SUBMODULARITY IN CONIC QUADRATIC MIXED 0-1 OPTIMIZATION

ALPER ATAMTÜRK AND ANDRÉS GÓMEZ

ABSTRACT. We describe strong convex valid inequalities for conic quadratic mixed 0-1 optimization. The inequalities exploit the submodularity of the binary restrictions and are based on the polymatroid inequalities over binaries for the diagonal case. We prove that the convex inequalities completely describe the convex hull of a single conic quadratic constraint as well as the rotated cone constraint over binary variables and unbounded continuous variables. We then generalize and strengthen the inequalities by incorporating additional constraints of the optimization problem. Computational experiments on mean-risk optimization with correlations, assortment optimization, and robust conic quadratic optimization indicate that the new inequalities strengthen the convex relaxations substantially and lead to significant performance improvements.

Keywords: Polymatroid, submodularity, second-order cone, nonlinear cuts, robust optimization, assortment optimization, value-at-risk.

October 2016; February 2018
1. Introduction

Submodular set functions play an important role in many fields and have received substantial interest in the literature as they can be minimized in polynomial time (Grötschel et al. [1981], Schrijver [2000], Orlin [2009]). Combinatorial optimization problems such as the min-cut problem, entropy minimization, matroids, binary quadratic function minimization with a non-positive matrix are special cases of submodular minimization (Fujishige [2005]). The utilization of submodularity, however, has been mainly restricted to 0-1 optimization problems although many practical problems involve continuous variables as well.

The goal in this paper is to exploit submodularity to derive valid inequalities for mixed 0-1 minimization problems with a conic quadratic objective:

\[
\min a'x + \sqrt{x'Qx} : x \in X \subseteq \{0, 1\}^n \times \mathbb{R}^m, \tag{1}
\]

or a conic quadratic constraint:

\[
a'x + \sqrt{x'Qx} \leq b, \quad x \in X \subseteq \{0, 1\}^n \times \mathbb{R}^m, \tag{2}
\]

where \(\Omega \in \mathbb{R}_+\) and \(Q\) is a symmetric positive semidefinite matrix. Formulations (1) and (2) are frequently used to model mean-risk problems. In particular, (1) is value-at-risk minimization and (2) is a probabilistic constraint for a random variable \(r'x\), with \(r \sim N(a, Q)\). They are also used to model conservative robust formulations with an appropriate value of \(\Omega\) if \(r\) is not normally distributed (Ben-Tal et al. [2009a]).

Introducing an auxiliary variable \(z\) to represent the square root term \(\sqrt{x'Qx}\) in (1)–(2), we write

\[
f(x) = \sqrt{x'Qx} \leq z, \quad x \in X \subseteq \{0, 1\}^n \times \mathbb{R}^m.
\]

The motivation for this study stems from the fact that \(f\) is submodular for the simplest nontrivial non-convex case: when \(Q\) is diagonal and \(m = 0\) (Shen et al. [2003]). Therefore, one may expect submodularity to play a significant role in analyzing and solving optimization problems with a general conic quadratic objective or constraint as submodularity is contained in a basic form.

Toward this goal we consider the conic quadratic mixed-binary set

\[
H_X = \left\{(x, y) \in X, z \in \mathbb{R}_+ : \sigma + \sum_{i=1}^n c_i x_i + \sum_{i=1}^m d_i y_i^2 \leq z^2 \right\},
\]

where \(X \subseteq D = \{0, 1\}^n \times \mathbb{R}_+^m, c \in \mathbb{R}_+^n, d \in \mathbb{R}_+^m\) and \(\sigma \geq 0\) and derive strong inequalities for it. Note that \(H_D\) is the mixed-integer epigraph of the function

\[
f(x, y) = \sqrt{\sigma + \sum_{i=1}^n c_i x_i + \sum_{i=1}^m d_i y_i^2}.
\]
The set $H_X$ arises frequently in mixed-integer optimization models, well beyond the natural extension to mixed 0-1 mean-risk minimization or chance constrained optimization with uncorrelated random variables. In particular, in Section 2 we describe applications on optimization with correlated random variables, inventory and scheduling problems, assortment optimization, ratio minimization, facility location problems, and robust conic quadratic mixed-binary optimization with discrete uncertainty.

Let $H_B$ denote the pure binary case of $H_D$ with $m = 0$, for which $f$ is submodular. While the convex hull of $H_B$, $\text{conv}(H_B)$, is a polyhedral set and well-understood, that is not the case for the mixed-integer set $H_D$. Note, however, that for a fixed $y$, $f$ is submodular in $x$. By exploiting this partial submodularity for the mixed-integer case, in this paper, we give a complete nonlinear inequality description of $\text{conv}(H_D)$.

We review the polymatroid inequalities for the pure binary case in Section 3. Moreover, we show that the resulting nonlinear inequalities are also strong for the rotated conic quadratic mixed 0-1 set

$$R_X = \left\{(x,y) \in X, (w,z) \in \mathbb{R}^2_+ : \sigma + \sum_{i=1}^n c_i x_i + \sum_{i=1}^m d_i y_i^2 \leq 4wz \right\}.$$

Observe that even for the binary case ($m = 0$), the definition of $R_X$ has the product of two continuous variables $w, z$ on the right-hand-side. Therefore, the existing polymatroid inequalities from the binary case cannot be directly applied to $R_X$. Several of the applications in Section 2 are modeled using the rotated cone set $R_X$.

**Literature review.** A major difficulty in developing strong formulations for mixed-integer nonlinear sets such as $H_X$ is that the corresponding convex hulls are not polyhedral, while most of the theory and methodology developed for mixed-integer optimization focus on the polyhedral case. Recently, there has been an increasing effort to generalize methods from the linear case to the nonlinear case, including Gomory cuts (Cezik and Iyengar 2005), MIR cuts (Atamtürk and Narayanan 2010), cut generating functions (Santana and Dev 2017), minimal valid inequalities (Kılınç-Karzan 2015), conic lifting (Atamtürk and Narayanan 2011), intersection cuts, disjunctive cuts, and lift-and-project cuts (Ceria and Soares 1999, Stubbs and Mehrotra 1999). Kılınç et al. (2010) and Bonami (2011) discuss the separation of split cuts using outer approximations and nonlinear programming. Additionally, some classes of nonlinear sets have been studied in detail: Belotti et al. (2015) study the intersection of a convex set and a linear disjunction, Modaresi and Vielma (2014) study intersections of a quadratic and a conic quadratic inequalities, Kılınç-Karzan and Yıldız (2015) study disjunctions on the second order cone, Burer and Kılınç-Karzan (2017) study the intersection of a non-convex quadratic and a conic quadratic inequality, Dadush et al. (2011a) and Dadush et al. (2011b) investigate the the Chvátal-Gomory closure of convex sets and Dadush et al. (2011c) investigate the split closure of a convex set. These inequalities are general and do not exploit any special structure.
Another stream of research for mixed-integer nonlinear optimization 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 problems, they tend to be far more effective than the general cuts that ignore any problem 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. Atamtürk and Joen (2017) give strong valid inequalities for mean-risk minimization with variable upper bounds.

Closely related to this paper, Atamtürk and Narayanan (2008) study $H_B$ in the context of mean-risk minimization. Yu and Ahmed (2017) study the generalization with a cardinality constraint, i.e., $H_Y$ where $Y = \{x \in \{0,1\}^n : \sum_{i=1}^n x_i \leq k\}$. However, more general sets have not been considered in the literature. More importantly perhaps, the valid inequalities derived for the pure-binary case have limited use for mixed-integer problems or even for pure-binary problems with correlated random variables (non-diagonal matrix $Q$).

**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. Define $N = \{1, \ldots, n\}$ and $M = \{1, \ldots, m\}$. Let $\text{conv}(X)$ denote the convex hull of $X$. Given a vector $a \in \mathbb{R}^n$ and $S \subseteq \{1, \ldots, n\}$, let $\text{diag}(a)$ denote the $n \times n$ diagonal matrix $A$ with $A_{ii} = a_i$, and let $a(S) = \sum_{i \in S} a_i$. Let $B = \{0,1\}^n$ and $C = \{0,1\}^n \times [0,1]^m$.

**Outline.** The rest of the paper is organized as follows. In Section 2, we discuss applications in which sets $H_X$ and $R_X$ arise naturally. In Section 3, we review the existing results for $H_B$ and $H_C$. In Section 4, we show that a nonlinear generalization of the polymatroid inequalities is sufficient to describe the convex hull of $H_D$. In Section 5, we study the bounded set $H_C$, give an explicit convex hull description for the case $n = m = 1$, and propose strong valid inequalities for the general case. In Section 6, we describe a strengthening procedure for the nonlinear polymatroid inequalities for any mixed-integer set $X$; the approach generalizes the lifting method of Yu and Ahmed (2017) for the pure-binary cardinality constrained case. In Section 7, we test the effectiveness of the proposed inequalities for a variety of problems discussed in Section 2. Section 8 concludes the paper.

### 2. Applications

In this section, we present six mixed 0-1 optimization problems in which sets $H_X$ and $R_X$ arise naturally.
2.1. Mean-risk minimization and chance constraints with uncorrelated random variables. Conic quadratic constraints are frequently used to model probabilistic optimization with Gaussian distributions (e.g. Birge and Louveaux 2011). In particular, if $a_i$, $c_i$ denote the mean and variance of random variables $p_i$, $i \in N$, and $b_i$, $d_i$ the mean and variance of random variables $q_i$, $i \in N$, and all variables are independent, then

$$\min_{(x,y,z) \in H_X} a'x + b'y + \Phi^{-1}(\alpha)z$$

corresponds to the value-at-risk minimization problem over $X$, where $\Phi$ is the c.d.f. of the standard normal distribution and $0.5 < \alpha < 1$. Alternatively, the chance constraint $\Pr(\tilde{p}'x + \tilde{q}'y \leq b) \geq \alpha$ is equivalent to $a'x + b'y + \Phi^{-1}(\alpha)z \leq b$, $(x, y, z) \in H_X$. Models with $H_X$ also arise in robust optimization problems with ellipsoidal uncertainty sets (Ben-Tal and Nemirovski 1998, 1999, Ben-Tal et al. 2009b).

2.2. Mean-risk minimization and chance constraints with correlated random variables. If $\tilde{p} \sim \mathcal{N}(a, Q)$, where $a$ is the mean vector and $Q \succeq 0$ is the covariance matrix, then the value-at-risk minimization or chance constrained optimization with 0–1 variables involve constraints of the form $\sqrt{x'Qx} \leq z$.

A standard technique in quadratic optimization consists in utilizing the diagonal entries of matrices to construct strong convex relaxations (e.g. Poljak and Wolkowicz 1995, Anstreicher 2012). In particular, for $x \in \{0, 1\}^n$, we have

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

with $c \in \mathbb{R}_+^n$ such that $Q - \text{diag}(c) \succeq 0$. This transformation is based on the ideal (convex hull) representation of the separable quadratic term $x'\text{diag}(c)x$ as $c'x$ for $x \in \{0, 1\}^n$. Using a similar idea and introducing a continuous variable $y \in \mathbb{R}_+$, we get

$$\sqrt{x'Qx} \leq z \iff (x, y, z) \in H_X \text{ and } \sqrt{x'(Q - \text{diag}(c))x} \leq y.$$

The approach presented here can also be used for mixed-binary sets $X$.

2.3. Robust Value-at-Risk with discrete uncertainty. We consider robust Value-at-Risk (VaR) optimization with discrete uncertainty, i.e., optimization problems of the form

$$\omega^* = \min_{x \in X} \max_{w \in W} a_0'x + \sum_{i=1}^p (a_i'x) w_i + \sqrt{x'Q_0x + \sum_{i=1}^p (x'Q_i x) w_i,}$$

where $W = \{w \in \{0, 1\}^p : Aw \leq b\}$ is the uncertainty set and $x \in X \subseteq \{0, 1\}^n$ are the decision variables of the robust problem. Problem (4) arises when modeling robustness with respect to a set of potential adverse events (e.g., natural disasters, disruptions, enemy attacks), in which case $W = \{w \in \{0, 1\}^p : \sum_{i=1}^p w_i \leq b\}$ for a positive integer $b$. In this setting $(a_0, Q_0) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$ denotes the nominal expected value and covariance matrix, and $(a_i, Q_i) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$ denotes the expected increase in costs and covariances if scenario $i = 1, \ldots, p$ is realized.

Although solving the inner maximization problem in (4) is $NP$-hard for a fixed value of $x$, Atamtürk and Gómez (2017) show that it can be approximated by a
constant factor by introducing additional variables $u \in \mathbb{R}_+^p$, $v \in \mathbb{R}^{p+1}_+$, $y \in \mathbb{R}^{p+1}_+$, and $\gamma \in \mathbb{R}_+^p$ as

$$
\omega_a = \min \frac{1}{4} u + a'_0 x + v_0 + b' \gamma \\
\text{s.t. } A'_i \gamma \geq a'_i x + v_i \\
\sqrt{x'(Q_i - \text{diag}(c_i))x} \leq y_i \\
(x, y_i, u, v_i) \in X, \\
\gamma \geq 0,
$$

where $A'_i$ denotes the transpose of the $i$-th column of $A$. Let $\omega_n = \min_{x \in X} a'_0 x + \sqrt{x'Q_0x}$ be the optimal objective value of the nominal problem and let

$$
\omega^* := \omega_n + \frac{\omega_a - \omega_n}{1.25}.
$$

If $W$ is an integral polytope, e.g., $W$ consists of a single cardinality constraint, then $\omega^* \leq \omega_a \leq \omega^*$. In other words, the approximate robust counterpart (RC) is more conservative than the exact robust problem (4), but by a small amount only. In particular, if $\omega_n \geq 0$, then (RC) is a 1.25-approximation of (4). The computations in Atamtürk and Gómez (2017) indicate that the empirical gap is much smaller, and our computations in Section 7 show that the solutions delivered by solving (RC) are not overly conservative.

2.4. Lot-sizing and scheduling problems. Inventory problems with economic order quantity involve expressions of the form $kDQ$, where $D$ is the demand, $Q$ is the lot size, and $k$ is a fixed cost for ordering inventory. In simple settings, the optimal lot size $Q^*$ can be expressed explicitly (Nahmias 2001), but in more complex settings, where the demand is a linear function of discrete variables, e.g., in joint location-inventory problems (Özsen et al. 2008, Atamtürk et al. 2012) this is not possible. In such cases, the order costs involve expressions of the form

$$
c'_x Q \leq z \Leftrightarrow (x, Q, z) \in R_B.
$$

The ratio (7) also arises in scheduling, specifically in the economic lot scheduling problem (Bollapragada and Rao 1999, Bulut and Taşgetiren 2014, Pesenti and Ükovich 2003, Sahinidis and Grossmann 1991). In this context, $c$ is the vector of setup costs/times and $Q$ denotes a production cycle length, thus $z$ in (7) corresponds to setup costs/times per unit time. Expression (7) also arises in the plant design and scheduling problems to model the profitability or productivity of the plant (Castro et al. 2005, 2009).

2.5. Binary linear fractional problems. Generalizing the models in Section 2.4, binary linear fractional problems are optimization problems with constraints of the
form
\[
\frac{c_0 + \sum_{i=1}^{n} c_i x_i}{a_0 + \sum_{i=1}^{n} a_i x_i} \leq z \Leftrightarrow c_0 + \sum_{i=1}^{n} c_i x_i^2 \leq z w, \quad w = a_0 + \sum_{i=1}^{n} a_i x_i
\]
\[
\Leftrightarrow (x, w, z) \in R_B, \text{ with } w = a_0 + \sum_{i=1}^{n} a_i x_i
\]
where \(a_i, c_i \geq 0\) for \(i = 0, \ldots, n\). Note that a lower bound on the ratio can also be expressed similarly by complementing variables. Binary fractional optimization arises in numerous applications including assortment optimization with mixtures of multinomial logits (Méndez-Díaz et al. 2014, Sen et al. 2015), WLAN design (Amaldi et al. 2011), facility location problems with market share considerations (Tawarmalani et al. 2002), and cutting stock problems (Gilmore and Gomory 1963), among others; see also the survey Borrero et al. (2016a) and the references therein.

Applications of binary linear fractional optimization are abundant in network problems. For example, given a graph \(G = (V, E)\), problems of the form
\[
\min \left\{ \frac{\sum_{(i,j) \in E} c_{ij} x_{ij}}{\sum_{i \in V} a_i w_i} : x_{ij} \geq |w_i - w_j|, (i, j) \in E, w \in W \subseteq \{0, 1\}^V, x \in \{0, 1\}^E \right\}
\] (8)
arise in the study of expander graphs (Davidoff et al. 2003); in particular, the optimal value of (8) with \(c = 1, a = 1\) and \(W = \{w \in \{0, 1\}^V : 1 \leq \sum_{i \in V} w_i \leq 0.5|V|\}\) corresponds to the Cheeger constant of the graph. See Hochbaum (2010), Hochbaum et al. (2013) for other fractional cut problems arising in image segmentation, and see Prokopyev et al. (2009) for a discussion of other ratio problems in graphs arising in facility location.

2.6. Sharpe ratio. Let \(a_i, c_i\) be the mean and variance of normally distributed independent random variables \(\tilde{p}_i\), \(i \in N\) as in Section 2.1. A natural alternative to mean-risk minimization for a risk-adverse decision maker is, given a budget \(b\), to maximize the probability of meeting the budget; that is,
\[
\max_{x \in X} \Pr(\tilde{p}' x \leq b).
\] (9)

Problems of the form (9) are considered in Nikolova et al. (2006) in the context of the stochastic shortest path problem.

Assuming there is a solution \(x \in X\) satisfying \(a' x \leq b\), note that
\[
\Pr(\tilde{p}' x \leq b) = \Pr\left(\frac{\tilde{p}' x - a' x}{\sqrt{c' x}} \leq \frac{b - a' x}{\sqrt{c' x}}\right) = \Phi\left(\frac{b - a' x}{\sqrt{c' x}}\right).
\]
Since \(\Phi\) is monotone non-decreasing and \(b - a' x \geq 0\) for any optimal solution, we see that (9) is equivalent to maximizing \(\frac{b - a' x}{\sqrt{c' x}}\). Observe that the resulting objective corresponds to maximizing the reward-to-volatility or Sharpe ratio (Sharpe 1994), a well-known risk measure often used in finance.
Maximizing the Sharpe ratio is equivalent to minimizing \( \sqrt{c'x} \). Therefore, we can restate (9) as

\[
\min z \\
\text{s.t. } w = b - a'x \leq wz \quad \text{(10)} \\
x \in X, \; w, z \geq 0. \quad \text{(11)}
\]

Constraint (10) is not conic quadratic. Note, however, for \( w, z \geq 0 \) we have

\[
\sqrt{c'x} \leq wz \iff \sqrt{4 (4\sqrt{c'x})^2 + (w - z)^2} \leq w + z.
\]

Then one gets a convex relaxation by replacing the non-convex term \( 4\sqrt{c'x} \) by its convex lower bound \( \sum_{i \in N} c_i x_i^4 \). The resulting conic quadratic representable convex inequality can be written as

\[
\sqrt{\sum_{i \in N} c_i x_i^4} \leq wz.
\]

As we will show in Section 4.2, a nonlinear version of the extended polymatroid inequalities corresponding to the submodular function \( \bar{h}(x) = 2 \sqrt{c'x} \) is sufficient to describe the convex hull of the set given by (10)–(11) for \( X = \{0,1\}^n \) (Remark 5).

3. Preliminaries

In this section we review the polymatroid inequalities for the binary case and an extension to the mixed 0-1 case. Given \( \sigma \geq 0 \) and \( c_i > 0 \), \( i \in N \), consider the set

\[
H_B = \left\{ (x, z) \in B \times \mathbb{R}_+: \sqrt{\sigma + \sum_{i \in N} c_i x_i} \leq z \right\}. \quad \text{(12)}
\]

Observe that \( H_B \) is the binary restriction of \( H_D \) obtained by setting \( y = 0 \) and it is the union of finite number of line segments; therefore, its convex hull is polyhedral. For a given permutation \((1), (2), \ldots, (n)\) of \( N \), let

\[
\sigma(k) = \sigma(k) + \sigma(k-1), \quad \text{and } \sigma(0) = \sigma, \\
\pi(k) = \sqrt{\sigma(k)} - \sqrt{\sigma(k-1)}, \quad \text{(13)}
\]

and define the polymatroid inequality as

\[
\sum_{i=1}^{n} \pi(i) x(i) \leq z - \sqrt{\sigma}. \quad \text{(14)}
\]

Let \( \Pi_\sigma \) be the set of such coefficient vectors \( \pi \) for all permutations of \( N \). The set function defining \( H_B \) is non-decreasing submodular; therefore, \( \Pi_\sigma \) form the extreme points of a polymatroid \( \text{[Edmonds 1970]} \) and the convex hull of \( H_B \) is given by the set of all polymatroid inequalities and the bounds of the variables \( \text{[Lovász 1983]} \).
Proposition 1 (Convex hull of $H_B$).

$$\text{conv}(H_B) = \{(x, z) \in [0, 1]^N \times \mathbb{R}_+ : \pi'x \leq z - \sqrt{\sigma}, \ \forall \pi \in \Pi_\sigma\}.$$ 

As shown by Edmonds (1970), the maximization of a linear function over a polymatroid can be solved by the greedy algorithm; therefore, a point $(\bar{x}, \bar{z}) \in [0, 1]^N \times \mathbb{R}_+$ can be separated from $\text{conv}(H_B)$ via the greedy algorithm by sorting $\bar{x}_i$, $i \in N$ in non-increasing order in $O(n \log n)$ time.

Proposition 2 (Separation). A point $(\bar{x}, \bar{z}) \not\in \text{conv}(H_B)$ such that $\bar{x}(1) \geq \bar{x}(2) \geq \cdots \geq \bar{x}(n)$ is separated from $\text{conv}(H_B)$ by inequality (14).

Atamtürk and Narayanan (2008) consider the mixed-integer version of $H_B$:

$$H_C = \left\{(x, y, z) \in C \times \mathbb{R}_+ : \sqrt{\sigma + \sum_{i \in N} c_i x_i + \sum_{i \in M} d_i y_i^2} \leq z \right\},$$

where $d_i > 0$, $i \in M$ and give valid inequalities for $H_C$ based on the polymatroid inequalities. Without loss of generality, the upper bounds of the continuous variables in $H_C$ are set to one by scaling.

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

$$\pi'x + \sqrt{\sigma + \sum_{i \in T} d_i y_i^2} \leq z, \ \pi \in \Pi_{\sigma + d(T)}$$

are valid for $H_C$.

Inequalities (15) 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

$$\xi'x + \sqrt{\sigma + \sum_{i \in T} d_i y_i^2} \leq z$$

with $\xi \in \mathbb{R}^n$. Although inequalities (15) are the strongest possible among inequalities that are linear in $x$ and conic quadratic in $y$, they may be weak or dominated by other classes of nonlinear inequalities. In this paper we introduce stronger and more general inequalities than (15) for $H_C$.

4. The case of unbounded continuous variables

In this section we focus on the case with unbounded continuous variables, i.e., on $H_D$, where $D = \{0, 1\}^N \times \mathbb{R}_+^n$. In this case, since the continuous variables have no upper bound, the only class of valid inequalities of type (15) are the polymatroid inequalities

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

themselves from the “binary-only” relaxation by letting $T = \emptyset$. Inequalities (16) ignore the continuous variables and are, consequently, weak for $H_D$. Here, we define a new class of nonlinear valid inequalities and prove that they are sufficient to define the convex hull of $H_D$. 

Consider the inequalities
\[
(\sqrt{\sigma} + \pi' x)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2, \quad \pi \in \Pi_\sigma.
\] (17)

**Proposition 4.** Inequalities (17) are valid for \( H_D \).

**Proof.** Consider the extended formulation of \( H_D \) given by
\[
\hat{H}_D = \left\{ (x,y) \in D, (z,s) \in \mathbb{R}^2_+ : s^2 + \sum_{i \in M} d_i y_i^2 \leq z^2, \sigma + \sum_{i \in N} c_i x_i \leq s^2 \right\}.
\]
The validity of inequalities (17) for \( H_D \) follows directly from the validity of the polymatroid inequality \( \sqrt{\sigma} + \pi' x \leq s, \pi \in \Pi_\sigma \) (Proposition 1) for \( \hat{H}_D \).

\( \square \)

**Remark 1.** For \( M = \emptyset \) inequalities (17) reduce to the polymatroid inequalities (14).

**Remark 2.** Although inequalities (17) are nonlinear in the original space of variables, they can be represented as linear inequalities in the extended formulation \( \hat{H}_D \). Such a linear representation is desirable when using (17) as cutting planes in branch-and-cut algorithms.

**Remark 3.** Since inequalities (17) correspond to 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).

Inequalities (17) are obtained simply by extracting a submodular component from function \( f \). The approach can be generalized to sets of the form
\[
U = \left\{ (x,s) \in K, (y,z) \in \mathbb{R}^{m+1}_+ : s^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \right\},
\]
where \( K = \left\{ x \in \bar{X} \subseteq \{0,1\}^n, s \geq 0 : \sqrt{h(x)} \leq s \right\} \) and \( h : \{0,1\}^n \rightarrow \mathbb{R}_+ \) is an arbitrary nonnegative function. Observe that since \( K \) is a finite union of line segments, \( \text{conv}(K) \) is a polyhedron, and valid inequalities for \( K \) of the form \( \gamma' x \leq s, \gamma \in \Gamma \), can be lifted into valid nonlinear inequalities for \( U \) of the form
\[
(\gamma' x)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2.
\] (18)

Proposition 5 below implies inequalities of the form (18) are sufficient to describe \( \text{conv}(U) \) if \( \gamma' x \leq s, \gamma \in \Gamma \), are sufficient to describe \( \text{conv}(K) \).

**Proposition 5.** The convex hull of \( U \) is described as
\[
\text{conv}(U) = \left\{ (x,s) \in \text{conv}(K), (y,z) \in \mathbb{R}^{m}_+ : s^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \right\}.
\]
Proof. Consider the optimization of an arbitrary linear function over $U$,

$$
\begin{align*}
\text{min} & \quad -a'x - b'y + rz \\
\text{s.t.} & \quad s^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \\
& \quad (x, s) \in K \\
& \quad y \in \mathbb{R}^m_+, z \geq 0
\end{align*}
$$

(BP)

and over its convex relaxation,

$$
\begin{align*}
\text{min} & \quad -a'x - b'y + rz \\
\text{s.t.} & \quad s^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \\
& \quad (x, s) \in \text{conv}(K) \\
& \quad y \in \mathbb{R}^m_+, z \geq 0
\end{align*}
$$

(P1)

We prove that for any linear objective both (BP) and (P1) are unbounded or (P1) has an optimal solution that is integer in $x$. Without loss of generality, we can assume that $r > 0$ (if $r < 0$ then both problems are unbounded, and if $r = 0$ then (P1) reduces to a linear program over an integral polyhedron by setting $z$ sufficiently large, and is equivalent to (BP)), $r = 1$ (by scaling), $b_i > 0$ (otherwise $y_i = 0$ in any optimal solution), and $d_i = 1$ for all $i \in M$ (by scaling $y_i$).

Eliminating the variable $z$ from (P1) we restate the problem as

$$
\begin{align*}
\text{(P2)} & \quad \min \left\{ -a'x - b'y + \sqrt{s^2 + \sum_{i \in M} y_i^2} : (x, s) \in \text{conv}(K), \ y \in \mathbb{R}^m_+ \right\}.
\end{align*}
$$

Note that if $y = 0$ in an optimal solution of (P2), then (P2) reduces to a linear optimization over $\text{conv}(K)$, which has an optimal integer solution. Thus we assume that $\sqrt{s^2 + \sum_{i \in M} y_i^2} > 0$, and in that case the objective function is differentiable and, by convexity of (P2), optimal solutions correspond to KKT points. Let $\mu \in \mathbb{R}^m_+$ be the dual variables for constraints $y \geq 0$. From the KKT conditions of (P2) with respect to $y$, we see that

$$
-\mu_k = b_k - \frac{y_k}{\sqrt{s^2 + \sum_{i \in M} y_i^2}}, \ \forall k \in M.
$$

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

$$
y_k = b_k \sqrt{s^2 + \sum_{i \in M} y_i^2}, \ \forall k \in M.
$$
Defining $\beta = \sum_{i=1}^{m} b_i^2$, we have
\[
\sum_{i \in M} b_i y_i = \beta \sqrt{s^2 + \sum_{i \in M} y_i^2}
\]
and
\[
\sum_{i \in M} y_i^2 = \beta \left( s^2 + \sum_{i \in M} y_i^2 \right).
\] (19)

Observe that if $\beta \geq 1$, equality (19) cannot be satisfied (unless $\beta = 1$ and $s = 0$), and the feasible (P2) is dual infeasible. Indeed, let $\lambda > 0$ and $\bar{y}_i = \lambda b_i$ for all $i \in M$, and observe that for any value of $s$
\[
\lim_{\lambda \to \infty} -b'\bar{y} + \sqrt{s^2 + \sum_{i \in M} \bar{y}_i^2} = \begin{cases} -\infty & \text{if } \beta > 1 \\ 0 & \text{if } \beta = 1. \end{cases}
\]

Thus, if $\beta > 1$, then both problems (BP) and (P2) are unbounded. Moreover, if $\beta = 1$, let
\[
(x^*, s^*) \in \arg \min_{(x, s) \in \text{conv}(K)} -a'x
\]
with minimal value of $s^*$; if $s^* = 0$, then $(x^*, \bar{y}, s^*)$ is an optimal solution of both (BP) and (P2) for any $\lambda > 0$, and if $s^* > 0$ then there does not exist an optimal solution for problems (BP) and (P2), but infima of the objective functions are attained at $x^*, s^*$ and $y = \bar{y}$ as $\lambda \to \infty$.

If $\beta < 1$, then we deduce from (19) that
\[
\sum_{i \in M} y_i^2 = \frac{\beta}{1 - \beta} s^2.
\]
Replacing the summands in the objective, we rewrite (P2) as
\[
\min -a'x + (1 - \beta) s \\
(P3) \quad \text{s.t.} \ (x, s) \in \text{conv}(K).
\]
As $\beta < 1$, (P3) has an optimal solution and it is integral in $x$.

Remark 4. From Proposition 5 we see that, with no constraints on the continuous variables, describing the mixed-integer set $\text{conv}(H_X)$ reduces to describing a polyhedral set. Moreover, strong inequalities from pure binary sets (e.g., Yu and Ahmed 2017) can be naturally lifted into strong inequalities for $H_X$.

Corollary 1. Inequalities (17) and bound constraints completely describe $\text{conv}(H_D)$.

Proof. Follows from Proposition 5 with $K = H_B$, where the convex hull of $H_B$ is given in Proposition 1 and substituting out the auxiliary variable $s$. □
4.1. **Comparison with inequalities in the literature.** As seen in this section inequalities (17) give the convex hull of $H_D$. Therefore, they are the strongest possible inequalities for $H_D$. It is of interest to study the relationships to inequalities previously given in the literature. It turns out that for the case of a single binary variable, they can be obtained as either split cuts or conic MIR inequalities based on a single disjunction. The equivalence does not hold in higher dimensions as neither split cuts nor conic MIR inequalities are sufficient to describe $\text{conv}(H_D)$ in those cases.

To see the equivalence, we now consider the special case of conic quadratic constraint with a single binary variable $x$:

$$H^1 = \left\{ (x, y, z) \in \{0, 1\} \times \mathbb{R}_{+}^{m+1} : \sqrt{\sigma + cx + \sum_{i \in M} d_i y_i^2} \leq z \right\}.$$

4.1.1. **Comparison with split cuts.** We first compare inequalities (17) with the split cuts given in Modaresi et al. (2016). Following the notation used by the authors, let

$$B = \left\{ (y, z) \in \mathbb{R}_{+}^{m+2} : \sqrt{\sigma + y_0^2 + \sum_{i \in M} d_i y_i^2} \leq z \right\},$$

be the base set, let $F = \{ y \in \mathbb{R}_{+}^{m+1} : 0 \leq y_0 \leq c \}$ be the forbidden set, and define $K = B \setminus \text{int}(F)$, where $\text{int}(F)$ denotes the interior of $F$. Letting $y_0 := \sqrt{c}x$, we see that $H^1$ and $K$ are equivalent.

First consider the case $\sigma = 0$. From Corollary 1 we see that

$$\text{conv}(H^1) = \left\{ (x, y, z) \in [0, 1] \times \mathbb{R}_{+}^{m+1} : \sqrt{cx^2 + \sum_{i \in M} d_i y_i^2} \leq z \right\}.$$

Moreover, from Corollary 5 of Modaresi et al. (2016), since $0 \notin (0, c)$, we find that $\text{conv}(K) = B$. Thus, the results coincide in that the convex hulls of $H^1$ and $K$ are the natural convex relaxations of the sets.

Now consider the case $\sigma > 0$. From Corollary 1 we see that that

$$\text{conv}(H^1) = \left\{ (x, y, z) \in [0, 1] \times \mathbb{R}_{+}^{m+1} : (\sqrt{\sigma + (\sqrt{c}\pm \sqrt{\sigma})x})^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \right\}.$$

Moreover, from Proposition 8 of Modaresi et al. (2016) we find that

$$\text{conv}(K) = \left\{ (y, z) \in \mathbb{R}_{+}^{m+2} : \left(\sqrt{\sigma + \frac{\sqrt{c} + \sqrt{\sigma}}{\sqrt{c}} y_0}\right)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \right\}.$$

(20)

Thus, the results coincide again.
4.1.2. Comparison with MIR inequalities. We now compare inequalities (17) with the simple nonlinear conic mixed-integer rounding inequality given in Atamtürk and Narayanan (2010). Letting \( a = \sqrt{\sigma} + \sqrt{\sigma} + c \) and \( b = \frac{\sqrt{\sigma}}{a} \), we can write

\[
H^1 = \left\{ (x, y, z) \in \{0, 1\} \times \mathbb{R}^{n+1}_+ : (x - b)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \frac{a^2}{a^2} \right\}.
\]

Note that if \( \sigma = 0 \) then \( b = 0 \) and the MIR inequalities reduces to the original inequality–which defines the convex hull of \( H^1 \). If \( \sigma > 0 \), then \( \lfloor b \rfloor = 0 \) and the simple mixed integer rounding inequality is

\[
((1 - 2b)x + b)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \frac{a^2}{a^2}
\]

\[
\Leftrightarrow \left( (1 - 2\sqrt{\frac{\sigma}{\sigma + \sigma + c}})x + \sqrt{\frac{\sigma}{a}} \right)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \frac{a^2}{a^2}
\]

\[
\Leftrightarrow \left( \frac{\sqrt{\sigma} + c - \sqrt{\sigma}}{a}x + \sqrt{\frac{\sigma}{a}} \right)^2 + \sum_{i \in M} d_i y_i^2 \leq z^2 \frac{a^2}{a^2},
\]

and multiplying both sides by \( a^2 \) we get (20).

4.2. Set \( R_X \) with rotated cone. Here we consider the set \( R_X \) and, more generally, sets of the form written in conic quadratic form

\[
U_R = \left\{ (x, s) \in K, (y, w, z) \in \mathbb{R}^{m+2}_+ : s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2 \leq (w + z)^2 \right\},
\]

where \( K = \left\{ x \in \tilde{X} \subseteq \{0, 1\}^n, s \geq 0 : \sqrt{h(x)} \leq s \right\} \) and \( h : \tilde{X} \to \mathbb{R}_+ \).

Observe that the approach discussed in Section 4 can be used for \( R_X \) and \( U_R \).

For example, using inequalities (17) for \( R_X \) results in the valid inequalities

\[
(\sqrt{\sigma} + \pi' x)^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2 \leq (w + z)^2, \quad \pi \in \Pi_\sigma.
\]

We can also write inequalities (21) in rotated cone form,

\[
(\sqrt{\sigma} + \pi' x)^2 + \sum_{i \in M} d_i y_i^2 \leq 4wz, \quad \pi \in \Pi_\sigma.
\]

Note, however, that the second-order cone constraint defining \( R_X \) and \( U_R \) has additional structure, namely the continuous nonnegative variables \( w \) and \( z \) in both sides of the inequality. Nevertheless, as Proposition 6 states, inequalities (21) are sufficient to characterize \( \text{conv}(R_X) \). The proof of Proposition 6 is provided in Appendix A.

**Proposition 6.** The convex hull of \( U_R \) is described as

\[
\text{conv}(U_R) = \left\{ (x, s) \in \text{conv}(K), (w, z) \in \mathbb{R}^2_+ : s^2 + \sum_{i \in M} d_i y_i^2 \leq 4wz \right\}.
\]
Remark 5. Consider again the set given by (10)–(11) in Section 2.6, and observe that it corresponds to $U_R$ with $m = 0$ and $K = \{ x \in X, s \in \mathbb{R}_+: 2 \sqrt{c'x} \leq s \}$. Thus, if $X = \{0, 1\}^n$, then $\text{conv}(U_R)$ is given by bounds constraints and inequalities

$$(\gamma'x)^2 \leq wz, \gamma \in \Pi(\bar{h}),$$

where $\Pi(\bar{h})$ is the set of extreme points of the polymatroid associated with the submodular function $\bar{h}(s) = 2 \sqrt{c'x}$.

5. The case of bounded continuous variables

In this section we study $H_C$ with bounded continuous variables, i.e., by scaling $C = \{0, 1\}^n \times [0, 1]^m$. We first give a description of $\text{conv}(H_C)$ for the case $n = m = 1$ and discuss the difficulties in obtaining the convex hull description for the general case (Section 5.1). Then we describe valid inequalities that can be efficiently implemented in branch-and-cut algorithms (Section 5.2).

5.1. Two variable case with a bounded continuous variable. In this section we study the three-dimensional set

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

where $\sigma \geq 0$ is a constant. First we give its convex hull description.

**Proposition 7.** The convex hull of $L$ is described as

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

where

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

**Proof.** A point $(x, y, z)$ belongs to $\text{conv}(L)$ if and only if there exist $x_1, x_2, y_1, y_2, z_1, z_2$, and $0 \leq \lambda \leq 1$ such that the system

$$x = (1-\lambda)x_1 + \lambda x_2 \quad (22)$$
$$y = (1-\lambda)y_1 + \lambda y_2 \quad (23)$$
$$z = (1-\lambda)z_1 + \lambda z_2 \quad (24)$$
$$z_1 \geq \sqrt{\sigma + dy_1^2} \quad (25)$$
$$z_2 \geq \sqrt{\sigma + c + dy_2^2} \quad (26)$$
$$0 \leq y_1, y_2 \leq 1, \quad x_1 = 0, \quad x_2 = 1 \quad (27)$$

is feasible. Observe that from (22) and (27) we can conclude that $\lambda = x$. Also observe that from (22), (25) and (26) we have that

$$z = (1-x)z_1 + xz_2$$

$$\Leftrightarrow z \geq (1-x)\sqrt{\sigma + dy_1^2} + x\sqrt{\sigma + c + dy_2^2}.$$
Therefore, the system is feasible if and only if
\[
  z \geq \min_{y_1, y_2} \left(1 - x\right) \sqrt{\sigma + dy_1^2} + x \sqrt{\sigma + c + dy_2^2}
\]
\[
  \text{s.t. } y = \left(1 - x\right)y_1 + xy_2
\]
\[
y_1 \leq 1
\]
\[
y_2 \leq 1
\]
\[
y_1 \geq 0
\]
\[
y_2 \geq 0,
\]
and let \(\gamma, \alpha\) and \(\beta\) be the dual variables of the optimization problem above. Note that the objective function is differentiable even if \(\sigma = 0\) since in that case the function \(\sqrt{\sigma + dy_1^2}\) reduces to the linear function \(\sqrt{dy_1}\). Moreover, the optimization problem is convex, and from KKT conditions for variables \(y_1\) and \(y_2\) we find that
\[
  -(1 - x) \frac{dy_1}{\sqrt{\sigma + dy_1^2}} = \gamma(1 - x) + \alpha_1 - \beta_1
\]
\[
  -x \frac{dy_2}{\sqrt{\sigma + c + dy_2^2}} = \gamma x + \alpha_2 - \beta_2
\]
\[
  \implies \frac{y_1}{\sqrt{\sigma + dy_1^2}} + \bar{\alpha}_1 - \bar{\beta}_1 = \frac{y_2}{\sqrt{\sigma + c + dy_2^2}} + \bar{\alpha}_2 - \bar{\beta}_2,
\]
where \(\bar{\alpha}, \bar{\beta}\) correspond to \(\alpha\) and \(\beta\) after scaling. We deduce from (29) and complementary slackness that \(y_1, y_2 > 0\) (unless \(y = 0\)) and that \(y_1 \leq y_2\): if \(y_1 = 0\) and \(y_2 > 0\) then \(\bar{\alpha}_1 = \bar{\beta}_2 = 0\), and (29) reduces to \(-\bar{\beta}_1 = y_2/\sqrt{\sigma + c + dy_2^2} + \bar{\alpha}_2\), which has no solution since the right-hand-side is positive; letting \(y_2 = 0\) and \(y_1 > 0\) results in a similar contradiction; and if \(0 < y_2 < y_1\) then \(\bar{\beta}_1 = \bar{\alpha}_2 = \bar{\beta}_2 = 0\) and (29) reduces to \(y_1/\sqrt{\sigma + dy_1^2} + \bar{\alpha}_1 = y_2/\sqrt{\sigma + c + dy_2^2}\), which has no solution since \(y_1 > y_2\) implies that \(y_1/\sqrt{\sigma + dy_1^2} > y_2/\sqrt{\sigma + c + dy_2^2}\).

Therefore, for 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
\[
y_1^* = y \frac{\sqrt{\sigma}}{x\sqrt{c + \sigma} + (1 - x)\sqrt{\sigma}}
\]
\[
y_2^* = y \frac{\sqrt{c + \sigma}}{x\sqrt{c + \sigma} + (1 - x)\sqrt{\sigma}}
\]
satisfy conditions (23) and (29). Moreover, if
\[
y_2^* \leq 1
\]
\[
\Leftrightarrow y \leq \frac{x\sqrt{c + \sigma} + (1 - x)\sqrt{\sigma}}{\sqrt{c + \sigma}} = x + (1 - x)\sqrt{\frac{\sigma}{c + \sigma}},
\]
then \(y_1^*, y_2^*\) also satisfy the bound constraints, and thus correspond to an optimal solution to the optimization problem. Replacing in (28), we find that
\[
z \geq \sqrt{\left(\sqrt{\sigma} + x(\sqrt{c + \sigma} - \sqrt{\sigma})\right)^2 + dy_2^2}
\]
when \( y \leq x + (1 - x)\sqrt{\frac{\sigma}{\sigma + c}} \). On the other hand, if \( y^*_2 > 1 \), an optimal solution to the optimization problem is given by \( \bar{y}_2 = 1 \) and \( \bar{y}_1 = \frac{y - x}{1 - x} \). Substituting in (28)

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

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

Note that inequality \( g_1(x, y) \leq z \) is a special case of inequalities (17). If \( \sigma = 0 \), then we get

\[
g_2(x, y) \leq z \text{ with } y \geq x \Leftrightarrow \sqrt{dy} + x(\sqrt{c + d} - \sqrt{d}),
\]

which is a special case of inequalities (15). However, inequality \( g_2(x, y) \leq z \) is not valid if \( \sigma > 0 \). In particular, it cuts off the feasible point \( (x, y, z) = (1, 0, \sqrt{\sigma + c}) \). Moreover, it can be shown that the inequality \( g_2(x, y) \leq z \) cuts off portions of \( \text{conv}(L_\sigma) \) whenever \( y \leq x + (1 - x)\sqrt{\frac{\sigma}{\sigma + c}} \).

**Example 1.** Consider the set \( L \) with \( \sigma = d = 1 \) and \( c = 2 \). Figure 1 shows functions \( g_1 \) and \( g_2 \) when \( x = 0.5 \) is fixed, and illustrates the comments above. We see that the function \( g_2 \) is always “above” the function \( g_1 \), and cuts the convex hull of \( L \) (the shaded region) whenever \( y \leq x + (1 - x)\sqrt{\frac{\sigma}{\sigma + c}} \).

![Figure 1. Functions \( g_1, g_2 \) with \( \sigma = d = 1, c = 2 \) (\( x = 0.5 \)).](image)

Unfortunately, Proposition 7 does not help to describe the convex hull of \( H_C \) with more than one bounded variable. Additionally, piece-wise convex functions like \( g(x, y) \) in Proposition 7 cannot be directly used with standard branch-and-bound algorithms. Thus, we now turn our attention to deriving inequalities that are valid and can be implemented as cutting surfaces in branch-and-cut algorithms, if not sufficient to describe \( \text{conv}(H_C) \) in general.
5.2. The general (multi-variable) case. To obtain valid inequalities for \( H_C \) we write the conic quadratic constraint in extended form for a subset \( T \subseteq M \) of the continuous variables:

\[
s^2 + \sum_{i \in M \setminus T} d_i y_i^2 \leq z^2
\]

\[
\sum_{i \in N} c_i x_i + \sum_{i \in T} d_i y_i^2 \leq s^2.
\]

\[
x \in \{0,1\}^n, y \in [0,1]^M, s \geq 0.
\]

Applying inequality (15) to (31) and eliminating the auxiliary variable \( s \), we obtain the inequalities

\[
\left( \sqrt{\sigma + \sum_{i \in T} d_i y_i^2 + \pi' x} \right)^2 + \sum_{i \in M \setminus T} d_i y_i^2 \leq z^2, \quad \pi \in \Pi_{\sigma+d(T)}.
\]

(32)

**Proposition 8.** For \( T \subseteq M \) inequalities (32) are valid for \( H_C \).

Note that inequalities (32) generalize or strengthen the previous valid inequalities proposed in this paper and other inequalities in the literature.

**Remark 6.** For \( T = \emptyset \) inequalities (32) coincide with inequalities (17). For \( T = M \) inequalities (32) coincide with inequalities (15). If \( T \subset M \), then inequalities (32) dominate inequalities (15).

**Example 1 (Continued).** We obtain from (32) the valid inequality

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

for \( L_\sigma \). As Figure 2 shows, the inequality provides a good approximation of \( L_\sigma \) for the example considered.

![Figure 2](image-url)

**Figure 2.** Functions \( g_1, g_2, g_3 \) with \( \sigma = d = 1, c = 2 \) (\( x = 0.5 \)).
6. Valid inequalities for general $H_X$

In this section we derive inequalities that exploit the structure for an arbitrary set $X \subseteq D$. We first describe a lifting procedure for obtaining valid inequalities for any mixed-binary set $X$, where computing each coefficient requires solving an integer optimization problem (Section 6.1). Then, we propose an approach based on linear programming to efficiently compute weaker valid inequalities (Section 6.2).

6.1. General mixed-binary set $X$. We now consider valid inequalities for $H_X$ where $X \subseteq D$. The inequalities described here have a structure similar to the non-linear extended polymatroid inequalities (17) and (32). For a given a permutation $((1), (2), \ldots, (n))$ of $N$ and $T \subseteq M$, let

$$h_k(x, y) = \sigma + \sum_{i=1}^{k-1} c(i)x(i) + \sum_{i \in T} d_i y_i^2$$

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

$$\rho(k) = \begin{cases} \sqrt{c(k)} + \bar{\sigma}(k) - \sqrt{\bar{\sigma}(k)} & \text{if } \bar{\sigma}(k) < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Consider the inequality

$$\left( \sqrt{\sigma + \sum_{i \in T} d_i y_i^2 + \sum_{i=1}^{n} \rho(i)x(i)} \right)^2 + \sum_{i \in M \setminus T} d_i y_i^2 \leq z^2 .$$

Proposition 9. For $T \subseteq M$ inequalities (35) are valid for $H_X$.

Proof. Let

$$H_X(T) = \left\{ (x, y) \in X, s \geq 0 : \sqrt{\sigma + \sum_{i \in N} c_i x_i + \sum_{i \in T} d_i y_i^2} \leq s \right\} ,$$

and consider the extended formulation of $H_X$ given by

$$\tilde{H}_X = \left\{ (x, y, s) \in H_X(T), z \geq 0 : \sqrt{s^2 + \sum_{i \in M \setminus T} d_i y_i^2} \leq z \right\} .$$

To prove the validity of (35) for $H_X$, it is sufficient to show that

$$\sqrt{\sigma + \sum_{i \in T} d_i y_i^2 + \sum_{i=1}^{n} \rho(i)x(i)} \leq s$$

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

$$\sqrt{\sigma + \sum_{i \in T} d_i y_i^2 + \sum_{i=1}^{k} \rho(i)x(i)} \leq \sqrt{\sigma + \sum_{i=1}^{k} c(i)x(i) + \sum_{i \in T} d_i y_i^2}$$

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


Base case: $k = 0$. Inequality (37) holds trivially.

Inductive step. Let $(\bar{x}, \bar{y}) \in X$, and suppose inequality (37) holds for $k - 1$. Observe that if $\bar{x}_{(k)} = 0$ or $\rho_{(k)} = 0$, then inequality (37) clearly holds for $k$. Therefore, assume that $\bar{x}_{(k)} = 1$ and $\bar{\sigma}_{(k)} < \infty$. We have

\[
\sqrt{\sigma + \sum_{i=1}^{k} c_i \bar{x}_{(i)} + \sum_{i \in T} d_i \bar{y}_i^2} = \sqrt{h_k(\bar{x}, \bar{y}) + c(k)}
\]

\[
= \sqrt{h_k(\bar{x}, \bar{y}) + \left( \sqrt{h_k(\bar{x}, \bar{y}) + c(k)} - \sqrt{h_k(\bar{x}, \bar{y})} \right)}
\]

\[
\geq \sqrt{h_k(\bar{x}, \bar{y}) + \left( \sqrt{\bar{\sigma}(k) + c(k)} - \sqrt{\bar{\sigma}(k)} \right)} \quad \quad (38)
\]

\[
\geq \sqrt{\sigma + \sum_{i \in T} d_i \bar{y}_i^2 + \sum_{i=1}^{k} \rho_{(i)} \bar{x}_{(i)}}, \quad \quad (39)
\]

where (38) follows from $\bar{\sigma}_{(k)} \geq h_k(\bar{x}, \bar{y})$ (by definition of $\bar{\sigma}_{(k)}$) and from the concavity of the square root function, and (39) follows from $\sqrt{h_k(\bar{x}, \bar{y})} \geq \sqrt{\sigma + \sum_{i \in T} d_i \bar{y}_i^2} + \sum_{i=1}^{k-1} \rho_{(i)} \bar{x}_{(i)}$ (induction hypothesis) and from the definition of $\rho_{(k)}$. □

**Remark 7.** For $T = \emptyset$ and $X = D$, inequalities (35) reduce to inequalities (17). For $T = \emptyset$ and $X \subset D$, then inequalities (35) dominate inequalities (17).

**Remark 8.** For $X = C$, then inequalities (35) reduce to inequalities (32). For $X \subset C$, inequalities (35) dominate inequalities (32).

**Remark 9.** For the case of the pure-binary set with defined by a cardinality constraint, i.e., $Y = \{x \in \{0, 1\}^n : \sum_{i=1}^{n} x_i \leq k\}$ and $\sigma = 0$, Yu and Ahmed (2017) give facets for $\text{conv}(H_Y)$. However, noting the computation burden of constructing them, they propose approximate lifted inequalities of the form $\sum_{i \leq k} \pi_{(i)} x_{(i)} + \sum_{i > k} \rho_{(i)} x_{(i)} \leq \bar{z}$, where $\pi$ are computed according to (13), and

\[
\rho_{(i)} = \sqrt{c(T_{(i)}) + c_i - \sqrt{c(T_{(i)})}}
\]

with $T_{(i)} = \arg \max \{c(T) : T \subseteq \{(1), \ldots, (i-1)\}, |T| = k-1\}$. Thus, their approximate lifted inequalities coincide with inequalities (35) and can be computed in $O(n \log n)$. If the set $X$ has additional constraints, then inequalities (35) are stronger than the approximate lifted inequalities of Yu and Ahmed (2017).

**Remark 10.** The strengthened extended polymatroid inequalities described in this section can be used with rotated cone constraints as well. In particular, for the set

\[
R_X = \left\{ (x, y) \in X, w \geq 0, z \geq 0 : \sigma + \sum_{i \in N} c_i x_i + \sum_{i \in M} d_i y_i^2 \leq 4wz \right\},
\]

we find that inequalities

\[
\left( \sqrt{\sigma + \sum_{i \in T} d_i y_i^2 + \sum_{i=1}^{n} \rho_{(i)} x_{(i)}} \right)^2 + \sum_{i \in M \setminus T} d_i y_i^2 \leq 4wz \quad \quad (40)
\]
are valid for $R_X$.

6.2. Computational efficiency. Note that computing each coefficient of inequality (35) requires solving a non-convex mixed 0-1 optimization problem (33), which may not be practical in most cases. However, observe from Remarks 7 and 8 that solving the optimization problem over any relaxation of $X$ that includes the bound constraints results in valid inequalities at least as strong as the ones resulting from using only the bound constraints.

In particular, assume in problem (33) that, for $i \in T$, $y_i$ has a finite upper bound (otherwise the problem is unbounded and $\rho_i = 0$) and $u_i = 1$ (by scaling). Moreover let $X_P$ be a polyhedron 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, Lubin et al. 2016].

Given $X_P$, the approximate coefficients

$$
\hat{\rho}(k) = \sqrt{c(k)} + \hat{\sigma}(k) - \sqrt{\hat{\sigma}(k)}, \quad \text{with}
$$

$$
\hat{\sigma}(k) = \sigma + \max \left\{ \sum_{i=1}^{k-1} c(i) x(i) + \sum_{i \in T} d_i y_i : (x, y) \in X_P, x_k = 1 \right\}
$$

can be computed efficiently by solving $n$ linear programs. 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.

7. Computational experiments

In this section we report computational experiments performed to test the effectiveness of the polymatroid inequalities in solving second order cone optimization with a branch-and-cut algorithm. In Section 7.1 we test the inequalities introduced in Sections 4 and 5 on problems with bounded continuous variables. In Section 7.2 we solve instances with general covariance matrices (see application in Section 2.2), in Section 7.3 we solve robust second order cone programs (see application in Section 2.3), and in Section 7.4 we solve binary linear fractional problems (see application in Section 2.5).

All experiments are done using CPLEX 12.6.2 solver on a workstation with a 2.93GHz Intel®Core™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, and all times reported include the time required to add cuts.

7.1. Instances with bounded continuous variables. In this section we test the effectiveness of the polymatroid inequalities (17) and (32) in solving optimization problems of the form

$$
\min \left\{ -a^T x - b^T y + \Omega z : (x, y, z) \in H_X \right\}
$$

(42)
with $X = \{0, 1\}^n \times [0, 1]^m$ 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 $\ell$ 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 $\Omega$ is the solution of

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

These instances have large integrality gaps with a single conic quadratic constraint.

Inequalities (17) are added as linear cuts in an extended formulation, as described in Remark 2. Inequalities (32) 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 (32). Thus, given a fractional solution $(\bar{x}, \bar{y})$, we add the linear underestimator $g(x, y) \leq z$, where

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

where

$$\eta = \sqrt{\sigma + \sum_{i \in T} d_i \bar{y}_i^2 + \pi' \bar{x}},$$
$$\zeta = \frac{\eta}{\sqrt{\sigma + \sum_{i \in T} d_i \bar{y}_i^2}},$$
$$\psi = \sqrt{\eta^2 + \sum_{i \in M \setminus T} d_i \bar{y}_i^2}.$$

A greedy heuristic is used to choose $T \subseteq M$ for inequalities (32): 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 (32), 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 (32) 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 ($\text{igap}$), the root gap improvement ($\text{rimp}$), the number

\footnote{This choice of $\Omega$ ensures that the linear and nonlinear components are well-balanced, resulting in challenging instances with large integrality gap.}
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 $igap = \frac{t_{opt} - t_{relax}}{|t_{opt}|} \times 100$, 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 $egap = \frac{t_{opt} - t_{bb}}{|t_{opt}|} \times 100$, where $t_{bb}$ is the objective value of the best lower bound at termination. The root improvement is computed as $rimp = \frac{t_{root} - t_{relax}}{t_{opt} - t_{relax}} \times 100$, 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 | 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 (17), which do not exploit the upper bounds of the continuous variables, closes 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 (32), which exploit the upper bounds of the continuous variables, close 99% on 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 only 0.7%. Thus, the inequalities perform very well for the sets they are designed for.

7.2. Mean-risk minimization with correlated random variables. In this section we test the effectiveness the polymatroid inequalities in instances with correlated random variables. In particular, we solve mean-risk optimization problems

$$\min_{x \in \{0,1\}^n} \left\{ -a'x + \Omega \sqrt{x'Qx} : \sum_{i=1}^{n} x_i \leq k \right\}, \quad (43)$$

where the matrix $Q$ is generated according to a factor model, i.e., $Q = FX' + D$ where $F \in \mathbb{R}^{r \times r}$ is the factor covariance matrix, $X \in \mathbb{R}^{n \times r}$ is the exposure matrix and $D \in \mathbb{R}^{n \times n}$ is diagonal matrix with the specific covariances. Observe that in such instances, we can set $\text{diag}(c) = D$ in equation (3).

In our experiments $F = GG'$, 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 $\delta \geq 0$ 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 $\Omega = \Phi^{-1}(\alpha)$, where $\Phi$ is the cumulative distribution function of the normal distribution. We let $n = 200$, $r = 40$ and $k$ equal to 10%, 15%, and 20% of the number of the variables. The effectiveness of inequalities (17) and (32) are compared with default CPLEX. The inequalities are added using an extended formulation as described in Remark 2.
Table 2. Experiments with general covariance matrices ($\delta = 0.5$).

| $k$  | $\alpha$ | $\text{gap}$ | $\text{time}$ | $\text{egap}$ | $\text{rimp nodes}$ | $\text{time}$ | $\text{egap}$ | $\text{rimp nodes}$ | $\text{time}$ | $\text{egap}$ |
|------|----------|---------------|--------------|--------------|-------------------|--------------|--------------|-------------------|--------------|--------------|
| 0.95 | 1.7      | 22.6          | 9.557        | 74           | 0.05              | 53.3         | 3.957        | 23               | 0.05         | 55.6         |
| 20   | 0.975    | 21.3          | 33.468       | 242          | 0.05              | 54.5         | 13.316       | 86               | 0.05         | 55.9         |
| 0.99 | 5.2      | 15.2          | 164.568      | 1.845        | 0.05              | 52.8         | 80.735       | 730              | 0.05         | 55.3         |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          | 19.7          | 69,198       | 720          | 0.0                | 53.2         | 32,660       | 280              | 0.0          | 55.6         |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          |               |             |              |                   |              |              |                  |              |              |

Table 3. Experiments with general covariance matrices ($\delta = 1.0$).

| $k$  | $\alpha$ | $\text{gap}$ | $\text{time}$ | $\text{egap}$ | $\text{rimp nodes}$ | $\text{time}$ | $\text{egap}$ | $\text{rimp nodes}$ | $\text{time}$ | $\text{egap}$ |
|------|----------|---------------|--------------|--------------|-------------------|--------------|--------------|-------------------|--------------|--------------|
| 0.95 | 2.9      | 21.6          | 64,283       | 927          | 0.05              | 55.1         | 14,984       | 165              | 0.05         | 50.1         |
| 20   | 0.975    | 15.5          | 240,224      | 3,975        | 0.43              | 44.4         | 189,826      | 3,390             | 0.43         | 50.9         |
| 0.99 | 9.0      | 6.4           | 378,116      | 7,200        | 2.20              | 35.7         | 477,553      | 7,200             | 1.90         | 43.1         |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          | 14.5          | 227,541      | 4,034        | 0.9                | 45.1         | 227,454      | 3,585             | 0.8           | 51.0         |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          |               |             |              |                   |              |              |                  |              |              |
|      |          |               |             |              |                   |              |              |                  |              |              |

Tables 2 and 3 present the results for different choices of the diagonal dominance parameter $\delta$. Observe that adding inequalities (17) or (35) closes the initial integrality gaps by 45% to 75%, resulting in significant performance improvement over default CPLEX. In particular, using inequalities (35) 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 $k \geq 30$ using inequalities (35) results in at least an order-of-magnitude speed-up over default CPLEX. As in the previous section, inequalities (35), exploiting the cardinality constraint, are more effective than (17). The impact of both inequalities increases with higher diagonal dominance.

7.3. Robust VaR instances. In this section we test the effectiveness of the polymatroid inequalities in the approximate robust counterpart (RC) of problem (4) with $W$ given as in (2.3).

In our computations, we model a decision-maker that seeks a path with minimal VaR and robust to interruptions in the arcs (corresponding to traffic incidents or

\[2\] 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.
attacks by an adversary. The feasible region $X$ is given by path constraints on a $40 \times 40$ grid network. There is a potential adverse event corresponding to each arc, and each event results in an increase in the nominal duration and variance of that arc: in particular, for $i = 1, \ldots, n$, $a_i \sim U[0, 2] e^i$, where $e^i$ is the vector which has value 1 in the $i$-th position and 0 elsewhere, and the $i$-th row and column of $Q_i$ are drawn from $U[0, 2]$ and $Q_i$ has 0 entries elsewhere. Each element of the nominal cost vector $a_0$ is drawn from $U[0, 1]$, and the squared roots of every diagonal element of $Q_0$ are also generated from $U[0, 1]$. The parameter $\Omega$ is set as in Section 7.2.

Table 4 shows the results for different values of $\alpha$ and the parameter controlling the robustness $b$. Observe that strengthened polymatroid cuts (40) results in a better root improvement of 55% – compared to 30–37% achieved by default CPLEX. Moreover, when using the strengthened inequalities, 37 instances are solved to optimality, while default CPLEX is able to solve only 22 instances. We also see that in these path instances, the polymatroid inequalities (21) that do not exploit additional problem constraints result in longer solution times than cpx (despite better root improvements). On the other hand, the strengthened polymatroid inequalities proposed in Section 6 are effective both in reducing the integrality gaps and solution times.

### Table 4. Experiments with robust conic instances.

| $b$ | $\alpha$ | $\text{igap}$ | $\text{cpx}$ | $\text{pimp}$ | $\text{nodes}$ | $\text{time}$ | $\text{gap}$ [$\%$] | $\text{pimp}$ | $\text{nodes}$ | $\text{time}$ | $\text{gap}$ [$\%$] | $\text{pimp}$ | $\text{nodes}$ | $\text{time}$ | $\text{gap}$ [$\%$] |
|-----|---------|---------------|--------------|-------------|----------------|-------------|----------------|--------------|----------------|-------------|----------------|--------------|----------------|-------------|----------------|
| 0.95 | 22.6    | 35.1          | 65,533       | 3,124       | 0.64          | 44.1        | 72,322        | 5,220        | 1.23          | 56.8        | 17,057        | 917          | 0.04          | 55.2        | 30,222        | 2,648        | 0.02          |
| 4   | 0.975   | 24.1          | 30.2         | 95,337      | 4,239         | 0.84        | 41.1          | 87,697       | 7,200        | 3.50        | 55.2        | 53,022        | 2,648        | 0.05          | 55.2        | 57,552        | 2,672        | 0.05          |
| 0.99 | 25.7    | 26.6          | 153,451      | 7,200       | 2.20          | 37.9        | 80,160        | 7,200        | 7.60          | 53.5        | 102,578       | 4,452        | 0.05          | 55.2        | 57,552        | 2,672        | 0.05          |
| Average | 30.6 | 104,117 | 4,854 | 1.28 | 41.1 | 80,060 | 6,540 | 4.13 | 55.2 | 57,552 | 2,672 | 0.05 | 55.2 | 57,552 | 2,672 | 0.05 |

Additionally, Figure 7.3 shows, for different values of the robustness parameter $b$, the objective value of the approximate robust counterpart $\bar{\omega}_n$ and the lower bound on the exact robust problem $\omega_\ell$, defined in (5) and (6), respectively. The figure also shows the nominal objective value corresponding to the robust solution $\bar{\omega}_n$, i.e., $\bar{\omega}_n = a_0^T \bar{x} + \sqrt{\bar{x}^T Q_0 \bar{x}}$ where $\bar{x}$ is an optimal solution of (RC). We see that the nominal values $\bar{\omega}_n$, corresponding to the objective value if none of the events are realized, are not substantially affected by changes in $b$. The computations suggest that the approximate robust formulation (RC) protects against adverse scenarios without leading to overly conservative solutions.

### 7.4. Binary fractional optimization instances.

We test the inequalities for rotated cone constraints (21) in a binary fractional problem arising in assortment...
optimization with cardinality constraint:

$$\max \sum_{j=1}^{m} \frac{\sum_{i=1}^{n} c_{ij} x_i}{a_{0j} + \sum_{i=1}^{n} a_{ij} x_i}$$

(subject to) \[ x \in \{0, 1\}^n. \]

The data is generated as in the assortment optimization problems considered in S¸en et al. (2015): \(a_{ij} \sim U[0, 1]\) for all \(i, j\), \(c_{ij} = a_{ij} r_{ij}\) with \(r_{ij} \sim U[1, 3]\), \(n = 200\), \(m = 20\) and \(a_{0j} = a_0\) for all \(j = 1, \ldots, m\) with \(a_0 \in \{5, 10\}\), and \(k \in \{10, 20, 50\}\).

Binary fractional problems (FP) are usually solved by linearizing the fractional terms (see Prokopyev et al. 2005, Bront et al. 2009, Méndez-Díaz et al. 2014, S¸en et al. 2015, Borrero et al. 2016b), which requires the addition of \(n \times m\) additional variables and big-M constraints. On the other hand, the rotated cone reformulation outlined in Section 2.5 requires adding only \(m\) additional variables and avoids big-M constraints altogether.

We test the classical big-M linear formulation used in Bront et al. (2009), Méndez-Díaz et al. (2014) (milo), the conic formulation without adding inequalities (conic) and the conic formulation strengthened with inequalities (conic+ (21)). Table 5 shows the results. Each row represents the average over five instances generated with the same parameters and for each combination of the parameters \(a_0\) and \(k\) and for each formulation, the root gap (\(\text{rgap}\)), the number of nodes explored (\(\text{nodes}\)), the time elapsed in seconds (\(\text{time}\)), and the end gap (\(\text{egap}\)) in brackets, the number of instances solved to optimality (\(#\)). The root gap is computed as

$$\text{rgap} = \frac{t_{\text{opt}} - t_{\text{root}}}{|t_{\text{opt}}|} \times \text{#}$$

For these instances the strengthened inequalities (21) perform very similarly to (21). Therefore, we only present the results with inequalities (21).
Table 5. Experiments with binary fractional optimization.

| a₀  | k   | milo      | conic     | conic+ [21] |
|-----|-----|-----------|-----------|-------------|
|     |     | rgap nodes | time      | egap[#]    | rgap nodes | time      | egap[#]    | rgap nodes | time      | egap[#]    |
| 10  | 5   | 18.0 51.180 | 7.200 17.000 | 2.7 123.655 | 7.200 1.900 | 0.0 118 | 54 | 0.05 | 0.1 46 | 19 | 0.05 |
| 50  | 0.9 621.742 | 6.010 0.51 | 4.9 55.155 | 7.200 4.500 | 0.1 15.465 | 263 | 0.05 |
|     |     | 23.3 231.220 | 6.803 20.41 | 3.2 67.628 | 4.991 2.15 | 0.1 5.210 | 112 | 0.05 |
| Average | 10  | 46.8 380.700 | 7.200 15.90 | 2.2 48.541 | 972 | 0.05 | 0.0 6 | 14 | 0.05 |
| 10  | 20  | 39.8 23.770 | 7.200 37.40 | 3.7 206.603 | 7.200 1.40 | 0.0 61 | 37 | 0.05 |
| 50  | 5.6 136.382 | 7.200 5.20 | 5.1 52.700 | 7.200 4.60 | 0.1 36.959 | 396 | 0.05 |
| Average | 30.7 | 180.284 | 7.200 19.50 | 4.3 102.615 | 5.124 | 2.05 | 0.0 12.342 | 149 | 0.05 |

100, where t_{opt} is the objective value of the best feasible solution at termination, and t_{root} is the objective value of the relaxation obtained after processing the root node (i.e., after user cuts and cuts added by CPLEX).

We see that the conic formulation with polymatroid inequalities results in substantially faster solution times than the other formulations. In particular, CPLEX with the classical big-M linear optimization formulation milo can only solve 1/30 instances after two hours of branch and bound, and the average end gaps are 20%; the conic formulation with extended polymatroid cuts is able to solve all instances to optimality in less than 3 minutes (on average). We see that root gaps for conic+ are very small in all instances (less than 0.1%), and optimality can be proven in instances with small cardinality parameter k after few branch-and-bound nodes (e.g., in instances with k = 10 and a₀ = 5 optimality is proven after 46 nodes, while conic requires 24,000 nodes to prove optimality).

8. Conclusions

We propose new convex valid inequalities for conic quadratic mixed 0-1 sets. The studied sets arise in a variety of risk-adverse decision-making problems (e.g., chance constrained optimization with correlated variables, robust optimization with ellipsoidal or discrete uncertainty sets) as well as in models of other problems commonly arising in operations research (e.g., lot sizing, scheduling, assortment, fractional linear optimization). The inequalities are derived by exploiting partial submodularity arising from the binary variables and generalize the polymatroid inequalities known for the binary case. They completely describe the convex hull of a single conic quadratic constraint as well as rotated cone constraint with unbounded continuous variables. We also show how to strengthen the inequalities with additional problem constraints. The experiments on testing the inequalities for different classes of problems indicate that in all cases the inequalities improve the performance of branch-and-cut solvers significantly.

Acknowledgment

A. Atamtürk is supported, in part, by grant FA9550-10-1-0168 from the Office of the 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.

Amaldi, E., Bosio, S., Malucelli, F., and Yuan, D. (2011). Solving nonlinear covering problems arising in wlan design. *Operations Research*, 59(1):173–187.

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 Gómez, A. (2017). Maximizing a class of utility functions over the vertices of a polytope. *Operations Research*, 65:433–445.

Atamtürk, A. and Joen, H. (2017). Lifted polymatroid inequalities for mean-risk optimization with indicator variables. BCOL Research Report 17.01, UC Berkeley.

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

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

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

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. (2009a). *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009b). *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.
Bollapragada, R. and Rao, U. (1999). Single-stage resource allocation and economic lot scheduling on multiple, nonidentical production lines. Management Science, 45:889–904.

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.

Borrero, J. S., Gillen, C., and Prokopyev, O. A. (2016a). Fractional 0–1 programming: applications and algorithms. Journal of Global Optimization, pages 1–28.

Borrero, J. S., Gillen, C., and Prokopyev, O. A. (2016b). A simple technique to improve linearized reformulations of fractional (hyperbolic) 0–1 programming problems. Operations Research Letters, 44:479–486.

Bront, J. J. M., Méndez-Díaz, I., and Vulcano, G. (2009). A column generation algorithm for choice-based network revenue management. Operations Research, 57(3):769–784.

Bulut, O. and Tasgetiren, M. F. (2014). An artificial bee colony algorithm for the economic lot scheduling problem. International Journal of Production Research, 52:1150–1170.

Burer, S. and Kılınç-Karzan, F. (2017). How to convexify the intersection of a second order cone and a nonconvex quadratic. Mathematical Programming, 162(1-2):393–429.

Castro, P. M., Barbosa-Póvoa, A. P., and Novais, A. Q. (2005). Simultaneous design and scheduling of multipurpose plants using resource task network based continuous-time formulations. Industrial & Engineering Chemistry Research, 44:343–357.

Castro, P. M., Westerlund, J., and Forssell, S. (2009). Scheduling of a continuous plant with recycling of byproducts: A case study from a tissue paper mill. Computers & Chemical Engineering, 33:347–358.

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., and Vielma, J. (2011a). On the chvátal-gomory closure of a compact convex set. Integer Programming and Combinatorial Optimization, pages 130–142.

Dadush, D., Dey, S. S., and Vielma, J. P. (2011b). The chvátal-gomory closure of a strictly convex body. Mathematics of Operations Research, 36(2):227–239.

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

Davidoff, G., Sarnak, P., and Valette, A. (2003). Elementary number theory, group theory and Ramanujan graphs, volume 55. Cambridge University Press.

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

Fujishige, S. (2005). Submodular functions and optimization, volume 58. Elsevier.

Gilmore, P. C. and Gomory, R. E. (1963). A linear programming approach to the cutting stock problem part ii. Operations Research, 11:863–888.

Grötschel, M., Lovász, L., and Schrijver, A. (1981). The ellipsoid method and its consequences in combinatorial optimization. Combinatorica, 1:169–197.

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.

Hochbaum, D. S. (2010). Polynomial time algorithms for ratio regions and a variant of normalized cut. *IEEE transactions on pattern analysis and machine intelligence*, 32:889–898.

Hochbaum, D. S., Lyu, C., and Bertelli, E. (2013). Evaluating performance of image segmentation criteria and techniques. *EURO Journal on Computational Optimization*, 1:155–180.

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

Kılınç-Karzan, F. (2015). On minimal valid inequalities for mixed integer conic programs. *Mathematics of Operations Research*, 41(2):477–510.

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

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

Méndez-Díaz, I., Miranda-Bront, J. J., Vulcano, G., and Zabala, P. (2014). A branch-and-cut algorithm for the latent-class logit assortment problem. *Discrete Applied Mathematics*, 164:246–263.

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.

Modaresi, S. and Vielma, J. P. (2014). Convex hull of two quadratic or a conic quadratic and a quadratic inequality. *Mathematical Programming*, pages 1–27.

Nahmias, S. (2001). *Production and Operations Analysis*. McGraw Hill.

Nikolova, E., Kelner, J., Brand, M., and Mitzenmacher, M. (2006). Stochastic shortest paths via quasi-convex maximization. *Algorithms–ESA 2006*, pages 552–563.

Orlin, J. B. (2009). A faster strongly polynomial time algorithm for submodular function minimization. *Mathematical Programming*, 118:237–251.

Özsen, L., Coullard, C. R., and Daskin, M. S. (2008). Capacitated warehouse location model with risk pooling. *Naval Research Logistics*, 55:295–312.

Pesenti, R. and Ukovich, W. (2003). Economic lot scheduling on multiple production lines with resource constraints. *International Journal of Production Economics*, 81:469–481.

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

Prokopyev, O. A., Kong, N., and Martinez-Torres, D. L. (2009). The equitable dispersion problem. *European Journal of Operational Research*, 197(1):59–67.

Prokopyev, O. A., Meneses, C., Oliveira, C. A., and Pardalos, P. M. (2005). On multiple-ratio hyperbolic 0–1 programming problems. *Pacific Journal of Optimization*, 1:327–345.

Sahinidis, N. and Grossmann, I. E. (1991). Minlp model for cyclic multiproduct scheduling on continuous parallel lines. *Computers & Chemical Engineering*, 15:85–103.
Appendix A

Proof of Proposition 6. Consider the optimization of an arbitrary linear function over the convex relaxation of $U_R$:

$$\begin{align*}
\min & \ a'x + b'y + pw + qz \\
\text{s.t.} & \ s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2 \leq (w + z)^2 \\
& \quad (x, s) \in \text{conv}(K) \\
& \quad y \in \mathbb{R}^m_+, w \geq 0, z \geq 0.
\end{align*}$$

(44)

Without loss of generality, we can assume that $p > 0$ and $q > 0$ (if $p < 0$ or $q < 0$ then the problem is unbounded, and if $p = 0$ or $q = 0$ then $(P_R)$ reduces to a linear program over an integral polyhedron). Moreover, observe that if $w = z$ in an optimal solution, then the problem reduces to a linear optimization over $U$ which has an optimal integral solution (Proposition 5). Thus, we can assume that $w \neq z$, in which case the left hand size of (44) is differentiable, and we infer from KKT conditions with respect to $w$ and $z$ that

$$\begin{align*}
-p &= -\lambda + \lambda \frac{w - z}{\sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2}} \\
-q &= -\lambda - \lambda \frac{w - z}{\sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2}}.
\end{align*}$$

(45) (46)
where $\lambda$ is the dual variable associated with constraint (44). We deduce from (45) that $w - z = \frac{\lambda - p}{\lambda} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2}$, and from (46) that

$$w - z = \frac{q - \lambda}{\lambda} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2}.$$  \hspace{1cm} (47)

Thus, we find that $\lambda = \frac{p + q}{2}$.

Moreover, we obtain from (47) that

$$(w - z)^2 = \left( \frac{q - \lambda}{\lambda} \right)^2 \left( s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2 \right) = \left( \frac{q - p}{q + p} \right)^2 \left( s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2 \right) = \beta \left( s^2 + \sum_{i \in M} d_i y_i^2 \right),$$

where $\beta = \left( \frac{q - p}{q + p} \right)^2$. Therefore, we have that

$$\sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2} = \sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}.$$  

Moreover, since in any optimal solution of $(P_R)$ constraint (44) is binding, we have

$$w + z = \sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}.$$

Multiplying equality (45) by $w$ in both sides, and multiplying equality (46) by $z$ in both sides, we find that

$$pw + qz = \lambda (w + z) - \lambda \frac{(w - z)^2}{\sqrt{s^2 + \sum_{i \in M} d_i y_i^2 + (w - z)^2}} = \lambda \sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2} - \lambda \frac{\beta \left( s^2 + \sum_{i \in M} d_i y_i^2 \right)}{\sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}} = \lambda \frac{s^2 + \sum_{i \in M} d_i y_i^2}{\sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}} = \lambda \frac{s^2 + \sum_{i \in M} d_i y_i^2}{\sqrt{1 + \beta} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}}.$$

Therefore, we see that problem $(P_R)$ reduces to
$$\min a^t x + b^t y + \frac{p + q}{2\sqrt{1 + \beta}} \sqrt{s^2 + \sum_{i \in M} d_i y_i^2}$$

s.t. \( (x, s) \in \text{conv}(K), y \in \mathbb{R}^m_+ \),

which is equivalent to the optimization over \( (x, s) \in K, y \in \mathbb{R}^M_+ \) (Proposition 5). \( \square \)