Strong Valid Inequalities for Orthogonal Disjunctions and Bilinear Covering Sets

Mohit Tawarmalani [1], Jean-Philippe P. Richard [2], Kwanghun Chung [2]

September 24, 2008

Abstract

In this paper, we develop a convexification tool that enables the construction of convex hulls for orthogonal disjunctive sets using convex extensions and disjunctive programming techniques. A distinguishing feature of our technique is that, unlike most applications of disjunctive programming, it does not require the introduction of new variables in the relaxation. We develop and apply a toolbox of results that help in checking the technical assumptions under which the convexification tool can be employed. We demonstrate its applicability in integer programming by deriving the intersection cut for mixed-integer polyhedral sets and the convex hull of certain mixed/pure-integer bilinear sets. We then develop a key result that extends the applicability of the convexification tool to relaxing nonconvex inequalities, which are not naturally disjunctive, by providing sufficient conditions for establishing the convex extension property over the non-negative orthant. We illustrate the utility of this result by deriving the convex hull of a continuous bilinear covering set over the non-negative orthant.

1 Introduction and Motivation

Finding globally optimal solutions to nonconvex problems is a challenging problem that has received much attention in the last few decades; see Neumaier [16] for a survey of the existing solution methods. Nonlinear branch-and-bound is one such method that has been implemented successfully in various global optimization software; see Adjiman et al. [1], Sahinidis and Tawarmalani [18], LINDO Systems Inc. [14], and Belotti et al. [7]. The branch-and-bound method typically bounds the nonconvex optimization problem by solving its convex relaxations over successively refined partitions (see Falk and Soland [11] and Horst and Tuy [13]). For factorable problems–problems involving functions that can be written as recursive sums and products of univariate functions–McCormick [15] proposed a composition theorem that allows automatic construction of convex relaxations provided that tight concave and/or convex envelopes are known for the intrinsic nonlinear terms. McCormick's relaxation is an instance of a commonly used technique for deriving convex relaxations for nonconvex problems that relaxes inequalities of the form $f(x) \geq r$ by $\bar{f}(x) \geq r$, where $\bar{f}(x)$ is a concave overestimator of the function $f(x)$. There is a significant amount of literature that develops techniques for deriving tight overestimators for various classes of functions; see Tawarmalani and Sahinidis [22] and Bliek et al. [8] for a more detailed treatment. However, the current literature rarely considers the right-hand-side of the inequality. More precisely, the above technique relaxes the hypograph of $f(x)$ instead of relaxing the appropriate upper-level set. As a result, the derived relaxations can be weak. For an illustration of the difference, consider the set S defined as:

$$S = \left\{(x, y, z) \in \mathbb{R}^3_+ \;\middle|\; xy + z \geq r\right\},$$

where $r > 0$. It can be easily seen that S is not convex since both $\left(\sqrt{r}, \sqrt{r}, 0\right)$ and $(0, 0, r)$ belong to S while their convex combination with a weight of $\frac{1}{2}$ on each point does not. Therefore, if the constraint

[1] Krannert School of Management, Purdue University. Email: mtawarma@purdue.edu. Corresponding Author.
[2] Department of Industrial and Systems Engineering, University of Florida.

defining S was to appear as one of the constraints in a problem, local optimization techniques would not be guaranteed to find a globally optimal solution for the problem. However, because this set belongs to the general family of factorable programs, it can be relaxed using McCormick's scheme. More generally, if traditional techniques were used to derive a convex relaxation of S, a concave overestimator $\bar{f}$ of the function $f(x,y,z) = xy+z$ would first be obtained. Observe that the concave envelope of this function over the non-negative orthant is infinite as long as x and y are both positive. The resulting convex relaxation of S is $\{(x,y,z) \in \mathbb{R}^3_+ \mid x,y > 0\} \cup \{(x,y,z) \in \mathbb{R}^3_+ \mid z \geq r,\, xy = 0\}$. If in addition, the concave overestimator is required to be upper-semicontinuous, as is typically the case, or even if the relaxation is required to be a closed set, then the relaxation would be $\mathbb{R}^3_+$. In other words, standard relaxation schemes will essentially drop the defining constraint.

In this paper, we propose a scheme that produces tighter convex approximations by considering the right-hand-side of the constraint. In particular, for the set S presented above, our scheme produces the following convex relaxation

$$\mathrm{RS} = \left\{(x,y,z) \in \mathbb{R}^3_+ \;\middle|\; \sqrt{\frac{xy}{r}} + \frac{z}{r} \geq 1\right\},$$

which is a much tighter approximation than $\mathbb{R}^3_+$. Considering this simple example, we can make three interesting observations. First, the relaxation, RS, is nonlinear. This is in contrast to current implementations of nonlinear branch-and-bound that typically construct linear relaxations for multivariate terms (see Tawarmalani and Sahinidis [25]). Second, the form of the nonlinear cut is surprising as it applies different functions to the different terms of the initial inequality. For S, the first term is modified using a square-root after being divided by r, while the second is simply divided by r. Third, RS is not only a convex relaxation of S, but it is in fact (as will be shown later) the convex hull of S. Surprisingly, the convex hull for these sets can be expressed in a simple form without introducing new variables while developing the concave envelope of the corresponding polynomial can be much harder.

The convex hull representation for bilinear covering sets arises from a general theory of orthogonal disjunctions that we develop in this paper. To provide an example, consider the set S again. We will show that the convex hull of S is determined by the points of S that either belong to the half-plane $(x,y,0)$, where $(x,y) \in \mathbb{R}^2_+$ or to the half-line $(0,0,z)$, where $z \in \mathbb{R}_+$. In other words, the set S satisfies the convex extension property (see Tawarmalani and Sahinidis [23]) in which the important subsets belong to orthogonal subspaces. Because such a convex extension property holds, it is natural to expect that one could build a higher dimensional description of the convex hull of S using disjunctive programming arguments; see Rockafellar [17] and Balas [4]. Disjunctive programming has been used to develop tight relaxations and cutting planes in integer, nonlinear, and robust optimization; see [3, 20, 21, 9, 24, 6, 2, 19]. Unlike this paper, the literature on disjunctive programming formulations, however, is mostly focused on naturally disjunctive sets. Cutting planes based on disjunctive formulations, are typically linear and derived by solving separation problems over extended formulations; see Cornuéjols and Lemaréchal [10]. One interesting observation in this paper is that, as long as the disjunctive terms are orthogonal and a few technical conditions are satisfied, there is no need to introduce additional variables. Furthermore, the convex hull of S can be easily expressed in closed-form using the representations of the convex hull of S in each of the two orthogonal subspaces, namely $\sqrt{\frac{xy}{r}} \geq 1$ and $\frac{z}{r} \geq 1$. We establish a much more general set of conditions under which the argument evoked above is correct, allowing the use of both right-hand-side and left-hand-side information in the derivation of convex relaxations for nonlinear programming. Our results rely on the ability to prove that a convex extension property holds over orthogonal disjunctions and the ability to derive closed form expressions of convex hulls (possibly in a higher dimensional space) over each of the subspaces. Our techniques are applicable to large families of problems and yield stronger convex approximations than those currently used in the nonlinear branch-and-bound solvers.

In Section 2, we describe a tool to obtain the convex hull of orthogonal disjunctive sets. The result can be evoked under certain technical conditions. We provide tools to verify these assumptions. We also provide counterexamples to show the need for some of the assumptions. The intersection/split cut for mixed integer linear sets is shown to be a special case of our general convexification tool. In

Section 3, we illustrate the application of the tool in nonlinear integer programming by convexifying a bilinear pure/mixed-integer set. Nonconvex inequalities in continuous variables are not naturally disjunctive. For such inequalities, we establish sufficient conditions under which the convex extension property holds over the non-negative orthant. We show that these sufficient conditions are satisfied by continuous bilinear covering sets and develop their convex hulls over the non-negative orthant. We summarize the contributions of this work in Section 4 and conclude with remarks and directions for future research.

2 Convexification of Orthogonal Disjunctive Sets

In this section, we first introduce and prove a general result that exposes the closed-form convex hull inequality description of the disjunctive union of a finite number of sets defined over subspaces that are orthogonal to each other. This result also applies to non-disjunctive sets provided that their convex hulls are entirely defined by their restrictions over a finite number of orthogonal subspaces. Next, we illustrate the utility of this result in finding convex hull descriptions. Simultaneously, we discuss the need for certain seemingly technical assumptions in the statement of the result. In particular, we discuss each one of the four assumptions of the theorem and describe, with examples, situations where they are satisfied. For some of the assumptions, we establish sufficient conditions that are simple to verify. We also show later that the cuts that yield the convex hull, under the specified technical conditions, continue to produce valid inequalities even when some of the conditions are not satisfied. Throughout, we demonstrate the generality and applicability of our convexification result by deriving new convex hull descriptions of various continuous, mixed, and pure integer bilinear covering sets, and providing an alternate derivation of the classic intersection cut derived in the integer programming literature.

In the following, given a set S, we represent its convex hull by $\operatorname{conv}(S)$, its closure by $\operatorname{cl}(S)$, and its projection on the space of z variables by $\operatorname{proj}_z S$. For a closed convex set, S, $0^+(S)$ denotes the set of its recession directions. When we display equations, we sometimes write $\min\begin{Bmatrix} f(z) \\ g(z) \end{Bmatrix}$ to denote $\min\{f(z), g(z)\}$.

Theorem 2.1. *Let $S \subseteq \mathbb{R}^{\sum_i d_i}$ and for all $i \in N = \{1, \ldots, n\}$, let $S_i \subseteq S$. Let the points z of S be written as $z = (z_1, \ldots, z_i, \ldots, z_n) \in S$, where $z_i \in \mathbb{R}^{d_i}$. Assume that:*
(A1) if $(z_1, \ldots, z_i, \ldots, z_n) \in S_i$, then $z_j = 0$ for $\forall j \neq i$,
(A2) for any $z \in S$, there exists $\chi_i \in \operatorname{conv}(S_i)$, $i \in I \subseteq N$, such that $z \in \operatorname{conv}\big(\bigcup_{i \in I} \chi_i\big)$,
(A3) $\operatorname{conv}(S_i) \subseteq \operatorname{proj}_z A_i \subseteq \operatorname{cl}\big(\operatorname{conv}(S_i)\big)$, where, for each $i \in \{1, \ldots, n\}$,

$$A_i = \left\{ \big(0, z_i, u_i, 0\big) \;\middle|\; \begin{array}{ll} t_i^{j_i}(z_i, u_i) \geq 1, & \forall j_i \in J_i, \\ v_i^{k_i}(z_i, u_i) \geq -1, & \forall k_i \in K_i, \\ w_i^{l_i}(z_i, u_i) \geq 0, & \forall l_i \in L_i \end{array} \right\}. \tag{1}$$

Assume that $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ are positively-homogenous functions, i.e., for $\lambda > 0$,

$$\lambda t_i^{j_i}\left(\frac{z_i, u_i}{\lambda}\right) = t_i^{j_i}(z_i, u_i),\ \lambda v_i^{k_i}\left(\frac{z_i, u_i}{\lambda}\right) = v_i^{k_i}(z_i, u_i),\ \lambda w_i^{l_i}\left(\frac{z_i, u_i}{\lambda}\right) = w_i^{l_i}(z_i, u_i).$$

(A4) $\operatorname{proj}_z C_i$ is a subset of the recession cone of $\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$, i.e., for all i,

$$\operatorname*{proj}_z C_i \subseteq 0^+\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)\right)$$

where

$$C_i = \left\{ \big(0, z_i, u_i, 0\big) \;\middle|\; \begin{array}{ll} t_i^{j_i}(z_i, u_i) \geq 0, & \forall j_i \in J_i, \\ v_i^{k_i}(z_i, u_i) \geq 0, & \forall k_i \in K_i, \\ w_i^{l_i}(z_i, u_i) \geq 0, & \forall l_i \in L_i \end{array} \right\}.$$

Let

$$X = \left\{ (z,u) \;\middle|\; \begin{array}{ll} \sum_{i\in N} t_i^{j_i}(z_i,u_i) \geq 1, & \forall j_i \in J_i, \\ \sum_{i\in I} v_i^{k_i}(z_i,u_i) \geq -1, & \forall I \subseteq N, \forall k_i \in K_i, \\ t_i^{j_i}(z_i,u_i) + v_i^{k_i}(z_i,u_i) \geq 0, & \forall i, \forall j_i \in J_i, \forall k_i \in K_i, \\ t_i^{j_i}(z_i,u_i) \geq 0, & \forall i, \forall j_i \in J_i, \\ w_i^{l_i}(z_i,u_i) \geq 0, & \forall i, \forall l_i \in L_i \end{array} \right\}. \tag{2}$$

Then, $\mathrm{conv}(S) \subseteq \mathrm{proj}_z X \subseteq \mathrm{cl}\,\mathrm{conv}(S)$*. If in addition,* $\forall i \in N$*,* $\mathrm{proj}_z A_i$ *is closed and* $\mathrm{proj}_z C_i = 0^+\big(\mathrm{cl}\,\mathrm{conv}(S_i)\big)$*, then* $\mathrm{proj}_z X = \mathrm{cl}\,\mathrm{conv}(S)$*.*

Before proving Theorem 2.1, we briefly comment on its assumptions, its practical importance, and its applicability. In Assumption (A2), we impose that any point in S can be expressed as a convex combination of points in some of the S_is. This implies that only the subsets S_is are needed when computing the convex hull of S. In Assumption (A1), we require that these subsets are orthogonal to each other and aligned along the principal axes. In Assumption (A3), we require that an inequality description of the convex hull of each one of the sets S_i be known. Note that this inequality description might make use of an extended formulation (using the additional variables u_i). The assumption that the right-hand-sides of all the inequalities are either 1, 0, or -1 is without loss of generality as inequalities with nonzero right-hand-sides can be rescaled to satisfy this assumption. Note also that in Theorem 2.1, we require that all inequalities be defined using positively-homogeneous functions. We will show later that this assumption is often not needed to prove the validity of the cuts derived in Theorem 2.1. In Assumption (A4), we impose, in essence, that the recession directions of each one of the sets A_i are also the recession directions for the closure convex hull of the union of the S_is. Under these four assumptions, we show that an inequality description of the convex hull of S can be obtained by combining in a systematic way the inequalities arising in the convex hull descriptions of the S_is. Note however that, for reasons that will be described later, this inequality description might describe a superset of the desired convex hull. However, the superset will never be larger than the closure convex hull of S, which is sufficient for all practical purposes.

Proof. Claim 1: We claim that $\mathrm{conv}(S) = \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. We first show that $\mathrm{conv}(S)$ contains $\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Clearly, for all i, $S_i \subseteq S$. Therefore, $S \supseteq \bigcup_{i=1}^n S_i$ and, so, $\mathrm{conv}(S) \supseteq \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Now, we show that (A2) implies that $\mathrm{conv}(S) \subseteq \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Let $z \in S$. There exists $I \subseteq N$ and $\chi_i \in \mathrm{conv}(S_i)$ such that $z \in \mathrm{conv}\big(\bigcup_{i\in I} \chi_i\big) \subseteq \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Claim 1 is thus proved and, therefore, we can use disjunctive programming techniques to compute the convex hull of S. Using these techniques, we now show that it is possible to construct, in a closed-form, a set X that contains $\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$ and is itself contained in $\mathrm{cl}\big(\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)\big)$.

For $T \subseteq N$, we define

$$R_T(\lambda_T) = \left\{ (z_T,u_T) \;\middle|\; \begin{array}{ll} \sum_{i\in T} t_i^{j_i}(z_i,u_i) \geq \lambda_T & \forall j_i \in J_i \\ \sum_{i\in I} v_i^{k_i}(z_i,u_i) \geq -\lambda_T & \forall I \subseteq T, \forall k_i \in K_i \\ t_i^{j_i}(z_i,u_i) + v_i^{k_i}(z_i,u_i) \geq 0 & \forall i, \forall j_i \in J_i, \forall k_i \in K_i \\ t_i^{j_i}(z_i,u_i) \geq 0 & \forall i, \forall j_i \in J_i \\ w_i^{l_i}(z_i,u_i) \geq 0 & \forall i, \forall l_i \in L_i \end{array} \right\}.$$

In the remainder of this proof, whenever T is a singleton, say $\{i\}$, we will denote it as i itself. Also, we define

$$Q = \left\{ (\lambda,z,u) \;\middle|\; \begin{array}{ll} \lambda_i \geq 0 & \forall i \in N \end{array} \right.$$

$$
\left.\begin{array}{ll}
(z_i,u_i)\in R_i(\lambda_i) & \forall i\in N\\
\sum_{i=1}^{n}\lambda_i=\lambda_{1,\ldots,n}=1 &
\end{array}\right\}.
$$

We next prove that $X = \mathrm{proj}_{z,u}\, Q$ and $\mathrm{conv}(S) \subseteq \mathrm{proj}_z\, Q \subseteq \mathrm{cl\,conv}(S)$. Clearly, together these results imply that $\mathrm{conv}(S) \subseteq \mathrm{proj}_z\, X \subseteq \mathrm{cl\,conv}(S)$. First, we prove that $X = \mathrm{proj}_{z,u}\, Q$. Given two disjoint subsets A and B of N, we consider

$$
W=\left\{(\lambda_A,\lambda_B,\lambda_{A\cup B},z_A,u_A,z_B,u_B)\;\middle|\;\begin{array}{l}
\lambda_A\ge 0\\
(z_A,u_A)\in R_A(\lambda_A)\\
\lambda_B\ge 0\\
(z_B,u_B)\in R_B(\lambda_B)\\
\lambda_A+\lambda_B=\lambda_{A\cup B}
\end{array}\right\},
$$

and

$$
P=\left\{(\lambda_{A\cup B},z_{A\cup B},u_{A\cup B})\;\middle|\;\begin{array}{l}
\lambda_{A\cup B}\ge 0\\
(z_{A\cup B},u_{A\cup B})\in R_{A\cup B}(\lambda_{A\cup B})
\end{array}\right\}.
$$

A straightforward sequential application of the following claim shows that when $\lambda_1,\ldots,\lambda_n$ are projected out from Q we obtain $R_N(1) = X$.

Claim 2: If $z_{A\cup B} = (z_A, z_B)$ and $u_{A\cup B} = (u_A, u_B)$, then P is the set obtained when λ_A and λ_B are projected out from W. Note that since A and B are disjoint and $z_{A\cup B} \in \mathbb{R}^{|\sum_{i\in A} d_i+\sum_{i\in B} d_i|} = \mathbb{R}^{|\sum_{i\in A} d_i|} \times \mathbb{R}^{|\sum_{i\in B} d_i|}$, the definitions of $z_{A\cup B}$ and, similarly, $u_{A\cup B}$ are dimensionally consistent. Claim 2 is verified by first substituting $\lambda_B = \lambda_{A\cup B} - \lambda_A$ and then projecting λ_A out using Fourier-Motzkin elimination; see Theorem 1.4 in [26]. We substitute $\lambda_B = \lambda_{A\cup B} - \lambda_A$ in W to obtain:

$$
\begin{array}{l}
\lambda_A\ge 0\\
(z_A,u_A)\in R_A(\lambda_A)\\
\lambda_{A\cup B}-\lambda_A\ge 0\\
(z_B,u_B)\in R_B(\lambda_{A\cup B}-\lambda_A).
\end{array}
$$

On the one hand, note that the inequalities

$$
t_i^{j_i}(z_i,u_i)+v_i^{k_i}(z_i,u_i)\ge 0, \tag{3}
$$

$$
t_i^{j_i}(z_i,u_i)\ge 0, \tag{4}
$$

$$
w_i^{l_i}(z_i,u_i)\ge 0 \tag{5}
$$

for all $i \in A\cup B$, $j_i \in J_i$, $k_i \in K_i$, and $l_i \in L_i$ remain untouched during projection since they are independent of λ_A. On the other hand, the inequalities containing λ_A can be rewritten as:

$$
\min\left\{\begin{array}{c}
\sum_{i\in A} t_i^{j_i}(z_i,u_i)\\
\lambda_{A\cup B}+\min_{B'\subseteq B}\sum_{i\in B'} v_i^{k_i}(z_i,u_i)
\end{array}\right\}\ge\lambda_A\ge\max\left\{\begin{array}{l}
\lambda_{A\cup B}-\sum_{i\in B} t_i^{j_i}(z_i,u_i)\\
-\min_{A'\subseteq A}\sum_{i\in A'} v_i^{k_i}(z_i,u_i)
\end{array}\right\}
$$

so that Fourier-Motzkin elimination is simple to perform. Observe that the constraints $\lambda_{A\cup B}-\lambda_A \ge 0$ and $\lambda_A \ge 0$ are represented in the above system respectively when $A' = \emptyset$ and $B' = \emptyset$. Projecting λ_A out of the system, we obtain:

$$
\sum_{i\in A\cup B} t_i^{j_i}(z_i,u_i)\ge\lambda_{A\cup B} \tag{6}
$$

$$\sum_{i\in A} t_i^{j_i}(z_i,u_i) + \sum_{i\in A'} v_i^{k_i}(z_i,u_i) \geq 0 \quad \forall A' \subseteq A, j_i \in J_i, k_i \in K_i \quad \text{(redundant)} \tag{7}$$

$$\sum_{i\in B} t_i^{j_i}(z_i,u_i) + \sum_{i\in B'} v_i^{k_i}(z_i,u_i) \geq 0 \quad \forall B' \subseteq B, j_i \in J_i, k_i \in K_i \quad \text{(redundant)} \tag{8}$$

$$\sum_{i\in A'\cup B'} v_i^{k_i}(z_i,u_i) \geq -\lambda_{A\cup B} \qquad \forall B' \subseteq B, A' \subseteq A. \tag{9}$$

Inequalities (3) for $i \in A'$ and (4) for $i \in A\backslash A'$ imply (7), showing that (7) is redundant. Similarly, Inequality (8) can be shown to be redundant. Observe that $\lambda_{A\cup B} \geq 0$ can be shown to be represented in (9) by selecting $A' = B' = \emptyset$. Therefore, the set obtained by projecting λ_A and λ_B out of W is given by (3), (4), (5), (6), and (9), which is exactly the definition of P. We have thus proved Claim 2. By applying this result sequentially, we obtain that $X = \operatorname{proj}_{z,u} Q$.

We now prove that $\operatorname{conv}(S) \subseteq \operatorname{proj}_z Q \subseteq \operatorname{cl}\big(\operatorname{conv}(S)\big)$. We first show that if $z \in \operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$, it can be extended to a point that belongs to Q by suitably defining (λ, u). If $z \in \operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$, then, by (A1), there exist λ_i and z_i' such that

$$z = (z_1, \ldots, z_i, \ldots, z_n) = \sum_{i=1}^n \lambda_i(0, z_i', 0),$$

where, for each i, $\lambda_i \geq 0$, $(0, z_i', 0) \in \operatorname{conv}(S_i)$, and the multipliers sum up to one, *i.e.*, $\sum_{i=1}^n \lambda_i = 1$. We reindex S_i so that the sets containing the points associated with non-zero multipliers are indexed from 1 to t. Then, $(z,u) = \sum_{i=1}^t \lambda_i(0, z_i', u_i', 0)$, where $(0, z_i', u_i', 0) \in A_i$, $\lambda_i > 0$ for $i = 1, \ldots, t$, and $\sum_{i=1}^t \lambda_i = 1$. Such a representation exists since z is expressible as a convex combination of points in $\operatorname{conv}(S_i)$ which can be extended to belong to A_i, the representation of a superset of $\operatorname{conv}(S_i)$, possibly in a higher dimensional space. Observe that $\lambda_i z_i' = z_i$ and $\lambda_i u_i' = u_i$. Observe further that $R_i(1)$ is the same as A_i, except that it is defined in a lower-dimensional space. Since $(z_i', u_i') \in R_i(1)$ for each $i \in \{1, \ldots, t\}$, it is clear that

$$\begin{array}{ll}
t_i^{j_i}(z_i', u_i') \geq 1 & \forall j_i \in J_i \\
v_i^{k_i}(z_i', u_i') \geq -1 & \forall k_i \in K_i \\
t_i^{j_i}(z_i', u_i') + v_i^{k_i}(z_i', u_i') \geq 0 & \forall j_i \in J_i, \forall k_i \in K_i \\
t_i^{j_i}(z_i', u_i') \geq 0 & \forall j_i \in J_i \\
w_i^{l_i}(z_i', u_i') \geq 0 & \forall l_i \in L_i.
\end{array}$$

After substituting $(z_i', u_i') = \left(\frac{z_i}{\lambda_i}, \frac{u_i}{\lambda_i}\right)$ for each $i \in \{1, \ldots, t\}$ and multiplying both sides of the inequalities by the positive value λ_i, we obtain:

$$\begin{array}{ll}
\lambda_i t_i^{j_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) \geq \lambda_i & \forall j_i \in J_i \\
\lambda_i v_i^{k_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) \geq -\lambda_i & \forall k_i \in K_i \\
\lambda_i t_i^{j_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) + \lambda_i v_i^{k_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) \geq 0 & \forall j_i \in J_i, \forall k_i \in K_i \\
\lambda_i t_i^{j_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) \geq 0 & \forall j_i \in J_i \\
\lambda_i w_i^{l_i}\left(\dfrac{z_i}{\lambda_i}, \dfrac{u_i}{\lambda_i}\right) \geq 0 & \forall l_i \in L_i.
\end{array}$$

Since $t_i^{j_i}$, $v_i^{k_i}$ and $w_i^{l_i}$ are positively-homogenous by (A3), and $\lambda_i > 0$, the above system of inequalities can be rewritten as:

$$t_i^{j_i}(z_i, u_i) \geq \lambda_i \qquad \forall j_i \in J_i$$

$$\begin{array}{ll} v_i^{k_i}(z_i,u_i) \geq -\lambda_i & \forall k_i \in K_i \\ t_i^{j_i}(z_i,u_i) + v_i^{k_i}(z_i,u_i) \geq 0 & \forall j_i \in J_i, \forall k_i \in K_i \\ t_i^{j_i}(z_i,u_i) \geq 0 & \forall j_i \in J_i \\ w_i^{l_i}(z_i,u_i) \geq 0 & \forall l_i \in L_i, \end{array}$$

which implies that $(z_i, u_i) \in R_i(\lambda_i)$. Therefore, it follows that, for each $i \in \{1,\ldots,t\}$, (λ_i, z_i, u_i) is such that $\lambda_i > 0$ and $(z_i, u_i) \in R_i(\lambda_i)$. Additionally, we set $(z_i, u_i) = 0$ for $t < i \leq n$. Since $t_i^{j_i}(0,0) = \lambda t_i^{j_i}\left(\frac{0}{\lambda}, \frac{0}{\lambda}\right)$ for $\lambda > 0$, it follows that $t_i^{j_i}(0,0) = 0$. Similarly, for all i, $j_i \in J_i$, $k_i \in K_i$, and $l_i \in L_i$, $t_i^{j_i}(0,0) = w_i^{l_i}(0,0) = v_i^{k_i}(0,0) = 0$. It follows that $(0,0) \in R_i(0)$. In other words, for each $i \in N$, (λ_i, z_i, u_i) is such that $\lambda_i \geq 0$ and $(z_i, u_i) \in R_i(\lambda_i)$. Therefore, $(\lambda, z, u) \in Q$. Now, we show that if $(\lambda, z, u) \in Q$ then $z \in \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$. Clearly, if $(\lambda, z, u) \in Q$ and $\lambda_i > 0$, then by positive homogeneity of $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$, it follows that $\frac{(z_i,u_i)}{\lambda_i} \in R_i(1)$. As before, then $\left(0, \frac{z_i}{\lambda_i}, \frac{u_i}{\lambda_i}, 0\right) \in A_i$. Assume without loss of generality, by reindexing S_i if necessary, that $\lambda_i > 0$ for $i = 1,\ldots,t$ and $\lambda_i = 0$ for $i = t+1,\ldots,n$. Then, it follows easily that $(z_1, u_1, \ldots, z_t, u_t, 0, 0) \in \operatorname{conv}\left(\bigcup_{i=1}^n A_i\right)$ since it can be expressed as a convex combination of points in $\bigcup_{i=1}^t A_i$. Since $\operatorname{proj}_z \operatorname{conv}\left(\bigcup_{i=1}^n A_i\right) \subseteq \operatorname{conv}\left(\bigcup_{i=1}^n \operatorname{proj}_z A_i\right)$ and, by (A3), $\operatorname{proj}_z A_i \subseteq \operatorname{cl}\operatorname{conv}(S_i)$, it follows that $(z_1, \ldots, z_t, 0) \in \operatorname{conv}\left(\bigcup_{i=1}^n \operatorname{cl}(\operatorname{conv}(S_i)\right) \subseteq \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$. Now, since $\lambda_{t+1} = 0$, then by (A4), $(0, z_{t+1}, 0) \in 0^+\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)\right)$. Therefore, $(z_1, \ldots, z_t, z_{t+1}, 0) \in \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$. By induction, $z \in \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$.

We now prove the last part of the theorem. For this, we assume that, for every i, $\operatorname{proj}_z A_i$ is closed and $\operatorname{proj}_z C_i = 0^+\left(\operatorname{cl}\operatorname{conv}(S_i)\right)$. Since the sets S_i are orthogonal, there do not exist vectors $\psi_i = (0, z_i, 0) \in \operatorname{proj}_z C_i$, not all zero, such that $\sum_{i=1}^n \psi_i = 0$. Define $T_i(\lambda_i) = \lambda_i \operatorname{cl}\operatorname{conv}(S_i)$ for $\lambda_i > 0$ and $T_i(0) = 0^+(\operatorname{cl}\operatorname{conv}(S_i))$. Then, by Theorem 9.8 in [17], it follows that $\bigcup_{i=1}^n \{z \mid \sum_{i=1}^n \lambda_i = 1, z_i \in T_i(\lambda_i)\}$, denoted hereafter as T, equals $\operatorname{cl}\operatorname{conv}(S)$. If $\bar{z} \in T$, then there exists a λ such that $\bar{z}_i \in T_i(\lambda_i)$. If $\lambda_i > 0$, then $\frac{\bar{z}_i}{\lambda_i} \in \operatorname{cl}\operatorname{conv}(S_i)$, and therefore, there exists u_i such that $\frac{(\bar{z}_i,u_i)}{\lambda_i} \in A_i$. On the other hand, if $\lambda_i = 0$, there exists u_i such that $(\bar{z}_i, u_i) \in C_i$. Since A_i and C_i (restricted to the space of z_i and u_i variables) are $R_i(1)$ and $R_i(0)$ respectively, it follows that $(\lambda, \bar{z}, u) \in Q$ and so $\bar{z} \in \operatorname{proj}_z X$ and $\operatorname{cl}\operatorname{conv}(S) \subseteq \operatorname{proj}_z X$. However, we already showed that $\operatorname{proj}_z X \subseteq \operatorname{cl}\operatorname{conv}(S)$ and, therefore, $\operatorname{proj}_z X$ is equal to $\operatorname{cl}\operatorname{conv}(S)$. □

We now discuss the result and the assumptions of Theorem 2.1 in more detail. Considering first the result of this theorem, one might initially think that the stronger result that $\operatorname{proj}_z X = \operatorname{conv}(S)$ holds. We show with examples that $\operatorname{proj}_z X$ can be different from $\operatorname{conv}(S)$ and from $\operatorname{cl}\operatorname{conv}(S)$. In that sense, the result of Theorem 2.1 is as tight as possible. We consider first an example where $\operatorname{conv}(S) \subsetneq \operatorname{proj}_z X$.

Example 2.2. *Consider the set* $S \subseteq \mathbb{R}^2_+$*, defined as* $S = S_1 \cup S_2$*, where* $S_1 = \left\{(z_1, 0) \,\middle|\, 1 \leq z_1 \leq 2\right\}$ *and* $S_2 = \left\{(0, z_2) \,\middle|\, z_2 \geq 1\right\}$*. It can be easily verified that* $\operatorname{conv}(S) = \left\{(z_1, z_2) \,\middle|\, z_1 + z_2 \geq 1, z_1 \geq 0, z_1 < 2, z_2 \geq 0\right\} \cup \left\{(2,0)\right\}$ *as is shown in Figure 1. Observe that* $\operatorname{conv}(S)$ *is not closed. We now apply the convexification tool of Theorem 2.1 to* S *and derive a set* X *that contains* $\operatorname{conv}(S)$ *but is no larger than* $\operatorname{cl}\operatorname{conv}(S)$*. First, we verify that the set* S *satisfies the assumptions of Theorem 2.1. Clearly, (A1) and (A2) hold by the definition of* S*. Next, it is easy to verify that* $\operatorname{conv}(S_1) = \left\{\left(z_1, 0\right) \,\middle|\, z_1 \geq 1, -\frac{1}{2}z_1 \geq -1\right\}$ *and* $\operatorname{conv}(S_2) = \left\{\left(0, z_2\right) \,\middle|\, z_2 \geq 1\right\}$*. Since* z_1*,* $-\frac{1}{2}z_1$*, and* z_2 *are linear, and, therefore, positively-homogeneous, (A3) clearly holds. Finally, since* $C_1 = \left\{(0,0)\right\} \subseteq 0^+\left(\operatorname{cl}\operatorname{conv}(S)\right)$ *and* $C_2 = \left\{(0, z_2) \,\middle|\, z_2 \geq 0\right\} \subseteq 0^+\left(\operatorname{cl}\operatorname{conv}(S_2)\right) \subseteq 0^+\left(\operatorname{cl}\operatorname{conv}(S)\right)$*, then (A4) also holds. Applying Theorem 2.1, we obtain that* $X = \left\{\left(z_1, z_2\right) \,\middle|\, z_1 + z_2 \geq 1, z_1 \leq 2, z_1 \geq 0, z_2 \geq 0\right\}$*. In fact, since, for each* i*,* $C_i = 0^+\left(\operatorname{cl}\operatorname{conv}(S_i)\right)$ *and* $\operatorname{conv}(S_i)$ *is closed, it follows from Theorem 2.1 and is apparent for this example that* $X = \operatorname{cl}\operatorname{conv}(S)$*. This example illustrates that* X *may contain* $\operatorname{conv}(S)$ *as a strict subset.* □

We now consider an example where $\operatorname{proj}_z X \subsetneq \operatorname{cl}\operatorname{conv}(S)$.

Example 2.3. *Consider the set* $S = \bigcup_{i=1}^n S_i$*, where* $S_i = \operatorname{proj}_z \left\{(0, z_i, u_i, 0) \in \mathbb{R}^{2n}_+ \,\middle|\, \sqrt{z_i u_i} \geq 1\right\} = \left\{(0, z_i, 0) \,\middle|\, z_i > 0\right\}$*. Clearly, (A1) and (A2) hold by the definition of* S*. Since* $\sqrt{z_i u_i}$ *is positively-homogeneous, (A3) is also satisfied. Observe that* $\operatorname{proj}_z C_i = \operatorname{proj}_z \left\{(0, z_i, u_i, 0) \in \mathbb{R}^{2n}_+ \,\middle|\, \sqrt{z_i u_i} \geq 0\right\} =$

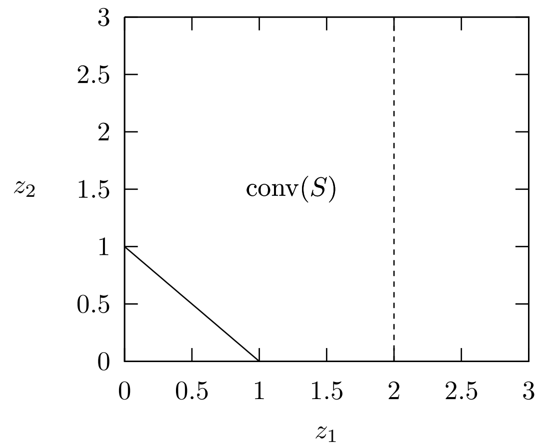

Figure 1: Illustration of Theorem 2.1 and that $\text{conv}(S) \subsetneq \text{proj}_z X$

$\{(0, z_i, 0) \mid z_i \geq 0\} \subseteq 0^+(\text{cl}\,\text{conv}(S))$. *Therefore, (A4) holds. Applying Theorem 2.1, we obtain that* $X = \{(z, u) \in \mathbb{R}^{2n}_+ \mid \sum_{i=1}^n \sqrt{z_i u_i} \geq 1\}$. *If, for any* i, $z_i > 0$ *then there exists* u *such that* $(z, u) \in X$. *Further, for all* u, *it is easy to see that* $(0, u) \notin X$. *Therefore,* $\text{proj}_z X = \{z \in \mathbb{R}^n_+ \mid \sum_{i=1}^n z_i > 0\}$. *This example illustrates that if* $\text{proj}_z A_i$ *is not closed then* $\text{proj}_z X$ *may not be closed either and that, in some cases,* $\text{proj}_z X \subsetneq \text{cl}\,\text{conv}(S)$. □

In the above example, we exploit the fact that $\text{proj}_z A_i$s are not closed to show that $\text{proj}_z X$ may not be closed either. Instead, if $\text{proj}_z A_i$s were closed for all i then, as shown in Theorem 2.1, $\text{proj}_z X$ would typically be closed as well.

We now turn our attention to Assumption (A1) in Theorem 2.1. Assumption (A1) requires that the sets S_i be oriented along orthogonal principal subspaces. A weaker assumption however suffices to prove the theorem. Consider L_i, for $i \in \{1, \ldots, n\}$, to be linear subspaces of $\mathbb{R}^{\sum_{i=1}^n d_i}$, where L_i has dimension d_i. Further, assume that a vector $z_i \in L_i$ cannot be expressed as a linear combination of vectors in $\{L_1, \ldots, L_{i-1}, L_{i+1}, \ldots, L_n\}$. In this case, it is possible to construct a matrix B whose columns form a basis for $\mathbb{R}^{\sum_{i=1}^n d_i}$ where the columns, that are indexed from $1 + \sum_{i=1}^{j-1} d_i$ to $\sum_{i=1}^{j} d_i$, form a basis for L_j. Then, define new variables s such that $s = B^{-1}z$. If $z \in S_j \subseteq L_j$, it follows that $s_k \neq 0$ only if $1 + \sum_{i=1}^{j-1} d_i \leq k \leq \sum_{i=1}^{j} d_i$. Therefore, the theorem now applies to the transformed space of s variables. This observation leads to the following simple derivation of the intersection cut in integer programming.

Example 2.4. *Consider a polyhedral cone* $P = \{x \mid Ax \leq b\}$, *where* $A \in \mathbb{R}^{n \times n}$ *is an invertible matrix. Let* X *be the set of points that satisfy the disjunction* $\pi^T x \leq \pi_0^1 \vee \pi^T x \geq \pi_0^2$, *where* $\pi_0^1 < \pi_0^2$. *We are interested in deriving the convex hull of* $P \cap X$. *Observe that this setting can be used to derive all intersection/split cuts (see Balas [5]). Introducing the slack variables* μ *and defining* $\gamma = \pi^T A^{-1}$, $\gamma_0^1 = \gamma b - \pi_0^2$, *and* $\gamma_0^2 = \gamma b - \pi_0^1$, *we reduce the above problem into one involving convexification of* $\mathcal{M} = \{\mu \mid \mu \geq 0, \gamma\mu \leq \gamma_0^1 \vee \gamma\mu \geq \gamma_0^2\}$. *We assume without loss of generality that, for each* i, $\gamma_i \neq 0$. *The reformulation of the problem in the space of the slack variables, after suitable translation, is an example of the orthogonalization discussed above. Here,* μ *corresponds to* $-s$ *and* x *corresponds to* z. *The matrix* B *equals* A^{-1} *and its columns are the extreme rays of* P. *Since* $\mu \geq 0$ *is the recession cone for* $\mathcal{M}$, *whenever it contains a feasible point, if* $\mu = 0$ *is feasible to* $\mathcal{M}$, *then* $\text{conv}(\mathcal{M}) = \{\mu \mid \mu \geq 0\}$. *Define* $p_i = \frac{\gamma_0^1}{\gamma_i}$ *and* $q_i = \frac{\gamma_0^2}{\gamma_i}$. *If* $\mu = 0$ *is not feasible to* $\mathcal{M}$, *then* $\gamma_0^1 < 0$ *and* $\gamma_0^2 > 0$. *It follows that, for each* i, *exactly one of* p_i *or* q_i *is greater than* 0. *Since* $\mu_i \geq 0$ *is a recession direction for* $\text{conv}(\mathcal{M})$ *and the extreme points of* $\mathcal{M}$ *have at most one non-zero, it follows that:*

$$\text{conv}(\mathcal{M}) = \bigcup_{i=1}^{n} \{(0, \ldots, 0, \mu_i, 0, \ldots, 0) \mid \mu_i \geq \max\{p_i, q_i\}\}.$$

Now, applying Theorem 2.1, it follows that:

$$\operatorname{conv}(\mathcal{M}) = \left\{ \mu \;\middle|\; \sum_{i=1}^{n} \frac{\mu_i}{\max\{p_i, q_i\}} \geq 1, \mu \geq 0 \right\}.$$

Substituting back μ, p_i, and q_i in the above, we obtain:

$$\operatorname{conv}(\mathcal{M}) = \left\{ x \;\middle|\; \sum_{i=1}^{n} \frac{(b-Ax)_i}{\max\left\{\frac{\pi^T A_{\cdot i}^{-1} b - \pi_0^2}{\pi^T A_{\cdot i}^{-1}}, \frac{\pi^T A_{\cdot i}^{-1} b - \pi_0^1}{\pi^T A_{\cdot i}^{-1}}\right\}} \geq 1, Ax \leq b \right\}. \qquad \square$$

We next discuss Assumption (A3). This assumption requires that the convex hulls of the sets S_i be known, possibly in a higher dimensional space, and that the functions $t_i^{j_i}$, for all $j_i \in J_i$, $v_i^{k_i}$, for all $k_i \in K_i$, and $w_i^{l_i}$, for all $l_i \in L_i$, used in the description of the convex hulls be positively-homogenous. In the ensuing example, we show that a simple transformation might suffice to convert the natural inequality description of $\operatorname{conv}(S_i)$ into one that uses positively-homogenous functions. We also illustrate that it is necessary to make the assumption that the functions are positively-homogenous.

Example 2.5. *Let $S = \bigcup_{i=1}^n S_i$, where $S_i = \left\{(0, x_i, y_i, 0) \in \mathbb{R}_+^{2n} \;\middle|\; x_i y_i \geq r\right\}$ and $r > 0$. Clearly, (A1) and (A2) hold by the definition of S. Since S_i is already closed and convex, $\operatorname{cl}\operatorname{conv}(S_i) = S_i$, i.e., $\operatorname{cl}\operatorname{conv}(S_i) = \left\{(0, x_i, y_i, 0) \in \mathbb{R}_+^{2n} \;\middle|\; \frac{1}{r} x_i y_i \geq 1\right\}$. The above representation of $\operatorname{cl}\operatorname{conv}(S_i)$ does not directly satisfy (A3) since $\frac{1}{r} x_i y_i$ is not a positively-homogenous function of (x_i, y_i). However, $\operatorname{cl}\operatorname{conv}(S_i)$ may be rewritten as $\operatorname{cl}\operatorname{conv}(S_i) = \left\{(0, x_i, y_i, 0) \in \mathbb{R}_+^{2n} \;\middle|\; \sqrt{\frac{1}{r} x_i y_i} \geq 1\right\}$, an expression that uses the function, $\sqrt{\frac{1}{r} x_i y_i}$, which is positively-homogenous in (x_i, y_i). With this representation, (A3) is satisfied. Since $C_i = \left\{(0, x_i, y_i, 0) \in \mathbb{R}_+^{2n} \;\middle|\; \sqrt{x_i y_i} \geq 0\right\} = 0^+(\operatorname{cl}\operatorname{conv}(S_i))$, (A4) is satisfied. Therefore, Theorem 2.1 implies that $X = \operatorname{cl}\operatorname{conv}(S) = \left\{(x, y) \in \mathbb{R}_+^{2n} \;\middle|\; \sum_{i=1}^n \sqrt{x_i y_i} \geq \sqrt{r}\right\}$. Observe finally that the transformation to positively-homogenous functions is necessary and not an artifact of the proof technique. In fact, if we use the original definition of $\operatorname{cl}\operatorname{conv}(S_i)$, when applying Theorem 2.1, and disregard the lack of positive-homogeneity, the resulting set would be $X' = \left\{(x, y) \in \mathbb{R}_+^{2n} \;\middle|\; \sum_{i=1}^n x_i y_i \geq r\right\}$. The set X' is nonconvex and does not even contain $\operatorname{conv}(S)$. To see this, let $r = 1$ and $n = 2$. Note that $(x_1, y_1, x_2, y_2) = (0.5, 0.5, 0.5, 0.5)$ is expressible as a convex combination of the two points in S, namely, $(1, 1, 0, 0) \in S_1$ and $(0, 0, 1, 1) \in S_2$. Therefore $(0.5, 0.5, 0.5, 0.5)$ belongs to $\operatorname{conv}(S)$. However, it does not satisfy the defining inequality of X' whereas it does satisfy the defining inequality of X.* $\square$

If $\lambda_i t_i^{j_i}\left(\frac{z_i}{\lambda_i}, \frac{u_i}{\lambda_i}\right) \leq t_i^{j_i}(z_i, u_i)$ for all $\lambda \in (0, 1]$, then X still outer-approximates $\operatorname{cl}\operatorname{conv}(S)$. Intuitively, while performing Fourier-Motzkin elimination, $\lambda_i t_i^{j_i}\left(\frac{z_i}{\lambda_i}, \frac{u_i}{\lambda_i}\right) \leq t_i^{j_i}(z_i, u_i)$ ensures that X is contained in the closure convex hull of the disjunctive union of S_i, whereas $\lambda_i t_i^{j_i}\left(\frac{z_i}{\lambda_i}, \frac{u_i}{\lambda_i}\right) \geq t_i^{j_i}(z_i, u_i)$ ensures that X is contained in $\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$. Similar statements can be made about $v_i^{k_i}(z_i, u_i)$ and $w_i^{l_i}(z_i, u_i)$. The latter of these conditions will be explored further in Proposition 3.8 to derive sufficient conditions that help verify a slightly relaxed version of (A2).

We now turn our attention to Assumption (A4). This assumption might appear quite technical and might also seem difficult to verify in practice. However, this is not the case. We show next that by simply requiring that the functions $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ are concave, in addition to being positively-homogenous, Assumption (A4) is automatically satisfied. Concavity of $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ is not an important restriction since the convexity of a positively-homogenous function's upper-level set implies concavity over the region of interest.

Proposition 2.6. *If, for all i, $j_i \in J_i$, $k_i \in K_i$, and $l_i \in L_i$, the functions $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$, as defined in Theorem 2.1, are concave in addition to being positively-homogeneous, and the sets S_i are not empty, then $\operatorname{proj}_z C_i \subseteq 0^+\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)\right)$, i.e., Assumption (A4) is satisfied. Moreover,* *if the upper-level set of a positively-homogenous function is convex, then the function is concave, wherever it is positive. More precisely, if* $W = \{(z,u) \mid t(z,u) \geq 1\}$ *is convex and* $t(z,u)$ *is positively-homogenous, then* $D = \{(z,u) \mid t(z,u) > 0\}$ *is convex and* $t(z,u)$ *is concave over* D*. If, in addition,* $\mathrm{cl}(D)$ *is locally simplicial or more specially, polyhedral, and* $t(z,u)$ *is continuous then* $t(z,u)$ *is concave over* $\mathrm{cl}(D)$*.*

Proof. Let $(0, z_i, 0) \in S_i$. By Assumption (A3), there exists u_i such that $(0, z_i, u_i, 0) \in A_i$. Consider $(0, z_i', u_i', 0) \in C_i$ and $\alpha > 0$. Then, by positive homogeneity and concavity of $t_i^{j_i}$, it follows that

$$t_i^{j_i}(z_i + \alpha z_i', u_i + \alpha u_i') \geq t_i^{j_i}(z_i, u_i) + t_i^{j_i}(\alpha z_i', \alpha u_i') = t_i^{j_i}(z_i, u_i) + \alpha t_i^{j_i}(z_i', u_i') \geq t_i^{j_i}(z_i, u_i) \geq 1.$$

The first inequality holds because of Theorem 4.7 in [17], the first equality because $t_i^{j_i}$s are positively-homogenous, the second inequality because $(0, z_i', u_i', 0) \in C_i$ and $\alpha > 0$, and the last inequality because $(0, z_i, u_i, 0) \in A_i$. Similarly, $v_i^{k_i}(z_i + \alpha z_i', u_i + \alpha u_i') \geq -1$ and $w_i^{l_i}(z_i + \alpha z_i', u_i + \alpha u_i') \geq 0$. Therefore, $(z_i + \alpha z_i', u_i + \alpha u_i') \in A_i$ and so, for all $\alpha > 0$, $(0, z_i + \alpha z_i', 0) \in \mathrm{cl}\,\mathrm{conv}(S_i) \subseteq \mathrm{cl}\,\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Since $(0, z_i, 0) \in \mathrm{cl}\,\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$, it follows by Theorem 8.3 in [17] that $(0, z_i', 0) \in 0^+\left(\mathrm{cl}\,\mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)\right)$.

If W is convex, then $W_K = \{(\lambda, x) \mid \lambda > 0, x = \lambda(z,u),\ t(z,u) \geq 1\}$ is the smallest convex cone containing $\{(1,x) \mid x \in W\}$. Exploiting the positive homogeneity of t, we may rewrite W_K as:

$$W_K = \{(\lambda, x) \mid \lambda > 0, t(x) \geq \lambda\}.$$

Now, D is the projection of W_K in the space of x and is therefore convex. Further, the hypograph of $t(z,u)$ over D is $\{(r,x) \mid r \leq t(x), x \in D\} = \{(r,x) \mid r \leq \lambda \leq t(x), \lambda > 0\}$, which is convex if W_K is convex. The last statement of the proposition follows from Theorems 10.3 and 20.5 in [17]. □

Even when some of the technical assumptions of Theorem 2.1 are not satisfied, it is typically the case that X yields an outer-approximation of $\mathrm{conv}(S)$. To see this, observe that Proposition 2.6 shows that the functions $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ are concave, if they are positively-homogenous, as is assumed in Theorem 2.1, and their upper-level sets are convex. However, if concavity of these functions is known, then the outer-approximation of $\mathrm{conv}(S)$ by $\mathrm{proj}_z X$ can be shown under relatively mild assumptions.

Proposition 2.7. *Let* $S \subseteq \mathbb{R}^{\sum_i d_i}$ *and, for all* $i \in N = \{1, \ldots, n\}$*, let* $S_i \subseteq S$*. Let the points* z *of* S *be written as* $z = (z_1, \ldots, z_i, \ldots, z_n) \in S$*, where* $z_i \in \mathbb{R}^{d_i}$*. Assume that Assumption (A1) of Theorem 2.1 holds. Further, assume that* $\mathrm{proj}_z A_i$*, where* A_i *is as defined in (1), yields an outer-approximation of* $\mathrm{conv}(S_i)$ *and that, for all* $i \in N$*,* $j_i \in \{1, \ldots, J_i\}$*,* $k_i \in \{1, \ldots, K_i\}$*, and* $l_i \in \{1, \ldots, L_i\}$*,* $t_i^{j_i}(0,0)$*,* $v_i^{k_i}(0,0)$*, and* $w_i^{l_i}(0,0)$ *are non-negative. Then,* $\mathrm{proj}_z(X)$*, where* X *is as defined in (2), outer-approximates* $\bigcup_{i=1}^n S_i$*. If, in addition, Assumption (A2) of Theorem 2.1 holds and* X *is convex (for example, if the functions* $t_i^{j_i}$*,* $v_i^{k_i}$*, and* $w_i^{l_i}$ *are concave), then* $\mathrm{proj}_z X \supseteq \mathrm{conv}(S)$*.*

Proof. If Assumption (A1) is satisfied, then the sets S_i, for $i \in N$, are orthogonal. It can be easily verified that, if $t_i^{j_i}(0,0)$, $v_i^{k_i}(0,0)$, and $w_i^{l_i}(0,0)$ are non-negative, then every constraint defining X is valid for all S_i, where $i \in N$. Therefore, $\mathrm{proj}_z X \supseteq \bigcup_{i=1}^n S_i$. If Assumption (A2) is satisfied as well, then Claim 1 in the proof of Theorem 2.1 holds. Therefore, $\mathrm{conv}(S) = \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right)$. Further, if X is convex, so is $\mathrm{proj}_z X$. Since $\mathrm{proj}_z X \supseteq \bigcup_{i=1}^n S_i$, it follows that $\mathrm{proj}_z X \supseteq \mathrm{conv}\left(\bigcup_{i=1}^n S_i\right) = \mathrm{conv}(S)$. □

When the constituent functions $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ are concave, the result of Proposition 2.7 could also be derived using disjunctive programming. We verify Proposition 2.7 using this approach, since it more clearly reveals the source of the difference between the outer-approximation of Proposition 2.7 and the convex hull identified in Theorem 2.1. For example, one can assert that $\sum_{i \in N} t_i^{j_i}(z_i, u_i) \geq 1$, by simply noticing that if $\lambda_i > 0$ for $i \in \{1, \ldots, t\}$ then:

$$\begin{aligned} 1 &= \sum_{i=1}^{t} \lambda_i \\ &\leq \sum_{i=1}^{t} \lambda_i t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) + \sum_{i=t+1}^{n} t_i^{j_i}(z_i, u_i) \\ &\leq \sum_{i=1}^{t} \lambda_i \left(t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) + \sum_{i' \in N, i' \neq i} t_{i'}^{j_{i'}}\left(\frac{0,0}{\lambda_i}\right)\right) + \sum_{i=t+1}^{n} t_i^{j_i}(z_i, u_i) \leq \sum_{i=1}^{n} t_i^{j_i}(z_i, u_i), \end{aligned} \tag{10}$$

where the first inequality follows by summing the inequalities $\lambda_i \leq \lambda_i t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right)$ for $i \in \{1, \ldots, t\}$ and $t_i^{j_i}(z_i, u_i) \geq 0$ for $i \in \{t+1, \ldots, n\}$, the second inequality follows since $t_{i'}^{j_{i'}}(0,0) \geq 0$, and the third inequality from the concavity of $\sum_{i=1}^{t} t_i^{j_i}(z_i, u_i)$. Similarly, $\sum_{i \in T} v_i^{j_i}(z_i, u_i) \geq -1$, by realizing, in addition, that $-\sum_{i \in T} \lambda_i \geq -1$.

Proposition 2.7 provides a simple proof of the validity of the constraints defining X for $\text{conv}(S)$. In fact, if the primary purpose of deriving X is to develop a convex outer-approximation, then Proposition 2.7 can often replace Theorem 2.1. For example, the convex hull description for the bilinear covering sets (derived in Proposition 3.9) can be shown to yield a convex-outerapproximation, if Proposition 2.7 is invoked instead of Theorem 2.1 in the proof of the result. Nevertheless, the insights gained from Theorem 2.1 are very useful. For example, we illustrate next that the search for a representation of $\text{conv}(S_i)$ using positively-homogenous functions can substantially improve the relaxation. This insight will play an important role in deriving relaxations for the bilinear covering set.

Example 2.8. *Consider* $S = \bigcup_{i=1}^{n} S_i$*, where, for each* $i \in \{1, \ldots, n\}$*, let*

$$S_i = \left\{(0, \ldots, 0, z_i, 0, \ldots, 0) \in \mathbb{R}_+^n \;\middle|\; \sqrt{z_i} \geq 1\right\}.$$

Proposition 2.7 shows that

$$X' = \left\{(z_1, \ldots, z_n) \in \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^{n} \sqrt{z_i} \geq 1\right\}$$

is a convex outer-approximation of $\text{conv}(S)$*. Note that the square-root function used in expressing* S_i *is concave, but not positively-homogenous. Instead, if* S_i*s are represented equivalently as*

$$S_i = \left\{(0, \ldots, 0, z_i, 0, \ldots, 0) \in \mathbb{R}_+^n \;\middle|\; z_i \geq 1\right\},$$

then Theorem 2.1 yields the convex hull of S*, which is*

$$X = \left\{(z_1, \ldots, z_n) \in \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^{n} z_i \geq 1\right\}.$$

Clearly, by construction, $X = \text{conv}(S) \subseteq X'$*. In this particular example, the inclusion of* X *in* X' *can also be verified using the subadditivity of the square-root function for non-negative variables. This example illustrates that it often helps to find representations of convex hulls of* S_i *using positively-homogenous functions, even when equivalent representations exist using concave functions.*

As discussed in Example 2.8, if one can find a description of $\text{conv}(S_i)$ that uses positively-homogenous functions then one can apply Theorem 2.1 to identify the convex hull of the orthogonal disjunctions, thus deriving a superior relaxation. In general, a positive homogenous description can be obtained by adding one homogenizing variable for each orthogonal disjunction and expressing A_i using the inequalities, $t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq 1$ for all $j_i \in J_i$, $v_i^{k_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq -1$ for all $k_i \in K_i$, and $w_i^{l_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq 1$ for all $l_i \in L_i$ along with the inequalities, $\lambda_i \geq 1$ and $-\lambda_i \geq -1$. However, this process suffers from the drawback that it introduces new variables in the relaxation. Instead, it may be possible to find a separating hyperplane without increasing the problem dimension and, thereby, circumvent the need to introduce new variables. Consider, for simplicity, the case of Theorem 2.1, where A_i is not an extended formulation, *i.e.*, it does not need the additional u_i variables. The case where A_i contains u_i variables can be handled similarly. Now, consider a point z' that does not belong to $\operatorname{cl}\operatorname{conv}(S)$. If it is possible to find, for all i, a $j'_i \in \operatorname{argmin}_j \left\{ t_i^j(z'_i) \;\middle|\; j = 1, \ldots, J_i \right\}$, a $k'_i \in \operatorname{argmin}_k \left\{ v_i^k(z'_i) \;\middle|\; k = 1, \ldots, K_i \right\}$ and an $l'_i \in \operatorname{argmin}_l \left\{ w_i^l(z'_i) \;\middle|\; l = 1, \ldots, L_i \right\}$ then using the closed-form expression of X in (2), one can identify an inequality that separates z' from X. For example, if an inequality of the form $\sum_{i \in N} t_i^{j_i}(z_i) \geq 1$ violates z'_i, *i.e.*, $\sum_{i \in N} t_i^{j_i}(z'_i) < 1$, then $\sum_{i \in N} t_i^{j'_i}(z'_i) < 1$ as well, since, by the definition of j'_i, $t_i^{j'_i}(z'_i) \leq t_i^{j_i}(z'_i)$ for all i. Such a technique will be useful in deriving a separating hyperplane for mixed-integer and pure-integer bilinear covering sets.

Now, we discuss another technique that can be used to find representations of the convex hull of each S_i that uses positively-homogenous functions but does not require additional variables. The main idea is that one can homogenize the inequality using an extra variable and then maximize the resulting function over the introduced variable to derive a positively-homogenous function describing the set. We illustrate this idea by deriving a positively-homogenous function that describes the following bilinear covering set:

$$Q = \left\{ (x, y) \in \mathbb{R}^2_+ \mid axy + bx + cy \geq r \right\}, \tag{11}$$

where a, b, and c are assumed to be non-negative. We assume without loss of generality that $r > 0$. Otherwise, $Q = \mathbb{R}^2_+$. We may also assume without loss of generality that $c \geq b$ and, consequently, assume that at least one of a and c is strictly positive. Then, for any feasible (x, y), it follows that $ax + c > 0$. Therefore, $Q = \left\{ (x, y) \in \mathbb{R}^2_+ \mid y \geq \frac{r - bx}{ax + c} \right\}$. First, we verify that the inequality is convex. Let $f(x) = \frac{r - bx}{ax + c}$. Since

$$\frac{\partial^2 f}{\partial x^2} = \frac{2a(bc + ar)}{(ax + c)^3}$$

is nonnegative if $x \geq 0$, Q is expressed as the intersection of the epigraph of a convex function with the non-negative orthant. Therefore, Q is convex. Also, note that the defining inequality of Q is not positively-homogenous. We show how the above inequality can be homogenized without introducing new variables in the formulation. To carry out this transformation, we first homogenize the defining inequality, $axy + bx + cy \geq r$, using an additional variable h, that is restricted to be positive. This is accomplished by rewriting the defining inequality of Q as $\frac{axy}{h} + bx + cy \geq rh$. Since h is positive, we can multiply throughout by h, and express the above inequality as: $axy + bxh + cyh \geq rh^2$. This is a positively-homogenous inequality which defines Q as long as h is positive. Therefore, Q can now be described by the inequalities:

$$axy + bxh + cyh \geq rh^2 \text{ and } h \geq 1.$$

In order for (x, y, h) to satisfy the first inequality above, h must be such that:

$$\frac{bx + cy - \sqrt{(bx + cy)^2 + 4arxy}}{2r} \leq h \leq \frac{bx + cy + \sqrt{(bx + cy)^2 + 4arxy}}{2r}.$$

It can be easily verified that the functions bounding h are positively-homogenous. In fact, when the bounding functions on h are derived from a positively-homogenous constraint, they must, in general, be positively-homogenous. This can be inferred because for each (x, y, h) that satisfies a positively-homogenous constraint and an arbitrary $\lambda > 0$, it must be that $(\lambda x, \lambda y, \lambda h)$ satisfies the constraint as well. The lower bounding function is nonpositive. Therefore, the set Q can be rewritten as:

$$\eta(x, y) = \frac{1}{2} \left(bx + cy + \sqrt{(bx + cy)^2 + 4arxy} \right) \geq r. \tag{12}$$

We have thus expressed Q as the upper-level set of a positively-homogenous function without introducing new variables. In fact, since Proposition 2.6 asserts that a positively-homogenous function whose upper-level set is convex, is concave, it follows from the convexity of Q that $\eta(x, y)$ must be concave over the non-negative quadrant. In other words, we have established the following result.

Proposition 2.9. *Let $Q = \big\{(x, y) \in \mathbb{R}^2_+ \mid axy + bx + cy \geq r\big\}$, where a, b, c are non-negative, and r is strictly positive. Then, Q has a convex description (upper level set of a concave function) that uses positively-homogenous functions. In particular, $Q = \big\{(x, y) \in \mathbb{R}^2_+ \mid \eta(x, y) \geq r\big\}$, where $\eta(x, y)$ is as defined in (12).* □

3 Convex Extension Property

In this section, we study the convex extension property which is the basis for Assumption (A2) in Theorem 2.1. The convex extension property clearly holds when S is defined as the union of orthogonal sets, S_i, for $i \in \{1, \dots, n\}$. However, it is also satisfied in other situations where it may not initially be suspected to hold. In this section, we show that the convex extension property holds for certain mixed, pure and continuous bilinear sets. In the process, we establish a general set of sufficient conditions that are useful in proving that the convex extensions property holds for many bilinear covering sets. We first formally define the notion of a convex extension for orthogonal disjunctive sets. This definition is adapted from Tawarmalani and Sahinidis [23].

Definition 3.1. *Let $S_i \subseteq S$ for $i \in N = \{1, \dots, n\}$. We say that S has the convex extension property for orthogonal disjunctive sets S_i if (A1) and a slightly relaxed form of (A2) hold. More specifically, S has the convex extension property if every point z in S can be expressed as a convex combination of points χ_i in $\operatorname{cl}\operatorname{conv}(S_i)$ and/or a conic combination of rays ψ_i in $0^+\big(\operatorname{cl}\operatorname{conv}(S_i)\big)$, i.e., for $i \in I \subseteq N$, there exist $\lambda_i \geq 0$ and $\mu_i \geq 0$, that satisfy $\sum_{i\in I} \lambda_i = 1$, such that*

$$z = \sum_{i\in I} \lambda_i \chi_i + \sum_{i\in I} \mu_i \psi_i. \tag{13}$$

The convex extension property in Definition 3.1 is more general than Assumption (A2) in Theorem 2.1, in that it allows the use of non-negative multiples of recession directions in the expression of z. Since $\chi_i + \frac{\mu_i}{\lambda_i}\psi_i \in \operatorname{cl}\operatorname{conv}(S_i)$, it may seem that the recession directions in (13) are not necessary. However, this is not true since λ_i may be zero even when μ_i is not. This technicality is often important in practical applications. Nevertheless, it can be observed that even if (A2) is replaced with (13), Theorem 2.1 holds with slight modifications, as discussed below. Instead of $\operatorname{conv}(S) = \operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$, as was proved in Claim 1 in the proof of Theorem 2.1, we can only establish that (13) implies

$$\operatorname{cl}\operatorname{conv}(S) = \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right). \tag{14}$$

In fact, (14) is equivalent to (13). On the one hand, since, for each $i \in \{1, \dots, n\}$, $S_i \subseteq S$ it follows that $\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right) \subseteq \operatorname{cl}\operatorname{conv}(S)$. On the other hand, since S_is are orthogonal, by Theorem 9.8 in [17],

$$\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right) = \bigcup\left\{\lambda_1 \operatorname{cl}\operatorname{conv}(S_1) + \cdots \lambda_n \operatorname{cl}\operatorname{conv}(S_n) \;\middle|\; \lambda_i \geq 0^+,\ \sum_{i=1}^n \lambda_i = 1\right\}, \tag{15}$$

where the notation $\lambda_i \geq 0^+$ means that $\lambda_i \operatorname{cl}\operatorname{conv}(S_i)$ is taken to be $0^+(\operatorname{cl}\operatorname{conv}(S_i))$ rather than $\{0\}$ when $\lambda_i = 0$. Observe that (13) is another way to represent the set on the right-hand-side of (15) since if $\lambda_i > 0$ then $\chi_i + \frac{\mu_i}{\lambda_i}\psi_i \in \operatorname{cl}\operatorname{conv}(S_i)$. Otherwise, $\psi_i \in 0^+(\operatorname{cl}\operatorname{conv}(S_i))$. Now, if we assume (13), or equivalently, (14), the proof of Theorem 2.1 shows that $\operatorname{cl}\operatorname{proj}_z X = \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$, and, therefore, by (14), $\operatorname{cl}\operatorname{proj}_z X = \operatorname{cl}\operatorname{conv}(S)$. In this case, the last statement of Theorem 2.1 can often be used to establish closedness of $\operatorname{proj}_z X$. Note that $\operatorname{proj}_z A_i$ is closed whenever $\operatorname{conv}(S_i)$ is closed. Therefore, if $\operatorname{conv}(S_i)$ is closed and $\operatorname{proj}_z C_i = 0^+(\operatorname{cl}\operatorname{conv} S_i)$, it follows that $\operatorname{proj}_z X = \operatorname{cl}\operatorname{conv}(S)$. Since most practical situations demand $\operatorname{cl}\operatorname{conv}(S)$, it suffices to establish (13) instead of Assumption (A2) in Theorem 2.1. Similarly, if Assumption (A2) is replaced with (13) in Proposition 2.7, it can be easily established that $\operatorname{cl}\operatorname{conv}(S) \subseteq \operatorname{cl}\operatorname{proj}_z X$. This is because $\operatorname{cl}\operatorname{conv}(S) = \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right) \subseteq \operatorname{cl}\operatorname{conv}\left(\operatorname{proj}_z X\right) = \operatorname{cl}\operatorname{proj}_z X$, where the first equality follows from the equivalence of (13) and (14), the first containment since $\bigcup_{i=1}^n S_i \subseteq \operatorname{proj}_z X$, and the last equality since $\operatorname{proj}_z X$ is convex.

We next present a nontrivial set for which it can be proved from first principles that the convex extension property holds for orthogonal disjunctive sets. This set appears in a nonconvex formulation of the trim-loss problem proposed by Harjunkoski et al. [12]. The model is designed to determine the best way to cut a finite number of large rolls of a raw-material into smaller products using a certain number of cutting patterns. Let I be the index set of products and the J be the index set of the cutting patterns that are to be chosen. The demand for a product i is known *a priori* and is denoted by $n_{i,\text{order}}$. For each $(i,j) \in I \times J$, let $n_{ij} \in \mathbb{Z}_+$ be the decision variable that specifies the number of products to type i produced in the cutting pattern j and, for each $j \in J$, let $m_j \in \mathbb{Z}_+$ be the number of times the cutting pattern j is used. The following bilinear constraints model that the demand for each product is met:

$$\sum_{j=1}^{J} m_j n_{ij} \geq n_{i,\text{order}}, \text{ for } i = 1, \dots, I, \tag{16}$$

In Proposition 3.2, we show that the bilinear integer sets defined by the constraint (16) satisfy the convex extension property for disjunctive orthogonal sets. We use this result along with Theorem 2.1 to obtain the convex hull of integer bilinear covering sets in Proposition 3.3.

Proposition 3.2. *Consider a bilinear integer knapsack set*

$$B^I = \left\{ (x_1, y_1, x_2, y_2) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 \;\middle|\; x_1 y_1 + x_2 y_2 \geq r \right\}.$$

where $r > 0$. Then, B^I has the convex extension property (13) with respect to the orthogonal disjunctive sets

$$\begin{aligned} B_1^I &= \left\{ (x_1, y_1, 0, 0) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 \;\middle|\; x_1 y_1 \geq r \right\}, \\ B_2^I &= \left\{ (0, 0, x_2, y_2) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 \;\middle|\; x_2 y_2 \geq r \right\}. \end{aligned}$$

Proof. Let $(x_1, y_1, x_2, y_2) \in B^I$. We show that there exist (i) certain subsets I and I' of $\{1, 2\}$, (ii) for each $i \in I$, a finite j_i, (iii) for each $i \in I'$, a finite j'_i, (iv) for each $i \in I$ and $j \in \{1, \dots, j_i\}$, a point $\chi_{i,j} \in B_i^I$, and (v) for each $i \in I'$ and $j \in \{1, \dots, j'_i\}$, a ray $\psi_{i,j}$ of B_i^I, such that

$$(x_1, y_1, x_2, y_2) = \sum_{i \in I} \sum_{j=1}^{j_i} \lambda_{i,j} \chi_{i,j} + \sum_{i \in I'} \sum_{j=1}^{j'_i} \mu_{i,j} \psi_{i,j}, \tag{17}$$

where the multipliers are such that (a) $\sum_{i \in I} \sum_{j=1}^{j_i} \lambda_{i,j} = 1$, (b) for each $i \in I$ and $j \in \{1, \dots, j_i\}$, $\lambda_{i,j} \geq 0$, and (c) for each $i \in I'$ and $j \in \{1, \dots, j'_i\}$, $\mu_{i,j} \geq 0$.

We assume without loss of generality that $x_1 \leq y_1 \leq y_2$ and $x_2 \leq y_2$ since the variables x_1, y_1, x_2, and y_2 can be renamed such that the largest variable is called y_2 and the largest variable in the other pair is called y_1. Note first that if $x_1 = 0$, it suffices to choose $I = \{2\}$, $I' = \{1\}$, $j_2 = 1$, $j'_1 = 1$ with $\chi_{2,1} = (0, 0, x_2, y_2)$ and $\psi_{1,1} = (0, 1, 0, 0)$ to show that (13) holds. Therefore, we assume in the remainder of this proof that $x_1 \geq 1$ and, consequently, $x_1 y_1 \geq 1$. We consider two cases.

Case 1: $x_2 \geq x_1 y_1$. In this case, we choose $I = \{1, 2\}$, $I' = \{2\}$, and $j_1 = j_2 = j'_2 = 1$. Consider the points $\chi_{1,1} = ((y_2 + 1)\, x_1, (y_2 + 1)\, y_1, 0, 0)$ and $\chi_{2,1} = (0, 0, x_2, y_2 + 1)$, and the ray $\psi_{2,1} = (0, 0, 1, 0)$. Clearly, $\chi_{1,1} \in B_1^I$, since $(y_2 + 1)^2 x_1 y_1 \geq x_1 y_1 + y_2^2 x_1 y_1 \geq x_1 y_1 + y_2^2 \geq x_1 y_1 + x_2 y_2 \geq r$. Similarly, $\chi_{2,1} \in B_2^I$, since $x_2\,(y_2 + 1) \geq x_2 y_2 + x_2 \geq x_2 y_2 + x_1 y_1 \geq r$. It is easily verified that

$$(x_1, y_1, x_2, y_2) = \frac{1}{y_2 + 1} \chi_{1,1} + \frac{y_2}{y_2 + 1} \chi_{2,1} + \frac{x_2}{y_2 + 1} \psi_{2,1}$$

which shows that (17) is feasible.

Case 2: $x_2 \le x_1y_1 - 1$. In this case, we choose $I = \{1,2\}$, $I' = \{1,2\}$, $j_2 = 1$, and $j_1 = j_1' = j_2' = 2$ with $\chi_{1,1} = (x_1+\alpha, y_1, 0, 0)$, $\chi_{1,2} = (x_1, y_1+\beta, 0, 0)$, $\chi_{2,1} = (0,0,x_2,y_2+\delta)$, $\psi_{1,1} = (1,0,0,0)$, $\psi_{1,2} = (0,1,0,0)$, $\psi_{2,1} = (0,0,1,0)$, and $\psi_{2,2} = (0,0,0,1)$, where $\alpha = \left\lceil \frac{x_2y_2}{y_1} \right\rceil$, $\beta = \left\lceil \frac{x_2y_2}{x_1} \right\rceil$, and $\delta = \left\lceil \frac{x_1y_1}{x_2} \right\rceil$. It follows from the way α, β, and δ are defined that $\chi_{1,1}$ and $\chi_{1,2}$ belong to B_I^1 whereas $\chi_{2,1}$ belongs to B_I^2. We need to prove that (17) has a feasible solution. Eliminating $\mu_{i,j}$ and using $\lambda_{2,1} = 1 - \lambda_{1,1} - \lambda_{1,2}$ to eliminate $\lambda_{2,1}$, (17) reduces to the following system:

$$\begin{aligned}
\lambda_{1,1}(x_1+\alpha) + \lambda_{1,2}(x_1) &\le x_1 \\
\lambda_{1,1}(y_1) + \lambda_{1,2}(y_1+\beta) &\le y_1 \\
(1-\lambda_{1,1}-\lambda_{1,2})\,x_2 &\le x_2 \quad \text{(redundant)} \\
(1-\lambda_{1,1}-\lambda_{1,2})\,(y_2+\delta) &\le y_2 \\
\lambda_{1,1}+\lambda_{1,2} &\le 1 \\
\lambda_{1,1} &\ge 0 \\
\lambda_{1,2} &\ge 0.
\end{aligned} \tag{18}$$

Projecting out $\lambda_{1,1}$ using Fourier-Motzkin elimination, we obtain

$$\max\left\{0, \frac{\alpha\delta - x_1y_2}{\alpha\,(y_2+\delta)}\right\} \le \lambda_{1,2} \le \min\left\{1, \frac{y_1}{y_1+\beta}, \frac{y_1y_2}{\beta\,(y_2+\delta)}\right\}.$$

Since $\beta\delta = \left\lceil \frac{x_2y_2}{x_1} \right\rceil \left\lceil \frac{x_1y_1}{x_2} \right\rceil \ge \frac{x_2y_2}{x_1}\frac{x_1y_1}{x_2} = y_1y_2$, it follows that:

$$\frac{y_1y_2}{\beta(y_2+\delta)} = \frac{1}{\frac{\beta}{y_1}\left(1+\frac{\delta}{y_2}\right)} \le \frac{1}{\frac{\beta}{y_1}+1} = \frac{y_1}{y_1+\beta} = \min\left\{1, \frac{y_1}{y_1+\beta}\right\}.$$

Moreover, since $\alpha\delta = \left\lceil \frac{x_2y_2}{y_1} \right\rceil \left\lceil \frac{x_1y_1}{x_2} \right\rceil \ge y_2x_1$, it follows that $0 \le \frac{\alpha\delta - x_1y_2}{\alpha(y_2+\delta)}$ and (18) is feasible if $\alpha\beta\delta \le \alpha y_1y_2 + \beta x_1y_2$. We consider two cases:

Case 2.1: $x_2 = 1$. In this case, $\alpha = \left\lceil \frac{y_2}{y_1} \right\rceil$, $\beta = \left\lceil \frac{y_2}{x_1} \right\rceil$, and $\delta = x_1y_1$. There exist $f_\alpha, f_\beta \in [0,1)$ such that $\alpha = \frac{y_2}{y_1} + f_\alpha$ and $\beta = \frac{y_2}{x_1} + f_\beta$. We observe that

$$\begin{aligned}
\alpha\beta\delta &= \left(\frac{y_2}{y_1}+f_\alpha\right)\left(\frac{y_2}{x_1}+f_\beta\right)x_1y_1 \\
&= y_1y_2\left(\frac{y_2}{y_1}+f_\alpha\right) + x_1y_2\left(\frac{y_1}{y_2}f_\alpha f_\beta + f_\beta\right) \\
&\le y_1y_2\left(\frac{y_2}{y_1}+f_\alpha\right) + x_1y_2\left(\frac{y_2}{x_1}+f_\beta\right) \\
&= \alpha y_1y_2 + \beta x_1y_2
\end{aligned}$$

where the inequality holds because $x_1 \le y_1 \le y_2$ implies that $x_1y_1f_\alpha f_\beta \le x_1y_1 \le y_2^2$.

Case 2.2: $x_2 \ge 2$. For $(u,v) \in \mathbb{Z}_+^2$, we define $\bar{l}(u,v) = u - l$ where l is the only integer in the interval $\{0,\ldots,v-1\}$ that is such that $u = qv + l$ for some $q \in \mathbb{Z}_+$, *i.e.*, l is the remainder when u is divided by v. Using this notation, it is easy to verify that $\alpha = \frac{x_2y_2+\bar{l}(x_2y_2,y_1)}{y_1}$, $\beta = \frac{x_2y_2+\bar{l}(x_2y_2,x_1)}{x_1}$, and $\delta = \frac{x_1y_1+\bar{l}(x_1y_1,x_2)}{x_2}$. Now observe that:

$$\begin{aligned}
\frac{\delta}{y_2} = \frac{x_1y_1+\bar{l}\,(x_1y_1,x_2)}{x_2y_2} &\le \frac{x_1y_1+x_2-1}{x_2y_2} \\
&= \frac{x_1y_1}{x_2y_2}\left(1+\frac{x_2-1}{x_1y_1}\right) \\
&\le \frac{x_1y_1}{x_2y_2}\left(1+\frac{x_2-1}{x_2+1}\right)
\end{aligned}$$

$$\begin{aligned} &= \frac{1}{x_2 y_2}\left(\frac{x_1 y_1}{1+\frac{1}{x_2}} + \frac{x_1 y_1}{1+\frac{1}{x_2}}\right) \\ &\le \frac{1}{x_2 y_2}\left(\frac{x_1 y_1}{1+\frac{y_1-1}{x_2 y_2}} + \frac{x_1 y_1}{1+\frac{x_1-1}{x_2 y_2}}\right) \\ &\le \frac{x_1 y_1}{x_2 y_2 + \bar{l}\,(x_2 y_2, y_1)} + \frac{x_1 y_1}{x_2 y_2 + \bar{l}\,(x_2 y_2, x_1)} = \frac{x_1}{\alpha} + \frac{y_1}{\beta}, \end{aligned}$$

where the first inequality holds because $\bar{l}\,(x_1 y_1, x_2) \le x_2 - 1$, the second inequality because $x_2 \le x_1 y_1 - 1$, the third inequality holds since $y_1 \le y_2$ implies $\frac{y_1-1}{y_2} \le 1$ and $x_1 \le y_2$ implies that $\frac{x_1-1}{y_2} \le 1$, and the fourth inequality holds since $y_1 - 1 \ge \bar{l}\,(x_2 y_2, y_1)$ and $x_1 - 1 \ge \bar{l}\,(x_2 y_2, x_1)$. Therefore, $\alpha\beta\delta \le \alpha y_1 y_2 + \beta x_1 y_2$.

In summary, for $(x_1, y_1, x_2, y_2) \in B^I$, (17) is feasible, and, therefore, (13) holds for B^I. □

We now apply the result of Proposition 3.2 in conjunction with Theorem 2.1 to obtain the following result that describes the convex hull of (16).

Proposition 3.3. *Let*

$$B^I = \left\{(x,y) \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \;\middle|\; \sum_{i=1}^n x_i y_i \ge r\right\}, \tag{19}$$

where $r > 0$ and, for each $i \in \{1, \ldots, n\}$, define:

$$B_i^I = \left\{(x,y) \in B^I \;\middle|\; (x_j, y_j) = (0,0), \forall j \ne i\right\}.$$

Let the convex hull of B_i^I be represented by:

$$\mathrm{conv}(B_i^I) = \left\{(0,0,x_i,y_i,0,0) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; l^j(x_i,y_i) \ge 1, \forall j \in J\right\},$$

where $l^j(x_i, y_i) = \alpha_j x_i + \beta_j y_i$. Then,

$$\mathrm{conv}(B^I) = \left\{(x,y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^n l^{j_i}(x_i,y_i) \ge 1, \forall j_i \in J\right\}. \tag{20}$$

Proof. We prove this result by applying Theorem 2.1. Let $z_i = (x_i, y_i)$. Assumption (A1) holds by the definition of B_i^I. The convex extension property, (13), follows from a sequential application of Proposition 3.2. Assumption (A3) is satisfied since $\mathrm{conv}\left(B_i^I\right)$ is closed and the functions $l^j(x_i, y_i)$ are positively-homogeneous. Further, since $0^+\left(\mathrm{cl\,conv}\left(B_i^I\right)\right) = \mathbb{R}_+^n \times \mathbb{R}_+^n$, it follows that

$$C_i = \left\{(0,0,x_i,y_i,0,0) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; l^j(x_i,y_i) \ge 0, \forall j \in J\right\} \subseteq 0^+\left(\mathrm{cl\,conv}\left(B_i^I\right)\right).$$

Therefore, (A4) holds. Now, by Theorem 2.1 and the discussion following Definition 3.1, it follows that

$$\mathrm{cl\,conv}(B^I) = X = \left\{(x,y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^n l^{j_i}(x_i,y_i) \ge 1, \forall j_i \in J\right\},$$

where the closure operation is not needed on X since it is a closed set, being an intersection of closed half-spaces. In fact, X is polyhedral, since there are only finitely many half-spaces in its expression. Now, consider the closed sets $B_i^{I'} = \left\{(x,y) \in \mathbb{Z}_+^{2n} \;\middle|\; x_i y_i \ge r\right\}$. Observe that $B_i^I \subseteq B_i^{I'} \subseteq B^I$. Now, by Corollary 9.8.1 in [17], $\mathrm{conv}\left(\bigcup_{i=1}^n B^{I'}\right)$ is closed. Since

$$\mathrm{conv}\left(B^I\right) \subseteq \mathrm{cl\,conv}(B^I) \subseteq \mathrm{cl\,conv}\left(\bigcup_{i=1}^n B_i^{I'}\right) = \mathrm{conv}\left(\bigcup_{i=1}^n B_i^{I'}\right) \subseteq \mathrm{conv}\left(B^I\right),$$

where the second containment holds since $B_i^I \subseteq B_i^{I'}$ and because the discussion following Definition 3.1 argues that $\text{cl}\,\text{conv}(B^I) = \text{cl}\,\text{conv}\left(\bigcup_{i=1}^n B_i^{I'}\right)$, the first equality since $\text{conv}\left(\bigcup_{i=1}^n B_i^{I'}\right)$ is closed, and the third containment since $B_i^{I'} \subseteq B^I$. Therefore, the equality holds throughout, and the result follows. □

Observe that, even though $\text{conv}(B^I)$ is closed, $\text{conv}\left(\bigcup_{i=1}^n B_i^I\right)$ is not closed. Proposition 3.3 shows that $\text{conv}(B^I)$ has exponentially many facets. In particular, if B_i^I has $|J|$ facets, there are $|J|^n$ inequalities in the description of $\text{conv}(B^I)$. We note, however, that separation is not difficult to perform as the coefficients of each pair of variables can be determined independently. Since there is an obvious pseudo-polynomial algorithm to compute the facets of $\text{conv}(B_i^I)$, it is clearly possible to separate the facets of $\text{conv}(B^I)$ in pseudo-polynomial time.

Example 3.4. *Consider the set*

$$B^I = \left\{(x,y) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 \;\middle|\; x_1y_1 + x_2y_2 \geq 10\right\}. \tag{21}$$

It is easily verified that for both $i \in \{1,2\}$

$$\text{conv}\left(B_i^I\right) = \left\{(0,x_i,y_i,0) \in \mathbb{R}_+^4 \;\middle|\; y_i \geq 1, 10x_i + 2y_i \geq 30, x_i + y_i \geq 7, 2x_i + 10y_i \geq 30, x_i \geq 1\right\}.$$

It follows from Proposition 3.3 that the convex hull of B^I has 25 inequalities and is represented by

$$\text{conv}(B^I) = \left\{(x,y) \in \mathbb{R}_+^2 \times \mathbb{R}_+^2 \;\middle|\; \left\{\begin{array}{rr} & y_1 \\ \frac{5}{15}x_1 & +\frac{1}{15}y_1 \\ \frac{1}{7}x_1 & +\frac{1}{7}y_1 \\ \frac{1}{15}x_1 & +\frac{5}{15}y_1 \\ x_1 & \end{array}\right\} + \left\{\begin{array}{rr} & y_2 \\ \frac{5}{15}x_2 & +\frac{1}{15}y_2 \\ \frac{1}{7}x_2 & +\frac{1}{7}y_2 \\ \frac{1}{15}x_2 & +\frac{5}{15}y_2 \\ x_2 & \end{array}\right\} \geq 1\right\}, \tag{22}$$

where each pair of coefficients for (x_1,y_1) can be matched with each pair of coefficients for (x_2,y_2).

Similarly, the convex hull characterization for a variety of bilinear sets can be obtained using the result of Theorem 2.1. In particular, we study now the mixed integer variant. We will study the continuous version in Proposition 3.9.

Proposition 3.5. *Let*

$$B^M = \left\{(x,y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^n a_i x_i y_i \geq r\right\}, \tag{23}$$

where $r > 0$, and, for each $i \in \{1,\dots,n\}$, $a_i > 0$. Define, for each $i \in \{1,\dots,n\}$,

$$B_i^M = \left\{(x,y) \in B^M \;\middle|\; (x_j,y_j) = (0,0), \forall j \neq i\right\}.$$

Let the convex hull of B_i^M be represented by:

$$\text{conv}\left(B_i^M\right) = \left\{(0,0,x_i,y_i,0,0) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; l^j(x_i,y_i) \geq 1, \forall j \in J_i\right\},$$

where $l^j(x_i,y_i) = \alpha_j x_i + \beta_j y_i$. Then,

$$\text{conv}\left(B^M\right) = \left\{(x,y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \;\middle|\; \sum_{i=1}^n l^{j_i}(x_i,y_i) \geq 1, \forall j_i \in J_i\right\}. \tag{24}$$

Proof. Because the verification of the convex extension property is the only technical part of the proof that is significantly different from that of B^I, we only discuss the proof of this property next. Because induction can be used, it suffices to prove the result when $n = 2$. Let $(x_1,y_1,x_2,y_2) \in B^M$.

We show that there exist (i) subsets I and I' of $\{1,2\}$, (ii) for each $i \in I$, a point $\chi_i \in B_i^M$, and (iii) for each $i \in I'$, a ray ψ_i of B_i^M, such that

$$(x_1, y_1, x_2, y_2) = \sum_{i\in I} \lambda_i \chi_i + \sum_{i \in I'} \mu_i \psi_i, \tag{25}$$

where the multipliers satisfy the following conditions: (a) $\sum_{i\in I} \lambda_i = 1$, (b) for all $i \in I$, $\lambda_i \geq 0$, and (c) for all $i \in I'$, $\mu_i \geq 0$.

Note first that, if $x_2 = 0$, it suffices to choose $I = \{1\}$, $I' = \{2\}$, $\chi_1 = (x_1, y_1, 0, 0)$, and $\psi_2 = (0, 0, 0, 1)$ to show that (13) holds. Similarly, if $y_2 = 0$, it suffices to choose $I = \{1\}$, $I' = \{2\}$, $\chi_1 = (x_1, y_1, 0, 0)$, and $\psi_2 = (0, 0, 1, 0)$ to show that (13) holds. We assume without loss of generality that $x_1y_1 \geq x_2y_2$ since the pair of variables (x_1, y_1) and (x_2, y_2) can be interchanged along with their respective coefficients a_1 and a_2. Therefore, in addition to the positivity of x_2 and y_2, we may also assume in the remainder of this proof that $x_1 \geq 1$ and $y_1 > 0$. Define $\chi_1 = \Big(x_1, y_1 + \frac{a_2x_2y_2}{a_1x_1}, 0, 0\Big)$, $\chi_2 = \Big(0, 0, x_2, y_2 + \frac{a_1x_1y_1}{a_2x_2}\Big)$, $\psi_1 = (x_1, 0, 0, 0)$, and $\psi_2 = (0, 0, x_2, 0)$. It can be easily verified that

$$(x_1, y_1, x_2, y_2) = \frac{a_1x_1y_1}{a_1x_1y_1 + a_2x_2y_2}(\chi_1 + \psi_2) + \frac{a_2x_2y_2}{a_1x_1y_1 + a_2x_2y_2}(\chi_2 + \psi_1)$$

which shows that the convex extension property (13) holds. □

Propositions 3.3 and 3.5 illustrate both the fact that the convex extension property used in Theorem 2.1 holds in surprising settings and that this property might not always be trivial to verify. We next present in Theorem 3.6 and Proposition 3.8 conditions under which the convex extension property over orthogonal disjunctive sets can be shown to hold. For example, these conditions are satisfied by bilinear covering sets we discuss later in this paper.

Theorem 3.6. *Consider a function* $g(z_1, \ldots, z_n) : \mathbb{R}_+^{\sum_i d_i} \mapsto \mathbb{R}$, *where* $z_i \in \mathbb{R}_+^{d_i}$, *and the set* $G = \Big\{z \in \mathbb{R}_+^{\sum_i d_i} \;\Big|\; g(z_1, \ldots, z_n) \geq r\Big\}$, *where* $r > 0$. *Let* $G_i = G \cap \Big\{(0, \ldots, 0, z_i, 0, \ldots, 0) \;\Big|\; z_i \in \mathbb{R}_+^{d_i}\Big\}$ *and* $g_i(z_i) = g(0, \ldots, 0, z_i, 0, \ldots, 0)$. *If there exist functions* $h_i : \mathbb{R}_+^{d_i} \mapsto \mathbb{R}^{k_i}$ *and* $f : \mathbb{R}^{\sum_i k_i} \mapsto \mathbb{R}$ *such that:*
(S1) $g(z) \leq f\big(h_1(z_1), \ldots, h_n(z_n)\big)$, *where* f *is a convex function,*
(S2) $f(y_1) > f(y_2)$ *whenever* $y_1 \geq y_2$ *and at least one component of* y_1 *is larger than the corresponding component of* y_2,
(S3) $g_i(z_i) = f(0, \ldots, 0, h_i(z_i), 0, \ldots, 0)$,
(S4) For all i, $h_i(0) = 0$ *and, for* $\lambda \in (0, 1]$, $\lambda h_i\big(\frac{z_i}{\lambda}\big) \geq h_i(z_i)$, *and*
(S5) For all i, $h_i(z_i) \leq 0$ *implies that* $(0, z_i, 0) \in 0^+\big(\operatorname{cl}\operatorname{conv} G_i\big)$,
are satisfied over $\mathbb{R}_+^{\sum_{i=1}^n d_i}$ *then the convex extension property, (13), holds for the set* G. *Assume that, for each* $i \in \{1, \ldots, n\}$, $\operatorname{conv}(G_i)$ *is closed. Define* $G_i' = \operatorname{conv}(G_i) + \sum_{i' \neq i} 0^+(\operatorname{conv} G_{i'})$. *If, for all* i, $G_i' \subseteq \operatorname{conv}(G)$ *then* $\operatorname{conv}(G)$ *is closed.*

Proof. Let $z \in G$ and $y(z) = \big(h_1(z_1), \ldots, h_n(z_n)\big)$. In the following, we sometimes denote $h_i(z_i)$ as $y_i(z)$ to emphasize that it is the i^{th} component of $y(z)$. Let $T = \{i \mid h_i(z_i) \leq 0\}$. Then, by (S5), for each $i \in T$, $(0, z_i, 0) \in 0^+\big(\operatorname{cl}\operatorname{conv} G_i\big)$. If $z - \sum_{i\in T}(0, z_i, 0) \in \operatorname{cl}\operatorname{conv}\big(\bigcup_{i=1}^n G_i\big)$, then so does z. We now show that $z' = z - \sum_{i\in T}(0, z_i, 0) \in \operatorname{cl}\operatorname{conv}\big(\bigcup_{i=1}^n G_i\big)$. Let δ be a subgradient of f at $y(z')$. Then, (S2) implies that $\delta > 0$. Otherwise, suppose that $\delta_i \leq 0$. Let e_i denote the i^{th} unit vector and choose $\epsilon > 0$. Observe that $f(y(z') - \epsilon e_i) \geq f(y(z')) - \epsilon\langle \delta, e_i\rangle = f\big(y(z')\big) - \epsilon\delta_i \geq f\big(y(z')\big)$, a contradiction to (S2). Clearly, for each $i \notin T$, $h_i(z_i') = h_i(z_i)$. By construction, for each $i \in T$, $z_i' = 0$ and, therefore, $h_i(z_i') = 0 \geq h_i(z_i)$. In other words, $y(z') = \max\{y(z), 0\}$. Observe that (S1) and (S2) together imply that $f(y(z')) \geq f(y(z)) \geq g(z) \geq r$.

First, consider the case where $\langle \delta, y(z')\rangle = 0$. Then, $y(z') = 0$. Consider any i and a $\bar{z}_i \in \mathbb{R}_+^{d_i}$. On the one hand, assume that $h_i(\bar{z}_i) > 0$. Then, $g(0, \bar{z}_i, 0) = g_i(\bar{z}_i) = f(0, h_i(\bar{z}_i), 0) > f(0) = f(y(z')) \geq r$, where the first equality follows from the definition of g_i, the second from (S3), and the first inequality from (S2) and $h_i(\bar{z}_i) > 0$. Therefore, $(0, \bar{z}_i, 0) \in G \subseteq \operatorname{cl}\operatorname{conv}(G)$. On the other hand, assume that $h_i(\bar{z}_i) \le 0$. Then, by (S5), we know that $(0, \bar{z}_i, 0) \in 0^+\big(\mathrm{cl\,conv}\, G_i\big)$. Since $g_i(0) = f(0) = f(y(z')) \ge r$, it follows that $0 \in G_i$. Combining $0 \in G_i$ and $(0, \bar{z}_i, 0) \in 0^+\big(\mathrm{cl\,conv}\, G_i\big)$, we can conclude that $(0, \bar{z}_i, 0) \in \mathrm{cl\,conv}(G_i) \subseteq \mathrm{cl\,conv}(G)$. In other words, regardless of the sign of $h_i(z_i)$, it follows that $(0, \bar{z}_i, 0) \in \mathrm{cl\,conv}(G)$. Since $\bar{z}_i$ was arbitrarily chosen in $\mathbb{R}^{d_i}_+$, it follows that $\mathbb{R}^{\sum_i d_i}_+ \subseteq \mathrm{cl\,conv}\left(\bigcup_{i=1}^n G_i\right) \subseteq \mathrm{cl\,conv}(G) \subseteq \mathbb{R}^{\sum_i d_i}_+$. Since equality holds throughout, (14), or equivalently (13) holds for G.

Now, consider the case when $\langle \delta, y(z') \rangle > 0$. Define $\lambda_i = \frac{\delta_i y_i(z')}{\langle \delta, y(z') \rangle}$. Since δ_i and $y_i(z')$ are non-negative, it follows that $\lambda_i \ge 0$ and $\sum_{i=1}^n \lambda_i = 1$. Define $I = \big\{ i \mid \lambda_i > 0 \big\}$ and observe that $|I| \ge 1$. The following chain of implication holds

$$i \notin I \Rightarrow y_i(z') = 0 \Rightarrow i \in T \Rightarrow z'_i = 0,$$

where the first implication follows since $\delta_i > 0$; the second because, for each $i \notin T$, $y_i(z') > 0$; and the third by the construction of z'. Therefore, $z' = \sum_{i \in I} z''_i$, where $z''_i = \big(0, \ldots, 0, z'_i, 0, \ldots, 0\big)$. For each $i \in I$, let $\chi_i = \frac{z''_i}{\lambda_i}$. Observe that $z' = \sum_{i \in I} \lambda_i \chi_i$, *i.e.*, z' can be expressed as a convex combination of χ_i for $i \in I$. The following shows that, for all $i \in I$, $\chi_i \in G_i$:

$$\begin{aligned} g(\chi_i) &= g_i\left(\frac{z'_i}{\lambda_i}\right) = f\big(y(\chi_i)\big) \ge f\left(\frac{1}{\lambda_i} y(z''_i)\right) \ge f\big(y(z')\big) + \delta_i \frac{\langle \delta, y(z') \rangle}{\delta_i y_i(z')} y_i(z'') - \sum_{j=1}^n \delta_j y_j(z') \\ &= f\big(y(z')\big) + \delta_i \frac{\langle \delta, y(z') \rangle}{\delta_i y_i(z')} y_i(z') - \sum_{j=1}^n \delta_j y_j(z') = f\big(y(z')\big) \ge r. \end{aligned}$$

The first equality follows from the definition of g_i, the second equality from (S3), the first inequality follows since f is non-decreasing by (S2) and $h_i\big(\frac{z'_i}{\lambda_i}\big) \ge \frac{1}{\lambda_i} h_i(z'_i)$, the second inequality because δ is a subgradient of f at $y(z')$, and the third equality because $y_i(z'') = h_i(z'_i) = y_i(z')$. Since $z = z' + \sum_{i \in T}(0, z_i, 0)$, where, for each $i \in T$, $(0, z_i, 0) \in 0^+\left(\mathrm{cl\,conv}(G_i)\right)$ it follows that (13) holds for G.

We now prove the last statement of the theorem. Consider an arbitrary $i \in N$. Clearly, G'_i, as defined in the statement of the theorem, is convex. We argue that it is closed as well. By Corollary 9.1.1 in [17], G'_i is closed if there do not exist $(0, z_i, 0) \in \mathrm{conv}(G_i)$ and, for $i' \in N \backslash \{i\}$, $(0, z_{i'}, 0) \in 0^+\left(\mathrm{conv}\, G_{i'}\right)$, not all zero, such that $\sum_{i=1}^n (0, z_i, 0) = 0$. But, the orthogonal vectors $(0, z_i, 0)$ sum to zero if and only if each of the vectors is zero. Therefore, G'_i is closed. Again by Corollary 9.1.1 in [17], $0^+(G'_i) = \sum_{i=1}^n 0^+\left(\mathrm{conv}\, G_i\right)$. Since the recession directions of G'_i are independent of i, it follows by Corollary 9.8.1 in [17] that $\mathrm{conv}\left(\bigcup_{i=1}^n G_i\right)$ is closed. Now,

$$\mathrm{conv}(G) \subseteq \mathrm{cl\,conv}(G) = \mathrm{cl\,conv}\left(\bigcup_{i=1}^n G_i\right) \subseteq \mathrm{cl\,conv}\left(\bigcup_{i=1}^n G'_i\right) = \mathrm{conv}\left(\bigcup_{i=1}^n G'_i\right) \subseteq \mathrm{conv}(G),$$

where the first equality follows from the equivalence of (13) and (14), the second containment follows since $G_i \subseteq G'_i$, the second equality follows since $\mathrm{conv}\left(\bigcup_{i=1}^n G'_i\right)$ is closed and the third containment follows since $G'_i \subseteq \mathrm{conv}(G)$. □

The main challenge in applying Theorem 3.6 in practical situations is verifying Assumption (S4). However, when $h_i(z_i)$ is derived from other functions using operations such as summations, minimizations, or maximizations, then (S4) can often be established easily by studying the same properties for the functions used in the derivation of $h_i(z_i)$. To see this, first note that the assumption is satisfied trivially by any linear function. If $h(z) = w\left(p_1(z), \ldots, p_K(z)\right)$, for all $k \in \{1, \ldots, K\}$, $p_k(z)$ satisfies (S4), w satisfies (S4), w is isotonic, *i.e.*, $w(y_1) \ge w(y_2)$ if $y_1 \ge y_2$, and $w(0, \ldots, 0) = 0$, then $h(z)$ satisfies (S4) as well. Clearly, $h(0) = w(p_1(0), \ldots, p_k(0)) = w(0, \ldots, 0) = 0$ and:

$$\lambda h\left(\frac{z}{\lambda}\right) = \lambda w\left(p_1\left(\frac{z}{\lambda}\right), \ldots, p_k\left(\frac{z}{\lambda}\right)\right) \ge \lambda w\left(\frac{1}{\lambda} p_1(z), \ldots, \frac{1}{\lambda} p_k(z)\right) \ge w\left(p_1(z), \ldots, p_k(z)\right) = h(z),$$

where the first inequality follows since w is isotonic and $p_k(z)$ obeys (S4); and the second inequality because w obeys (S4). If w satisfies (S4) only over the non-negative orthant, then $p_k(z)$ must be non-negative as well. In particular, $\sum_{k=1}^{K} p_k(z)$ satisfies the assumption as long as, for all k, $p_k(z)$ satisfies the assumption. Now, consider $h(z) = \operatorname{op}_y p(y,z)$, where op is an operator such as min or max that satisfies $\operatorname{op}_y f_1(y) \geq \operatorname{op}_y f_2(y)$ if, for all y, $f_1(y) \geq f_2(y)$ and $\lambda \operatorname{op}_y f(y) \geq \operatorname{op}_y \lambda f(y)$ for $\lambda \in (0,1]$. In addition, assume that $\lambda p\left(y, \frac{z}{\lambda}\right) \geq p(y,z)$ for $\lambda \in (0,1]$. Then,

$$\lambda h\left(\frac{z}{\lambda}\right) = \lambda \operatorname*{op}_y p\left(y, \frac{z}{\lambda}\right) \geq \lambda \operatorname*{op}_y \frac{1}{\lambda} p(y,z) \geq \operatorname*{op}_y p(y,z) = h(z),$$

for $\lambda \in (0,1]$. In particular, if $h(z) = \min\bigl(p_1(z), \ldots, p_K(z)\bigr)$ and, for all $\lambda \in (0,1]$, $p_k(z) \leq \lambda p_k\left(\frac{z}{\lambda}\right)$ then $h(z) \leq \lambda h\left(\frac{z}{\lambda}\right)$.

The following corollary of Theorem 3.6 discusses the case where f is the summation operator and $h_i(z_i) = g_i(z_i)$. Such a setup can be used to show that convex extensions property holds for many bilinear covering sets. In addition, we also prove that $\operatorname{conv}(G)$ is closed if the function $g(\cdot)$ eventually increases in each one of the principal directions of the non-negative orthant.

Corollary 3.7. *Consider a function* $g(z_1, \ldots, z_n) : \mathbb{R}_+^{\sum_i d_i} \mapsto \mathbb{R}$, *where* $z_i \in \mathbb{R}_+^{d_i}$, *and the set* $G = \left\{z \in \mathbb{R}_+^{\sum_i d_i} \mid g(z_1, \ldots, z_n) \geq r\right\}$, *where* $r > 0$. *Let* $G_i = G \cap \left\{(0, \ldots, 0, z_i, 0, \ldots, 0) \;\middle|\; z_i \in \mathbb{R}_+^{d_i}\right\}$ *and* $g_i(z_i) = g(0, \ldots, 0, z_i, 0, \ldots, 0)$. *If*
(B1) $g(z) \leq \sum_{i=1}^{n} g_i(z_i)$,
(B2) For all i, $g_i(0) = 0$ *and, for* $\lambda \in (0,1]$, $\lambda g_i\bigl(\frac{z_i}{\lambda}\bigr) \geq g_i(z_i)$, *and*
(B3) For all i, $g_i(z_i) \leq 0$ *implies that* $(0, z_i, 0) \in 0^+\bigl(\operatorname{cl}\operatorname{conv} G_i\bigr)$,
are satisfied over $\mathbb{R}_+^{\sum_{i=1}^n d_i}$ *then the convex extension property, (13), holds for the set* G. *Let* e_i^d *be the* d^{th} *principal axis in the space of* z_i *variables. Assume that, for all* i, $\operatorname{conv}(G_i)$ *is closed. Assume further that there exists a* $\gamma \geq 1$ *such that, for all* $\gamma' \geq \gamma$, $i \in N$, $d \in \{1, \ldots, d_i\}$, *and* $z \geq 0$, *it holds that* $g\left(z + \gamma' e_i^d\right) \geq g(z)$. *Then,* $\operatorname{conv}(G)$ *is closed.*

Proof. Choose f to be the summation operator and $h_i(z_i) = g_i(z_i)$. Then, the first part of the result follows from Theorem 3.6. The rest of the result follows if G_i', as defined in the statement of Theorem 3.6, is contained in $\operatorname{conv}(G)$. Consider a $\bar{z}$ which can be expressed as $\bar{z}_i + \sum_{i' \neq i}(0, \bar{z}_{i'}, 0)$, where $\bar{z}_i \in \operatorname{conv}(G_i)$ and for all $i' \neq i$, $\bar{z}_{i'} \geq 0$. By Caratheodory's theorem, there exist, for $d \in \{1, \ldots, d_i + 1\}$, $\tilde{z}_d$ and $\lambda_d \geq 0$, such that $\sum_{d=1}^{d_i+1} \lambda_d \tilde{z}^d = \bar{z}_i$, $\sum_{d=1}^{d_i+1} \lambda_d = 1$, and $\tilde{z}^d \in G_i$ for all d. Let $D = \sum_{i' \neq i} d_{i'}$. Then, define $m = \min\{z_{i'd} D \mid i' \neq i, d = 1, \ldots, d_{i'}, z_{i'd} > 0\}$. For each $i' \neq i$ and $d' \in \{1, \ldots, d_{i'}\}$, define $\tilde{z}_{i'}^{dd'} = \tilde{z}^d + D\gamma \frac{z_{i'd'}}{m} e_{i'}^{d'}$. On one hand, for all (i', d') with $z_{i'd'} > 0$, $D\frac{z_{i'd'}}{m} \geq 1$. Therefore, it follows that $g\left(\tilde{z}_{i'}^{dd'}\right) \geq g\left(\tilde{z}^d\right) \geq r$. On the other hand, if $z_{i'd'} = 0$ then $\tilde{z}_{i'}^{dd'} = \tilde{z}^d \in G$. It follows that $\tilde{z}_{i'}^{dd'} \in G$ for all (i', d, d'). Therefore, $\bar{z}$ can be written as a convex combination of points in G as follows:

$$\sum_{d=1}^{d_i+1} \left(\lambda_d \frac{\gamma - 1}{\gamma} \tilde{z}^d + \lambda_d \frac{1}{D\gamma} \sum_{i' \neq i} \sum_{d'=1}^{d_{i'}} \tilde{z}_{i'}^{dd'} \right) = \sum_{d=1}^{d_i+1} \lambda_d \tilde{z}^d + \sum_{d=1}^{d_i+1} \lambda_d \sum_{i' \neq i} \sum_{d'=1}^{d_{i'}} \bar{z}_{i'd'} e_{i'}^{d'} = \bar{z}_i + \sum_{i' \neq i} (0, \bar{z}_{i'}, 0).$$

Observe that the multipliers are non-negative since $\gamma \geq 1$ and

$$\sum_{d=1}^{d_i+1} \lambda_d \left(\frac{\gamma - 1}{\gamma} + \frac{1}{D\gamma} \sum_{i' \neq i} \sum_{d=1}^{d_{i'}} 1 \right) = 1.$$

Therefore, the result follows. ∎

Theorem 2.1 also points to an interesting set of sufficient conditions that can be used to verify the convex extension property. The primary difference from the conditions in Theorem 3.6 is that Proposition 3.8 does not impose a structure on the original set S. Instead, it constructs a set X whose projection in the z-space is contained within $\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^{n} S_i\right)$, using a construction similar to Theorem 2.1, and then leaves it to the user to verify that X outerapproximates S. This technique may be useful when S is defined by more than one inequality. Also, note that the special case of Theorem 3.6, discussed in Corollary 3.7, also follows from Proposition 3.8.

Proposition 3.8. *For a set S and its subsets $S_i \subseteq S$ for $i \in N = \{1, \dots, n\}$, let $z_i \in \mathbb{R}^{d_i}$ and $z = (z_1, \dots, z_i, \dots, z_n) \in S \subseteq \mathbb{R}^{\sum_i d_i}$. Assume that (A1) and (A4) are satisfied as in Theorem 2.1 and the sets A_i and X are as defined in (1) and (2) respectively. If, in addition, the following assumptions are satisfied:*
(N1) $S_i \subseteq \operatorname{proj}_z A_i \subseteq \operatorname{cl}\big(\operatorname{conv}(S_i)\big)$,
(N2) $t_i^{j_i}$, $v_i^{k_i}$, and $w_i^{l_i}$ are such that for all $0 < \lambda \leq 1$,

$$\lambda t_i^{j_i}\left(\frac{z_i, u_i}{\lambda}\right) \geq t_i^{j_i}(z_i, u_i),\ \lambda v_i^{k_i}\left(\frac{z_i, u_i}{\lambda}\right) \geq v_i^{k_i}(z_i, u_i),\ \lambda w_i^{l_i}\left(\frac{z_i, u_i}{\lambda}\right) \geq w_i^{l_i}(z_i, u_i),$$

(N3) $S \subseteq \operatorname{cl}\operatorname{proj}_z X$.
Then, (13) holds for S.

Proof. Here, Fourier-Motzkin elimination shows, as it did in the proof of Theorem 2.1, that $X = \operatorname{proj}_{z,u} Q$. We will now show that $\operatorname{proj}_z X = \operatorname{proj}_z Q \subseteq \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right)$. The proof is again similar to that for Theorem 2.1 except that the positive homogeneity is replaced by the weaker inequalities assumed in (N2). Even then, if $(\lambda, z, u) \in Q$ and $0 < \lambda_i \leq 1$, it follows that $\frac{z_i, u_i}{\lambda_i} \in R_i(1)$ since the inequalities are satisfied in the same manner as:

$$t_i^{j_i}(z_i, u_i) \geq \lambda_i \text{ and } \lambda_i t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq t_i^{j_i}(z_i, u_i) \Rightarrow \lambda_i t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq t_i^{j_i}(z_i, u_i) \geq \lambda_i \Rightarrow t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right) \geq 1.$$

Clearly, $\operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right) \subseteq \operatorname{cl}\operatorname{conv}(S)$ and we have assumed that $S \subseteq \operatorname{proj}_z X$. Observe that $\operatorname{cl}\operatorname{conv}(S) \subseteq \operatorname{cl}\operatorname{conv}(\operatorname{proj}_z X) \subseteq \operatorname{cl}\operatorname{conv}\left(\bigcup_{i=1}^n S_i\right) \subseteq \operatorname{cl}\operatorname{conv}(S)$ and, therefore, equality holds throughout. □

Observe that Assumptions (N1) and (N2) are less restrictive than (A3) in Theorem 2.1 since $\operatorname{proj}_z A_i$ may be a nonconvex subset of $\operatorname{conv}(S_i)$ and the positive homogeneity is relaxed. Here, it is not necessary to use $t_i^{j_i}(z_i, u_i)$, $v_i^{k_i}(z_i, u_i)$ and $w_i^{l_i}(z_i, u_i)$ as the underestimators in Assumption (N2). Rather, any function of (z_i, u_i) that underestimates $\lambda_i t_i^{j_i}\left(\frac{z_i, u_i}{\lambda_i}\right)$, $\lambda_i v_i^{k_i}\left(\frac{z_i, u_i}{\lambda_i}\right)$, and $\lambda_i w_i^{l_i}\left(\frac{z_i, u_i}{\lambda_i}\right)$ for all $\lambda_i \in (0, 1]$ suffices. As long as the set C_i defined using these functions inner-approximates the recession cone of $\operatorname{cl}\operatorname{conv}(S)$, a suitable set X can be derived by projecting out the λ variables and Assumption (N3) can be posed in terms of this set. Instead of exploring this extension further, we will retain in the remainder of this paper that $t_i^{j_i}(z_i, u_i)$, $v_i^{k_i}(z_i, u_i)$, and $w_i^{l_i}(z_i, u_i)$ are themselves the underestimating functions since it keeps the notation simpler while still conveys the main ideas.

We now discuss the application of Theorem 3.6 to convexifying bilinear covering sets. The bilinear covering sets we consider generalize the bilinear set discussion in Proposition 2.9. In fact, the bilinear covering set reduces to Q, as defined in (11) when restricted to any one of n orthogonal subspaces. As long as the convex extension property holds, since Proposition 2.9 provides the defining inequality for the convex hull in each of the orthogonal subspaces, we can use Theorem 2.1 to find the convex hull description of the bilinear covering set over the non-negative orthant. We formalize this argument in the following proposition.

Proposition 3.9. *Consider a bilinear covering set:*

$$B^R = \left\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^n (a_i x_i y_i + b_i x_i + c_i y_i) \geq r\right\}.$$

where, for each $i \in \{1, \dots, n\}$, a_i, b_i and c_i are non-negative and r is strictly positive. Let

$$\eta_i(x_i, y_i) = \frac{1}{2}\left(b_i x_i + c_i y_i + \sqrt{(b_i x_i + c_i y_i)^2 + 4 a_i r x_i y_i}\right).$$

Then,

$$\operatorname{conv}\big(B^R\big) = X = \left\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^n \eta_i(x_i, y_i) \geq r\right\}. \tag{26}$$

Proof. We may assume without loss of generality that, for each i, at least one of a_i, b_i, or c_i is positive. First, we use Corollary 3.7 to show that the convex extension property (13) holds for B^R. Let $z_i = (x_i, y_i)$ and $g_i(z_i) = a_i x_i y_i + b_i x_i + c_i y_i$. Clearly, $g_i(0) = 0$ and for $0 < \lambda \leq 1$,

$$\lambda g_i\left(\frac{z_i}{\lambda}\right) = \frac{a_i x_i y_i}{\lambda} + b_i x_i + c_i y_i \geq g_i(z_i).$$

Therefore, Assumption (B2) is satisfied. Let $B_i^R = \left\{(0, x_i, y_i, 0) \in \mathbb{R}_+^{2n} \;\middle|\; g_i(x_i, y_i) \geq 0\right\}$. Observe that, if $(x_i', y_i') \geq 0$, then $g_i(x_i + x_i', y_i + y_i') \geq g_i(x_i, y_i)$. Therefore, if $z_i' = (x_i', y_i') \geq 0$ then $(0, z_i', 0) \in 0^+\big(\operatorname{cl}\operatorname{conv} B_i^R\big)$ and, consequently, Assumption (B3) is satisfied. It follows that the convex extension property holds for B^R. In fact, since g is non-decreasing and $\operatorname{cl}\operatorname{conv}(B_i^R) = B_i^R$ it follows from the last statement of Corollary 3.7 that $\operatorname{conv}(B_i^R)$ is closed as well. By Proposition 2.9, it follows that the convex hull of B_i^R is defined by $\eta_i(x_i, y_i) \geq r$. Observe that $\eta_i(x_i, y_i)$ is a positively-homogenous function. Therefore, Assumption (A3) is satisfied. Finally, $\eta_i(x_i, y_i)$ is concave by Proposition 2.6 and since for sufficiently large z_i, $g_i(x_i, y_i) \geq r$, it follows that $B_i^R \neq \emptyset$ and, therefore, by Proposition 2.6 that Assumption (A4) is satisfied as well. Then, by Theorem 2.1 and the discussion following Definition 3.1, the set X in (26) is $\operatorname{cl}\operatorname{conv}(B^R)$. But, as argued earlier, $\operatorname{cl}\operatorname{conv}(B^R) = \operatorname{conv}(B^R)$, and the result follows. □

Consider the special case of Proposition 3.9 where $b_i = c_i = 0$. In this case, the convex hull inequality takes the following simple form: $\sum_{i=1}^n \sqrt{a_i x_i y_i} \geq \sqrt{r}$. First, the validity of the inequality can be verified using the following argument:

$$\sum_{i=1}^n \sqrt{a_i x_i y_i} \geq \sqrt{\sum_{i=1}^n a_i x_i y_i} \geq \sqrt{r},$$

where the first inequality follows by the subadditivity of square-root over non-negative real numbers. Second, by Example 2.5, the above inequality defines the closure convex hull of the disjunctive union of $\{(x_i, y_i) \mid a_i x_i y_i \geq r\}$ over the non-negative orthant and, therefore, it must also be the closure convex hull of $\sum_{i=1}^n a_i x_i y_i \geq r$ over the same set. Note that in the argument, we did not employ Theorem 3.6. Instead, we replaced it with a proof that the convex hull of the disjunctive union of orthogonal restrictions of the set includes the original set. This illustrates a different technique, similar to the proof technique of Proposition 3.8, that may sometimes be useful in establishing the convex extension property.

However, the above technique for establishing validity fails for another special case of Proposition 3.9, where the defining inequality is $ax_1y_1 + bx_2 \geq r$ with $a > 0$, $b > 0$, and $r > 0$. A simpler variant of this set was mentioned in the introduction of the paper. By Proposition 3.9, its convex hull over the non-negative orthant is defined by

$$\sqrt{\frac{ax_1y_1}{r}} + \frac{bx_2}{r} \geq 1. \tag{27}$$

Note that the right-hand-side r participates differently with different subsets of variables in this convex hull inequality. One could use subadditivity of the square-root function to instead derive the following valid inequality

$$\sqrt{\frac{ax_1y_1}{r}} + \sqrt{\frac{bx_2}{r}} \geq 1. \tag{28}$$

However, as expected, (28) is not as tight as (27). This can be seen by considering a point (x_1, y_1, x_2) that is feasible to (27). If $\frac{bx_2}{r} \geq 1$, it follows that $\sqrt{\frac{bx_2}{r}} \geq 1$. Otherwise, $\frac{bx_2}{r} < 1$, in which case

$$\sqrt{\frac{ax_1y_1}{r}} + \sqrt{\frac{bx_2}{r}} > \sqrt{\frac{ax_1y_1}{r}} + \frac{bx_2}{r} \geq 1.$$

Therefore, (x_1, y_1, x_2) is feasible to (28) as well. Observe that the subadditivity of the square-root function is not sufficient to prove the convex extension property for this bilinear covering set, and, thus, cannot replace Theorem 3.6. Without realizing the convex extension property *a priori*, even the form of the inequality (27) is not obvious. The key to deriving this convex hull is thus to realize that the convex hull is formed by restricting attention to orthogonal subspaces. The first subspace spans the (x_1, y_1) variables and the second subspace spans x_2. Then, Theorem 2.1 quickly reveals the structure of the convex hull. Note that for this example, $\sqrt{\frac{bx_2}{r}} \geq 1$ as well as $\frac{bx_2}{r} \geq 1$ define the convex hull of the set restricted to $(0, 0, x_2)$. However, as the insight from Theorem 2.1 suggests, it is preferable to choose the latter representation since it uses a positively-homogenous function.

The construction of Proposition 3.9 can be carried out as long as it is possible to invoke Theorem 3.6 to establish the convex extension property and Theorem 2.1 to convexify the orthogonal disjunctions. This idea can be exploited to develop tighter relaxations when the variables are restricted to belong to the hypercube by suitably altering the inequality outside the hypercube so that Theorem 3.6 can still be invoked. This technique for deriving relaxations will be pursued in future research.

4 Conclusions

In this paper, we developed a convexification tool for orthogonal disjunctions that does not introduce new variables. As an application, we provided a simple derivation of intersection cuts for mixed-integer polyhedral sets. The convexification tool was also shown to be useful in deriving cuts for a variety of nonconvex constraints; those that satisfy a key convex extension property. Verifying the convex extension property can be an arduous task. To address this difficulty, we provided a general set of conditions that are sufficient to establish the convex extension property. We used the convexification tool to find the convex hull representation for a bilinear covering set. Future work will concentrate on applying these results to other classes of problems, and on incorporating the findings in relaxation constructors within a branch-and-bound algorithm for global optimization.

References

[1] C. S. Adjiman, I. P. Androulakis, and C. A. Floudas. A global optimization method, αBB, for general twice-differentiable constrained NLPs–II. Implementation and computational results. *Computers & Chemical Engineering*, 22:1159–1179, 1998.

[2] A. Atamturk. Strong formulations of robust mixed 0-1 programming. *Mathematical Programming*, pages 235–250, 2006.

[3] E. Balas. Disjunctive programming: Cutting planes from logical conditions. In O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, editors, *Nonlinear Programming*, pages 279–312. Academic Press, NY, 1975.

[4] E. Balas. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89(1-3):3–44, 1998. Original manuscript was published as a technical report in 1974.

[5] E. Balas. Intersection cuts - a new type of cutting planes for integer programming. *Operations Research*, 19:19–39, 1971.

[6] E. Balas and M. Perregaard. A precise correspondence between lift-and-project cuts, simple disjunctive cuts, and mixed-integer gomory cuts for 0-1 programming. *Mathematical Programming*, 94:221–245, 2003.

[7] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Waechter. Branching and bounds tightening techniques for non-convex MINLP. Available at `http://www.optimization-online.org/DB_HTML/2008/08/2059.html`, 2008.

[8] C. Bliek, C. Jermann, and A. Neumaier, editors. *Global Optimization and Constraint Satisfaction*, Fifth Annual Workshop on Global Constraint Optimization and Constraint Satisfaction, COCOS, 2002. Springer.

[9] S. Ceria and J. Soares. Convex programming for disjunctive convex optimization. *Mathematical Programming*, 86A:595–614, 1999.

[10] G. Cornuéjols and C. Lemaréchal. A convex-analysis perspective on disjunctive cuts. *Mathematical Programming*, 106:567–586, 2006.

[11] J. E. Falk and R. M. Soland. An algorithm for separable nonconvex programming problems. *Management Science*, 15:550–569, 1969.

[12] I. Harjunkoski, T. Westerlund, R. Porn, and H. Skrifvars. Different transformations for solving non-convex trim-loss problems by MINLP. *European Journal of Operational Research*, 105: 594–603, 1998.

[13] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer Verlag, Berlin, Third edition, 1996.

[14] LINDO Systems Inc. LINGO 11.0 optimization modeling software for linear, nonlinear, and integer programming. Available at `http://www.lindo.com`, 2008.

[15] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I - convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976.

[16] A. Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica*, 13:271–369, 2004.

[17] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[18] N. V. Sahinidis and M. Tawarmalani. *BARON*. The Optimization Firm, LLC, Urbana-Champaign, IL, 2005. Available at `http://www.gams.com/dd/docs/solvers/baron.pdf`.

[19] N. W. Sawaya and I. E. Grossmann. A cutting plane method for solving linear generalized disjunctive programming problems. *Computers & Chemical Engineering*, 29:1891–1913, 2005.

[20] S. Sherali and S. Sen. Cuts from combinatorial disjunctions. *Operations Research*, pages 928–933, 1985.

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

[22] M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*. Kluwer, Dordrecht, The Netherlands, 2002.

[23] M. Tawarmalani and N. V. Sahinidis. Convex extensions and envelopes of lower semi-continuous functions. *Mathematical Programming*, 93:247–263, 2002.

[24] M. Tawarmalani and N. V. Sahinidis. Semidefinite relaxations of fractional programs via novel techniques for constructing convex envelopes of nonlinear functions. *Journal of Global Optimization*, 20:137–158, 2001.

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

[26] G. M. Ziegler. *Lectures on Polytopes*. Springer, NY, 1998.