# Conic Mixed-Binary Sets: Convex Hull Characterizations and Applications

Fatma Kılınç-Karzan
Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA, fkilinc@andrew.cmu.edu

Simge Küçükyavuz
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA, simge@northwestern.edu

Dabeen Lee
Discrete Mathematics Group, Institute for Basic Science (IBS), Daejeon 34126, Republic of Korea, dabeenl@ibs.re.kr

We consider a general conic mixed-binary set where each homogeneous conic constraint involves an affine function of independent continuous variables and an epigraph variable associated with a nonnegative function, $f_j$, of common binary variables. Sets of this form naturally arise as substructures in a number of applications including mean-risk optimization, chance-constrained problems, portfolio optimization, lot-sizing and scheduling, fractional programming, variants of the best subset selection problem, and distributionally robust chance-constrained programs. When all of the functions $f_j$'s are submodular, we give a convex hull description of this set that relies on characterizing the epigraphs of $f_j$'s. Our result unifies and generalizes an existing result in two important directions. First, it considers *multiple general convex cone* constraints instead of a single second-order cone type constraint. Second, it takes *arbitrary nonnegative functions* instead of a specific submodular function obtained from the square root of an affine function. We close by demonstrating the applicability of our results in the context of a number of broad problem classes.

*Key words*: Conic mixed-binary sets, conic quadratic optimization, convex hull, submodularity, extended polymatroid inequalities, fractional binary optimization, best subset selection, distributionally robust optimization

---

## 1. Introduction

In this paper we study the following set

$$\mathcal{S}(f,\mathbb{K}) := \{(x,z) \in \mathbb{R}^m \times \{0,1\}^n : \ \exists y \in \mathbb{R} \text{ s.t. } y \geq f(z),\ Ax + By \in \mathbb{K}\}, \tag{1}$$

where $f : \{0,1\}^n \to \mathbb{R}_+$ is a nonnegative function, $\mathbb{K}$ is a convex cone containing the origin, and $A, B$ are any matrices of appropriate dimension. Throughout, for $a \in \mathbb{Z}_+$, we let $[a] := \{1,\dots,a\}$.

The set $\mathcal{S}(f,\mathbb{K})$ we consider is a generalization of the two fundamental sets studied by Atamtürk and Gómez (2020) defined as follows:

$$\mathcal{H} := \left\{ (x,z) \in \mathbb{R}^m_+ \times \{0,1\}^n : \ \sqrt{\sigma + \sum_{i\in[n]} c_i z_i + \sum_{j\in[m-1]} d_j x_j^2} \leq x_m \right\}, \tag{2}$$

$$\mathcal{R} := \left\{ (x,z) \in \mathbb{R}^m_+ \times \{0,1\}^n : \ \sigma + \sum_{i\in[n]} c_i z_i + \sum_{j\in[m-2]} d_j x_j^2 \le 4 x_{m-1} x_m \right\}, \tag{3}$$

where $c \in \mathbb{R}^n_+$, $d \in \mathbb{R}^{m-1}_+$, and $\sigma \in \mathbb{R}_+$. Atamtürk and Gómez (2020) demonstrate that the set $\mathcal{H}$ naturally arises in mean-risk minimization and chance-constrained programs, and the set $\mathcal{R}$ appears as a substructure in robust conic quadratic interdiction, lot-sizing and scheduling, queueing system design, binary linear fractional problems, Sharpe ratio maximization, and portfolio optimization. The main contribution of Atamtürk and Gómez (2020) is deriving the convex hulls of $\mathcal{H}$ and $\mathcal{R}$.

In Section 2, we prove that the convex hull of the set $\mathcal{S}(f,\mathbb{K})$ for any nonnegative function $f$ and arbitrary cone $\mathbb{K}$ is given by the following:

$$\widehat{\mathcal{S}}(f,\mathbb{K}) := \{ (x,z) \in \mathbb{R}^m \times [0,1]^n : \ \exists y \in \mathbb{R}_+ \text{ s.t. } (y,z) \in \operatorname{conv}(\operatorname{epi}(f)),\ Ax + By \in \mathbb{K} \}, \tag{4}$$

where given a function $f$, $\operatorname{epi}(f)$ denotes its epigraph, i.e., $\operatorname{epi}(f) := \{(y,z) \in \mathbb{R} \times \{0,1\}^n : y \ge f(z)\}$ and $\operatorname{conv}(\operatorname{epi}(f))$ denotes its convex hull. Therefore, the convex hull of $\mathcal{S}(f,\mathbb{K})$ is given by the inequalities describing $\operatorname{conv}(\operatorname{epi}(f))$ and the homogeneous conic constraint $Ax + By \in \mathbb{K}$. This characterization highlights, in particular, that the challenge in developing the convex hull of $\mathcal{S}(f,\mathbb{K})$ is solely determined by the complexity of the convex hull characterization of the epigraph of the function $f$. In fact, we prove a more general result that covers this characterization as a special case. In particular, we consider a set with multiple conic constraints, each of which involves an affine function of independent continuous variables and an epigraph variable associated with a nonnegative function of common binary variables.

There are a number of cases where characterizing the convex hull of $\operatorname{conv}(\operatorname{epi}(f))$ is easy. For example, when $f$ is a nonnegative submodular function, imposing the condition that $(y,z) \in \operatorname{conv}(\operatorname{epi}(f))$ is equivalent to applying the *extended polymatroid inequalities* (Atamtürk and Narayanan 2008, Lovász 1983). We review the extended polymatroid inequalities and their generalization to arbitrary set functions, namely the polar inequalities, (Atamtürk and Narayanan 2020) in Section 3. In Section 4, we discuss how our results can be applied to more general functions and non-homogeneous conic constraints.

The set $\mathcal{S}(f,\mathbb{K})$ naturally arises in a number of other applications including fractional binary optimization, best subset selection, and distributionally robust optimization. We will discuss these in Section 5. Furthermore, the sets $\mathcal{H}$ and $\mathcal{R}$ studied by Atamtürk and Gómez (2020) are indeed special cases of $\mathcal{S}(f,\mathbb{K})$ where $\mathbb{K}$ is taken to be the direct product of a second order cone (SOC) and the nonnegative orthant, and the function $f(z)$ is restricted to be the square root of an affine function of $z$. We elaborate this connection in Section 5.1. Consequently, even in the case of a single conic constraint and a single function $f$, our convex hull characterization immediately generalizes the results from Atamtürk and Gómez (2020) in two directions: (1) by allowing general convex cones $\mathbb{K}$ as opposed to the standard SOC, and (2) by allowing a general nonnegative function $f(z)$ as opposed to the specific one studied by Atamtürk and Gómez (2020). We discuss applications of our framework with multiple conic constraints and multiple functions in Section 5.2. In particular, this generalization allows us cover the fractional optimization models that appear in a broad range of applications including modeling multinomial logit (MNL) choice models in assortment optimization, set covering, market share based facility location, stochastic service systems, bi-clustering, and optimization of boolean query for databases. We discuss another application of our generalization in best subset selection in machine learning in Section 5.3. In this application, the function $f$ is an exponential function, which demonstrates the applicability of our result beyond the square root function. Furthermore, our results in this context pave the way to use standard solvers for this problem all the while exploiting the submodular structure as opposed to the approach of Gómez and Prokopyev (2020) based on parameterizing the fractional objective function and applying a customized Newton-type method. We also consider an application in distributionally robust chance-constrained programming under Wasserstein ambiguity in Section 5.4 and show that our results can be used to strengthen the mixed-integer conic reformulation for the case of mixed-binary decision variables, thereby generalizing an existing strengthening that assumes pure binary decisions in the original chance constraint.

## 2. Convex hull characterization

We next examine a generalization of our set defined by

$$\mathcal{S}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]}) := \left\{ (x,z) \in \mathbb{R}^{mp} \times \{0,1\}^n : \ \exists y \in \mathbb{R}^p \text{ s.t. } \begin{array}{l} y_j \geq f_j(z), \quad \forall j \in [p], \\ A^j x^j + B^j y_j \in \mathbb{K}_j, \quad \forall j \in [p] \end{array} \right\}, \quad (5)$$

where the functions $f_j : \{0,1\}^n \to \mathbb{R}_+$ for $j \in [p]$ are nonnegative functions and $\mathbb{K}_j$ for $j \in [p]$ are cones. $A^j$ and $B^j$ for $j \in [p]$ are matrices of appropriate dimensions. The lengths of the continuous vectors $x^1, \dots, x^p$ may be different, but we focus on the setting of equal lengths for simplicity. Our theoretical developments can be easily extended to the case of heterogeneous lengths.

Note that the following is a convex relaxation of $\mathcal{S}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]})$:

$$\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]}) := \left\{ (x,z) \in \mathbb{R}^{mp} \times [0,1]^n : \ \exists y \in \mathbb{R}^p \text{ s.t. } \begin{array}{l} (y,z) \in \operatorname{conv}\left(\mathcal{G}\left(\{f_j\}_{j\in[p]}\right)\right), \\ A^j x^j + B^j y_j \in \mathbb{K}_j, \ \forall j \in [p] \end{array} \right\}, \quad (6)$$

$$\text{where} \qquad \mathcal{G}\left(\{f_j\}_{j\in[p]}\right) := \left\{ (y,z) \in \mathbb{R}^p \times \{0,1\}^n : \ y_j \geq f_j(z), \quad \forall j \in [p] \right\}. \quad (7)$$

We note that the functions $f_j$ for $j \in [p]$ take the same binary variables $z$. The constraint $y_j \geq f_j(z)$ gives rise to the epigraph of $f_j$ for each $j$, and therefore, $\mathcal{G}\left(\{f_j\}_{j\in[p]}\right)$ can be viewed as the "intersection" of the epigraphs. Our main result establishes that, under minor assumptions, the convex hull of the set $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ is indeed given precisely by $\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$.

To prove our convex hull result, we first establish a technical proposition that applies to more general mixed-integer sets. We consider sets of the form

$$\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]}) := \left\{(x,y,z)\in\mathbb{R}^{mp}\times\mathbb{R}^p\times\{0,1\}^n:\quad (y,z)\in\mathcal{G},\quad A^jx^j+B^jy_j\in\mathbb{K}_j,\ \forall j\in[p]\right\},\quad (8)$$

where $\mathcal{G}\subseteq\mathbb{R}^p_+\times\{0,1\}^n$ and $\mathbb{K}_j$'s for $j\in[p]$ are cones. We further assume that $\mathcal{G}|_{z=\bar z}:=\mathcal{G}\cap\{(y,z): z=\bar z\}$ for any fixed $\bar z\in\{0,1\}^n$ is convex, which means that $\mathcal{G}|_{z=\bar z}$ is the face of $\operatorname{conv}(\mathcal{G})$ defined by $z=\bar z$. Note that $\mathcal{G}\left(\{f_j\}_{j\in[p]}\right)$ is indeed contained in $\mathbb{R}^p_+\times\{0,1\}^n$, as $f_j$ for $j\in[p]$ are non-negative functions. Moreover, $\mathcal{G}\left(\{f_j\}_{j\in[p]}\right)|_{z=\bar z}$ for a fixed $\bar z\in\{0,1\}^n$ is defined by linear constraints only, so it is convex. Observe that $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ is precisely the projection of $\mathcal{Q}(\mathcal{G}\left(\{f_j\}_{j\in[p]}\right),\{\mathbb{K}_j\}_{j\in[p]})$ onto the $(x,z)$-space. Given this relation, to derive our promised convex hull result for the set $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ in the original space, we first establish the following proposition for sets of the form $\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})$ defined in the extended space.

Proposition 1. *Consider $\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})$ defined as in (8) for some $\mathcal{G}\subseteq\mathbb{R}^p_+\times\{0,1\}^n$ such that $\mathcal{G}|_{z=\bar z}$ for any fixed $\bar z\in\{0,1\}^n$ is convex. Suppose $\mathbb{K}_j$ is a convex cone for each $j\in[p]$. Then,*

$$\operatorname{conv}(\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]}))=\widehat{\mathcal{Q}}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]}),$$

*where*

$$\widehat{\mathcal{Q}}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]}) := \left\{(x,y,z)\in\mathbb{R}^{mp}\times\mathbb{R}^p\times[0,1]^n:\quad (y,z)\in\operatorname{conv}(\mathcal{G}),\quad A^jx^j+B^jy_j\in\mathbb{K}_j,\ \forall j\in[p]\right\}.$$

**Proof of Proposition 1.** For brevity, define $\mathcal{Q}:=\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})$ and $\widehat{\mathcal{Q}}:=\widehat{\mathcal{Q}}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})$. The containment $\operatorname{conv}(\mathcal{Q})\subseteq\widehat{\mathcal{Q}}$ is immediate because $\mathcal{Q}\subseteq\widehat{\mathcal{Q}}$ and $\widehat{\mathcal{Q}}$ is convex. To prove the reverse containment, we show that the recessive directions of $\widehat{\mathcal{Q}}$ are also recessive directions of $\operatorname{conv}(\mathcal{Q})$ and that the extreme points of $\widehat{\mathcal{Q}}$ are contained in $\mathcal{Q}$.

Let $(d_x,d_y,d_z)$ be a recessive direction of $\widehat{\mathcal{Q}}$. Then, $A^jd_x^j+B^j(d_y)_j\in\mathbb{K}_j$ for $j\in[p]$ and $(d_y,d_z)$ is a recessive direction of $\operatorname{conv}(\mathcal{G})$. Suppose for a contradiction that $(d_x,d_y,d_z)$ is not a recessive direction of $\operatorname{conv}(\mathcal{Q})$. Then there exists $(\bar x,\bar y,\bar z)\in\operatorname{conv}(\mathcal{Q})$ such that $(\bar x,\bar y,\bar z)+(d_x,d_y,d_z)\notin\operatorname{conv}(\mathcal{Q})$. As $(\bar x,\bar y,\bar z)\in\operatorname{conv}(\mathcal{Q})$, it can be written as a convex combination of some points, denoted $(x^i,y^i,z^i)$ for $i\in I$, in $\mathcal{Q}$. Hence, $(\bar x,\bar y,\bar z)=\sum_{i\in I}\alpha_i(x^i,y^i,z^i)$ for some $\alpha\geq 0$ with $\sum_{i\in I}\alpha_i=1$. Since $(x^i,y^i,z^i)\in\mathcal{Q}$, we have $A^j(x^i)^j+B^j(y^i)_j\in\mathbb{K}_j$ for $j\in[k]$, and therefore, $A^j(x^i+d_x)^j+B^j(y^i+d_y)_j\in\mathbb{K}_j$ for $j\in[k]$. Since $(d_y,d_z)$ is a recessive direction of $\operatorname{conv}(\mathcal{G})$, $(y^i+d_y,z^i+d_z)\in\operatorname{conv}(\mathcal{G})$. Moreover, $\operatorname{conv}(\mathcal{G})\subseteq\mathbb{R}^p_+\times[0,1]^n$ implies that $d_z=0$, so $(y^i+d_y,z^i+d_z)\in\operatorname{conv}(\mathcal{G})\cap\{(y,z):z=z^i\}=\mathcal{G}|_{z=z^i}\subseteq\mathcal{G}$. Therefore, $(x^i+d_x,y^i+d_y,z^i+d_z)\in\mathcal{Q}$ for $i\in I$, which implies that $(\bar x,\bar y,\bar z)+(d_x,d_y,d_z)=$

$\sum_{i\in I}\alpha_i(x^i+d_x, y^i+d_y, z^i+d_z)\in \operatorname{conv}(\mathcal{Q})$, a contradiction. Therefore, $(d_x,d_y,d_z)$ is a recessive direction of $\operatorname{conv}(\mathcal{Q})$.

Next, consider an extreme point $(\hat{x},\hat{y},\hat{z})$ of $\widehat{\mathcal{Q}}$. Then, $(\hat{y},\hat{z})\in\operatorname{conv}(\mathcal{G})$ and $A^j\hat{x}^j+B^j\hat{y}_j\in\mathbb{K}_j$ holds for all $j\in[p]$. Also, as $\operatorname{conv}(\mathcal{G})\subseteq\mathbb{R}^p_+\times[0,1]^n$, $\hat{y}\in\mathbb{R}^p_+$. We claim that $(\hat{y},\hat{z})$ must be in $\mathcal{G}$. We will prove this by showing that $(\hat{y},\hat{z})$ must be an extreme point of $\operatorname{conv}(\mathcal{G})$. Assume for contradiction that there exist distinct points $(\bar{y},\bar{z})\in\operatorname{conv}(\mathcal{G})$ and $(\tilde{y},\tilde{z})\in\operatorname{conv}(\mathcal{G})$ such that $(\hat{y},\hat{z})=\frac{1}{2}(\bar{y},\bar{z})+\frac{1}{2}(\tilde{y},\tilde{z})$. From $\operatorname{conv}(\mathcal{G})\subseteq\mathbb{R}^p_+\times[0,1]^n$, we deduce that $\bar{y},\tilde{y}\in\mathbb{R}^p_+$. Moreover, if for some index $j\in[p]$ we have $\hat{y}_j=0$, we must also have $\bar{y}_j=\tilde{y}_j=0$. For each $j\in[p]$, define $\bar{x}^j:=\frac{\bar{y}_j}{\hat{y}_j}\hat{x}^j$ whenever $\hat{y}_j>0$, and $\bar{x}^j:=\hat{x}^j$ whenever $\hat{y}_j=0$. Consider any $j\in[p]$. Then, when $\hat{y}_j=0$, we have $\bar{y}_j=0$ as well as $\bar{x}^j=\hat{x}^j$, and thus $(\bar{x}^j,\bar{y}_j)=(\hat{x}^j,\hat{y}_j)$ and hence $A^j\bar{x}^j+B^j\bar{y}_j=A^j\hat{x}^j+B^j\hat{y}_j\in\mathbb{K}_j$. When $\hat{y}_j>0$, by definition we have $(\bar{x}^j,\bar{y}_j)=\frac{\bar{y}_j}{\hat{y}_j}(\hat{x}^j,\hat{y}_j)$. Moreover, because $A^j\hat{x}^j+B^j\hat{y}_j\in\mathbb{K}_j$ holds, $\frac{\bar{y}_j}{\hat{y}_j}\in\mathbb{R}_+$ and $\mathbb{K}_j$ is a cone, we deduce that $A^j\bar{x}^j+B^j\bar{y}_j\in\mathbb{K}_j$ as well. Hence, in either case we conclude that $(\bar{x},\bar{y},\bar{z})\in\widehat{\mathcal{Q}}$. Similarly, for each $j\in[p]$, define $\tilde{x}^j:=\frac{\tilde{y}_j}{\hat{y}_j}\hat{x}^j$ whenever $\hat{y}_j>0$, and $\tilde{x}^j:=\hat{x}^j$ whenever $\hat{y}_j=0$. As before, we deduce that $(\tilde{x},\tilde{y},\tilde{z})\in\widehat{\mathcal{Q}}$. Finally, for each $j\in[p]$ such that $\hat{y}_j>0$ we deduce

$$\frac{1}{2}(\bar{x}^j+\tilde{x}^j)=\frac{1}{2}\left(\frac{\bar{y}_j}{\hat{y}_j}\hat{x}^j+\frac{\tilde{y}_j}{\hat{y}_j}\hat{x}^j\right)=\frac{\bar{y}_j+\tilde{y}_j}{2\hat{y}_j}\hat{x}^j=\hat{x}^j,$$

where the last equation follows from $(\hat{y},\hat{z})=\frac{1}{2}(\bar{y},\bar{z})+\frac{1}{2}(\tilde{y},\tilde{z})$. Also, for each $j\in[p]$ such that $\hat{y}_j=0$, we have $\bar{x}^j=\tilde{x}^j=\hat{x}^j$. Thus, $(\hat{x},\hat{y},\hat{z})=\frac{1}{2}(\bar{x},\bar{y},\bar{z})+\frac{1}{2}(\tilde{x},\tilde{y},\tilde{z})$. This contradicts the fact that $(\hat{x},\hat{y},\hat{z})$ is an extreme point of $\widehat{\mathcal{Q}}$. Therefore, $(\hat{y},\hat{z})$ is an extreme point of $\operatorname{conv}(\mathcal{G})$, and thus we must have $(\hat{y},\hat{z})\in\mathcal{G}$. Hence, we have shown that every extreme point of $\widehat{\mathcal{Q}}$ is contained in $\mathcal{Q}$, as required. □

Proposition 1 is instrumental in proving the following main theorem that gives the convex hull characterization of $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$.

THEOREM 1. *For each $j\in[p]$, let $f_j:\{0,1\}^n\to\mathbb{R}_+$ be a nonnegative function and $\mathbb{K}_j$ be a convex cone containing the origin. Then, the convex hull of $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ defined in (5) is described by $\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ as defined in (6).*

**Proof of Theorem 1.** First, we observe that $\mathcal{G}:=\mathcal{G}\left(\{f_j\}_{j\in[p]}\right)\subseteq\mathbb{R}^p_+\times\{0,1\}^n$ as $f_j$'s are nonnegative functions and that $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})=\operatorname{Proj}_{x,z}\left(\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})\right)$. Moreover, $\mathcal{G}|_{z=\bar{z}}$ for any $z=\bar{z}$ is convex. Then, Proposition 1 implies that

$$\operatorname{conv}\left(\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})\right)=\left\{(x,y,z)\in\mathbb{R}^{mp}\times\mathbb{R}^p\times[0,1]^n:\ \begin{array}{l}(y,z)\in\operatorname{conv}(\mathcal{G}),\\ A^jx^j+B^jy_j\in\mathbb{K}_j,\ \forall j\in[p]\end{array}\right\}.$$

Finally, recall that the convex hull and projection operations commute, i.e., the projection of the convex hull of a set is equal to the convex hull of the projection of a set. Therefore,

$$\operatorname{conv}\left(\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})\right)=\operatorname{conv}\left(\operatorname{Proj}_{x,z}\left(\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})\right)\right)$$

$$= \operatorname{Proj}_{x,z}\left(\operatorname{conv}\left(\mathcal{Q}(\mathcal{G},\{\mathbb{K}_j\}_{j\in[p]})\right)\right)$$
$$= \left\{(x,z)\in\mathbb{R}^{mp}\times[0,1]^n:\ \exists y\in\mathbb{R}^p \text{ s.t. } \begin{array}{l}(y,z)\in\operatorname{conv}(\mathcal{G}),\\ A^jx^j+B^jy_j\in\mathbb{K}_j,\ \forall j\in[p]\end{array}\right\}$$
$$= \widehat{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]}).\quad \square$$

The interesting applications of Theorem 1 arise whenever it is relatively easy to give the explicit convex hull description of the intersections of the epigraphs of nonnegative functions $f_j:\{0,1\}^n\to\mathbb{R}_+$. One such case is when $f_j$'s are nonnegative and submodular, where $\operatorname{conv}(\operatorname{epi}(f_j))$'s are described by the *extended polymatroid inequalities*. When $f_j$'s are general set functions, a recent work of Atamtürk and Narayanan (2020) discusses the *polar inequalities* that generalize the extended polymatroid inequalities. Hence, the polar inequalities are valid for $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ for any set of nonnegative functions $f_j$'s. We will briefly review these inequalities in the next section.

Next, we focus on the $p=1$ case, introduced in Section 1, where we consider a single set function $f$ and a cone $\mathbb{K}$. To simplify our notation in this case, we define

$$\mathcal{S}(f,\mathbb{K}):=\mathcal{S}\left(\{f\},\{\mathbb{K}\}\right)=\left\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ \exists y\in\mathbb{R} \text{ s.t. } y\ge f(z),\ Ax+By\in\mathbb{K}\right\}.$$

Consider the set

$$\overline{\mathcal{S}}(f,\mathbb{K}):=\left\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ Ax+Bf(z)\in\mathbb{K}\right\}. \tag{9}$$

Then, clearly $\overline{\mathcal{S}}(f,\mathbb{K})\subseteq\mathcal{S}(f,\mathbb{K})$, and we arrive at the following immediate corollary of Theorem 1.

COROLLARY 1. *Let $f:\{0,1\}^n\to\mathbb{R}_+$ be a nonnegative function and $\mathbb{K}$ be a convex cone containing the origin. Then,* $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))\subseteq\widehat{\mathcal{S}}(f,\mathbb{K})$*. If $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$ further holds, then* $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))=\widehat{\mathcal{S}}(f,\mathbb{K})$.

**Proof of Corollary 1.** By Theorem 1 applied to the $p=1$ case, we deduce $\operatorname{conv}(\mathcal{S}(f,\mathbb{K}))=\widehat{\mathcal{S}}(f,\mathbb{K})$. Since $\overline{\mathcal{S}}(f,\mathbb{K})\subseteq\mathcal{S}(f,\mathbb{K})$, it follows that $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))\subseteq\widehat{\mathcal{S}}(f,\mathbb{K})$. If we have $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$, then $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))=\widehat{\mathcal{S}}(f,\mathbb{K})$ indeed holds, as required. $\square$

In general, the condition $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$ does not always hold. To illustrate, take a point $(\bar{x},\bar{z})\in\overline{\mathcal{S}}(f,\mathbb{K})$. Since $\overline{\mathcal{S}}(f,\mathbb{K})\subseteq\mathcal{S}(f,\mathbb{K})$, we know that $(\bar{x},\bar{z})\in\mathcal{S}(f,\mathbb{K})$. Moreover, by definition of $\mathcal{S}(f,\mathbb{K})$, we deduce that there exists $\bar{y}\in\mathbb{R}$ such that $\bar{y}\ge f(\bar{z})$ and $A\bar{x}+B\bar{y}\in\mathbb{K}$. Then, because $f$ is a nonnegative function, we have $\bar{y}\ge 0$. Moreover, using the fact that $\mathbb{K}$ is a cone we conclude that $(\alpha\bar{x},\bar{z})\in\mathcal{S}(f,\mathbb{K})$ for any $\alpha\ge 1$. However, $A\bar{x}+Bf(\bar{z})\in\mathbb{K}$ does not guarantee that $A(\alpha\bar{x})+Bf(\bar{z})\in\mathbb{K}$ if $\alpha>1$, which means that $(\alpha\bar{x},\bar{z})$ is not necessarily contained in $\overline{\mathcal{S}}(f,\mathbb{K})$. This in turn indicates that $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))$ is not equal to $\widehat{\mathcal{S}}(f,\mathbb{K})$ in general.

Nevertheless, there are some important examples, wherein the condition $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$ indeed holds and thus $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))=\widehat{\mathcal{S}}(f,\mathbb{K})$. The following proposition provides a necessary and sufficient condition for $\operatorname{conv}(\overline{\mathcal{S}}(f,\mathbb{K}))=\widehat{\mathcal{S}}(f,\mathbb{K})$.

Proposition 2. *$\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$ if and only if the following condition is satisfied:*

$$\text{every } (x,z) \text{ satisfying } Ax+Bf(z)\in\mathbb{K} \text{ also satisfies } A(\alpha x)+Bf(z)\in\mathbb{K} \text{ for any } \alpha\geq 1. \qquad (\star)$$

**Proof of Proposition 2.** **(⇒)** Take $(x,z)$ such that $Ax+Bf(z)\in\mathbb{K}$. By definition, we have $(x,z)\in\overline{\mathcal{S}}(f,\mathbb{K})$. Then, by $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$, we have $(x,z)\in\mathcal{S}(f,\mathbb{K})$. Now, since $(\alpha x,z)\in\mathcal{S}(f,\mathbb{K})$ for any $\alpha\geq 1$, once again using $\overline{\mathcal{S}}(f,\mathbb{K})=\mathcal{S}(f,\mathbb{K})$, we deduce $(\alpha x,z)\in\overline{\mathcal{S}}(f,\mathbb{K})$. Hence, $A(\alpha x)+Bf(z)\in\mathbb{K}$ for any $\alpha\geq 1$.

**(⇐)** It suffices to prove that $\mathcal{S}(f,\mathbb{K})\subseteq\overline{\mathcal{S}}(f,\mathbb{K})$. To this end, take $(x,z)\in\mathcal{S}(f,\mathbb{K})$. Then there exists $y\geq 0$ such that $y\geq f(z)$ and $Ax+By\in\mathbb{K}$. If $y=0$, then $f(z)=0$ as $f$ is a nonnegative function. This implies that $Ax+f(z)=Ax+By\in\mathbb{K}$, so $(x,z)\in\overline{\mathcal{S}}(f,\mathbb{K})$. If $y>0$, it follows from $A(f(z)/y)x+Bf(z)=(f(z)/y)(Ax+By)\in\mathbb{K}$. In this case, since $y/f(z)\geq 1$, the premise of this direction implies that $Ax+Bf(z)=A(y/f(z))(f(z)/y)x+Bf(z)\in\mathbb{K}$, and therefore, $(x,z)\in\overline{\mathcal{S}}(f,\mathbb{K})$. Thus, $\mathcal{S}(f,\mathbb{K})\subseteq\overline{\mathcal{S}}(f,\mathbb{K})$, as required. □

We highlight a useful sufficient condition that implies Condition ($\star$).

Remark 1. Let $\mathbb{K}$ be a convex cone containing the origin. Suppose $\mathbb{K}$, $A$ and $B$ are such that

$$\text{for any } (x,z) \text{ satisfying } Ax+Bf(z)\in\mathbb{K}, \text{ we also have } Ax\in\mathbb{K}.$$

For any $(x,z)$ with $Ax+Bf(z)\in\mathbb{K}$ and any $\alpha\geq 1$, we have $(\alpha-1)Ax\in\mathbb{K}$, so $A(\alpha x)+Bf(z)=(\alpha-1)Ax+(Ax+Bf(z))\in\mathbb{K}$. Therefore, Condition ($\star$) is immediately satisfied. ∎

We discuss in Section 5 some applications where Condition ($\star$) holds. In particular, using Remark 1, we will show in Section 5.1 that $\mathcal{H}$ and $\mathcal{R}$ defined in (2) and (3) satisfy Condition ($\star$) and consider other applications in Sections 5.2 – 5.4.

Corollary 1 and Proposition 2 can be extended to the case of multiple functions. Consider

$$\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]}):=\left\{(x,z)\in\mathbb{R}^{mp}\times\{0,1\}^n:\ A^jx^j+B^jf_j(z),\quad\forall j\in[p]\right\}, \qquad (10)$$

which is a subset of $\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ defined in (5).

Corollary 2. *Let $f_j:\{0,1\}^n\to\mathbb{R}_+$ be a nonnegative function and $\mathbb{K}_j$ be a convex cone containing the origin for $j\in[p]$. Then, $\operatorname{conv}(\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]}))\subseteq\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$. If $\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})=\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ further holds, then*

$$\operatorname{conv}(\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]}))=\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]}).$$

**Proof of Corollary 2.** Similar to the proof of Corollary 1. □

Moreover, we can characterize when $\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})=\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$.

Proposition 3. $\overline{\mathcal{S}}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})=\mathcal{S}(\{f_j\}_{j\in[p]},\{\mathbb{K}_j\}_{j\in[p]})$ *if and only if the following holds:*

$$\begin{aligned}&\text{every } (x,z) \text{ satisfying } A^j x^j + B^j f_j(z)\in\mathbb{K}_j \text{ for all } j\in[p]\\ &\quad\text{also satisfies } A^j\alpha_j x^j + B^j f_j(z)\in\mathbb{K}_j \text{ for any } \alpha_j\geq 1 \text{ for all } j\in[p].\end{aligned}\tag{$\star\star$}$$

**Proof of Proposition 3.** Similar to the proof of Proposition 2. □

In Section 5.2, inspired by fractional binary programming, we study a generalization of $\mathcal{H}$ and $\mathcal{R}$ that takes multiple conic quadratic constraints, whose convex hull can be characterized based on Corollary 2.

We close this section by noting that the structure of the set $\mathcal{S}(f,\mathbb{K})$ allows us to easily embed constraints on the continuous variables $x$ as well.

Remark 2 (Additional constraints on continuous variables). Let $\mathbb{C}$ be a convex cone containing the origin, and consider the set along with its transformation given by

$$\begin{aligned}&\left\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ \exists y\in\mathbb{R}\text{ s.t. } y\geq f(z),\ \tilde{A}x+\tilde{B}y\in\tilde{\mathbb{K}},\ Cx\in\mathbb{C}\right\}\\ &=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ \exists y\in\mathbb{R}\text{ s.t. } y\geq f(z),\ Ax+By\in\mathbb{K}\},\end{aligned}$$

where we set $\mathbb{K}=\tilde{\mathbb{K}}\times\mathbb{C}$, $A=[\tilde{A};C]$, and $B=[\tilde{B};0]$. Note that through this representation, we deduce that the additional conic constraints on the continuous variables $x$ can easily be embedded into our desirable form of the set $\mathcal{S}(f,\mathbb{K})$. ∎

## 3. The extended polymatroid inequalities and the polar inequalities

In this section, we discuss two classes of inequalities, namely, the extended polymatroid inequalities and the polar inequalities, which can be used to describe $\operatorname{conv}(\operatorname{epi}(f))$ completely or partially (and thus these give either $\operatorname{conv}(\mathcal{S}(f,\mathbb{K}))$ or remain valid for it) when $f$ has desirable structure.

Given a set function $f:2^{[n]}\to\mathbb{R}$, where $2^{[n]}$ is the power set of $[n]$, let the *associated polyhedron* of $f$ be defined as

$$\mathcal{P}_f:=\{\pi\in\mathbb{R}^n:\ \pi(V)\leq f(V),\ \forall V\subseteq[n]\},\tag{11}$$

where $\pi(V):=\sum_{i\in V}\pi_i$ and $\pi(\emptyset)=0$. By slight abuse of notation, throughout, we refer to a set function $f:2^{[n]}\to\mathbb{R}$ also as $f:\{0,1\}^n\to\mathbb{R}$, where $f(V):=f(\mathbf{1}_V)$ for $V\subseteq[n]$ and $\mathbf{1}_V$ denotes the characteristic vector of $V$. When $f$ is a submodular function, i.e., $f$ satisfies

$$f(S_1)+f(S_2)\geq f(S_1\cup S_2)+f(S_1\cap S_2),\quad \forall S_1,S_2\subseteq[n],$$

$\mathcal{P}_f$ is called the *extended polymatroid of* $f$. Note that $\mathcal{P}_f$ is nonempty if and only if $f(\emptyset)\geq 0$. In general, $f$ need not satisfy $f(\emptyset)\geq 0$. Nevertheless, we can take $f-f(\emptyset)$ instead so that $(f-f(\emptyset))(\emptyset)=0$. Hence, $\mathcal{P}_{f-f(\emptyset)}$ is always nonempty. Hereinafter, we use notation $\tilde{f}$ to denote $f-f(\emptyset)$ for any set function $f$. In particular, $\tilde{f}$ is submodular when $f$ is submodular.

The associated polyhedron $\mathcal{P}_f$ is instrumental in generating valid inequalities for epi($f$) due to a close polarity relation between $\mathcal{P}_f$ and conv(epi($f$)). (We refer the reader to Nemhauser and Wolsey (1988, Chapter I.4.5) for a review of how the concept of polarity is used to obtain facets of a polyhedron.) From this relation, Atamtürk and Narayanan (2020) show that the so-called *polar inequalities*

$$y - f(\emptyset) \geq \pi^\top z, \ \forall \pi \in \mathcal{P}_{\tilde{f}} \tag{12}$$

are valid for $(y,z) \in \text{conv}(\text{epi}(f))$. Atamtürk and Narayanan (2020) also prove that inequalities (12) are facet-defining for conv(epi($f$)) if and only if $\pi$ is an extreme point of $\mathcal{P}_{\tilde{f}}$. Therefore, the extreme points of $\mathcal{P}_{\tilde{f}}$ characterize facet-defining polar inequalities. In the case of a general set function $f$, the inclusion relationship in (12) is strict (see Example 1, Atamtürk and Narayanan 2020), which means that the polar inequalities may not be sufficient to describe the convex hull of epi($f$).

When $f$ is submodular, the polar inequalities are precisely the *extended polymatroid inequalities* (Atamtürk and Narayanan 2008). Moreover, it is well-known that when $f$ is submodular, the extended polymatroid inequalities indeed provide a complete description of conv(epi($f$)).

THEOREM 2 (**Lovász (1983), Atamtürk and Narayanan (2008, Theorem 1)**). *Let* $f : \{0,1\}^n \to \mathbb{R}$ *be a submodular function. Then,*

$$\text{conv}(\text{epi}(f)) = \left\{(y,z) \in \mathbb{R} \times [0,1]^n : \ y - f(\emptyset) \geq \pi^\top z, \ \forall \pi \in \mathcal{P}_{\tilde{f}}\right\}.$$

Edmonds (1970) provides the following explicit characterization of the extreme points of $\mathcal{P}_{\tilde{f}}$.

THEOREM 3 (**Edmonds (1970)**). *Let* $f : \{0,1\}^n \to \mathbb{R}$ *be a submodular function. Then* $\pi \in \mathbb{R}^n$ *is an extreme point of* $\mathcal{P}_{\tilde{f}}$ *if and only if there exists a permutation* $\sigma$ *of* $[n]$ *such that* $\pi_{\sigma(t)} = f(V_t) - f(V_{t-1})$, *where* $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ *for* $t \in [n]$ *and* $V_0 := \emptyset$ *by definition.*

We note that when $f$ is not submodular, there are extreme points $\pi$ of $\mathcal{P}_f$ are not necessarily of the form given in Therorem 3. The proof of Theroem 3 yields an $O(n \log n)$ time algorithm for separating a violated extended polymatroid inequality given a point $(\bar{y}, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n$, which amounts to solving $\max_\pi \left\{\bar{z}^\top \pi : \ \pi \in \mathcal{P}_{\tilde{f}}\right\}$ (Atamtürk and Narayanan 2008, Section 2).

These results on submodular functions lead us to the following corollary of Theorem 1.

COROLLARY 3. *Suppose* $f : \{0,1\}^n \to \mathbb{R}_+$ *is a nonnegative submodular function and* $\mathbb{K}$ *is a convex cone containing the origin. Then,* $\text{conv}(\mathcal{S}(f,\mathbb{K}))$ *is given by the extended polymatroid inequalities for* $f - f(\emptyset)$ *and the homogeneous conic constraint* $Ax + By \in \mathbb{K}$.

**Proof of Corollary 3.** This is a direct consequence of Theorems 1 and 2. □

COROLLARY 4. *For each $j \in [p]$, let $f_j : \{0,1\}^n \to \mathbb{R}_+$ be a nonnegative submodular function and $\mathbb{K}_j$ be a convex cone containing the origin. Then, the convex hull of $\mathcal{S}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]})$ defined in (5) is described by $\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]})$ as defined in (6), and moreover*

$$\widehat{\mathcal{S}}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]}) = \left\{ (x,z) \in \mathbb{R}^{mp} \times [0,1]^n : \ \exists y \in \mathbb{R}^p \ s.t. \ \begin{array}{l} (y_j, z) \in \operatorname{conv}(\operatorname{epi}(f_j)), \\ A^j x^j + B^j y_j \in \mathbb{K}_j, \end{array} \ \forall j \in [p] \right\}.$$

**Proof of Corollary 4.** Let $\mathcal{G} := \{(y,z) \in \mathbb{R}^p \times \{0,1\}^n : \ y_j \ge f_j(z), \ \ \forall j \in [p]\}$. Then, $\mathcal{G} \subseteq \mathbb{R}^p_+ \times \{0,1\}^n$ as $f_j$'s are nonnegative functions. The result follows from Theorem 1 and the following fact. When $f_j$ is a nonnegative submodular function for each $j \in [p]$, from Theorem 2 of Baumann et al. (2013) (see also, Proposition 1, Kılınç-Karzan et al. 2019) we deduce that

$$\operatorname{conv}(\mathcal{G}) = \{(y,z) \in \mathbb{R}^p \times [0,1]^n : \ (y_j, z) \in \operatorname{conv}(\operatorname{epi}(f_j)), \ \ \forall j \in [p]\}. \quad \square$$

We refer the reader to Edmonds (1970) and Lovász (1983) for a list of basic submodular functions as well as operations preserving submodularity.

REMARK 3. We close this section by emphasizing that our result holds not only for a function of the form $\sqrt{\sigma + \sum_{i\in[n]} c_i z_i}$ for $z \in \{0,1\}^n$, but also for general submodular functions, such as $g(\sigma + \sum_{i\in[n]} c_i z_i)$ where $g$ is concave. In particular, the constant elasticity of substitution function, $\left(\sum_{i\in[n]} c_i^p z_i^p\right)^{1/p}$ for any $c \in \mathbb{R}^n_+$ and $p \ge 1$ is submodular, and this property will be used in Section 5.4 when we consider a norm constraint in the mixed-integer conic reformulation of a distributionally robust optimization problem. ∎

## 4. Extensions

Our results are applicable in the cases where the conic constraint is non-homogeneous, or the function $f(x)$ does not satisfy the nonnegativity assumption, or where we have only a partial convex hull description of the epigraph of $f(x)$ available.

REMARK 4 (NON-HOMOGENEOUS CONIC CONSTRAINTS). Our results are still of interest when $\mathcal{S}(f,\mathbb{K})$ has a non-homogeneous constraint, i.e., $Ax + By + C \in \mathbb{K}$ for some $C \ne 0$ instead of $Ax + By \in \mathbb{K}$. Indeed, by adding a new variable and an affine constraint, we can always rewrite this set using a homogeneous conic constraint. That is,

$$\begin{aligned} &\{(x,z) \in \mathbb{R}^m \times \{0,1\}^n : \ \ \exists y \in \mathbb{R} \text{ s.t. } y \ge f(x), \ Ax + By + C \in \mathbb{K}\} \\ &\quad = \{(x,z) \in \mathbb{R}^m \times \{0,1\}^n : \ \ \exists y, v \in \mathbb{R} \text{ s.t. } y \ge f(x), \ Ax + By + Cv \in \mathbb{K}, \ v = 1\}. \end{aligned} \tag{13}$$

Here, $\{(x,v,z) \in \mathbb{R}^m \times \mathbb{R} \times \{0,1\}^n : \exists y \in \mathbb{R} \text{ s.t. } y \ge f(x), \ Ax + By + Cv \in \mathbb{K}\}$ is of the form $\mathcal{S}(f,\mathbb{K})$, and the set in (13) is obtained from the intersection of this set and an affine hyperplane defined by $v = 1$ after projecting out $v$. Therefore,

$$\left\{(x,z) \in \mathbb{R}^m \times \{0,1\}^n : \ \exists (y,v) \in \mathbb{R}^2_+ \text{ s.t. } (y,z) \in \operatorname{conv}(\operatorname{epi}(f)), \ Ax + By + Cv \in \mathbb{K}, \ v = 1\right\} \tag{14}$$

is a valid convex relaxation for the set in (13) by Theorem 1. ∎

While Remark 4 is useful, due to the presence of the affine constraint $v=1$ in (14), (14) may not provide a convex hull description of the set in (13). This is demonstrated in the following example.

EXAMPLE 1. Note that

$$\{(x,z)\in\mathbb{R}^2\times\{0,1\}^2:\ \sqrt{x_1^2+z_1+z_2}\le x_2-1\}=\{(x,z)\in\mathbb{R}^2\times\{0,1\}^2:\ (\sqrt{z_1+z_2},x_1,x_2-1)\in\mathbb{L}^3\}$$

is of the form $\overline{\mathcal{S}}(f,\mathbb{K})$ where $f(z)=\sqrt{z_1+z_2}$. In addition, $f$ is submodular, $f(\emptyset)=0$, and

$$\mathcal{P}_f=\left\{\pi\in\mathbb{R}^2:\ \pi_1\le 1,\ \pi_2\le 1,\ \pi_1+\pi_2\le\sqrt{2}\right\},$$

with extreme points $(\pi_1,\pi_2)=(1,\sqrt{2}-1)$ and $(\pi_1,\pi_2)=(\sqrt{2}-1,1)$. Then, (14) is given by

$$\left\{(x,z)\in\mathbb{R}^2\times\{0,1\}^2:\ \exists(y,v)\in\mathbb{R}_+\times\mathbb{R}\ \text{ s.t. }\ \begin{array}{l} y\ge(\sqrt{2}-1)z_1+z_2,\quad y\ge z_1+(\sqrt{2}-1)z_2,\\ \sqrt{x_1^2+y^2}\le x_2-v,\quad v=1\end{array}\right\}. \quad (15)$$

In Appendix A, we show that (15) has an extreme point where either $z_1$ or $z_2$ is fractional. ∎

REMARK 5 (ARBITRARY SET FUNCTIONS). Consider the set $\overline{\mathcal{S}}(h,\mathbb{K})$ where $h$ is neither nonnegative nor nonpositive. Let $h_{\min}:=\min_{z\in\{0,1\}^n}h(z)$. Then, the function $f(z):=h(z)-h_{\min}$ is a nonnegative function and we have

$$\begin{aligned}\overline{\mathcal{S}}(h,\mathbb{K})&=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\quad Ax+Bh(z)\in\mathbb{K}\}\\&=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\quad Ax+Bf(z)+Bh_{\min}\in\mathbb{K}\}.\end{aligned}$$

Now, we can apply the transformations from Remark 4 and use Corollary 1 to arrive at the desired representation. ∎

REMARK 6 (SUPERMODULAR FUNCTIONS). Suppose the function of interest is supermodular. Let $h:\{0,1\}^n\to\mathbb{R}$ be a supermodular function and $\mathbb{K}$ be a convex cone containing the origin. If $h$ is nonnegative, then one would want to apply Corollary 1 on the set $\overline{\mathcal{S}}(h,\mathbb{K})=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ Ax+Bh(z)\in\mathbb{K}\}$ and convexify it by introducing an auxiliary variable $y$ and replacing $Ax+Bh(z)\in\mathbb{K}$ with the constraints $(y,z)\in\operatorname{conv}(\operatorname{epi}(h))$ and $Ax+By\in\mathbb{K}$. However, we do not know how to describe $\operatorname{conv}(\operatorname{epi}(h))$ when $h$ is supermodular. Nevertheless, we can still utilize our technique of transforming the homogeneous conic constraint into a non-homogeneous one and applying Remark 4. To this end, define $h_{\max}:=\max_{z\in\{0,1\}^n}\{h(z)\}$ and $f(z):=-h(z)+h_{\max}$. Then, we know that $f$ is a nonnegative submodular function, and we have

$$\begin{aligned}\overline{\mathcal{S}}(h,\mathbb{K})&=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\quad Ax+B(h_{\max}-f(z))\in\mathbb{K}\}\\&=\{(x,z)\in\mathbb{R}^m\times\{0,1\}^n:\ \ \exists v\in\mathbb{R}\text{ s.t. }Ax+Bh_{\max}v-Bf(z)\in\mathbb{K},\ v=1\}.\end{aligned}$$

Therefore, as $f(z)$ is submodular, we can add the extended polymatroid inequalities for $f$ to strengthen the set $\overline{\mathcal{S}}(h,\mathbb{K})$.

We note that this transformation applied to a supermodular function is indeed easy as computing $h_{\max}$ amounts to minimizing a submodular function, which can be done in polynomial time. ∎

Remark 7. As discussed in Section 3, for general (not necessarily submodular) $f$, Atamtürk and Narayanan (2020) introduced the class of polar inequalities that are valid for $\widehat{\mathcal{S}}(f,\mathbb{K})$ for any nonnegative function $f$, which can be used to strengthen the continuous relaxation. ∎

## 5. Applications

In this section, we present several optimization problems in which sets of the form $\mathcal{S}(f,\mathbb{K})$ appear as a substructure. In this respect, we highlight two forms of objective functions that immediately lead to our desired structure:

(i) $\min\ \sqrt{f(z)+\|Dx+d\|_2^2}$,
(ii) $\min\ \frac{\|Dx+d\|_2^2}{f(z)}$,

where $f(z)$ is a nonnegative function, in addition to norm constraints of the form $\|(z;Dx)\|\le t$.

In Section 5.1, we discuss the implications of our results in the context of applications from Atamtürk and Gómez (2020) which have the objectives of form (i) where $f$ is submodular, and in Section 5.2 we explore their use in the case of fractional programming. In Section 5.3, we point out connections with variants of best subset selection problem (Gómez and Prokopyev 2020) in which the problems have objectives of the form (ii) and the function $f$ is supermodular. Finally, in Section 5.4, we highlight the connection of our work with substructures exploited within the context of distributionally robust chance constrained problems involving binary variables.

### 5.1. Recovering the results of Atamtürk and Gómez (2020)

Recall the following sets studied by Atamtürk and Gómez (2020):

$$\mathcal{H}=\left\{(x,z)\in\mathbb{R}_+^m\times\{0,1\}^n:\ \sqrt{\sigma+\sum_{i\in[n]}c_iz_i+\sum_{j\in[m-1]}d_jx_j^2}\le x_m\right\},$$

$$\mathcal{R}=\left\{(x,z)\in\mathbb{R}_+^m\times\{0,1\}^n:\ \sigma+\sum_{i\in[n]}c_iz_i+\sum_{j\in[m-2]}d_jx_j^2\le 4x_{m-1}x_m\right\}.$$

In this section, we let the function $f:\{0,1\}^n\to\mathbb{R}_+$ to be defined by $f(z):=\sqrt{\sigma+\sum_{i\in[n]}c_iz_i}$. Note that $f$ is submodular, provided that $\sigma+\sum_{i\in[n]}c_iz_i$ is nonnegative for $z\in\{0,1\}^n$. Hence, by Theorem 2, $\operatorname{conv}(\operatorname{epi}(f))$ is described by the extended polymatroid inequalities for $f-\sqrt{\sigma}$ and $\mathbf{0}\le z\le\mathbf{1}$. Motivated by this observation, Atamtürk and Gómez (2020) prove the following results for the sets $\mathcal{H}$ and $\mathcal{R}$ by analyzing KKT conditions of a generic linear optimization problem over the domain $\mathcal{H}$ or $\mathcal{R}$.

Proposition 4 (**Atamtürk and Gómez (2020, Proposition 5)**). *Let* $f(z) = \sqrt{\sigma + \sum_{i\in[n]} c_i z_i}$ *where* $\sigma + \sum_{i\in[n]} c_i z_i \geq 0$ *for* $z \in \{0,1\}^n$, *and let* $\mathcal{H}$ *be defined as in* (2). *Then*

$$\operatorname{conv}(\mathcal{H}) = \left\{ (x,z) \in \mathbb{R}_+^m \times \mathbb{R}^n : \ \exists y \in \mathbb{R}_+ \ \text{ s.t. } (y,z) \in \operatorname{conv}(\operatorname{epi}(f)), \ \sqrt{y^2 + \sum_{i\in[m-1]} d_i x_i^2} \leq x_m \right\}.$$

Proposition 5 (**Atamtürk and Gómez (2020, Proposition 6)**). *Let* $f(z) = \sqrt{\sigma + \sum_{i\in[n]} c_i z_i}$ *where* $\sigma + \sum_{i\in[n]} c_i z_i \geq 0$ *for* $z \in \{0,1\}^n$, *and let* $\mathcal{R}$ *be defined as in* (3). *Then*

$$\operatorname{conv}(\mathcal{R}) = \left\{ (x,z) \in \mathbb{R}_+^m \times [0,1]^n : \ \exists y \in \mathbb{R}_+ \ \text{ s.t. } (y,z) \in \operatorname{conv}(\operatorname{epi}(f)), \ y^2 + \sum_{i\in[m-2]} d_i x_i^2 \leq 4 x_{m-1} x_m \right\}.$$

We next show that these results are simple corollaries of Theorem 1. Recall $\mathbb{L}^{k+1} = \{(\xi, y) \in \mathbb{R}^k \times \mathbb{R} : \ y \geq \|\xi\|_2\}$ denotes the SOC in $\mathbb{R}^{k+1}$. Throughout, let $e_i$ be the unit vector of appropriate dimension with a one in the $i$th component, and let $\operatorname{Diag}(\cdot)$ represent a diagonal matrix with the specified diagonal entries.

Corollary 5. *Theorem 1 implies Proposition 4.*

**Proof of Corollary 5.** Note that the mixed-integer set $\mathcal{H}$ considered in Proposition 4 satisfies

$$\begin{aligned}
\mathcal{H} &= \left\{ (x,z) \in \mathbb{R}_+^m \times \{0,1\}^n : \ \sqrt{(f(z))^2 + \sum_{j\in[m-1]} d_j x_j^2} \leq x_m \right\} \\
&= \left\{ (x,z) \in \mathbb{R}_+^m \times \{0,1\}^n : \ \left[ f(z); \ \sqrt{d_1} x_1; \ \ldots; \ \sqrt{d_{m-1}} x_{m-1}; \ x_m \right] \in \mathbb{L}^{m+1} \right\} \\
&= \left\{ (x,z) \in \mathbb{R}_+^m \times \{0,1\}^n : \ \tilde{A}x + \tilde{B} f(z) \in \mathbb{L}^{m+1} \right\},
\end{aligned}$$

where $\tilde{B} := e_1 \in \mathbb{R}^{m+1}$ and $\tilde{A} := [0^\top; \ \operatorname{Diag}(\sqrt{d_1}; \ \ldots; \ \sqrt{d_{m-1}}; \ 1)] \in \mathbb{R}^{(m+1)\times m}$. Note that if $(x,z)$ satisfies $\tilde{A}x + \tilde{B} f(z) \in \mathbb{L}^{m+1}$, then due to the structure of $\tilde{A}, \tilde{B}$ and the cone $\mathbb{L}^{m+1}$, we deduce that $x$ satisfies $\tilde{A}x \in \mathbb{L}^{m+1}$ as well. Then, by Remark 1, Condition ($\star$) holds. Therefore, by Proposition 2,

$$\mathcal{H} = \left\{ (x,z) \in \mathbb{R}_+^m \times \{0,1\}^n : \ \exists y \in \mathbb{R} \text{ s.t. } y \geq f(z), \ \tilde{A}x + \tilde{B}y \in \mathbb{L}^{m+1} \right\}.$$

The result then follows by applying Theorems 1–3, Corollary 4, and Remark 2 so that we can handle the nonnegativity constraint on the continuous variables $x \in \mathbb{R}_+^m$ (which is nothing but a very simple conic constraint) in addition to the SOC constraint $\tilde{A}x + \tilde{B}y \in \mathbb{L}^{m+1}$. □

Corollary 6. *Theorem 1 implies Proposition 5.*

**Proof of Corollary 6.** Note that $y^2 + \sum_{i\in[m-2]} d_i x_i^2 \leq 4 x_{m-1} x_m$ is a rotated SOC constraint given by

$$\left[ y; \sqrt{d_1} x_1; \ldots; \sqrt{d_{m-2}} x_{m-2}; \ x_{m-1} - x_m; \ x_{m-1} + x_m \right]^\top \in \mathbb{L}^{m+2}.$$

By defining $\tilde{B} := e_1 \in \mathbb{R}^{m+2}$ and $\tilde{A} := [0^\top;\ \mathrm{Diag}(\sqrt{d_1};\ \ldots;\ \sqrt{d_{m-2}});\ 0^\top;\ 0^\top] + e_m[0,\ldots,0,1,-1] + e_{m+1}[0,\ldots,0,1,1] \in \mathbb{R}^{(m+2)\times m}$, this constraint is equivalent to $\tilde{A}x + \tilde{B}y \in \mathbb{L}^{m+2}$. Then,

$$\begin{aligned}\mathcal{R} &= \left\{(x,z) \in \mathbb{R}^m_+ \times \{0,1\}^n :\ (f(z))^2 + \sum_{j\in[m-2]} d_j x_j^2 \le 4x_{m-1}x_m\right\}\\ &= \left\{(x,z) \in \mathbb{R}^m_+ \times \{0,1\}^n :\ \tilde{A}x + \tilde{B}f(z) \in \mathbb{L}^{m+2}\right\}\\ &= \left\{(x,z) \in \mathbb{R}^m_+ \times \{0,1\}^n :\ \exists y \in \mathbb{R} \text{ s.t. } y \ge f(z),\ \tilde{A}x + \tilde{B}y \in \mathbb{L}^{m+2}\right\},\end{aligned}$$

where the last line follows from Proposition 2 (exactly as in the case of Corollary 5). Furthermore, by applying the transformation from Remark 2 so that we can handle the nonnegativity constraint on the continuous variables $x \in \mathbb{R}^m_+$ in addition to the SOC constraint $\tilde{A}x + \tilde{B}y \in \mathbb{L}^{m+2}$, we conclude that the result follows from Theorem 1, Corollary 4, and Theorems 2 and 3. $\square$

### 5.2. Multiple conic quadratic constraints

Recall that Theorem 1 may take multiple functions into account at the same time. Based on Theorem 1, we can characterize the convex hull of a set defined by multiple conic quadratic constraints, generalizing the results on $\mathcal{H}$ and $\mathcal{R}$. For two finite sets $H, R$ of indices, the following set is defined by $|H|$ conic quadratic constraints of the type used in $\mathcal{H}$ and $|R|$ constraints of the type used in $\mathcal{R}$:

$$\mathcal{M} := \left\{(x,z) :\ \begin{array}{l} (x,z) \in \mathbb{R}^{m(|H|+|R|)}_+ \times \{0,1\}^n, \\ \sqrt{\sigma_\ell + \sum_{i\in[n]} c_{\ell,i}z_i + \sum_{j\in[m-1]} d_{\ell,j}x_{\ell,j}^2} \le x_{\ell,m},\quad \forall \ell \in H \\ \sigma_\ell + \sum_{i\in[n]} c_{\ell,i}z_i + \sum_{j\in[m-2]} d_{\ell,j}x_{\ell,j}^2 \le 4x_{\ell,m-1}x_{\ell,m},\quad \forall \ell \in R \end{array}\right\}. \tag{16}$$

We are interested in $\mathrm{conv}(\mathcal{M})$. As in the proofs of Corollary 5 and 6, we can rewrite $\mathcal{M}$ as

$$\mathcal{M} := \left\{(x,z) \in \mathbb{R}^{m(|H|+|R|)}_+ \times \{0,1\}^n :\ \begin{array}{l} \tilde{A}^\ell x^\ell + \tilde{B}^\ell f_\ell(z) \in \mathbb{L}^{m+1},\quad \forall \ell \in H \\ \tilde{A}^\ell x^\ell + \tilde{B}^\ell f_\ell(z) \in \mathbb{L}^{m+1},\quad \forall \ell \in R \end{array}\right\},$$

where $x^\ell = (x_{\ell,1},\ldots,x_{\ell,m})^\top$ and $\tilde{A}^\ell, \tilde{B}^\ell$ are defined as in the proofs of Corollary 5 and 6 for $\ell \in H \cup R$. Notice that $\mathcal{M}$ is of the form $\overline{\mathcal{S}}(\{f_j\}_{j\in[p]}, \{\mathbb{K}_j\}_{j\in[p]})$ defined as in (10). Moreover, as both $\mathcal{H}$ and $\mathcal{R}$ satisfy Condition ($\star$), $\mathcal{M}$ satisfies the condition ($\star\star$). Therefore, we can apply Corollary 2 and Proposition 3 to obtain the following proposition characterizing the convex hull of $\mathcal{M}$.

Proposition 6. *For $\ell \in H \cup R$, let $f_\ell(z) = \sqrt{\sigma_\ell + \sum_{i\in[n]} c_{\ell,i}z_i}$ where $\sigma_\ell + \sum_{i\in[n]} c_{\ell,i}z_i \ge 0$ for $z \in \{0,1\}^n$, and let $\mathcal{M}$ be defined as in (16). Then*

$$\mathrm{conv}(\mathcal{M}) = \left\{(x,z) :\ \begin{array}{l} (x,z) \in \mathbb{R}^{m(|H|+|R|)}_+ \times [0,1]^n, \\ \exists y_\ell \in \mathbb{R}_+ \text{ s.t. } (y_\ell, z) \in \mathrm{conv}(\mathrm{epi}(f_\ell)),\ \sqrt{y_\ell^2 + \sum_{j\in[m-1]} d_{\ell,j}x_{\ell,j}^2} \le x_{\ell,m},\quad \forall \ell \in H \\ \exists y_\ell \in \mathbb{R}_+ \text{ s.t. } (y_\ell, z) \in \mathrm{conv}(\mathrm{epi}(f_\ell)),\ y_\ell^2 + \sum_{j\in[m-2]} d_{\ell,j}x_{\ell,j}^2 \le 4x_{\ell,m-1}x_{\ell,m},\quad \forall \ell \in R \end{array}\right\}.$$

We remark that Proposition 6 generalizes Propositions 4 and 5. For the rest of this section, we list some applications of Proposition 6.

Remark 8 (fractional binary programs). We can use $\mathcal{M}$ to model optimization problems of the following form:

$$\min_{z\in\mathcal{X}} \sum_{\ell\in R} \frac{a_{\ell,0}+\sum_{i\in[n]} a_{\ell,i}z_i}{b_{\ell,0}+\sum_{i\in[n]} b_{\ell,i}z_i}, \tag{17}$$

where $\mathcal{X}\subseteq\{0,1\}^n$ and $a_{\ell,0}, a_{\ell,i}, b_{\ell,0}, b_{\ell,i}$ for $i\in[n]$ are all nonnegative numbers. The fractional optimization model (17) is used in a wide range of application domains including modeling multinomial logit (MNL) choice models in assortment optimization, set covering, market share based facility location, stochastic service systems, bi-clustering, and optimization of boolean query for databases (see, Borrero et al. 2017, and references therein). We can reformulate (17) by introducing an auxiliary variable for each fraction. Note that

$$\frac{a_{\ell,0}+\sum_{i\in[n]} a_{\ell,i}z_i}{b_{\ell,0}+\sum_{i\in[n]} b_{\ell,i}z_i} \le u_\ell \quad\Leftrightarrow\quad a_{\ell,0}+\sum_{i\in[n]} a_{\ell,i}z_i \le u_\ell v_\ell,\quad v_\ell = b_{\ell,0}+\sum_{i\in[n]} b_{\ell,i}z_i.$$

Then (17) is equivalent to

$$\min_{z\in\{0,1\}^n,\ u,v\in\mathbb{R}^\ell} \left\{ \sum_{\ell\in R} u_\ell : \begin{array}{l} 4a_{\ell,0}+\sum_{i\in[n]} 4a_{\ell,i}z_i \le 4u_\ell v_\ell, \quad \forall \ell\in R, \\ v_\ell = b_{\ell,0}+\sum_{i\in[n]} b_{\ell,i}z_i, \quad \forall \ell\in R \end{array} \right\}.$$

In particular, $4a_{\ell,0}+\sum_{i\in[n]} 4a_{\ell,i}z_i \le 4u_\ell v_\ell$ for $\ell\in R$ give rise to a set of the form $\mathcal{M}$. ∎

### 5.3. Best subset selection

Gómez and Prokopyev (2020) study the following general model for the *Best Subset Selection* (BSS) problem:

$$\min_{\beta\in\mathbb{R}^n, z\in\{0,1\}^n} \left\{ \frac{\|a-U\beta\|_2^2}{g(\sum_{i\in[n]} z_i)} :\ -Mz_i \le \beta_i \le Mz_i,\ \forall i\in[n] \right\}, \tag{18}$$

where $a\in\mathbb{R}^k$, $U\in\mathbb{R}^{k\times n}$ are data, $M\in\mathbb{R}_+$ corresponds to big-M value, and $g:\mathbb{R}_+\to\mathbb{R}_+$ is a non-increasing convex function. The best subset selection problem in linear regression is to find a sparse subset of regressors that best fits the data. To this end, in Problem (18), the regression variables are $\beta$ and each binary variable $z_i$ determines whether a regressor $\beta_i$ is selected. While mean squared error (MSE) is a popular criterion to measure the goodness of fit, other criteria such as Akaike Information Criterion (AIC), corrected AIC (AICc), and Bayesian Information Criterion (BIC) have also been proposed. The latter three criteria have desirable properties that address some shortcomings of the MSE criterion, but they involve (non-convex) logarithmic terms in the objective function that call for advanced solution methods. In particular, for AIC and BIC, we have $g(z) = e^{-\alpha \sum_{i\in[n]} z_i}$, and for AICc we have $g(z) = e^{-2\alpha/(\alpha - \sum_{i\in[n]} z_i)}$ for an appropriate choice of $\alpha \geq 0$. We refer the reader to Gómez and Prokopyev (2020) for more details.

Gómez and Prokopyev (2020) work with the following reformulation of (18):

$$\min_{\beta\in\mathbb{R}^n, z\in\{0,1\}^n, s\in\mathbb{R}_+} \left\{ \frac{\|a - U\beta\|_2^2}{s} : \; s \leq g(\sum_{i\in[n]} z_i), \; -Mz_i \leq \beta_i \leq Mz_i, \; \forall i \in [n] \right\}. \tag{19}$$

In (19), $s \leq g(\sum_{i\in[n]} z_i)$ is equivalent to $-s \geq -g(\sum_{i\in[n]} z_i)$, and since $-g(\sum_{i\in[n]} z_i)$ is submodular, the feasible region of (19) can be strengthened by applying the extended polymatroid inequalities for $-g$. However, the objective of (19) is a still fractional function. Consequently, Gómez and Prokopyev (2020) apply a customized Newton-type method after parameterizing the fraction.

In contrast, we observe here that (18) admits a SOC reformulation with a linear objective and whose feasible region contains a substructure that fits our framework. Define $h(z) := g(\sum_{i\in[n]} z_i)$. By introducing a new variable $t \in \mathbb{R}_+$ to capture the objective function, we observe that (18) is equivalent to the following problem:

$$\min_{t\in\mathbb{R}_+, \beta\in\mathbb{R}^n, z\in\{0,1\}^n} \left\{ t : \; t \cdot h(z) \geq \|a - U\beta\|_2^2, \; -Mz_i \leq \beta_i \leq Mz_i, \; \forall i \in [n] \right\}.$$

Note that the nonlinear constraint $t \cdot h(z) \geq \|a - U\beta\|_2^2$ in this formulation is equivalent to requiring

$$[2a - 2U\beta; \; t - h(z); \; t + h(z)] \in \mathbb{L}^{k+2}.$$

Furthermore, when we define the function $h(z) := g(\sum_{i\in[n]} z_i)$ based on a nonnegative non-increasing convex function $g$, we deduce that $h(z)$ is nonnegative and supermodular. By using the transformation from Remark 6, we define $h_{\max} := \max_{z\in\{0,1\}^n}\{h(z)\}$ (note that $h_{\max} \geq 0$) and $f(z) := -h(z) + h_{\max}$, and arrive at the equivalent conic constraint

$$[2a - 2U\beta; \; t + f(z) - h_{\max}; \; t - f(z) + h_{\max}] \in \mathbb{L}^{k+2},$$

where $f$ is a nonnegative submodular function. In order to homogenize this constraint as we did in Remark 4, we introduce another decision variable $v \in \mathbb{R}$ and the constraint $v = 1$. By letting $x = [t; v; \beta]$ and $m = n + 2$, our conic constraint becomes $\tilde{A}x + \tilde{B}f(z) \in \mathbb{L}^{k+2}$, where we select

$$\tilde{A} = \begin{bmatrix} 0 & 2a & -2U \\ 1 & -h_{\max} & 0 \\ 1 & h_{\max} & 0 \end{bmatrix} \quad \text{and} \quad \tilde{B} = \begin{bmatrix} 0_k \\ 1 \\ -1 \end{bmatrix}.$$

Hence, we arrive at the equivalent problem

$$\min_{x\in\mathbb{R}^m, z\in\{0,1\}^n} \left\{ x_1 : \; \tilde{A}x + \tilde{B}f(z) \in \mathbb{L}^{k+2}, \; x_1 \in \mathbb{R}_+, \; x_2 = 1, \; -Mz_i \leq x_{i+2} \leq Mz_i, \; \forall i \in [n] \right\},$$

where $f(z) = -g(\sum_{i\in[n]} z_i) + \max_{\bar{z}\in\{0,1\}^n}\{g(\sum_{i\in[n]} \bar{z}_i)\}$ is a nonnegative submodular function. Consider a point $(x,z)$ that satisfies $\tilde{A}x + \tilde{B}f(z) \in \mathbb{L}^{k+2}$. Then, $x = [t; v; \beta]$ and $t \cdot (h_{\max}v - f(z)) \geq \|a - U\beta\|_2^2$ holds. Note that this constraint along with $t \geq 0$ implies that $h_{\max}v - f(z) \geq 0$. Moreover, since $f$ is a nonnegative function, we have $h_{\max}v - 0 \geq h_{\max}v - f(z)$. Then, $t \cdot (h_{\max}v - 0) \geq t \cdot (h_{\max}v - f(z)) \geq \|a - U\beta\|_2^2$ holds, i.e., $x$ satisfies $\tilde{A}x \in \mathbb{L}^{k+2}$ as well. Then, Remark 1 implies that Condition ($\star$) holds and we can thus apply Proposition 2 to arrive at the equivalent formulation

$$\min_{x\in\mathbb{R}^m, y\in\mathbb{R}_+, z\in\{0,1\}^n} \left\{ x_1 : \ y \geq f(z), \ \tilde{A}x + \tilde{B}y \in \mathbb{L}^{k+2}, \ x_1 \in \mathbb{R}_+, \ x_2 = 1, \ -Mz_i \leq x_{i+2} \leq Mz_i, \ \forall i \in [n] \right\}.$$

Our developments in this application highlight that one can exploit the submodularity structure in this problem all the while using the standard optimization solvers without the need to develop specialized algorithms such as Newton-type methods or parametrization of the fractional objective.

### 5.4. Distributionally robust chance-constrained programs under Wasserstein ambiguity

Distributionally robust chance-constrained programming (DR-CCP) under Wasserstein ambiguity is formulated as

$$\min_{(x,z)} \left\{ c^\top [x; z] : \sup_{\mathbb{P}\in\mathcal{F}_N(\theta)} \mathbb{P}[\xi \notin \mathcal{W}(x,z)] \leq \epsilon, \ (x,z) \in \mathcal{X} \right\}. \tag{20}$$

Here, $c \in \mathbb{R}^{n+m}$ is a cost vector, $x \in \mathbb{R}^m$ is a vector of continuous decision variables, $z$ is a vector of $n$ binary decision variables, $\mathcal{X} \subset \mathbb{R}^{n+m}$ is a compact domain for the decision variables, $\mathcal{W}(x,z) \subseteq \mathbb{R}^K$ is a decision-dependent safety set, $\xi \in \mathbb{R}^K$ is a vector of $K$ random variables with distribution $\mathbb{P}^*$, and $\epsilon \in (0,1)$ is the risk tolerance for the random variable $\xi$ falling outside the safety set $\mathcal{W}(x,z)$. Because the distribution $\mathbb{P}^*$ is often unavailable, independent and identically distributed (i.i.d.) samples $\{\xi_i\}_{i\in[N]}$ are drawn from $\mathbb{P}^*$ to approximate $\mathbb{P}^*$ using the empirical distribution $\mathbb{P}_N$ on these samples. To address the ambiguity in the true distribution, distributionally robust optimization model (20) considers the worst-case probability of violating the safety constraints over a set of distributions on $\mathbb{R}^K$, given by $\mathcal{F}_N(\theta)$ , that contains the empirical distribution $\mathbb{P}_N$ where $\theta$ is a parameter that governs the size of the ambiguity set, and the degree of conservatism of (20).

Chen et al. (2018) and Xie (2019) show that DR-CCP can be formulated as a mixed-integer conic program for certain $\mathcal{W}(x,z)$, thus enabling their solution with standard optimization solvers. However, these MIP reformulations are difficult to solve in certain cases for which the continuous relaxations provide weak lower bounds. Ho-Nguyen et al. (2020a,b) consider a mixed-integer *linear* substructure of the mixed-integer conic formulation of DR-CCP. The authors propose valid linear inequalities and other enhancements that strengthen the continuous relaxation bounds and scale up the sizes of the problem that can be solved.

While previous research focused on the linear constraints, the strengthening of the mixed-integer conic reformulation of DR-CCP considered by Xie (2019) focuses on constraints of the form

$\|[\eta_1 x; \eta_1 z; \eta_2]\|_* \le t$, where $t$ is a continuous epigraph variable, $\eta_1, \eta_2 \in \{0,1\}$ are constants with $\eta_1 + \eta_2 \ge 1$, and $\|\cdot\|_*$ is the dual of a norm $\|\cdot\|$. To obtain a strengthening from this conic constraint, we consider the case when $\eta_1 = 1$, i.e., when there is so-called left-hand side uncertainty. The norm is used to measure the Wasserstein distance and it is typical to consider an $\ell_q$-norm, whose dual norm is an $\ell_p$-norm with $\frac{1}{p} + \frac{1}{q} = 1$. Xie (2019) pays specific attention to the pure-binary case, i.e., $m = 0$ where the continuous variables $x$ are not present inside the norm. In this particular case, Xie (2019) observes that the function $\|[\eta_1 z; \eta_2]\|_p$ is a submodular function and the corresponding extended polymatroid inequalities can be applied to strengthen the set $\{(z,t) : \ \|[\eta_1 z; \eta_2]\|_p \le t\}$ and this approach results in significant computational benefit. When the continuous variables $x$ are present, although $\|[\eta_1 x; \eta_1 z; \eta_2]\|_p$ is not even a set function. Nevertheless, our framework applies to constraints of the form $\|[\eta_1 x; \eta_1 z; \eta_2]\|_p \le t$. To elaborate, let $\mathbb{K}_p^{m+2}$ denote the $p$-th order cone in $\mathbb{R}^{m+2}$, and define $f(z) := \|[z; \eta_2]\|_*$. Note that as $z_j^p = z_j$ for all $j \in [n]$ and $\eta_2^{1/p} = \eta_2$ as well, and thus from Remark 3 we deduce that $f$ is a nonnegative submodular function. Moreover,

$$\begin{aligned}
&\left\{(x,t,z) \in \mathbb{R}_+^m \times \mathbb{R}_+ \times \{0,1\}^n : \ \|[\eta_1 x; \eta_1 z; \eta_2]\|_p \le t\right\} \\
&\quad = \left\{(x,t,z) \in \mathbb{R}_+^m \times \mathbb{R}_+ \times \{0,1\}^n : \ [f(z); \ x_1; \ \ldots; \ x_m; \ t] \in \mathbb{K}_p^{m+2}\right\} \\
&\quad = \left\{(x,t,z) \in \mathbb{R}_+^m \times \mathbb{R}_+ \times \{0,1\}^n : \ \exists y \in \mathbb{R} \text{ s.t. } \ y \ge f(z), \ [f(z); \ x_1; \ \ldots; \ x_m; \ t] \in \mathbb{K}_p^{m+2}\right\} \\
&\quad = \left\{(x,t,z) \in \mathbb{R}_+^m \times \mathbb{R}_+ \times \{0,1\}^n : \ \exists y \in \mathbb{R} \text{ s.t. } \ y \ge f(z), \ \tilde{A}(x;t) + \tilde{B}y \in \mathbb{K}_p^{m+2}\right\},
\end{aligned}$$

where $\tilde{B} := e_1 \in \mathbb{R}^{m+2}$ and $\tilde{A} := [0^\top; \ \mathrm{Diag}(1; \ \ldots; \ 1)] \in \mathbb{R}^{(m+2)\times(m+1)}$. Here the second equation follows from Proposition 2 because Condition ($\star$) holds (for any $(x,z)$ satisfying $\tilde{A}x + \tilde{B}f(z) \in \mathbb{K}_p^{m+2}$, due to the structure of $\tilde{A}, \tilde{B}$ and the cone $\mathbb{K}_p^{m+2}$, we have $x$ satisfies $\tilde{A}x \in \mathbb{K}_p^{m+2}$ as well, implying that we can apply Remark 1 to deduce that Condition ($\star$) holds). Therefore, we can apply Corollary 1. Consequently, our results indicate that it is possible to exploit submodularity in the DR-CCP context even when l we have both continuous and binary decision variables. In particular, if $(x,z)$ is a mixed-binary decision vector, then by using Theorem 1, Corollaries 1 and 4, and Theorems 2 and 3 we can strengthen the resulting reformulation of DR-CCP under Wasserstein ambiguity.

## Acknowledgments

This research is supported, in part, by ONR grant N00014-19-1-2321, the Institute for Basic Science (IBS-R029-C1), and NSF grant CMMI 1454548.

## References

Alper Atamtürk and Andrés Gómez. Submodularity in conic quadratic mixed 0-1 optimization. *Operations Research*, 68:609–630, 2020.

Alper Atamtürk and Vishnu Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008.

Alper Atamtürk and Vishnu Narayanan. Submodular function minimization and polarity. *Mathematical Programming, to appear, arXiv:1912.13238v3*, 2020.

Frank Baumann, Sebastian Berckey, and Christoph Buchheim. Exact algorithms for combinatorial optimization problems with submodular objective functions. In Michael Jünger and Gerhard Reinelt, editors, *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*, pages 271–294. Springer, 2013.

Juan S. Borrero, Colin Gillen, and Oleg A. Prokopyev. Fractional 0-1 programming: applications and algorithms. *Journal of Global Optimization*, 69:255–282, 2017.

Zhi Chen, Daniel Kuhn, and Wolfram Wiesemann. Data-driven chance constrained programs over Wasserstein balls. *arXiv:1809.00210*, 2018.

Jack Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Structures and Their Applications*, pages 69–87. Gordon and Breach, 1970.

Andrés Gómez and Oleg Prokopyev. A mixed-integer fractional optimization approach to best subset selection. *INFORMS Journal on Computing, to appear*, 2020. http://www.optimization-online.org/DB_FILE/2018/09/6795.pdf.

Nam Ho-Nguyen, Fatma Kılınç-Karzan, Simge Küçükyavuz, and Dabeen Lee. Strong formulations for distributionally robust chance-constrained programs with left-hand side uncertainty under Wasserstein ambiguity. *arXiv:2007.06750*, 2020a.

Nam Ho-Nguyen, Fatma Kılınç-Karzan, Simge Küçükyavuz, and Dabeen Lee. Distributionally robust chance-constrained programs with right-hand side uncertainty under Wasserstein ambiguity. *Mathematical Programming, to appear, arXiv:2003.12685v2*, 2020b.

Fatma Kılınç-Karzan, Simge Küçükyavuz, and Dabeen Lee. Joint chance-constrained programs and the intersection of mixing sets through a submodularity lens. *arXiv:1910.01353*, 2019.

László Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art: Bonn 1982*, pages 235–257. Springer, Berlin, Heidelberg, 1983.

George L. Nemhauser and Laurence A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience, USA, 1988. ISBN 047182819X.

Weijun Xie. On distributionally robust chance constrained programs with Wasserstein distance. *Mathematical Programming, to appear*, 2019. https://doi.org/10.1007/s10107-019-01445-5.

## Appendix A: Example 1

Suppose for a contradiction that every extreme point of (15) has binary $z_1$ and $z_2$ components. Consider a point $(\bar{x},\bar{z})$ in (15) with $(\bar{z}_1,\bar{z}_2,\bar{y})=(1/2,1/2,\sqrt{2}/2)$ and $\bar{x}_2=1+\sqrt{\bar{x}_1^2+1/2}$. Because $\bar{z}$ is not binary, $(\bar{x},\bar{z})$ should not be an extreme point of (15), and thus it must be a convex combination of extreme points of (15). First, we show that the extreme points that give rise to $(\bar{x},\bar{z})$ must satisfy $(z_1,z_2,y)=(0,0,0)$ or $(z_1,z_2,y)=(1,1,\sqrt{2})$. Let $I_{0,0}$ be the extreme points of (15) with $(z_1,z_2)=(0,0)$, given by $(x;z;y)=(x_{1,i},x_{2,i},0,0,0)$ for $i\in I_{0,0}$. Similarly define $I_{1,1}$ as the set of extreme points of (15) with $(z_1,z_2)=(1,1)$, given by $(x_{1,i},x_{2,i},1,1,\sqrt{2})$ for $i\in I_{1,1}$; $I_{1,0}$ as the set of extreme points of (15) with $(z_1,z_2)=(1,0)$, given by $(x_{1,i},x_{2,i},1,0,1)$ for $i\in I_{1,0}$, and $I_{0,1}$ as the set of extreme points of (15), given by $(x_{1,i},x_{2,i},0,1,1)$ for $i\in I_{0,1}$. We know that $(\bar{x}_1,\bar{x}_2,\bar{z}_1,\bar{z}_2,\bar{y})$ is a convex combination of these extreme points with the associated multipliers $\alpha^i_{j,k}\ge 0, i\in I_{j,k}, j,k\in\{0,1\}$ and $\sum_{j,k\in\{0,1\}}\sum_{i\in I_{j,k}}\alpha^i_{j,k}=1$. Let $\alpha_{j,k}=\sum_{i\in I_{j,k}}\alpha^i_{j,k}$, for $j,k\in\{0,1\}$. Then, since $\bar{z}_1=1/2$ and $\bar{z}_2=1/2$, we must have $\alpha_{1,1}+\alpha_{1,0}=\alpha_{1,1}+\alpha_{0,1}=1/2$, and

$$\frac{\sqrt{2}}{2}=\alpha_{1,1}\sqrt{2}+\alpha_{1,0}+\alpha_{0,1}.$$

Substituting $\alpha_{1,0}=\alpha_{0,1}=1/2-\alpha_{1,1}$ and solving the above equation for $\alpha_{1,1}$, we obtain $\alpha_{1,1}=1/2=\alpha_{0,0}$ and $\alpha_{1,0}=\alpha_{0,1}=0$, as desired.

Moreover,

$$(\bar{x}_1,\bar{x}_2,\bar{z}_1,\bar{z}_2,\bar{y})=\sum_{i\in I_{0,0}}\alpha^i_{0,0}(x_{1,i},x_{2,i},0,0,0)+\sum_{j\in I_{1,1}}\alpha^j_{1,1}(x_{1,j},x_{2,j},1,1,\sqrt{2}),$$

and $\sum_{i\in I_{0,0}}\alpha^i_{0,0}=\sum_{j\in I_{1,1}}\alpha^j_{1,1}=1/2$. Then, $\bar{x}_1=\sum_{i\in I_{0,0}}\alpha^i_{0,0}x_{1,i}+\sum_{j\in I_{1,1}}\alpha^j_{1,1}x_{1,j}$. Note that

$$\begin{aligned}\bar{x}_2=\sum_{i\in I_{0,0}}\alpha^i_{0,0}x_{2,i}+\sum_{j\in I_{1,1}}\alpha^j_{1,1}x_{2,j}&\ge\sum_{i\in I_{0,0}}\alpha^i_{0,0}(1+|x_{1,i}|)+\sum_{j\in I_{1,1}}\alpha^j_{1,1}(1+\sqrt{x_{1,j}^2+2})\\&\ge 1+\sum_{i\in I_{0,0}}\alpha^i_{0,0}|x_{1,i}|+\sum_{j\in I_{1,1}}\alpha^j_{1,1}\sqrt{x_{1,j}^2+2}.\end{aligned}\tag{21}$$

Then, after subtracting 1 from each side of (21) and taking the square, we obtain

$$\begin{aligned}
(\bar{x}_2 - 1)^2 &\geq \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| + \sum_{j\in I_{1,1}} \alpha_{1,1}^j \sqrt{x_{1,j}^2 + 2} \right)^2 \\
&= \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \right)^2 + 2 \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \sum_{j\in I_{1,1}} \alpha_{1,1}^j \sqrt{x_{1,j}^2 + 2} \\
&\quad + \sum_{j\in I_{1,1}} (\alpha_{1,1}^j)^2 (x_{1,j}^2 + 2) + \sum_{j,k\in I_{1,1}: j\neq k} \alpha_{1,1}^j \alpha_{1,1}^k \sqrt{x_{1,j}^2 + 2}\sqrt{x_{1,k}^2 + 2} \\
&> \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \right)^2 + 2 \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \sum_{j\in I_{1,1}} \alpha_{1,1}^j |x_{1,j}| \\
&\quad + \sum_{j\in I_{1,1}} (\alpha_{1,1}^j)^2 (x_{1,j}^2 + 2) + \sum_{j,k\in I_{1,1}: j\neq k} \alpha_{1,1}^j \alpha_{1,1}^k \sqrt{x_{1,j}^2 + 2}\sqrt{x_{1,k}^2 + 2} \\
&\geq \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \right)^2 + 2 \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| \sum_{j\in I_{1,1}} \alpha_{1,1}^j |x_{1,j}| + \left( \sum_{j\in I_{1,1}} \alpha_{1,1}^j |x_{1,j}| \right)^2 + 2 \left( \sum_{j\in I_{1,1}} \alpha_{1,1}^j \right)^2 \\
&= \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| + \sum_{j\in I_{1,1}} \alpha_{1,1}^j |x_{1,j}| \right)^2 + \frac{1}{2}
\end{aligned} \tag{22}$$

where the second and third inequalities follow from

$$\sqrt{x_{1,j}^2 + 2} > |x_{1,j}|, \quad \text{and} \quad \sqrt{x_{1,j}^2 + 2}\sqrt{x_{1,k}^2 + 2} \geq |x_{1,j} x_{1,k}| + 2.$$

Moreover, note that as $\bar{x}_2 = 1 + \sqrt{\bar{x}_1^2 + 1/2}$ and $\bar{x}_1 = \sum_{i\in I_{0,0}} \alpha_{0,0}^i x_{1,i} + \sum_{j\in I_{1,1}} \alpha_{1,1}^j x_{1,j}$, we have

$$(\bar{x}_2 - 1)^2 = \bar{x}_1^2 + 1/2 = \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i x_{1,i} + \sum_{j\in I_{1,1}} \alpha_{1,1}^j x_{1,j} \right)^2 + \frac{1}{2}.$$

Therefore, we deduce a contradiction from (22) because for convex combination weights $\alpha \geq 0$ it is not possible to satisfy

$$(\bar{x}_2 - 1)^2 = \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i x_{1,i} + \sum_{j\in I_{1,1}} \alpha_{1,1}^j x_{1,j} \right)^2 + \frac{1}{2} \;>\; \left( \sum_{i\in I_{0,0}} \alpha_{0,0}^i |x_{1,i}| + \sum_{j\in I_{1,1}} \alpha_{1,1}^j |x_{1,j}| \right)^2 + \frac{1}{2}.$$

This in turn implies that (15) has an extreme point with a fractional $z$ component.