2,122	0.2[4]	57.3	2,729,459	1,557	0.2[4]
	0.99	5.6	6.0	15,334,782	6,738	1.8[1]	30.0	6,110,571	6,149	1.7[1]	42.6	5,222,829	5,799	1.2[1]
Average			**12.1**	**11,843,847**	**4,257**	**0.8[8]**	**49.9**	**3,333,261**	**2,844**	**0.6[10]**	**58.3**	**2,680,821**	**2,472**	**0.5[10]**
	0.95	1.0	8.9	270,852	72	0.0[5]	93.3	249	2	0.0[5]	93.3	98	2	0.0[5]
50	0.975	1.6	8.0	3,882,494	1,045	0.0[5]	81.3	316,625	221	0.0[5]	84.4	198,916	92	0.0[5]
	0.99	2.8	7.9	14,835,539	4,600	0.3[3]	57.3	4,695,268	3,480	0.2[3]	64.3	983,894	1,537	0.2[4]
Average			**8.3**	**6,329,628**	**1,906**	**0.1[13]**	**77.3**	**1,670,714**	**1,234**	**0.1[13]**	**80.7**	**394,293**	**544**	**0.1[14]**

Table 2 presents the results for each value of k and α. We see that for instances with $k = 50$, using inequalities (5) or (18) results in gap improvement of more than 75% and faster solutions times than default CPLEX. In particular, using inequalities (18) results in solutions times that are four times faster than default CPLEX on average. As expected, for instances with tighter cardinality constraints, inequalities (18), which exploit the cardinality constraints, are more effective than inequalities (5) in reducing the solution times as well as end gaps. On the other hand, when the cardinality constraint is loose, the effectiveness of both classes of inequalities improve.

6.3. **Instances with non-diagonal quadratic term and cardinality constraint.** Although the inequalities in this paper are developed for the diagonal case of the conic inequalities (1), they can, nevertheless, be used for the general non-diagonal case as well through a relaxation. Consider an optimization problem of the form

$$\min_{x \in \{0,1\}^n} \left\{ -a'x + \Omega\sqrt{x'Qx} : \sum_{i=1}^{n} x_i \leq k \right\}, \tag{26}$$

with $Q = D + Q_0$, where $Q_0 \succeq 0$, $D \succeq 0$ and D is diagonal. Given a general matrix $Q \succeq 0$, matrices Q_0 and D can be computed using the smallest eigenvalue (Frangioni and Gentile 2006) or solving an SDP (Frangioni and Gentile 2007). Alternatively, in many large-scale instances Q is a covariance matrix built through a *factor model*, in which case D is the diagonal matrix with the specific variances, $Q_0 = XFX'$, where $X \in \mathbb{R}^{n\times r}$ is the exposure matrix and $F \in \mathbb{R}^{r\times r}$ is the factor covariance matrix. Either way, given Q_0 and D, problem (26) can be reformulated as

$$\min_{(x,y)\in\{0,1\}^n\times\mathbb{R}_+} \left\{ -a'x + \Omega\sqrt{\sum_{i=1}^{n} D_{ii}x_i + y^2} : \sum_{i=1}^{n} x_i \le k, \sqrt{x'Q_0x} \le y \right\},$$

and the polymatroid inequalities can be applied to the diagonal objective.

In the computational experiments we generate the data using a factor model. Let $F = GG'$, with $G \in \mathbb{R}^{r\times r}$ and $G_{ij} \sim U[-1, 1]$, $X_{ij} \sim U[0, 1]$ with probability 0.2 and $X_{ij} = 0$ otherwise, $D_{ii} \sim U[0, \delta\bar{q}]$, where $\delta \ge 0$ is a diagonal dominance parameter and $\bar{q} = \frac{1}{N}\sum_{i\in N} Q_{0ii}$, and $a_i \sim U[0.85\sqrt{Q_{ii}}, 1.15\sqrt{Q_{ii}}]$. The parameter Ω is set as in Section 6.2. We let $n = 200$, $r = 40$ and k equal to 10%, 15%, and 20% of the number of the variables. The effectiveness of inequalities (8) and (18) are compared with default CPLEX. The inequalities are added using an extended formulation as described in Remark 2.

TABLE 3. Experiments with the non-diagonal case ($\delta = 0.5$).

k	α	igap	cpx				inequality (8)				inequality (18)			
			rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]	rimp	nodes	time	egap[#]
	0.95	1.7	22.6	9,557	74	0.0[5]	53.3	3,957	23	0.0[5]	55.6	2,367	17	0.0[5]
20	0.975	3.0	21.3	33,468	242	0.0[5]	53.5	13,316	86	0.0[5]	55.9	5,839	40	0.0[5]
	0.99	5.2	15.2	164,568	1,845	0.0[5]	52.8	80,735	730	0.0[5]	55.3	23,577	269	0.0[5]
Average			**19.7**	**69,198**	**720**	**0.0[15]**	**53.2**	**32,669**	**280**	**0.0[15]**	**55.6**	**10,594**	**109**	**0.0[15]**
	0.95	0.8	15.5	7,115	57	0.0[5]	53.3	1,656	11	0.0[5]	52.4	1,159	9	0.0[5]
30	0.975	1.3	14.9	18,901	135	0.0[5]	53.1	2,800	20	0.0[5]	54.0	2,095	15	0.0[5]
	0.99	2.3	5.7	76,675	1,005	0.0[5]	61.1	8,265	48	0.0[5]	62.1	5,131	30	0.0[5]
Average			**12.0**	**34,230**	**399**	**0.0[15]**	**55.8**	**4,240**	**26**	**0.0[15]**	**56.2**	**2,795**	**18**	**0.0[15]**
	0.95	0.4	23.3	2,910	18	0.0[5]	48.5	611	6	0.0[5]	50.5	577	6	0.0[5]
40	0.975	0.7	20.0	4,216	30	0.0[5]	54.3	884	7	0.0[5]	55.5	839	7	0.0[5]
	0.99	1.1	13.5	46,030	514	0.0[5]	55.9	2,493	18	0.0[5]	56.7	2,144	14	0.0[5]
Average			**18.9**	**17,719**	**187**	**0.0[15]**	**52.9**	**1,329**	**10**	**0.0[15]**	**54.2**	**1,187**	**9**	**0.0[15]**

Tables 3 and 4 present the results for different choices of the diagonal dominance parameter δ[2]. Observe that adding inequalities (8) or (18) closes the initial gaps by 45% to 75%, resulting in significant performance improvement over default CPLEX. In particular, using inequalities (18) for instances with $k = 20$ leads to seven times speed-up with $\delta = 0.5$ and two times speed-up with $\delta = 1$) and lower end gaps. Moreover, for instances with $k \ge 30$ using inequalities (18) results in at least an order-of-magnitude speed-up over default CPLEX. As in the previous section, inequalities (18), exploiting the cardinality constraint, are more effective than (8). The impact of both inequalities increases with higher diagonal dominance.

[2]Intuitively, if $\delta = 0.5$ then the factors explain 80% of the variance in the problem; if $\delta = 1.0$, then the factors explain 66% of the variance in the problem.

Table 4. Experiments with the non-diagonal case ($\delta = 1.0$).

k	α	igap	cpx rimp	cpx nodes	cpx time	cpx egap[#]	inequality (8) rimp	inequality (8) nodes	inequality (8) time	inequality (8) egap[#]	inequality (18) rimp	inequality (18) nodes	inequality (18) time	inequality (18) egap[#]
	0.95	2.9	21.6	64,283	927	0.0[5]	55.1	14,984	165	0.0[5]	59.1	6,233	68	0.0[5]
20	0.975	5.0	15.5	240,224	3,975	0.4[3]	44.4	189,826	3,390	0.4[3]	50.9	102,053	1,915	0.1[4]
	0.99	9.0	6.4	378,116	7,200	2.2[0]	35.7	477,553	7,200	1.9[0]	43.1	430,707	5,966	0.6[2]
Average			**14.5**	**227,541**	**4,034**	**0.9[8]**	**45.1**	**227,454**	**3,585**	**0.8[8]**	**51.0**	**179,664**	**2,650**	**0.2[11]**
	0.95	1.1	17.1	32,629	316	0.0[5]	77.2	1,082	12	0.0[5]	78.2	682	10	0.0[5]
30	0.975	2.0	12.5	150,756	2,046	0.1[4]	72.9	12,202	107	0.0[5]	75.5	4,896	39	0.0[5]
	0.99	3.5	10.5	258,866	3,679	0.5[3]	67.8	115,507	1,510	0.1[4]	70.6	59,106	511	0.0[5]
Average			**13.4**	**147,417**	**2,014**	**0.2[12]**	**72.6**	**42,930**	**543**	**0.0[14]**	**74.8**	**21,561**	**187**	**0.0[15]**
	0.95	0.6	23.9	6,522	64	0.0[5]	72.3	270	9	0.0[5]	74.8	192	8	0.0[5]
40	0.975	1.0	24.0	31,022	414	0.0[5]	71.0	823	12	0.0[5]	72.1	695	11	0.0[5]
	0.99	1.6	17.6	122,568	2,907	0.2[3]	73.9	4,416	37	0.0[5]	75.1	2,543	26	0.0[5]
Average			**21.8**	**53,371**	**1,128**	**0.1[13]**	**72.4**	**1,836**	**19**	**0.0[15]**	**74.0**	**1,143**	**15**	**0.0[15]**

Acknowledgement

The authors are supported, in part, by grant FA9550-10-1-0168 from the Office of Assistant Secretary of Defense for Research and Engineering.

References

Ahmed, S. and Atamtürk, A. (2011). Maximizing a class of submodular utility functions. *Mathematical programming*, 128:149–169.

Aktürk, M. S., Atamtürk, A., and Gürel, S. (2009). A strong conic quadratic reformulation for machine-job assignment with controllable processing times. *Operations Research Letters*, 37:187–191.

Aktürk, M. S., Atamtürk, A., and Gürel, S. (2010). Parallel machine match-up scheduling with manufacturing cost considerations. *Journal of Scheduling*, 13:95–110.

Alizadeh, F. and Goldfarb, D. (2003). Second-order cone programming. *Mathematical programming*, 95:3–51.

Anstreicher, K. M. (2012). On convex relaxations for quadratically constrained quadratic programming. *Mathematical programming*, 136:233–251.

Atamtürk, A., Berenguer, G., and Shen, Z.-J. (2012). A conic integer programming approach to stochastic joint location-inventory problems. *Operations Research*, 60:366–381.

Atamtürk, A. and Bhardwaj, A. (2015). Supermodular covering knapsack polytope. *Discrete Optimization*, 18:74–86.

Atamtürk, A. and Narayanan, V. (2008). Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36:618–622.

Atamtürk, A. and Narayanan, V. (2009). The submodular knapsack polytope. *Discrete Optimization*, 6:333–344.

Atamtürk, A. and Narayanan, V. (2010). Conic mixed-integer rounding cuts. *Mathematical Programming*, 122:1–20.

Atamtürk, A. and Narayanan, V. (2011). Lifting for conic mixed-integer programming. *Mathematical programming*, 126:351–363.

Belotti, P., Góez, J. C., Pólik, I., Ralphs, T. K., and Terlaky, T. (2015). A conic representation of the convex hull of disjunctive sets and conic cuts for integer second order cone optimization. In *Numerical Analysis and Optimization*, pages 1–35. Springer.

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust optimization*. Princeton University Press.

Ben-Tal, A. and Nemirovski, A. (1998). Robust convex optimization. *Mathematics of operations research*, 23:769–805.

Ben-Tal, A. and Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations research letters*, 25:1–13.

Ben-Tal, A. and Nemirovski, A. (2001). On polyhedral approximations of the second-order cone. *Mathematics of Operations Research*, 26:193–205.

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Bonami, P. (2011). Lift-and-project cuts for mixed integer convex programs. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 52–64. Springer.

Ceria, S. and Soares, J. (1999). Convex programming for disjunctive convex optimization. *Mathematical Programming*, 86:595–614.

Çezik, M. T. and Iyengar, G. (2005). Cuts for mixed 0-1 conic programming. *Mathematical Programming*, 104:179–202.

Dadush, D., Dey, S. S., and Vielma, J. P. (2011). The split closure of a strictly convex body. *Operations Research Letters*, 39:121–126.

Edmonds, J. (1970). Submodular functions, matroids, and certain polyhedra. In Guy, R., Hanani, H., Sauer, N., and Schönenheim, J., editors, *Combinatorial Structures and Their Applications*, pages 69–87. Gordon and Breach.

Frangioni, A. and Gentile, C. (2006). Perspective cuts for a class of convex 0–1 mixed integer programs. *Mathematical Programming*, 106:225–236.

Frangioni, A. and Gentile, C. (2007). Sdp diagonalizations and perspective cuts for a class of nonseparable miqp. *Operations Research Letters*, 35:181–185.

Günlük, O. and Linderoth, J. (2010). Perspective reformulations of mixed integer nonlinear programs with indicator variables. *Mathematical programming*, 124:183–205.

Hijazi, H., Bonami, P., and Ouorou, A. (2013). An outer-inner approximation for separable mixed-integer nonlinear programs. *INFORMS Journal on Computing*, 26:31–44.

Kılınç, M., Linderoth, J., and Luedtke, J. (2010). Effective separation of disjunctive cuts for convex mixed integer nonlinear programs. *Optimization Online*.

Kılınç-Karzan, F. and Yıldız, S. (2015). Two-term disjunctions on the second-order cone. *Mathematical Programming*, 154:463–491.

Lobo, M. S., Vandenberghe, L., Boyd, S., and Lebret, H. (1998). Applications of second-order cone programming. *Linear algebra and its applications*, 284:193–228.

Lubin, M., Yamangil, E., Bent, R., and Vielma, J. P. (2016). Polyhedral approximation in mixed-integer convex optimization. *arXiv preprint arXiv:1607.03566*.

Modaresi, S., Kılınç, M. R., and Vielma, J. P. (2016). Intersection cuts for nonlinear integer programming: Convexification techniques for structured sets. *Mathematical Programming*, 155:575–611.

Poljak, S. and Wolkowicz, H. (1995). Convex relaxations of (0, 1)-quadratic programming. *Mathematics of Operations Research*, 20:550–561.

Shen, Z.-J. M., Coullard, C., and Daskin, M. S. (2003). A joint location-inventory model. *Transportation science*, 37:40–55.

Stubbs, A. R. and Mehrotra, S. (1999). A branch-and-cut method for 0-1 mixed convex programming. *Mathematical Programming*, 86:515–532.

Tawarmalani, M. and Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103:225–249.

Vielma, J. P., Dunning, I., Huchette, J., and Lubin, M. (2015). Extended formulations in mixed integer conic quadratic programming. *arXiv preprint arXiv:1505.07857*.

Appendix A. Convex hull of L^2_σ

A point (x, y, z) belongs to $\text{conv}(L^2_\sigma)$ if and only if there exist $x_1, x_2, y_1, y_2, z_1, z_2, \lambda$ such that the system

$$x = (1-\lambda)x_1 + \lambda x_2 \tag{27}$$

$$y = (1-\lambda)y_1 + \lambda y_2 \tag{28}$$

$$z = (1-\lambda)z_1 + \lambda z_2 \tag{29}$$

$$z_1 \geq \sqrt{\sigma + dy_1^2} \tag{30}$$

$$z_2 \geq \sqrt{\sigma + c + dy_2^2} \tag{31}$$

$$0 \leq y_1, y_2 \leq 1, \; x_1 = 0, \; x_2 = 1 \tag{32}$$

is feasible. Observe that from (27) and (32) we can conclude that $\lambda = x$. Also observe that from (27), (30) and (31) we have that

$$\begin{aligned} & z = (1-x)z_1 + xz_2 \\ \Leftrightarrow\, & z \geq (1-x)\sqrt{\sigma + dy_1^2} + x\sqrt{\sigma + c + dy_2^2}. \end{aligned}$$

Therefore, the system is feasible if and only if

$$\begin{aligned}
z \geq \min_{y_1,y_2}\ & (1-x)\sqrt{\sigma + dy_1^2} + x\sqrt{\sigma + c + dy_2^2} && (33)\\
\text{s.t. } & y = (1-x)y_1 + xy_2 && (\gamma)\\
& y_1 \leq 1 && (\alpha_1)\\
& y_2 \leq 1 && (\alpha_2)\\
& y_1 \geq 0 && (\beta_1)\\
& y_2 \geq 0, && (\beta_2)
\end{aligned}$$

and let γ, α and β be the dual variables of the optimization problem above. From KKT conditions for variables y_1 and y_2 we find that

$$\begin{aligned}
-(1-x)\frac{dy_1}{\sqrt{\sigma + dy_1^2}} &= \gamma(1-x) + \alpha_1 - \beta_1\\
-x\frac{dy_2}{\sqrt{\sigma + c + dy_2^2}} &= \gamma x + \alpha_2 - \beta_2\\
\implies \quad \frac{y_1}{\sqrt{\sigma + dy_1^2}} + \bar{\alpha}_1 - \bar{\beta}_2 &= \frac{y_2}{\sqrt{\sigma + c + dy_2^2}} + \bar{\alpha}_2 - \bar{\beta}_2, \qquad (34)
\end{aligned}$$

where $\bar{\alpha}, \bar{\beta}$ correspond to α and β after scaling. We can deduce from (34) and complementary slackness that $y_1, y_2 > 0$ (unless $y = 0$) and that $y_1 \leq y_2$. Therefore, in an optimal solution either $0 < y_1, y_2 < 1$ (and $\bar{\alpha} = \bar{\beta} = 0$) or $y_2 = 1$ (and $\bar{\alpha}_2 \geq 0$). If $\bar{\alpha} = \bar{\beta} = 0$, then

$$\begin{aligned}
y_1^* &= y\frac{\sqrt{\sigma}}{x\sqrt{c+\sigma} + (1-x)\sqrt{\sigma}} \qquad \text{and}\\
y_2^* &= y\frac{\sqrt{c+\sigma}}{x\sqrt{c+\sigma} + (1-x)\sqrt{\sigma}}
\end{aligned}$$

satisfy conditions (34) and (28). Moreover, if

$$\begin{aligned}
& y_2^* \leq 1\\
\Leftrightarrow\ & y \leq \frac{x\sqrt{c+\sigma} + (1-x)\sqrt{\sigma}}{\sqrt{c+\sigma}} = x + (1-x)\sqrt{\frac{\sigma}{c+\sigma}},
\end{aligned}$$

then y_1^*, y_2^* also satisfy bound constraints, and thus correspond to an optimal solution to the optimization problem. Replacing in (33), we find that

$$z \geq \sqrt{\left(\sqrt{\sigma} + x(\sqrt{c+\sigma} - \sqrt{\sigma})\right)^2 + dy^2}$$

when $y \leq x + (1-x)\sqrt{\frac{\sigma}{\sigma+c}}$. On the other hand, if $y_1^* > 1$, an optimal solution to the optimization problem is given by $\bar{y}_2 = 1$ and $\bar{y}_1 = \frac{y-x}{1-x}$. Replacing in (33)

$$z \geq \sqrt{\sigma(1-x)^2 + d(y-x)^2} + x\sqrt{\sigma + c + d}$$

when $y \geq x + (1-x)\sqrt{\frac{\sigma}{\sigma+c}}$.