

Iowa State University

From the Selected Works of Danial Davarnia

August, 2017

Simultaneous convexification of bilinear functions over polytopes with application to network interdiction

Danial Davarnia, *University of Florida*

Jean-Philippe Richard, *University of Florida*

Mohit Tawarmalani, *Purdue University*



Available at: <https://works.bepress.com/danial-davarnia/2/>

SIMULTANEOUS CONVEXIFICATION OF BILINEAR FUNCTIONS OVER POLYTOPES WITH APPLICATION TO NETWORK INTERDICTION*

DANIAL DAVARNIA[†], JEAN-PHILIPPE P. RICHARD[†], AND MOHIT TAWARMALANI[‡]

Abstract. We study the simultaneous convexification of graphs of bilinear functions $g^k(\mathbf{x}; \mathbf{y}) = \mathbf{y}^\top A^k \mathbf{x}$ over $\mathbf{x} \in \Xi = \{\mathbf{x} \in [0, 1]^n \mid E\mathbf{x} \geq \mathbf{f}\}$ and $\mathbf{y} \in \Delta_m = \{\mathbf{y} \in \mathbb{R}_+^m \mid \mathbf{1}^\top \mathbf{y} \leq 1\}$. We propose a constructive procedure to obtain a linear description of the convex hull of the resulting set. This procedure can be applied to derive convex and concave envelopes of certain bilinear functions, to study unary expansions of integer variables in mixed integer bilinear sets, and to obtain convex hulls of sets with complementarity constraints. Exploiting the structure of Ξ , the procedure naturally yields stronger linearizations for bilinear terms in a variety of practical settings. In particular, we demonstrate the effectiveness of the approach by strengthening the traditional dual formulation of network interdiction problems and report encouraging preliminary numerical results.

Key words. bilinear functions, envelopes, convex hulls, cutting planes, network interdiction

AMS subject classifications. 52A27, 90C26, 90C30, 90C35, 90C57, 90C90, 91A43

DOI. 10.1137/16M1066166

1. Introduction. Bilinear functions occur in the modeling of various practical problems in engineering and management. Pooling is one such example that encompasses several applications in chemical engineering, including the operation of refinery processes and wastewater treatment; see [47] for a detailed description of applications, and see [7, 31] for surveys. Bidimensional packing problems, which play important roles in industrial cutting and packing applications, can also be modeled as bilinear programs (BLPs); see [14]. Similarly, BLPs find applications in the paper industry through trim-loss problems [20]. Other areas of application for BLPs can be found in sports [37], modular designs [2], supply chain management [33], and sharp separation process in networks with multicomponent streams [36]. Additional applications of mixed integer BLPs in various production, location-allocation, and product distribution problems can be found in [1].

Network interdiction problems (NIPs) are Stackelberg games played on a network where an interdictor (leader) seeks to impair to the greatest extent possible the operation of the network otherwise controlled by an interdictee (follower). NIPs have applications in homeland security [21], health care [6], and border control [32]. Interdiction problems defined on general networks are often formulated as linear bilevel programs [11, 24] and are known to be strongly NP-hard; see [17]. In the literature, these problems are often reformulated as single level mathematical programs by taking the dual of the follower problem; see [41] for instance. The resulting model is a mixed integer nonlinear program (MINLP) with linear constraints and a bilinear objective function that is often reformulated as a mixed integer linear program (MILP) using McCormick envelopes.

*Received by the editors March 16, 2016; accepted for publication (in revised form) May 3, 2017; published electronically August 17, 2017.

<http://www.siam.org/journals/siopt/27-3/M106616.html>

Funding: This work was supported by NSF CMMI grants 1234897 and 1235236.

[†]Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611-6595 (d.davarnia@ufl.edu, richard@ise.ufl.edu).

[‡]Krannert School of Management, Purdue University, West Lafayette, IN 47907-2076 (mtawarma@purdue.edu).

In this paper, we study the simultaneous convexification of graphs of bilinear functions $g^k(\mathbf{x}; \mathbf{y}) = \mathbf{y}^\top A^k \mathbf{x}$ over $\mathbf{x} \in \Xi = \{\mathbf{x} \in [0, 1]^n \mid E\mathbf{x} \geq \mathbf{f}\}$ and $\mathbf{y} \in \Delta_m = \{\mathbf{y} \in \mathbb{R}_+^m \mid \mathbf{1}^\top \mathbf{y} \leq 1\}$, where bold lowercase letters denote vectors. Although a description of the convex hull of the corresponding set can be derived using a specialization of the reformulation-linearization technique [38, 40] or disjunctive programming [9], these approaches necessitate the addition of new variables. We introduce a constructive procedure to obtain a convex hull description in the original space of variables. Developing techniques to obtain convex relations for MINLPs in the space of original variables is an active vein of research; see [13, 18, 22, 43, 45] and the references therein for various examples of these techniques. One of the advantages of our approach is that it provides valuable insight into the structure of strong inequalities in the space of original variables. For instance, it allows us to improve the traditional linearization approach for single bilinear terms that is often used in NIPs. Further, it leads us to develop inequalities that strengthen this improved linearization even further by considering multiple bilinear terms simultaneously.

The problem of finding relaxation procedures for bilinear terms and functions has been investigated in the past. In fact, the derivation of such relaxations is at the core of the development of branch-and-bound algorithms for global optimization; see [46] for an exposition. Since branch-and-bound algorithms require the construction of convex relaxations over successively refined partitions of the feasible region, the domain over which bilinear functions are defined is of particular significance. In this context, a classical method to obtain a linear programming relaxation of a BLP is to add a new variable z for each bilinear term xy and then relax the requirement $z = xy$. McCormick [28] develops a polyhedral relaxation for the set defined by $z = xy$ when variables (x, y) are constrained to a box. Al-Khayyal and Falk [3] show that this relaxation is in fact the convex hull of the set. Meyer and Floudas [29] generalize these results for trilinear monomials over a box domain. In [30], the same authors study properties of convex envelopes of edge-concave functions and propose an algorithm to compute facets of their envelopes over hyperrectangles in $\mathbb{R}^3$. Luedtke, Namazifar, and Linderoth [27] show that McCormick relaxations for multilinear terms can give rise to relaxations that are substantially larger than the convex hull. Stronger relaxations that consider the entire bilinear functions have been investigated; see, for example, [44, 46]. Nguyen, Richard, and Tawarmalani [34] provide a convex hull description of the set $z = x^a y^b$, where all variables (including z) belong to a box. While the above studies focus on situations where variables are constrained to belong to simple polyhedra, typically hyperrectangles, others consider special-structured functions over more general polytopes. As an example of the latter, [45] derives convex envelopes of functions that are extendable from vertices through polyhedral subdivisions. In [26], the authors obtain envelopes of bivariate functions over polytopes by computing supporting hyperplanes through the solution of a convex subproblem. Sherali and Alameddine [39] derive an explicit characterization of the convex envelopes of bivariate bilinear functions over D-polytopes, i.e., polytopes with no finite upward-sloping edge. The authors also propose a partitioning scheme for non-D-polytopes based on a triangular decomposition of the feasible region, over which the convex envelope can be obtained. Other relaxation techniques that utilize semidefinite programming [5, 10] and Lagrangian duals [12] have also been investigated.

The results we develop in this paper derive tight relaxations of bilinear functions constructively in the space of the original problem variables when (i) multiple bilinear terms in multiple constraints are considered simultaneously, and (ii) one of the polytopes in the domain is general. Our convexification results can therefore be used

to derive inequalities that are cognizant of the structure of the problem constraints; i.e., they do not only apply to box constraints. As examples of the advantages of the approach, we provide simple, explicit convex hull characterizations for sets involved in the unary expansion of integer variables in mixed integer BLPs, we obtain explicit descriptions for the envelopes of general bilinear functions over certain polytopes and derive a convex hull description for a certain relaxation of NIPs. The latter results are evaluated numerically and lead to significant improvements in the quality of their relaxations.

The remainder of this paper is organized as follows. In section 2, we describe our convexification procedure. In section 3, we present applications of the proposed technique. We then show in section 4 how it can be used to strengthen the linearization of the dual formulation of NIPs. In section 5, we present numerical results evaluating the strength of inequalities derived in section 4. We conclude the paper in section 6 with remarks and directions of future research.

Notation. We use the following notation throughout the paper. Given a set $X \subseteq \mathbb{R}^n$, we denote its convex hull by $\text{conv}(X)$. Further, given a set $Z \subseteq \mathbb{R}^n \times \mathbb{R}^m$ with variables $(\mathbf{x}; \mathbf{y})$, we use $\text{proj}_{\mathbf{x}}(Z)$ to represent the projection of Z onto the space of variables $\mathbf{x}$. We use $\mathbf{e}^j$ to denote a unit vector of suitable dimension whose components are zero except for the j th entry, which is equal to 1. Finally, given $u \in \mathbb{R}$, we write u^+ to denote $\max\{u, 0\}$.

2. Convexification procedure. For $M := \{1, \dots, m\}$, $N := \{1, \dots, n\}$, $K := \{1, \dots, \kappa\}$, and $T := \{1, \dots, \tau\}$, we consider

$$\mathcal{S} = \{(\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \Xi \times \Delta_m \times \mathbb{R}^\kappa \mid \mathbf{y}^\top A^k \mathbf{x} = z_k \quad \forall k \in K\},$$

where $\Xi = \{\mathbf{x} \in [0, 1]^n \mid E\mathbf{x} \geq \mathbf{f}\}$ and $\Delta_m = \{\mathbf{y} \in \mathbb{R}_+^m \mid \mathbf{1}^\top \mathbf{y} \leq 1\}$. In this definition, $A^k \in \mathbb{R}^{m \times n}$, $E \in \mathbb{R}^{\tau \times n}$, and $\mathbf{f} \in \mathbb{R}^\tau$. We refer to the j th row (resp., j th column) of A^k by A_j^k (resp., by $A_{\cdot j}^k$).

It is easy to verify that $\mathcal{S} \neq \emptyset$ if and only if $\Xi \neq \emptyset$. In the remainder of this paper, we therefore assume that $\Xi \neq \emptyset$. We are interested in studying the convex hull of $\mathcal{S}$, which is a polytope; see the discussion following Proposition 2.1 or Corollary 2.7 in [43]. If the bilinear constraints in the description of $\mathcal{S}$ contain linear and constant terms, say $\mathbf{y}^\top A^k \mathbf{x} + \mathbf{b}^k \mathbf{x} + \mathbf{c}^k \mathbf{y} + d_k = z_k$ for coefficient vectors of suitable dimension, we can use an affine transformation to reformulate them as $\mathbf{y}^\top A^k \mathbf{x} = \bar{z}_k$, where $\bar{z}_k = z_k - \mathbf{b}^k \mathbf{x} - \mathbf{c}^k \mathbf{y} - d_k$. Because variables z_k and $\bar{z}_k$ are unrestricted in sign and appear in a single constraint, it is easy to verify that a convex hull description of the set with variables $(\mathbf{x}, \mathbf{y}, \bar{\mathbf{z}})$, together with the equality constraints $\bar{z}_k = z_k - \mathbf{b}^k \mathbf{x} - \mathbf{c}^k \mathbf{y} - d_k$, is sufficient to describe the convex hull of the original set.

We refer to valid inequalities of $\text{conv}(\mathcal{S})$ that are expressed as/dominated by conic combinations of inequalities defining Ξ and Δ_m as *vertical*. In this section, we describe a procedure that generates all nonvertical facet-defining inequalities of $\text{conv}(\mathcal{S})$ in the original space of variables. To motivate this procedure, we first derive a nonvertical facet-defining inequality for the convex hull of a particular instance.

Example 1. Consider

$$\hat{\mathcal{S}}^1 = \{(x_1, x_2, y_1, y_2, z_1, z_2) \in \Xi^1 \times \Delta_2 \times \mathbb{R}^2 \mid (2.1a), (2.1b)\},$$

where $\Xi^1 = \{\mathbf{x} \in [0, 1]^2 \mid (2.1c)\}$, and where

$$(2.1a) \quad 2x_1y_1 - x_2y_1 - 5x_2y_2 = z_1,$$

$$(2.1b) \quad x_2y_2 = z_2,$$

$$(2.1c) \quad x_1 + x_2 \geq 1,$$

$$(2.1d) \quad x_1 \geq 0,$$

$$(2.1e) \quad x_2 \geq 0,$$

$$(2.1f) \quad x_1 \leq 1,$$

$$(2.1g) \quad x_2 \leq 1.$$

It can be verified using PORTA [15] that a linear description of $\text{conv}(\hat{S}^1)$ is given by the inequalities describing Ξ^1 and Δ_2 , together with

$$(2.2a) \quad 2y_1 - z_1 - 5z_2 \geq 0,$$

$$(2.2b) \quad -x_2 + y_1 - y_2 - z_1 - 4z_2 \geq -1,$$

$$(2.2c) \quad 2x_1 - x_2 + y_1 - 3y_2 - z_1 - 2z_2 \geq -1,$$

$$(2.2d) \quad y_2 - z_2 \geq 0,$$

$$(2.2e) \quad z_2 \geq 0,$$

$$(2.2f) \quad -x_2 - y_2 + z_2 \geq -1,$$

$$(2.2g) \quad 3x_2 - 2y_1 + z_1 + 2z_2 \geq 0,$$

$$(2.2h) \quad -2x_1 + x_2 - 2y_1 + z_1 + 4z_2 \geq -2,$$

$$(2.2i) \quad y_1 + z_1 + 5z_2 \geq 0,$$

$$(2.2j) \quad -2x_1 - y_1 + z_1 + 5z_2 \geq -2.$$

We next describe a constructive procedure that leads to the derivation of (2.2h). First, we aggregate (2.1a), (2.1b) and bound inequalities $x_2 \geq 0$ and $1 - x_1 \geq 0$ using weights -1 , -4 , $(1 - y_1 - y_2)$, and $2(1 - y_1 - y_2)$, respectively. As a result, we obtain the bilinear aggregated inequality

$$(2.3) \quad 2x_1y_2 - 2x_1 + x_2 - 2y_1 - 2y_2 + z_1 + 4z_2 \geq -2.$$

Second, we relax the term $2x_1y_2$ by $2y_2$ to obtain (2.2h). In the above derivation, (i) we use real numbers as multipliers for the selected bilinear equalities, (ii) we use combinations of variables $\mathbf{y}$ as multipliers for the selected inequalities in the description of Ξ , and (iii) we choose the weights of the selected inequalities so that 3 (one less than the number of inequalities being aggregated) bilinear terms (namely, x_1y_1 , x_2y_1 , and x_2y_2) are canceled, i.e., their coefficients reduce to zero in the aggregated inequality (2.3).

Building on Example 1, we next present a procedure to derive a family of valid inequalities for $\text{conv}(\mathcal{S})$ that we call *class- $l^\pm$ extended cancel-and-relax (EC $\mathcal{C}\mathcal{R}$)* inequalities. To obtain these inequalities, we select a subset of the constraints in the description of $\mathcal{S}$ and aggregate them with proper weights as follows:

1. We select $l \in K$ to be the lowest index among the bilinear constraints used in the aggregation. We refer to this constraint as the *base* equality. We also select a sign indicator $+$ or $-$ to specify whether weight 1 or -1 is used for the base equality during aggregation.

2. We select L and $\bar{L}$ as disjoint subsets of $K \setminus \{1, \dots, l\}$. Then, for each $k \in L$ (resp., $k \in \bar{L}$), we multiply $\mathbf{y}^\top A^k \mathbf{x} - z_k = 0$ by β_k^+ (resp., $-\beta_k^-$), where $\beta_k^+ \geq 0$ (resp., $\beta_k^- \geq 0$).
3. We select $I_1, \dots, I_m$ and $\bar{I}$ as subsets of T whose intersection is empty. Then, for each $j \in M$ and for each $t \in I_j$ (resp., $t \in \bar{I}$), we multiply $E_t \mathbf{x} \geq f_t$ by $\gamma_t^j y_j$, where $\gamma_t^j \geq 0$ (resp., by $\theta_t(1 - \sum_{i \in M} y_i)$, where $\theta_t \geq 0$).
4. We select J and $\bar{J}$ as disjoint subsets of N . Then, for each index $i \in J$, we multiply $x_i \geq 0$ by $\lambda_i(1 - \sum_{j \in M} y_j)$, where $\lambda_i \geq 0$, and for each $i \in \bar{J}$, we multiply $1 - x_i \geq 0$ by $\mu_i(1 - \sum_{j \in M} y_j)$, where $\mu_i \geq 0$.

We compactly record these sets as $[L, \bar{L} | I_1, \dots, I_m, \bar{I} | J, \bar{J}]$, which we call an *assignment*. We describe this assignment as being class- $l^\pm$, where l is the index of the base inequality and $\pm$ is its sign indicator. We next aggregate all aforementioned weighted constraints. During the aggregation, we require that weights β_k , γ_t^j , θ_t , λ_i , and μ_i be chosen in such a way that

- (i) at least $|L| + |\bar{L}| + \sum_{j \in M} |I_j| + |\bar{I}| + |J| + |\bar{J}|$ bilinear terms are canceled (i.e., their coefficient becomes zero),
- (ii) at least one bilinear term that appears in the base equality l is canceled, and
- (iii) for each $i \in J \cup \bar{J}$, at least one bilinear term among $x_i y_1, x_i y_2, \dots, x_i y_m$ is canceled.

The desired EC&R inequality is then obtained by relaxing the remaining bilinear terms $x_i y_j$ in the aggregated inequality using either $x_i y_j \geq 0$ or $y_j - x_i y_j \geq 0$, depending on the sign of their coefficients. We refer to the resulting linear inequality as a class- $l^\pm$ EC&R inequality.

We refer to the above aggregation and relaxation process as the *EC&R procedure* or *EC&R* for short. A simpler variant of this scheme, coined cancel-and-relax (C&R), is given in [35] for the purpose of deriving the convex hulls of certain separable complementarity sets. Such sets are faces of instances of $\mathcal{S}$, where $m = \kappa = 1$ and $A^1 \mathbf{x}$ is nonnegative for $\mathbf{x} \in \Xi$.

Example 1 (continued). We next show that similar to (2.2h), inequalities (2.2a)–(2.2g) and (2.2i)–(2.2j) can also be obtained using EC&R. For this instance, there are four classes of EC&R inequalities, namely class-1 $^\pm$ and class-2 $^\pm$. Each inequality in class- $l^\pm$ is determined by an assignment $[L, \bar{L} | I_1, I_2, \bar{I} | J, \bar{J}]$. Table 1 describes how inequalities (2.2a)–(2.2j) can be derived using EC&R. Recall that once the aggregated bilinear inequality is obtained, the remaining bilinear terms $x_i y_j$ are uniquely relaxed into either 0 or y_j , depending on the sign of their coefficients.

The previous example suggests that all nonvertical facet-defining inequalities of $\text{conv}(\hat{\mathcal{S}}^1)$ can be obtained using EC&R. The next section is dedicated to proving that this result holds for all instances of $\mathcal{S}$.

2.1. Derivation of EC&R procedure. In this section, through a sequence of intermediary results, we will prove Theorem 2.7 (stated below for the reader's convenience).

THEOREM 2.7. *A linear description of $\text{conv}(\mathcal{S})$ is given by the inequalities defining Ξ and Δ_m , together with all class- $l^\pm$ EC&R inequalities for $l \in K$.*

The description of the convex hull of $\mathcal{S}$ given in Theorem 2.7 is typically not minimal; i.e., not all EC&R inequalities are facet-defining for $\text{conv}(\mathcal{S})$. Further, Theorem 2.7 does not preclude the aggregation of all bilinear constraints and all linear constraints in the derivation of strong inequalities for $\text{conv}(\mathcal{S})$. For instance, in Example 1, (2.2c) requires all constraints to be aggregated. Intuitively, we can

TABLE 1
Facet-defining ECR inequalities for $\widehat{S}^1$.

Ineq.	Class; assignment	Weights
(2.2a)	1^+ ; $[\{(2.1b)\}, \emptyset \emptyset, \emptyset, \emptyset \emptyset, \emptyset]$	5
(2.2b)	1^+ ; $[\{(2.1b)\}, \emptyset \emptyset, \emptyset, \emptyset \emptyset, \{2\}]$	$4, (1 - y_1 - y_2)$
(2.2c)	1^+ ; $[\{(2.1b)\}, \emptyset \emptyset, \{(2.1c)\}, \emptyset \{(2.1d)\}, \{(2.1g)\}]$	$2, 2y_2, 2(1 - y_1 - y_2), (1 - y_1 - y_2)$
(2.2d)	2^+ ; $[\emptyset, \emptyset \emptyset, \emptyset, \emptyset \emptyset, \emptyset]$	NA
(2.2e)	2^- ; $[\emptyset, \emptyset \emptyset, \emptyset, \emptyset \emptyset, \emptyset]$	NA
(2.2f)	2^- ; $[\emptyset, \emptyset \emptyset, \emptyset, \emptyset \emptyset, \{(2.1g)\}]$	$(1 - y_1 - y_2)$
(2.2g)	1^- ; $[\emptyset, \{(2.1b)\} \{(2.1c)\}, \emptyset, \emptyset \{(2.1e)\}, \emptyset]$	$-2, 2y_1, 3(1 - y_1 - y_2)$
(2.2h)	1^- ; $[\emptyset, \{(2.1b)\} \emptyset, \emptyset, \emptyset \{(2.1e)\}, \{(2.1f)\}]$	$-4, (1 - y_1 - y_2), 2(1 - y_1 - y_2)$
(2.2i)	1^- ; $[\emptyset, \{(2.1b)\} \emptyset, \emptyset, \emptyset \emptyset, \emptyset]$	-5
(2.2j)	1^- ; $[\emptyset, \{(2.1b)\} \emptyset, \emptyset, \emptyset \emptyset, \{(2.1f)\}]$	$-5, 2(1 - y_1 - y_2)$

interpret the requirement that the number of cancellations be on par with the number of constraints being aggregated as a mechanism to balance the weakness induced by aggregation. It is therefore our conjecture that inequalities produced by aggregating few constraints play an important role in computation. Example 1 supports this conjecture in that only one of its nontrivial facet-defining inequalities requires all constraints to be aggregated. This observation is computationally useful, as there is a relatively small number of these inequalities, and they can be separated efficiently. In the numerical experiments we report in section 5, we observe that inequalities derived using few aggregations are indeed most useful at reducing gaps.

We first show an ancillary result that allows the use of disjunctive programming [8] in studying $\text{conv}(\mathcal{S})$.

PROPOSITION 2.1. *Let $\delta = (\mathbf{x}; \mathbf{y}; \mathbf{z})$ be an extreme point of $\text{conv}(\mathcal{S})$. Then $\mathbf{y} \in \{0, 1\}^m$.*

Proof. It is clear that $\delta \in \mathcal{S}$, as it is an extreme point of $\text{conv}(\mathcal{S})$. Assume by contradiction that $y_j \in (0, 1)$ for some $j \in M$. First, assume that $\mathbf{1}^\top \mathbf{y} < 1$. Then, consider $\delta^1 = \delta + \bar{\delta}$ and $\delta^2 = \delta - \bar{\delta}$, where $\bar{\delta} = (\mathbf{0}; \epsilon \mathbf{e}^j; \sum_{k \in K} \epsilon (A_j^k \mathbf{x} - A_j^k \mathbf{x}) \mathbf{e}^k)$ for a sufficiently small but positive ϵ . It is clear that δ^1 and δ^2 belong to $\mathcal{S}$. This yields the desired contradiction as $\delta = \frac{1}{2}\delta^1 + \frac{1}{2}\delta^2$. Second, assume that $\mathbf{1}^\top \mathbf{y} = 1$. Therefore, there exists $i \in M \setminus \{j\}$ such that $y_i \in (0, 1)$. We construct $\delta^1 = \delta + \bar{\delta}$ and $\delta^2 = \delta - \bar{\delta}$, where $\bar{\delta} = (\mathbf{0}; \epsilon (\mathbf{e}^j - \mathbf{e}^i); \sum_{k \in K} \epsilon (A_j^k \mathbf{x} - A_i^k \mathbf{x}) \mathbf{e}^k)$ for a sufficiently small but positive ϵ . It is clear that δ^1 and δ^2 belong to $\mathcal{S}$, yielding the desired contradiction. $\square$

Proposition 2.1 implies that a variant of $\mathcal{S}$ where variables $\mathbf{y}$ are binary has the same convex hull as that of $\mathcal{S}$. As a result, a description of $\text{conv}(\mathcal{S})$ can be obtained in higher dimension using the special structure RLT (reformulation linearization technique), described in [38], or disjunctive programming [8]. In this paper, we refine these results by presenting a constructive procedure to produce $\text{conv}(\mathcal{S})$ in the original space of variables. Understanding how strong inequalities for $\text{conv}(\mathcal{S})$ can be constructed in the original space of variables yields insight into useful families of cuts for this set. It also allows for families of strong inequalities to be derived for more complicated sets, as we illustrate in section 4 for the case of NIPs.

It follows from Proposition 2.1 that $\text{conv}(\mathcal{S}) = \text{conv}(\bigcup_{\omega \in \{0, e^1, \dots, e^m\}} \mathcal{S}(\omega))$, where

$$\mathcal{S}(\omega) = \left\{ (x; y; z) \in \mathbb{R}^{n+m+\kappa} \left| \begin{array}{l} y = \omega \\ \omega^\top A^k x - z_k = 0 \quad \forall k \in K \\ Ex \geq f \\ 0 \leq x \leq 1 \end{array} \right. \right\}.$$

This provides a direct proof that $\text{conv}(\mathcal{S})$ is a polytope since it is now expressed as the convex hull of a union of polytopes. A disjunctive programming formulation [8] for this set is easily obtained. In this formulation, it can be seen that the convex multipliers ζ_j , for each disjunct $j \neq 0$, and variables y_j are equal, yielding the following.

PROPOSITION 2.2. Define $\mathcal{Q} = \{(x; y; z; u; v) | (2.4)\}$, where

$$\begin{aligned} \pm A_j^k u^j \mp v_k^j &\geq 0 & \forall (k, j) \in K \times M, \\ \mp \left(z_k - \sum_{j \in M} v_k^j \right) &\geq 0 & \forall k \in K, \\ Eu^j &\geq fy_j & \forall j \in M, \\ E \left(x - \sum_{j \in M} u^j \right) &\geq f \left(1 - \sum_{j \in M} y_j \right), \\ 0 \leq u^j &\leq 1y_j & \forall j \in M, \\ 0 \leq x - \sum_{j \in M} u^j &\leq 1 \left(1 - \sum_{j \in M} y_j \right). \end{aligned} \quad (2.4)$$

Then, $\text{conv}(\mathcal{S}) = \text{proj}_{(x; y; z)} \mathcal{Q}$.

In the description of $\mathcal{Q}$, variables u^j and v_k^j play the roles of x and z_k in disjunct $\mathcal{S}(e^j)$ scaled by ζ_j , i.e., $u^j = \zeta_j x$ and $v_k^j = \zeta_j z_k$. Equalities are formulated as pairs of inequalities of opposite directions. We introduce $\bar{\pi} = (\{\alpha^{j+}\}_{j \in M}; \{\alpha^{j-}\}_{j \in M}; \beta^+; \beta^-; \{\gamma^j\}_{j \in M}; \theta; \{\eta^j\}_{j \in M}; \{\rho^j\}_{j \in M}; \lambda; \mu)$ to compactly record the dual variables associated with the constraints of $\mathcal{Q}$, where $\alpha^{j\pm}, \beta^\pm \in \mathbb{R}_+^\kappa$, $\gamma^j, \theta \in \mathbb{R}_+^\tau$, and $\eta^j, \rho^j, \lambda, \mu \in \mathbb{R}_+^n$.

PROPOSITION 2.3. For $\pi = (\beta^+; \beta^-; \{\gamma^j\}_{j \in M}; \theta; \{\eta^j\}_{j \in M}; \{\rho^j\}_{j \in M}; \lambda; \mu) \in \mathbb{R}_+^{(m+1)(\tau+2n)+2\kappa}$, define

$$\mathcal{C} = \left\{ \pi \left| \sum_{k \in K} A_{ji}^k (\beta_k^+ - \beta_k^-) + \sum_{t \in T} E_{ti} (\gamma_t^j - \theta_t) + \eta_i^j - \rho_i^j - \lambda_i + \mu_i = 0 \quad \forall (i, j) \in N \times M \right. \right\}.$$

Then,

$$(2.5) \quad \sum_{i \in N} q_i(\pi) x_i + \sum_{j \in M} r_j(\pi) y_j + \sum_{k \in K} s_k(\pi) z_k \geq t(\pi),$$

where

$$\begin{aligned} q_i(\pi) &= \sum_{t \in T} E_{ti} \theta_t + \lambda_i - \mu_i, \\ r_j(\pi) &= \sum_{t \in T} f_t (\theta_t - \gamma_t^j) + \sum_{i \in N} (\rho_i^j - \mu_i), \\ s_k(\pi) &= -(\beta_k^+ - \beta_k^-), \\ t(\pi) &= \sum_{t \in T} f_t \theta_t - \sum_{i \in N} \mu_i \end{aligned}$$

is facet-defining for $\text{conv}(\mathcal{S})$ only if π is an extreme ray of $\mathcal{C}$.

Proof. From Proposition 2.2, we know that $\text{conv}(\mathcal{S}) = \text{proj}_{(\mathbf{x}; \mathbf{y}; \mathbf{z})} \mathcal{Q}$. We refer to the cone obtained by formulating the projection of $\mathcal{Q}$ relative to variables $\mathbf{u}^j$ and $\mathbf{v}$ as the projection cone of $\mathcal{Q}$ and denote it by $\bar{\mathcal{C}}$; see [8]. It is clear that $\bar{\mathcal{C}}$ is a subset of $\mathbb{R}_+^{(m+1)(\tau+2n+2\kappa)}$. The projection of variables v_k^j yields the constraint

$$(2.6) \quad -\alpha_k^{j+} + \alpha_k^{j-} = -\beta_k^+ + \beta_k^- \quad \forall (k, j) \in K \times M.$$

Using these relations, the projection of variables u_i^j yields the constraint

$$(2.7) \quad \sum_{k \in K} A_{ji}^k (\beta_k^+ - \beta_k^-) + \sum_{t \in T} E_{ti} (\gamma_t^j - \theta_t) + \eta_i^j - \rho_i^j - \lambda_i + \mu_i = 0 \quad \forall (i, j) \in N \times M.$$

Any weight vector $\bar{\pi}$ satisfying (2.6) and (2.7) gives rise to valid inequality (2.5), which is facet-defining for $\text{conv}(\mathcal{S})$ only if $\bar{\pi}$ is an extreme ray of $\bar{\mathcal{C}}$; see [9]. Select such $\bar{\pi}$, and define π to be the vector composed of all components of $\bar{\pi}$ except $\alpha^{j\pm}$. We show that π is an extreme ray of $\mathcal{C}$. To this end, we prove that $\bar{\pi}$ satisfies $\beta_k^+ \beta_k^- = 0$ for $k \in K$. Assume by contradiction that there exists $k \in K$ such that $\beta_k^+ > 0$ and $\beta_k^- > 0$. Construct $\hat{\pi} = \bar{\pi} + (\mathbf{0}; \mathbf{0}; \epsilon \mathbf{e}^k; -\epsilon \mathbf{e}^k; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0})$ and $\tilde{\pi} = \bar{\pi} + (\mathbf{0}; \mathbf{0}; -\epsilon \mathbf{e}^k; \epsilon \mathbf{e}^k; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0})$. It is simple to verify that $\hat{\pi}$ and $\tilde{\pi}$ belong to $\bar{\mathcal{C}}$ for a sufficiently small but positive ϵ . Note also that $\hat{\pi}$ and $\tilde{\pi}$ are not multiples of $\bar{\pi}$, as β_k^+ , β_k^- , and ϵ are all positive numbers. This contradicts the fact that $\bar{\pi}$ is an extreme ray of $\bar{\mathcal{C}}$ as $\bar{\pi} = \frac{1}{2}\hat{\pi} + \frac{1}{2}\tilde{\pi}$. A similar argument yields that $\alpha_k^{j\pm} = \beta_k^\pm$ for all $(k, j) \in K \times M$. Using (2.6), these relations imply that $\alpha_k^{j\pm} = \beta_k^\pm$ for all $(k, j) \in K \times M$. This shows the result. $\square$

In what follows, we describe a finite collection of points of $\mathcal{C}$ that subsumes its extreme rays. This collection, in turn, provides a description of $\text{conv}(\mathcal{S})$, which might not be minimal. We refer to the vectors of $\mathcal{C}$ that correspond to vertical inequalities of $\text{conv}(\mathcal{S})$ as *vertical*. Our next goal is to characterize nonvertical extreme rays of $\mathcal{C}$. We first give necessary conditions.

PROPOSITION 2.4. *If π is a nonvertical extreme ray of $\mathcal{C}$, then (i) $\beta_k^+ \beta_k^- = 0$ for $k \in K$, and (ii) there exists $l \in K$ such that $\beta_k^+ = \beta_k^- = 0$ for $k \in \{1, \dots, l-1\}$ with either $\beta_l^+ > 0$ or $\beta_l^- > 0$.*

Proof. Case (i) follows from an argument similar to that given in the proof of Proposition 2.3. We next show (ii). Assume by contradiction that $\beta^+ = \beta^- = \mathbf{0}$. Then, $(\{\gamma^j\}_{j \in M}; \theta; \{\eta^j\}_{j \in M}; \{\rho^j\}_{j \in M}; \lambda; \mu)$ belongs to the projection cone of the set

$$\begin{aligned} E\mathbf{u}^j &\geq \mathbf{f}y_j & \forall j \in M, \\ E\left(\mathbf{x} - \sum_{j \in M} \mathbf{u}^j\right) &\geq \mathbf{f}\left(1 - \sum_{j \in M} y_j\right), \\ \mathbf{0} \leq \mathbf{u}^j &\leq \mathbf{1}y_j & \forall j \in M, \\ \mathbf{0} \leq \mathbf{x} - \sum_{j \in M} \mathbf{u}^j &\leq \mathbf{1}\left(1 - \sum_{j \in M} y_j\right), \end{aligned}$$

which is itself the disjunctive programming formulation of the convex hull of the set

$$\tilde{\mathcal{P}} = \{(\mathbf{x}; \mathbf{y}) \in \Xi \times (\Delta_m \cap \{0, 1\}^m)\}.$$

Since $\text{conv}(\tilde{\mathcal{P}}) = \Xi \times \Delta_m$, we conclude that π corresponds to a vertical inequality of $\text{conv}(\mathcal{S})$, yielding the desired contradiction. $\square$

It follows from Proposition 2.4 that nonvertical extreme rays can be rescaled so that either $\beta_l^+ = 1$ or $\beta_l^- = 1$ for some $l \in K$, and $\beta_k^+ = \beta_k^- = 0$ for $k \in \{1, \dots, l-1\}$. Define $K_l = K \setminus \{1, \dots, l\}$. Since $\mathcal{C}$ is a cone contained in the positive orthant, such a ray corresponds to an extreme point of either

$$\mathcal{C}^{l+} = \left\{ \pi^l \left| \sum_{k \in K_l} A_{ji}^k (\beta_k^+ - \beta_k^-) + \sum_{t \in T} E_{ti} (\gamma_t^j - \theta_t) + \eta_i^j - \rho_i^j - \lambda_i + \mu_i = -A_{ji}^l \forall (i, j) \in N \times M \right. \right\},$$

or

$$\mathcal{C}^{l-} = \left\{ \pi^l \left| \sum_{k \in K_l} A_{ji}^k (\beta_k^+ - \beta_k^-) + \sum_{t \in T} E_{ti} (\gamma_t^j - \theta_t) + \eta_i^j - \rho_i^j - \lambda_i + \mu_i = A_{ji}^l \forall (i, j) \in N \times M \right. \right\},$$

where $\pi^l \in \mathbb{R}_+^{(m+1)(\tau+2n)+2(\kappa-l)}$ is defined similarly to π , but without elements β_k^+ and β_k^- for $k \in \{1, \dots, l\}$. Conversely, it is easily verified that each extreme point of $\mathcal{C}^{l+}$ (resp., $\mathcal{C}^{l-}$) corresponds to an extreme ray of $\mathcal{C}$ that is scaled so that $\beta_l^+ = 1$ and $\beta_l^- = 0$ (resp., $\beta_l^- = 1$ and $\beta_l^+ = 0$), and is such that $\beta_k^+ = \beta_k^- = 0$ for $k \in \{1, \dots, l-1\}$.

Using a proof technique similar to that used in [35], we give a characterization of extreme points of $\mathcal{C}^{l+}$ that correspond to nonvertical extreme rays of $\mathcal{C}$. We refer to such extreme points as *nonvertical*.

PROPOSITION 2.5. *Define*

$$\widehat{\mathcal{C}}^{l+} = \left\{ \hat{\pi}^l \in \mathbb{R}_+^q \left| \begin{array}{l} \exists L, \bar{L} \subseteq K_l, \\ \exists I_1, I_2, \dots, I_m, \bar{I} \subseteq T, \\ \exists J, \bar{J} \subseteq N, \\ \exists G \subseteq N \times M, \\ \text{s.t.} \\ |G| = |L| + |\bar{L}| + \sum_{j \in M} |I_j| + |\bar{I}| + |J| + |\bar{J}|, \\ \beta_i^+ = 0 \quad \forall i \in K_l \setminus L \\ \beta_i^- = 0 \quad \forall i \in K_l \setminus \bar{L} \\ \gamma_i^j = 0 \quad \forall j \in M, i \in T \setminus I_j \\ \theta_i = 0 \quad \forall i \in T \setminus \bar{I} \\ \lambda_i = 0 \quad \forall i \in N \setminus J \\ \mu_i = 0 \quad \forall i \in N \setminus \bar{J} \\ (2.5), \\ \text{the coefficient matrix of (2.5) is nonsingular} \end{array} \right. \right\},$$

where $q = \tau(m+1) + 2(n+\kappa-l)$, and $\hat{\pi}^l$ is defined similarly to π^l , except that components η^j and ρ^j are eliminated, and where

$$(2.8) \quad \begin{aligned} & \sum_{k \in L} A_{ji}^k \beta_k^+ - \sum_{k \in \bar{L}} A_{ji}^k \beta_k^- + \sum_{t \in I_j} E_{ti} \gamma_t^j - \sum_{t \in \bar{I}} E_{ti} \theta_t \\ & - \sum_{t \in J \cap \{i\}} \lambda_t + \sum_{t \in \bar{J} \cap \{i\}} \mu_t = -A_{ji}^l \quad \forall (i, j) \in G. \end{aligned}$$

If $\bar{\pi}^l = (\hat{\pi}^l; \{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M})$ is a nonvertical extreme point of $\mathcal{C}^{l+}$, then $\hat{\pi}^l \in \widehat{\mathcal{C}}^{l+}$. Conversely, if $\hat{\pi}^l \in \widehat{\mathcal{C}}^{l+}$, there exists a nonvertical extreme point of $\mathcal{C}^{l+}$ of the form $\bar{\pi}^l = (\hat{\pi}^l; \{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M})$ for some $(\{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M}) \in \mathbb{R}_+^{2mn}$.

The idea of the proof, which is given in the appendix, is as follows. Any nonvertical extreme point of $\mathcal{C}^{l+}$ is associated with a basic feasible solution (bfs) of this set. The basis corresponding to this bfs has a triangular block representation (through rearranging columns and rows) with two blocks on the diagonal. One of these blocks is an identity matrix whose columns are composed of basic variables among η^j and ρ^j for $j \in M$. The other diagonal block is a nonsingular square matrix whose columns are composed of the remaining basic variables. We record the indices of these basic variables in sets $L, \bar{L}, I_j, \bar{I}, J$, and $\bar{J}$. We record the indices of the rows of this nonsingular matrix in G . Examining the correspondence between bases of $\mathcal{C}^{l+}$ and the aforementioned nonsingular matrices, we obtain the characterization given in Proposition 2.5.

Nonvertical extreme points of $\mathcal{C}^{l-}$ admit a characterization similar to that of Proposition 2.5, except that the system of equations in $\widehat{\mathcal{C}}^{l-}$ has right-hand side A_{ji}^l instead of $-A_{ji}^l$. The results of Propositions 2.3 and 2.5 are the basis for our claim that nonvertical facet-defining inequalities of $\text{conv}(\mathcal{S})$ can be obtained using EC&R. However, in EC&R, we do not directly verify that the coefficient matrix in the description of $\widehat{\mathcal{C}}^{l\pm}$ is nonsingular. Instead, we only impose easily verifiable necessary conditions for nonsingularity.

COROLLARY 2.6. *Assume that $\hat{\pi}^l \in \widehat{\mathcal{C}}^{l\pm}$. Then, (i) $L \cap \bar{L} = \bigcap_{j \in M} I_j \cap \bar{I} = J \cap \bar{J} = \emptyset$; (ii) if $\hat{\pi}^l \neq \mathbf{0}$, there exists $(i, j) \in G$ such that $A_{ji}^l \neq 0$; and (iii) for each $i \in N$ with $\lambda_i > 0$ or $\mu_i > 0$, there exists $j \in M$ such that $(i, j) \in G$.*

Proof. For (i) we only show that $\bigcap_{j \in M} I_j \cap \bar{I} = \emptyset$, as the proofs for $L \cap \bar{L} = \emptyset$ and $J \cap \bar{J} = \emptyset$ are similar. Assume by contradiction that there exists $i \in \bigcap_{j \in M} I_j \cap \bar{I}$. Then, the coefficient columns of γ_i^j , for $j \in M$, and θ_i are linearly dependent, yielding a contradiction. For (ii), assume by contradiction that $A_{ji}^l = 0$ for all $(i, j) \in G$. Then, the right-hand side of (2.5) is the zero vector. Since the coefficient matrix of this system is nonsingular, its unique solution is $\hat{\pi}^l = \mathbf{0}$. This is a contradiction to the fact that $\hat{\pi}^l \neq \mathbf{0}$. For (iii), assume by contradiction that there exists $i \in N$ with $\lambda_i > 0$ or $\mu_i > 0$ such that $(i, j) \notin G$ for all $j \in M$. Therefore, the coefficient column of variable λ_i or μ_i in (2.5) is zero, a contradiction. $\square$

We are now ready to prove the main result of this section.

THEOREM 2.7. *A linear description of $\text{conv}(\mathcal{S})$ is given by the inequalities defining Ξ and Δ_m , together with class- $l^\pm$ EC&R inequalities for $l \in K$.*

Proof. Consider a facet-defining inequality of $\text{conv}(\mathcal{S})$. Assuming it is not vertical, we show that it can be obtained using EC&R. It follows from Proposition 2.3 that the coefficient vector of this facet-defining inequality corresponds to an extreme ray of $\mathcal{C}$. Proposition 2.4 shows that we can rescale this ray so that $\beta_l^\pm = 1$ for some $l \in K$ and $\beta_k^\pm = 0$ for $k \in \{1, \dots, l-1\}$. The studied ray corresponds to an extreme point of $\mathcal{C}^{l\pm}$, which we call $\hat{\pi}^l$. Proposition 2.5 shows that nontrivial components of $\hat{\pi}^l$, which we refer to as $\hat{\pi}^l$, correspond to an element of $\widehat{\mathcal{C}}^{l\pm}$. Select $L, \bar{L}, I_1, \dots, I_m, \bar{I}, J$, and $\bar{J}$ accordingly. We take the components of $\hat{\pi}^l$ not corresponding to these sets to be zero. We claim that the given inequality can be obtained as a class- $l^\pm$ EC&R inequality with assignment $[L, \bar{L} | I_1, \dots, I_m, \bar{I} | J, \bar{J}]$ and weights $\hat{\pi}^l$ (correspondingly ordered). To this end, we show that the weights satisfy EC&R requirements. First observe that, since $\hat{\pi}^l$ is an element of $\widehat{\mathcal{C}}^{l\pm}$, matrix (2.5) is nonsingular. It therefore satisfies condition (i) of Corollary 2.6. This implies that L and $\bar{L}$ are disjoint, $I_1, I_2, \dots, \bar{I}$ do not have common elements, and J and $\bar{J}$ are disjoint, as specified in the EC&R

procedure. Further, weights $\hat{\pi}^l$ are clearly nonnegative since they are extracted from solutions of $\hat{\mathcal{C}}^{l\pm}$. Second, observe that in $\hat{\mathcal{C}}^{l\pm}$, (2.5) is satisfied for each $(i, j) \in G$. This constraint corresponds to the projection of variable u_i^j . Since variable u_i^j can be viewed as $x_i y_j$, (2.5) expresses that bilinear terms $x_i y_j$ for $(i, j) \in G$ cancel out in the aggregated bilinear inequality obtained from elements of $[L, \bar{L}|I_1, \dots, I_m, \bar{I}|J, \bar{J}]$ with weights $\hat{\pi}^l$. Since $|G| = |L| + |\bar{L}| + \sum_{j \in M} |I_j| + |\bar{I}| + |J| + |\bar{J}|$, the number of cancellations is as specified in requirement (i) of EC&R. Third, observe that condition (ii) of Corollary 2.6 implies that if there exists a constraint among those corresponding to $[L, \bar{L}|I_1, \dots, I_m, \bar{I}|J, \bar{J}]$ that is aggregated with a nonzero weight, then at least one bilinear term appearing in the base equality is canceled during aggregation. This is expressed as requirement (ii) of EC&R. Similarly, condition (iii) of Corollary 2.6 shows that at least one bilinear term among $x_i y_1, x_i y_2, \dots, x_i y_m$ is canceled during aggregation for each $i \in J \cup \bar{J}$. This is given as requirement (iii) of EC&R. Finally, observe that $\hat{\pi}^l$ is only part of the vector $\bar{\pi}^l$. To obtain the complete set of dual weights, it remains to show that the remaining components $(\{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M})$ of $\bar{\pi}^l$ can also be obtained by EC&R. These values correspond to the projection of variables u_i^j for $(i, j) \notin G$. This corresponds to relaxing the remaining bilinear terms $x_i y_j$ with either $x_i y_j \geq 0$ or $y_j - x_i y_j \geq 0$, since η_i^j and ρ_i^j are the dual weights for constraints $x_i y_j \geq 0$ and $y_j - x_i y_j \geq 0$, respectively. $\square$

Determining all EC&R weights for a given set $\mathcal{S}$ typically requires the solution of a large number of linear systems of inequalities. In particular, it is not hard to create families of problems where the number of possible EC&R weights increases exponentially with the number of variables and constraints in the problem. This is expected, as the set of these weights corresponds to a super set of all extreme rays of $\mathcal{C}$. One possible way to access specific EC&R weights is to solve a cut-generating LP (CGLP) associated with the problem given in Proposition 2.3; see [9] for a discussion on CGLPs. A practical difficulty in using this approach stems from the fact that the dimension of the CGLP grows substantially with the number of variables y , which dictates the number of disjuncts. For this reason, we next focus in sections 3 and 4 on problems with specific structures for which we are able to explicitly obtain the full collection of combination weights without recourse to the solution of a CGLP.

2.2. Variations and properties. In this section, we establish certain properties of EC&R inequalities and study specific set structures for which the application of EC&R takes on a specialized and simpler form.

We first discuss a connection between EC&R and a simple technique to obtain convex relaxations of $\mathcal{S}$. Denote the bilinear constraints in the description of $\mathcal{S}$ by $g^k(\mathbf{x}; \mathbf{y}) = z_k$ for $k \in K$. For any $\boldsymbol{\xi} \in \mathbb{R}^k$, assume that the concave (resp., convex) envelope of $\bar{z} = \sum_{k \in K} \xi_k g^k(\mathbf{x}; \mathbf{y})$ over $\Xi \times \Delta_m$ is described by $\bar{z} \leq \min_{i \in H^\xi} h_i^\xi(\mathbf{x}; \mathbf{y})$ (resp., $\bar{z} \geq \max_{j \in \bar{H}^\xi} \bar{h}_j^\xi(\mathbf{x}; \mathbf{y})$), where functions $h(\mathbf{x}; \mathbf{y})$ (resp., $\bar{h}(\mathbf{x}; \mathbf{y})$) are linear. Then, it is clear that the collection of all *envelope inequalities* $\sum_{k \in K} \xi_k z_k \leq \min_{i \in H^\xi} h_i^\xi(\mathbf{x}; \mathbf{y})$ and $\sum_{k \in K} \xi_k z_k \geq \max_{j \in \bar{H}^\xi} \bar{h}_j^\xi(\mathbf{x}; \mathbf{y})$ for all $\boldsymbol{\xi} \in \mathbb{R}^k$ yields the convex hull of $\mathcal{S}$; see [43] for instance. The EC&R procedure shows that only a finite number of weights $\boldsymbol{\xi}$ are required to obtain the convex hull of $\mathcal{S}$. Further, the EC&R procedure specifies how these combination weights can be constructed. In particular, consider an EC&R inequality with weights $\beta^\pm$ for bilinear constraints. Then, it is easy to verify that this inequality is implied by the set of envelope inequalities, as described above for $\boldsymbol{\xi} = \beta^+ - \beta^-$. This follows from the fact that the EC&R inequality is valid for the convex hull of $\sum_{k \in K} \xi_k z_k = \sum_{k \in K} \xi_k g^k(\mathbf{x}; \mathbf{y})$ over $\Xi \times \Delta_m$, which

is described by the envelope inequalities corresponding to ξ . Since the collection of all EC&R inequalities describes the convex hull of $\mathcal{S}$, we conclude that the set of all envelope inequalities with $\xi = \beta^+ - \beta^-$ for all EC&R weights $\beta^\pm$ gives the convex hull of $\mathcal{S}$. We mention, however, that the collection of all envelope inequalities obtained above can be much larger than the collection of all EC&R inequalities, even when only EC&R-derived weights $\xi = \beta^+ - \beta^-$ are used.

As mentioned before, not all EC&R inequalities are facet-defining for $\text{conv}(\mathcal{S})$. We next address the question of assessing the strength of an EC&R inequality.

PROPOSITION 2.8. *Consider a class- $l^\pm$ EC&R inequality $h(\mathbf{x}; \mathbf{y}; \mathbf{z}) \geq h_0$. A point $(\bar{\mathbf{x}}; \bar{\mathbf{y}}; \bar{\mathbf{z}}) \in \mathcal{S}$ satisfies $h(\mathbf{x}; \mathbf{y}; \mathbf{z}) \geq h_0$ at equality if and only if for each linear constraint in Ξ that is used in the aggregation, either $\bar{\mathbf{x}}$ is binding, or its dual multiplier $\bar{y}_j$ or $1 - \sum_{j \in M} \bar{y}_j$ is zero. We refer to these constraints as “complementary slackness.” This definition also includes the bound constraints $y_j(1 - x_i) \geq 0$ and $y_j x_i \geq 0$ associated with the linearization step in EC&R.*

Further, let $\mathcal{X}$ be the set of points $(\bar{\mathbf{x}}; \bar{\mathbf{y}}; \bar{\mathbf{z}}) \in \mathcal{S}$, where $(\bar{\mathbf{x}}; \bar{\mathbf{y}})$ satisfies complementary slackness. Then, $h(\mathbf{x}; \mathbf{y}; \mathbf{z}) \geq h_0$ defines a face of $\text{conv}(\mathcal{S})$ of dimension equal to $\dim(\mathcal{X})$.

Proof. For the first result, complementary slackness forces the constraints in the aggregation to be equalities, and therefore the resulting EC&R inequality is tight. Conversely, if one of the complementary slackness constraints is violated, the corresponding weighted inequality participates in the aggregation as a strict inequality, and therefore the resulting EC&R inequality cannot be tight. The second result follows directly from the first. $\square$

Proposition 2.8 yields a procedure to determine the strength of an EC&R inequality by characterizing the points that are tight for the inequalities involved in the aggregation. Since it suffices to include the extreme points of $\mathcal{S}$ in $\mathcal{X}$, we can systematically obtain this set as follows. We consider each of the $m+1$ vertices of the simplex Δ_m , and identify the linear constraints in Ξ that need to be tight to satisfy complementary slackness. We obtain a linear set of constraints on variables $\mathbf{x}$ (a face of Ξ) and a linear set of equalities for variables $\mathbf{z}$. We can include this polytope or its extreme points in $\mathcal{X}$. Repeating this procedure for all $m+1$ vertices of Δ_m , we obtain a description of $\mathcal{X}$ as a disjunctive union of $m+1$ polytopes. Its dimension determines the dimension of the face defined by the EC&R inequality.

The set $\mathcal{X}$ has a specific interpretation when $\mathcal{S}$ contains a single bilinear constraint. In this case, the convex hull of $\mathcal{S}$ corresponds to the convex and concave envelopes of the bilinear function over its domain $\Xi \times \mathcal{W}$, together with the defining inequalities of the domain. It is well known that the projection of the facets of the convex (resp., concave) envelope over the domain gives rise to a polyhedral subdivision of the domain; see [45] for instance. Proposition 2.8 shows that the element of the polyhedral subdivision corresponding to a facet-defining EC&R inequality can be obtained as the convex hull of points $(\mathbf{x}; \mathbf{y}) \in \text{proj}_{(\mathbf{x}; \mathbf{y})} \mathcal{X}$. In the very special case where the bilinear set of interest is defined by $z = xy$ with $(x, y) \in [0, 1]^2$, we obtain, for instance, that one of the two inequalities describing the convex side of the McCormick envelope, $z \geq x + y - 1$, is the EC&R inequality obtained by combining $z = xy$ and $x \leq 1$ with weights 1 and $(1 - y)$, respectively. It follows from Proposition 2.8 that this inequality is tight when $x = 1$ or $y = 1$, i.e., it is tight over the simplex with vertices $(1, 0)$, $(0, 1)$, and $(1, 1)$. This recovers the well-known fact that this simplex is part of the polyhedral partition defining the convex envelope of xy over $[0, 1]^2$.

EC&R can be specialized and/or simplified if additional structure is assumed on the set $\mathcal{S}$. In the next few remarks, we describe certain setups for which such specializations/simplifications can be developed.

Remark 1. Assume that we wish to model epigraphs of some bilinear functions in $\mathcal{S}$ rather than their graphs, i.e., $\mathbf{y}^\top A^k \mathbf{x} \geq z_k$ for $k \in \bar{K} \subseteq K$. Then, Theorem 2.7 can be adapted to state that during EC&R, when an inequality $k \in \bar{K}$ is chosen as the base inequality, it is sufficient to compute the corresponding class- k^+ , and when inequality k is used during aggregation, it is included in L but not in $\bar{L}$.

Remark 2. Assume that $\mathcal{S}$ contains a single variable y_1 , i.e., $m = 1$. In this case, we may ignore the bound constraints on x variables in the aggregation step and consider them instead in the relaxation step. This amounts to reducing the assignment defining class- $l^\pm$ EC&R inequalities to $[L, \bar{L}|I, \bar{I}]$ and to omit the sets J and $\bar{J}$. We then require the number of canceled bilinear terms to be at least $|L| + |\bar{L}| + |I| + |\bar{I}|$. Further, when relaxing the aggregated inequality, we now use McCormick upper bounds $x_i y_1 \leq \min\{x_i, y_1\}$ and McCormick lower bounds $x_i y_1 \geq \max\{0, x_i + y_1 - 1\}$, depending on the sign of coefficients. In this variant, each aggregated bilinear inequality can be relaxed into multiple linear inequalities. The validity of this variation follows from the fact that under the given assumption, the columns of $\mathcal{C}^{l^\pm}$ associated with variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ contain a single nonzero element, and therefore play roles similar to $\boldsymbol{\eta}$ and $\boldsymbol{\rho}$.

Remark 3. EC&R can be applied to construct the convex hull of sets that contain complementarity constraints. In particular, consider the set

$$(2.9) \quad \mathcal{S}^c = \{(\mathbf{x}; \mathbf{y}) \in \Xi \times \Delta_m \mid (\mathbf{c}^k \mathbf{y}) (\mathbf{b}^k \mathbf{x}) = 0 \quad \forall k \in K\},$$

where we assume that the nonnegativity of complementarity multipliers $\mathbf{c}^k \mathbf{y}$ and $\mathbf{b}^k \mathbf{x}$ is implied by the constraints in the description of Ξ and Δ_m . Replacing the zeros in the right-hand side of the complementarity constraints with variables $\mathbf{z}$ converts $\mathcal{S}^c$ into an instance of $\mathcal{S}$, which we denote by $\bar{\mathcal{S}}^c$. It is clear that $\text{conv}(\bar{\mathcal{S}}^c)$ can be constructed using EC&R. Since $\text{conv}(\mathcal{S}^c)$ is the face of $\text{conv}(\bar{\mathcal{S}}^c)$ where $\mathbf{z} = \mathbf{0}$, it is therefore sufficient to replace $\mathbf{z}$ by $\mathbf{0}$ in the EC&R inequalities constructed for $\text{conv}(\bar{\mathcal{S}}^c)$ to obtain all nonvertical facet-defining inequalities of $\text{conv}(\mathcal{S}^c)$. The cancel-and-relax procedure of [35] applies to the special case of $\mathcal{S}^c$ where $\kappa = 1$, $m = 1$, and $h_1 = 0$. In particular, all inequalities obtained from C&R can be derived as EC&R inequalities using the procedure described above.

Remark 4. EC&R can be used to construct the convex hull, $\mathcal{P}$, of graphs of functions $g^k(\mathbf{x}; \mathbf{w}) = \mathbf{w}^\top A^k \mathbf{x}$ over $\mathbf{x} \in \Xi$ and $\mathbf{w} \in \mathcal{W}$, where $\mathcal{W}$ is a polytope with extreme points $\mathbf{w}^j$ for $j \in \{1, \dots, m+1\}$ that is not required to be a simplex. We write $\mathcal{W}$ as $T\Delta_m$, where T is an affine transformation and Δ_m is a simplex. Then we construct the convex hull, $\bar{\mathcal{P}}$, of $h^k(\mathbf{x}; \mathbf{y}) = \mathbf{y}^\top (T^\top A^k) \mathbf{x}$ over $\Xi \times \Delta_m$ using EC&R, and add $\mathbf{w} = T\mathbf{y}$ to recover the convex hull of $g^k(\mathbf{x}; \mathbf{w})$ over $\Xi \times \mathcal{W}$. We conclude that $\mathcal{P} = T\bar{\mathcal{P}}$ and $\bar{\mathcal{P}}$ can be obtained using EC&R.

3. Explicit derivation of hulls. In this section, we apply EC&R to construct an explicit description for the convex hull of the graphs of structured bilinear functions over certain polytopes. To this end, we use the following result for the case where Ξ is a Cartesian product of polytopes.

PROPOSITION 3.1. *Consider*

$$\mathcal{S}_\times = \{(\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \Xi \times \Delta_m \times \mathbb{R}^\kappa \mid \mathbf{y}^\top A^k \mathbf{x} = z_k \quad \forall k \in K\},$$

where $\Xi = \Xi^1 \times \Xi^2 \times \cdots \times \Xi^{\varrho}$, and where $\Xi^q = \{\mathbf{x}^q \in [0, 1]^{n^q} \mid E^q \mathbf{x}^q \geq \mathbf{f}^q\}$ for $q \in Q = \{1, \dots, \varrho\}$. Then, a linear description of $\text{conv}(\mathcal{S}_{\times})$ in a higher dimension is given by the inequalities describing Δ_m and Ξ , $z_k = \sum_{q \in Q} z_k^q$ for $k \in K$, together with all class-1 $^{\pm}$ EC&R inequalities of

$$\mathcal{S}_{\times}^q = \left\{ (\mathbf{x}^q; \mathbf{y}; \mathbf{z}^q) \in \Xi^q \times \Delta_m \times \mathbb{R}^{\kappa} \mid \begin{array}{l} \mathbf{y}^{\top} A^{q,k} \mathbf{x}^q = z_k^q \\ \forall k \in K \end{array} \right\}$$

for all $q \in Q$. In this definition $A^{q,k}$ is the matrix composed of columns of A^k corresponding to variables $\mathbf{x}^q$.

Further, when $\kappa = 1$, a linear description of $\text{conv}(\mathcal{S}_{\times})$ in the original space of variables is given by the inequalities describing Δ_m and Ξ , together with $z_1 \leq \sum_{q=1}^{\varrho} h^q(\mathbf{x}^q; \mathbf{y})$ and $z_1 \geq \sum_{q=1}^{\varrho} \bar{h}^q(\mathbf{x}^q; \mathbf{y})$, where $z_1^q \leq h^q(\mathbf{x}^q; \mathbf{y})$ and $z_1^q \geq \bar{h}^q(\mathbf{x}^q; \mathbf{y})$ are EC&R inequalities of $\mathcal{S}_{\times}^q$.

The proof of Proposition 3.1 is given in the appendix. We use the result of Proposition 3.1 to obtain a linear description for the convex hull of

$$\widehat{\mathcal{S}} = \left\{ (\mathbf{x}, \mathbf{y}, z) \in [0, 1]^n \times \Delta_m \times \mathbb{R} \mid \sum_{i \in N} \sum_{j \in M} a_{ij} x_i y_j = z \right\},$$

where $a_{ij} \in \mathbb{R}$ for $(i, j) \in N \times M$.

PROPOSITION 3.2. A linear description of $\text{conv}(\widehat{\mathcal{S}})$ is given by $\mathbf{1}^{\top} \mathbf{y} \leq 1$, bounds on variables $\mathbf{x}$ and $\mathbf{y}$, together with inequalities

$$(3.1a) \quad z \leq \sum_{i \in N} h^{l^i}(x_i; \mathbf{y}) \quad \forall (l^1, \dots, l^n) \in \prod_{i=1}^n (\{0\} \cup M_i^+ \cup M_i^-),$$

$$(3.1b) \quad z \geq \sum_{i \in N} \bar{h}^{l^i}(x_i; \mathbf{y}) \quad \forall (l^1, \dots, l^n) \in \prod_{i=1}^n (\{0\} \cup M_i^+ \cup M_i^-),$$

where $M_i^+ = \{j \in M \mid a_{ij} > 0\}$ and $M_i^- = \{j \in M \mid a_{ij} < 0\}$ for $i \in N$, and

(3.2a)

$$h^{l^i}(x_i; \mathbf{y}) = \begin{cases} \sum_{j \in M} a_{ij}^+ y_j & \text{if } l^i = 0, \\ \sum_{j \in M} (a_{ij} - a_{il^i})^+ y_j + a_{il^i} x_i & \text{if } l^i \in M_i^+, \\ \sum_{j \in M} (a_{ij} - a_{il^i})^+ y_j + a_{il^i} \sum_{j \in M} y_j + a_{il^i} x_i - a_{il^i} & \text{if } l^i \in M_i^-, \end{cases}$$

(3.2b)

$$\bar{h}^{l^i}(x_i; \mathbf{y}) = \begin{cases} \sum_{j \in M} a_{ij}^- y_j & \text{if } l^i = 0, \\ \sum_{j \in M} (a_{ij} - a_{il^i})^- y_j + a_{il^i} x_i & \text{if } l^i \in M_i^-, \\ \sum_{j \in M} (a_{ij} - a_{il^i})^- y_j + a_{il^i} \sum_{j \in M} y_j + a_{il^i} x_i - a_{il^i} & \text{if } l^i \in M_i^+. \end{cases}$$

Proof. We consider $\widehat{\mathcal{S}}^i = \{(x_i, \mathbf{y}, z^i) \in [0, 1] \times \Delta_m \times \mathbb{R} \mid \sum_{j \in M} a_{ij} x_i y_j = z^i\}$ and use Theorem 2.7 to obtain a linear description of $\text{conv}(\widehat{\mathcal{S}}^i)$ via class-1 $^+$ and class-1 $^-$ EC&R inequalities. These inequalities are generated from $\sum_{j \in M} a_{ij} x_i y_j = z^i$ with weights ± 1 , respectively. Since there are no linear side constraints on variables x_i , we express the assignment for each EC&R inequality as $[J, \bar{J}]$ after omitting L , $\bar{L}$, I , and $\bar{I}$. We first derive class-1 $^+$ inequalities. Since there is a single variable x_i , the EC&R assignment is either $[\emptyset, \emptyset]$, $[\{i\}, \emptyset]$, or $[\emptyset, \{i\}]$. For the first assignment,

the resulting aggregated bilinear inequality is $\sum_{j \in M} a_{ij} x_i y_j \geq z^i$, which leads to the first case of (3.2a) after relaxation. For the second assignment, inequality $x_i(1 - \sum_{j \in M} y_j) \geq 0$ is aggregated with a proper positive weight to $\sum_{j \in M} a_{ij} x_i y_j \geq z^i$ so that at least one bilinear term among those of the base inequality is canceled. The weight for $x_i(1 - \sum_{j \in M} y_j) \geq 0$ must therefore be equal to a_{il^i} for some $l^i \in M_i^+$. The resulting aggregated inequality is then $\sum_{j \in M} (a_{ij} - a_{il^i}) x_i y_j + a_{il^i} x_i \geq z^i$, which leads to the second case of (3.2a) when relaxed. Similarly, for the last assignment, inequality $(1 - x_i)(1 - \sum_{j \in M} y_j) \geq 0$ is aggregated with weight $-a_{il^i}$ for some $l^i \in M_i^-$ to $\sum_{j \in M} a_{ij} x_i y_j \geq z^i$ so that at least one bilinear term among those of the base inequality is canceled. The resulting aggregated inequality is then $\sum_{j \in M} (a_{ij} - a_{il^i}) x_i y_j + a_{il^i} \sum_{j \in M} y_j + a_{il^i} x_i - a_{il^i} \geq z^i$, which leads to the third case of (3.2a) after relaxation. Therefore, Proposition 3.1 and its following argument for the case where $\kappa = 1$ imply (3.1a). The result for (3.1b) is proven similarly. $\square$

As discussed in [19], a common way of reformulating the product of a nonnegative continuous variable and a nonnegative integer variable is to replace the integer variable by an expansion involving binary variables. In particular, given a bilinear term $x_1 w_1 = z_1$ where $x_1 \in [0, 1]$ and $w_1 \in \{0, 1, \dots, m\}$ for some $m \in \mathbb{Z}$, the unary reformulation is

$$\mathcal{U} = \left\{ (x_1; \mathbf{y}; z_1) \in [0, 1] \times \{0, 1\}^m \times \mathbb{R} \mid \sum_{j=1}^m j x_1 y_j = z_1, \mathbf{1}^\top \mathbf{y} \leq 1 \right\}.$$

It is common in the literature to relax bilinear terms $x_1 y_j$ in the description of $\mathcal{U}$ with their McCormick upper and lower bounds. This leads to an MILP reformulation of the original problem. The resulting relaxation does not describe $\text{conv}(\mathcal{U})$ when $m \geq 2$; see Proposition 2.2 in [19]. We denote by $\text{rlx}(\mathcal{U})$ the continuous relaxation of $\mathcal{U}$ where the binary restrictions on variables $\mathbf{y}$ are removed. EC&R allows for the derivation of the convex hull of $\mathcal{U}$, since it is easy to verify that $\text{conv}(\mathcal{U}) = \text{conv}(\text{rlx}(\mathcal{U}))$, and since $\text{rlx}(\mathcal{U})$ conforms to the structure of $\hat{\mathcal{S}}$ studied in Proposition 3.2 where $n = 1$ and $a_{1j} = j$. We obtain the following.

COROLLARY 3.3. *A linear description of $\text{conv}(\mathcal{U})$ is given by $y_j \geq 0$ for $j \in M$, $\mathbf{1}^\top \mathbf{y} \leq 1$, $0 \leq x_1 \leq 1$, $z_1 \leq \sum_{j=1}^m j y_j$, $z_1 \leq \sum_{j=l+1}^m (j-l) y_j + l x_1$ for all $l \in M$, $z_1 \geq 0$, and $z_1 \geq \sum_{j=1}^{l-1} (j-l) y_j + l \sum_{j=1}^m y_j + l x_1 - l$ for all $l \in M$.*

Proof. Since $n = 1$, (3.2a) reduces to $h^l = \sum_{j=1}^m j y_j$ for $l = 0$, and $h^l = \sum_{j=l+1}^m (j-l) y_j + l x_1$ for $l \in M_1^+ = M$. Similarly, (3.2b) reduces to $\bar{h}^l = 0$ for $l = 0$, and $\bar{h}^l = \sum_{j=1}^{l-1} (j-l) y_j + l \sum_{j=1}^m y_j + l x_1 - l$ for $l \in M_1^+ = M$. $\square$

4. Convex relaxations for network interdiction problems. The literature classifies NIPs based on a variety of factors including their objective function, the amount of certainty in the data, and the effect of interdiction on arcs and nodes; see [16] for a comprehensive treatment. In this section, we study a deterministic variant of NIPs defined on a network $G(V, E)$ with node set V and arc set E , and where the follower operates the network so as to maximize total flow rewards. This problem can be formulated as

$$(\mathcal{P}) \quad w^* = \min_{\mathbf{y} \in \mathcal{Y} \cap \{0, 1\}^m} \left\{ \max_{\mathbf{x} \in \mathbb{R}_+^n} \left\{ \mathbf{r}^\top \mathbf{x} \mid N\mathbf{x} = \mathbf{f}, \mathbf{x} \leq \mathbf{u} \circ (\mathbf{1} - \mathbf{y}) \right\} \right\},$$

where $n = |V|$ and $m = |E|$. In this formulation, the outer minimization and inner maximization problems are referred to as *leader* and *follower* problems, respectively.

Variable y_{ij} , for $(i, j) \in E$, represents the binary decision of the leader of whether or not arc (i, j) is destroyed/interdicted. Set $\mathcal{Y}$, which we assume to be a polyhedron, describes the set of restrictions on the leader's interdiction actions. Variable x_{ij} denotes the amount of flow transported through arc (i, j) . Further, $r_{ij} \in \mathbb{R}$ and $u_{ij} \in \mathbb{R}_+$ indicate the reward collected by the follower per unit of flow and the capacity associated with arc (i, j) , respectively. The first constraint of the follower problem imposes the typical flow-balance requirements, where $f_i \in \mathbb{R}$ is the supply/demand of node $i \in V$, and $N \in \{-1, 0, 1\}^{n \times m}$ is the node-arc incidence matrix of the network. The second constraint models the fact that an interdicted arc cannot transport flow. In this section, we use symbol $\circ$ to denote the Hadamard (componentwise) product of vectors.

The generic definition we give for $(\mathcal{P})$ captures many variants of NIPs studied in the literature, including those where the follower problem is chosen to be a maximum flow, a shortest path, or a maximum reliable path. Detailed descriptions of these special cases can be found in [23, 32, 49]. Polyhedral approaches for solving NIPs have been pursued in the literature with an aim to derive valid inequalities that can be implemented inside branch-and-cut schemes; see [4, 42] for instance. The cuts they derive apply when the follower solves a maximum flow problem. In contrast, the results we present here apply to general NIPs of the form $(\mathcal{P})$, including those for which interdiction variables $\mathbf{y}$ are continuous; see [25]. Such problems are rarely studied in the literature, probably because traditional linearization techniques relax but do not reformulate these models into MILPs. Other polyhedral results for NIPs are given in [17]. These results, however, are obtained for the KKT reformulation of $(\mathcal{P})$, which is an MILP with complementarity constraints.

The reformulation of $(\mathcal{P})$ we study here, which is sometimes referred to as *dual reformulation*, is the single level mathematical program obtained by replacing the follower problem by its linear programming dual, assuming interdiction variables are fixed; see [41] for instance. Introducing $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ as the dual variables for the first and second constraints of the follower problem, we can write the *dual reformulation* of $(\mathcal{P})$ as

$$(\mathcal{Q}) \quad w^* = \min \left\{ \mathbf{f}^\top \boldsymbol{\theta} + \mathbf{u}^\top \boldsymbol{\gamma} - \mathbf{u}^\top (\boldsymbol{\gamma} \circ \mathbf{y}) \mid \begin{array}{l} N^\top \boldsymbol{\theta} + I \boldsymbol{\gamma} \geq \mathbf{r} \\ \mathbf{0} \leq \boldsymbol{\gamma} \leq \dot{\boldsymbol{\gamma}} \\ \boldsymbol{\theta} \leq \boldsymbol{\theta} \leq \dot{\boldsymbol{\theta}} \\ \mathbf{y} \in \mathcal{Y} \cap \{0, 1\}^m \end{array} \right\},$$

where bounds $\dot{\boldsymbol{\gamma}}$, $\boldsymbol{\theta}$, and $\dot{\boldsymbol{\theta}}$ on variables $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ are added under ensuing Assumption 1; see, for instance, [17], where it is shown that $\|\mathbf{r}\|_1$ is a valid choice for $\dot{\boldsymbol{\gamma}}$, $-\boldsymbol{\theta}$, and $\dot{\boldsymbol{\theta}}$.

ASSUMPTION 1. *The follower problem is feasible for all $\mathbf{y} \in \mathcal{Y} \cap \{0, 1\}^m$.*

If the follower problem is infeasible for some $\mathbf{y} \in \mathcal{Y} \cap \{0, 1\}^m$, then it is optimal for the leader to choose such $\mathbf{y}$. For this reason, Assumption 1 is often encountered in the literature; see [41] for instance. We also note that, for many specific problems, Assumption 1 is either trivially satisfied [49] or achieved through suitable modeling, where artificial arcs that cannot be interdicted are introduced in the problem [25, 48]. We will illustrate such a modeling approach in section 5. Finally, we mention that computational strategies to verify Assumption 1 can be devised; see [17] for an example.

The dual reformulation $(\mathcal{Q})$ is commonly expressed in higher dimension through the introduction of a variable z_{ij} for each bilinear term $\gamma_{ij}y_{ij}$ that appears in the

objective function

$$\begin{aligned}
 (4.1a) \quad & \min \quad \sum_{i \in V} f_i \theta_i + \sum_{(i,j) \in E} u_{ij} \gamma_{ij} - \sum_{(i,j) \in E} u_{ij} z_{ij} \\
 (4.1b) \quad & \text{s.t.} \quad \gamma_{ij} y_{ij} = z_{ij} & \forall (i,j) \in E, \\
 (4.1c) \quad & \theta_i - \theta_j + \gamma_{ij} \geq r_{ij} & \forall (i,j) \in E, \\
 (4.1d) \quad & 0 \leq \gamma_{ij} \leq \hat{\gamma}_{ij} & \forall (i,j) \in E, \\
 (4.1e) \quad & 0 \leq \theta_i \leq \hat{\theta}_i & \forall i \in V, \\
 (4.1f) \quad & 0 \leq y_{ij} \leq 1 & \forall (i,j) \in E, \\
 (4.1g) \quad & \mathbf{y} \in \mathcal{Y}, \\
 (4.1h) \quad & y_{ij} \in \{0, 1\} & \forall (i,j) \in E.
 \end{aligned}$$

In the above model, variables θ are translated so that their lower bounds equal zero. Parameters $\hat{\theta}$ and $\mathbf{r}$ are redefined accordingly, and the constant term in the objective (due to the transformation of variables θ) is ignored. We refer to this model as $(\bar{Q})$ and to its feasible region as $\mathcal{D}$.

Many studies in the literature (see [41] for instance) replace (4.1b) with $z_{ij} \leq \gamma_{ij}$ and $z_{ij} \leq \hat{\gamma}_{ij} y_{ij}$ for $(i,j) \in E$ since the objective coefficients of z_{ij} are negative. These inequalities are sometimes referred to as *linearized inequalities*. The resulting model, which we call $(\bar{Q}^{tr})$, is a reformulation of $(\bar{Q})$, as linearized inequalities produce the exact value of the bilinear term $\gamma_{ij} y_{ij}$ when y_{ij} is binary. Since $(\bar{Q}^{tr})$ is an MILP, it can be solved with commercial linear programming-based branch-and-bound software. We have shown in [17] that the linear programming relaxation of $(\bar{Q}^{tr})$ can be very weak. In fact for the problems with nonnegative rewards, when the bounds $\hat{\gamma}_{ij}$ are large relative to r_{ij} , the optimal value of this relaxation becomes negative even though w^* is clearly nonnegative, as r is nonnegative and Assumption 1 holds.

In this section, we derive polyhedral results that provide relaxations of $(\bar{Q})$ that are stronger than those obtained through traditional linearization. The cuts we derive have the advantage of being valid for both discrete and continuous NIPs, since the convex hulls of the relaxations we study do not depend on interdiction variables being binary or continuous. We first observe that the linearized inequalities, which are used to relax (4.1b), are the facets of the convex hull of the following relaxation of $\mathcal{D}$:

$$\mathcal{D}_{(i,j)}^{(1)} = \left\{ (\gamma_{ij}, y_{ij}, z_{ij}) \in \mathbb{R}^3 \left| \begin{array}{l} \gamma_{ij} y_{ij} = z_{ij} \\ 0 \leq \gamma_{ij} \leq \hat{\gamma}_{ij} \\ 0 \leq y_{ij} \leq 1 \end{array} \right. \right\},$$

where $(i,j) \in E$. Relaxation $\mathcal{D}_{(i,j)}^{(1)}$ captures few of the interactions between variables. We next study a stronger relaxation of $\mathcal{D}$ associated with each arc (i,j) in view of obtaining reformulations with stronger linear programming relaxations.

4.1. Relaxations of single bilinear terms. For arc (i,j) of E , we consider the following relaxation of $\mathcal{D}$:

$$\mathcal{D}_{(i,j)} = \left\{ (\gamma; \theta; y_{ij}, z_{ij}) \in \mathbb{R}^{m+n} \times \mathbb{R}^2 \left| \begin{array}{l} \gamma_{ij} y_{ij} = z_{ij} \\ (4.1c) - (4.1e) \\ 0 \leq y_{ij} \leq 1 \end{array} \right. \right\}.$$

Set $\mathcal{D}_{(i,j)}$ contains the single bilinear term associated with arc (i,j) , together with all linear side constraints of $\mathcal{D}$, except for the budget requirement. Clearly, $\mathcal{D}_{(i,j)}$

contains more side constraints of $\mathcal{D}$ than $\mathcal{D}_{(i,j)}^{(1)}$. As a result, identifying facet-defining inequalities of its convex hull yields stronger linear programming relaxations of $\bar{\mathcal{Q}}$.

Next, we will assume without loss of generality that $(i, j) = (1, 2)$. Because the structure of $\mathcal{D}_{(1,2)}$ conforms to $\mathcal{S}$ with $m = \kappa = 1$, Theorem 2.7 implies the following.

COROLLARY 4.1. *A linear description of $\text{conv}(\mathcal{D}_{(1,2)})$ is given by the linear side inequalities of $\mathcal{D}_{(1,2)}$, together with class-1 $^\pm$ EC&R inequalities.*

In applying Corollary 4.1, we can use the simpler variant of EC&R presented in Remark 2, in which sets J and $\bar{J}$ are not included in the assignment but the full set of McCormick inequalities is used to relax the bilinear terms remaining in the aggregated bilinear inequality. Since $L = \bar{L} = \emptyset$, as there is only one bilinear equality, the assignment reduces to $[I, \bar{I}]$. In particular, I (resp., $\bar{I}$) refers to the constraints (4.1c) that are multiplied with y_{12} (resp., $1 - y_{12}$).

Deriving all EC&R inequalities requires identifying all possible combinations of constraint weights that meet the EC&R conditions. We show next that constraints used in the derivation of facet-defining EC&R inequalities are aggregated with weight 1.

PROPOSITION 4.2. *Consider a class-1 $^\pm$ facet-defining EC&R inequality of $\text{conv}(\mathcal{D}_{(1,2)})$ with assignment $[I, \bar{I}]$. Then, the aggregation weights of the constraints in $[I, \bar{I}]$ are equal to 0 or 1.*

Proof. We only show the result for class-1 $^+$, as the proof for class-1 $^-$ is similar. Weights used in the derivation of a facet-defining EC&R inequality of $\text{conv}(\mathcal{D}_{(1,2)})$ correspond to an extreme point of $\mathcal{C}^{1+}$. Rearranging the columns of the coefficient matrix of the system defining $\mathcal{C}^{1+}$, we obtain

$$(4.2) \quad \left[\begin{array}{c|c|c|c|c|c|c|c|c|c} N & -N & I & I & -I & -I & 0 & 0 & 0 & 0 \\ \hline I & -I & 0 & 0 & 0 & 0 & I & I & -I & -I \end{array} \right],$$

where N is the node-arc incidence matrix of the network, the first and second row blocks correspond to bilinear terms $\theta_i y_{12}$ and $\gamma_{ij} y_{12}$, respectively, while the first and second columns correspond to the weights of constraints multiplied by y_{12} and $1 - y_{12}$, respectively. Further, the right-hand side of the system defining $\mathcal{C}^{1+}$ is $-\mathbf{e}^{12} \in \mathbb{R}^{m+n}$, where $\mathbf{e}^{12}$ is the vector whose components are all zero except for that corresponding to $\gamma_{12} y_{12}$, which is equal to 1. Let B be a basis for (4.2). It follows from Cramer's rule that all elements of B^{-1} belong to $\{0, -1, 1\}$, since (4.2) is totally unimodular (TU). Therefore, the components of $-B^{-1}\mathbf{e}^{12}$ belong to $\{0, -1, 1\}$. We conclude that the components of basic feasible solutions to $\mathcal{C}^{1+}$ are equal to 0 or 1. $\square$

Remark 5. When sets I or $\bar{I}$ contain constraints that are aggregated with weight 0, we can reduce the assignment by dropping those constraints. In the remainder of this section, we only consider assignments $[I, \bar{I}]$ whose aggregation weights are equal to 1.

Our next goal is to provide, in Proposition 4.6, a characterization of all class-1 $^\pm$ EC&R inequalities. We first obtain class-1 $^\pm$ EC&R inequalities with assignment $[\emptyset, \emptyset]$. Using the fact that each bilinear term in the aggregated inequality is relaxed using McCormick inequalities, as stated after Corollary 4.1, a direct application of EC&R yields the following.

COROLLARY 4.3. *Class-1 $^+$ EC&R inequalities for $\mathcal{D}_{(1,2)}$ with assignment $[\emptyset, \emptyset]$ are $z_{12} \leq \gamma_{12}$ and $z_{12} \leq \hat{\gamma}_{12} y_{12}$, while class-1 $^-$ EC&R inequalities with assignment $[\emptyset, \emptyset]$ are $z_{12} \geq 0$ and $z_{12} \geq \gamma_{12} + \hat{\gamma}_{12} y_{12} - \hat{\gamma}_{12}$.*

The inequalities described in Corollary 4.3 are the well-known linearized inequalities used in the description of $(\bar{Q}^{tr})$. Before characterizing class-1 $^\pm$ EC&R inequalities with nonempty assignment $[I, \bar{I}]$, we introduce the following.

DEFINITION 4.4. *Given a directed subgraph $\mathbf{S}$ of a directed graph $G(V, E)$, we refer to $\bar{\mathbf{S}}$ as the undirected graph obtained by transforming the arcs of $\mathbf{S}$ into edges.*

DEFINITION 4.5. *In $G(V, E)$, we refer to $\mathbf{P}$ as a $(1, 2)$ -path if it is a directed subgraph of G that contains arc $(1, 2)$, and $\bar{\mathbf{P}}$ is a path. Similarly, we refer to $\mathbf{P}$ as a $(1, 2)$ -cycle if it is a directed subgraph of G that contains arc $(1, 2)$, and $\bar{\mathbf{P}}$ is a cycle.*

Given a $(1, 2)$ -path/cycle $\mathbf{P}$ of G , we refer to the number of its arcs as $|\mathbf{P}|$. We record the fact that (4.1c), for $(i, j) \in E$, belongs to I or $\bar{I}$ by including the pair $\langle i, j \rangle$ in the corresponding set. Given a $(1, 2)$ -path (resp., $(1, 2)$ -cycle) $\mathbf{P}$, we refer to $\bar{\mathbf{P}}$ as the unique directed path (resp., cycle) that can be obtained from $\mathbf{P}$ by reversing (if necessary) the orientation of arcs $(i, j) \neq (1, 2)$ of $\mathbf{P}$. We say that arc $(i, j) \neq (1, 2)$ of $\mathbf{P}$ is *backward* if it is reversed in $\bar{\mathbf{P}}$. Otherwise, we say that it is *forward*.

PROPOSITION 4.6. *Consider a facet-defining class-1 $^\pm$ EC&R inequality of $\text{conv}(\mathcal{D}_{(1,2)})$ with assignment $[I, \bar{I}]$, where $I \cup \bar{I} \neq \emptyset$. Define $\mathbf{P}$ to be the directed subgraph of $G(V, E)$ induced by the arcs (i, j) for which $\langle i, j \rangle$ belongs to $I \cup \bar{I}$. Then $\mathbf{P}$ is a $(1, 2)$ -path or a $(1, 2)$ -cycle satisfying (i) $\langle 1, 2 \rangle \in \bar{I}$ for class 1 $^+$, and $\langle 1, 2 \rangle \in I$ for class 1 $^-$, and (ii) forward arcs of $\mathbf{P}$ belong to I if and only if $\langle 1, 2 \rangle \in I$, and backward arcs of $\mathbf{P}$ belong to I if and only if $\langle 1, 2 \rangle \in \bar{I}$.*

Proof. Since I and $\bar{I}$ are not both empty by assumption, requirement (ii) of EC&R implies that $\gamma_{12}y_{12}$ be canceled during aggregation since it is the only bilinear term contained in base equality $\gamma_{12}y_{12} = z_{12}$. This shows that $\mathbf{P}$ contains arc $(1, 2)$. Since γ_{12} appears only in the dual constraint (4.1c) associated with arc $(1, 2)$, condition (i) must be satisfied for $\gamma_{12}y_{12}$ to be canceled. We now show that $\mathbf{P}$ has a single (weakly connected) component. Assume by contradiction that there exists a component, $\mathbf{P}_1$, disjoint from the one that contains arc $(1, 2)$. Similarly, we refer to the collection of remaining components as $\mathbf{P}_2$. It follows from Remark 5 that the weights of the constraints corresponding to arcs in $\mathbf{P}_1$ and $\mathbf{P}_2$ are equal to 1. Since the EC&R inequality under consideration is facet-defining, weights correspond to components of a bfs of the system of equations that describes $\mathcal{C}^{1^\pm}$. The equations defining basic variables can be written as

$$(4.3) \quad \left[\begin{array}{ccc|ccc|c} N_1 & A_1 & 0 & 0 & 0 & 0 & C_1 \\ B_1 & 0 & \bar{A}_1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & N_2 & A_2 & 0 & C_2 \\ 0 & 0 & 0 & B_2 & 0 & \bar{A}_2 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & C_3 \end{array} \right] \begin{bmatrix} 1 \\ \nu_1 \\ \bar{\nu}_1 \\ 1 \\ \nu_2 \\ \bar{\nu}_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \pm e^{12} \\ 0 \end{bmatrix},$$

where the columns and rows of the basis matrix have been suitably reordered. In (4.3), the first row block (resp., third row block) corresponds to bilinear terms $\theta_i y_{12}$ for nodes i that belong to $\mathbf{P}_1$ (resp., $\mathbf{P}_2$), while the second row block (resp., fourth row block) corresponds to bilinear terms $\gamma_{ij} y_{12}$ for arcs (i, j) that belong to $\mathbf{P}_1$ (resp., $\mathbf{P}_2$). Similarly, the first column block (resp., fourth column block) represents positive or negative multiples of dual network constraints corresponding to arcs in $\mathbf{P}_1$ (resp., $\mathbf{P}_2$), while the second, third, fifth, and sixth column blocks contain positive or negative multiples of columns of identity matrices. The last row block corresponds to the

remaining bilinear terms, while the last column block corresponds to dual constraints that have weights 0 and are added to complete the basis. Further, e^{12} is a unit vector whose elements are all zeros except that corresponding to $\gamma_{12}y_{12}$, which is equal to 1. It is now easy to verify that the linear combination of the columns of the basis matrix with weights $(\mathbf{1}; \nu_1; \bar{\nu}_1; \mathbf{0}; \mathbf{0}; \mathbf{0}; \mathbf{0})$ yields the zero vector. This shows that the columns are linearly dependent, which is a contradiction.

Requirement (i) of EC&R implies that at least $|I| + |\bar{I}|$ bilinear terms cancel during aggregation. Counting the cancellation of $\gamma_{12}y_{12}$, there must be at least $|I| + |\bar{I}| - 1 = |P| - 1$ bilinear terms canceled. In the description of $\mathcal{D}_{(1,2)}$ variables γ_{ij} appear in a single dual network constraint, and therefore no combination of dual network constraints leads to the cancellation of terms $\gamma_{ij}y_{12}$ for $(i, j) \neq (1, 2)$. As a result, only bilinear terms $\theta_i y_{12}$ can be canceled. To cancel $\theta_i y_{12}$, at least two dual constraints that contain variable θ_i must be used during aggregation. This implies that nodes of P leading to cancellation must have degree at least 2 in P . We next consider all possible structures for $\bar{P}$.

First, assume that $\bar{P}$ contains no cycles, i.e., $\bar{P}$ is a tree. Let $\deg(i)$ represent degree of node i in $\bar{P}$, and let l and k be the number of leaf and nonleaf nodes of $\bar{P}$, respectively. We write

$$2|\bar{P}| = \sum_{i: \text{node of } \bar{P}} \deg(i) = \sum_{i: \text{leaf}} \deg(i) + \sum_{i: \text{nonleaf}} \deg(i) \geq l + 2k \geq l + 2(|\bar{P}| - 1),$$

where the last inequality follows from the cancellation requirement of EC&R. The above chain of relations shows that $l \leq 2$. Since $\bar{P}$ is a tree, $l \geq 2$ as $|\bar{P}| \geq 1$. We conclude that $l = 2$, which shows that $\bar{P}$ is a path. Since P contains arc $(1, 2)$, it is a $(1, 2)$ -path. Finally, to cancel bilinear term $\theta_i y_{12}$ during aggregation, the two dual constraints containing θ_i must be members of I or $\bar{I}$ so that this bilinear term appears with opposite signs and equal weights in the aggregated inequality, which is captured by condition (ii).

Second, assume that $\bar{P}$ contains cycles. We first show that it has exactly one cycle and that this cycle contains arc $(1, 2)$. Assume by contradiction that there exists a cycle $\bar{P}_2$ in $\bar{P}$ that does not contain $(1, 2)$. We denote the directed subgraph of P associated with $\bar{P}_2$ by P_2 . We refer to the subgraph of P obtained by removing P_2 as P_1 . After suitable reordering of its columns and rows, the basis matrix of $\mathcal{C}^{1\pm}$ can be written as

$$(4.4) \quad \left[\begin{array}{cccccc|c} N_0 & N_1 & N_2 & A & 0 & 0 & 0 \\ B_0 & 0 & 0 & 0 & \bar{A}_0 & 0 & 0 \\ 0 & B_1 & 0 & 0 & 0 & \bar{A}_1 & 0 \\ \hline 0 & 0 & B_2 & 0 & 0 & 0 & \bar{A}_2 \end{array} \right].$$

In (4.4), the first row block corresponds to bilinear terms $\theta_i y_{12}$, while the third row block (resp., fourth row block) corresponds to bilinear terms $\gamma_{ij} y_{12}$ for arcs (i, j) that belong to P_1 (resp., P_2). The second row block corresponds to bilinear terms $\gamma_{ij} y_{12}$ for arcs that do not belong to P_1 or P_2 . Further, N_1 and N_2 are matrices whose columns are ± 1 multiples of the columns of node-arc incidence matrices corresponding to P_1 and P_2 , respectively. The first column block corresponds to dual constraints with zero weights that are added to complete the basis. Matrix B_2 is square and composed of columns of I and $-I$, and so is $\bar{A}_2$. Since (4.4) is a basis, the submatrix obtained by eliminating its last column and last row blocks must be nonsingular. However, since

N_2 corresponds to the node-arc incidence matrix of a cycle, its columns are linearly dependent. This is a contradiction. This shows that all cycles in $\bar{P}$ contain arc $(1, 2)$. Furthermore, observe that if $\bar{P}$ has two distinct cycles containing $(1, 2)$, then it also has a cycle that does not contain $(1, 2)$, which is a contradiction. We conclude that $\bar{P}$ has a single cycle and that this cycle contains arc $(1, 2)$.

To complete the proof we next argue that $\bar{P}$ does not contain leaves. Assume by contradiction that it contains l leaves where $l \geq 1$. Let k represent the number of nonleaf nodes in $\bar{P}$. On the one hand, in order to cancel a sufficient number of bilinear terms during aggregation we must have that $k \geq |\bar{P}| - 1$. On the other hand, consider any spanning tree $\bar{T}$ of $\bar{P}$. It holds that $l + k = |\bar{T}| + 1 \leq |\bar{P}|$, where equality holds because a tree has one less arc than it has nodes, and the inequality follows from the fact that $\bar{T} \subsetneq \bar{P}$. Therefore, we conclude that $|\bar{P}| - 1 \leq k \leq |\bar{P}| - l$. Since $l \geq 1$, equality must hold throughout, showing that $l = 1$. Therefore, $\bar{P}$ is the union of a cycle containing arc $(1, 2)$ and a single path. We refer to the node where they meet as t . Clearly, the degree of node t is 3. Since Proposition 4.2 establishes the fact that the aggregation weights of the three constraints associated with arcs connected to t are equal to 1, the corresponding bilinear term $\theta_t y_{12}$ does not cancel. Consequently, the cancellation requirement of EC&R is violated, yielding the desired contradiction. We conclude that $l = 0$, and therefore $\bar{P}$ is a single cycle that contains arc $(1, 2)$, showing that P is a $(1, 2)$ -cycle. It is then easy to verify that in order for bilinear terms corresponding to the nodes of this $(1, 2)$ -cycle to cancel, condition (ii) must be satisfied. $\square$

The following partial converse to Proposition 4.6 can easily be established.

LEMMA 4.7. *For any $(1, 2)$ -path or $(1, 2)$ -cycle in $G(V, E)$, we can construct a class-1 $^\pm$ EC&R inequality with assignment $[I, \bar{I}]$ satisfying conditions (i) and (ii) of Proposition 4.6.*

We next describe the form of the aggregated bilinear inequalities obtained for $(1, 2)$ -paths/cycles of $G(V, E)$. In particular, the class-1 $^\pm$ inequalities corresponding to $(1, 2)$ -cycle P with assignment $[I, \bar{I}]$ are obtained by relaxing the aggregated bilinear inequality

$$(4.5) \quad \sum_{\langle i, j \rangle \in I \setminus \{(1, 2)\}} \gamma_{ij} y_{12} - \sum_{\langle i, j \rangle \in \bar{I} \setminus \{(1, 2)\}} \gamma_{ij} y_{12} + \left(\sum_{\langle i, j \rangle \in I} r_{ij} - \sum_{\langle i, j \rangle \in \bar{I}} r_{ij} \right) y_{12} + \sum_{\langle i, j \rangle \in \bar{I}} (\theta_i - \theta_j + \gamma_{ij}) \mp z_{12} \geq \sum_{\langle i, j \rangle \in \bar{I}} r_{ij},$$

using McCormick inequalities. Observe that we use McCormick inequalities to linearize $\gamma_{ij} y_{12}$ terms for $(i, j) \neq (1, 2)$ since such terms appear in (4.5), while they do not appear in (4.1b). The coefficient of variable z_{12} reflects the class sign. Each bilinear inequality (4.5) generated from a $(1, 2)$ -cycle P can be relaxed into $2^{|\bar{P}|-1}$ linear inequalities. We refer to these inequalities as *cycle inequalities*. Similarly, the class-1 $^\pm$ inequalities corresponding to $(1, 2)$ -path P and assignment $[I, \bar{I}]$ are obtained by

relaxing the aggregated bilinear inequality

$$(4.6) \quad \sum_{i \in \{l_1, l_2\}} \text{sgn}(i) \theta_i y_{12} + \sum_{\langle i, j \rangle \in I \setminus \{1, 2\}} \gamma_{ij} y_{12} - \sum_{\langle i, j \rangle \in \bar{I} \setminus \{1, 2\}} \gamma_{ij} y_{12} \\ + \left(\sum_{\langle i, j \rangle \in \bar{I}} r_{ij} - \sum_{\langle i, j \rangle \in I} r_{ij} \right) y_{12} + \sum_{\langle i, j \rangle \in \bar{I}} (\theta_i - \theta_j + \gamma_{ij}) \mp z_{12} \geq \sum_{\langle i, j \rangle \in \bar{I}} r_{ij}.$$

In (4.6), l_1 and l_2 are leaf nodes of P , $\text{sgn}(i) = 1$ if $\langle i, j \rangle \in I$ or $\langle j, i \rangle \in \bar{I}$ for some j , and $\text{sgn}(i) = -1$ otherwise. Each bilinear inequality (4.6) generated from a $(1, 2)$ -path P can be relaxed into $2^{|P|+1}$ linear inequalities, which we refer to as *path inequalities*.

Corollary 4.3, Proposition 4.6, and Lemma 4.7 yield the following result.

THEOREM 4.8. *A description of $\text{conv}(\mathcal{D}_{(1,2)})$ is given by the linear inequalities in the description of $\mathcal{D}_{(1,2)}$, those in Corollary 4.3, and all the path and cycle inequalities.*

Example 2. Consider the network represented in Figure 1.

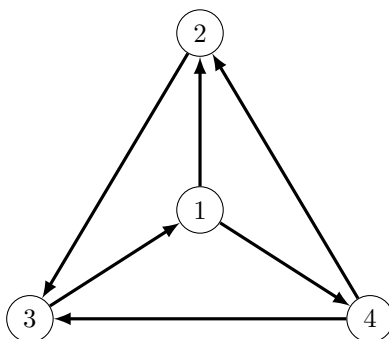


FIG. 1. Network of Example 2.

To derive all class-1 $^\pm$ EC&R inequalities with nonempty assignments, we enumerate all $(1, 2)$ -paths/cycles in the network and compute their corresponding EC&R assignments for class-1 $^+$; see Table 2. Assignments for class-1 $^-$ are obtained by flipping the role of I and $\bar{I}$. To illustrate the derivation procedure, we obtain the aggregated bilinear inequalities for one path and one cycle in the list. First, consider the $(1, 2)$ -path composed of arcs $(1, 2) - (4, 2) - (4, 3)$. To obtain the corresponding assignment for class-1 $^+$, condition (i) of Proposition 4.6 requires that $\langle 1, 2 \rangle \in \bar{I}$. Further, since arc $(4, 2)$ is backward and $(4, 3)$ is forward, condition (ii) assigns $\langle 4, 2 \rangle$ to I and $\langle 4, 3 \rangle$ to $\bar{I}$. We obtain the assignment $[I, \bar{I}] = [\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle, \langle 4, 3 \rangle\}]$. Using (4.6), we write

$$-\theta_1 y_{12} + \theta_3 y_{12} + \gamma_{42} y_{12} - \gamma_{43} y_{12} \\ + (r_{12} + r_{43} - r_{42}) y_{12} + (\theta_1 - \theta_2 + \gamma_{12}) + (\theta_4 - \theta_3 + \gamma_{43}) - z_{12} \geq r_{12} + r_{43}.$$

This bilinear inequality can be relaxed into 2^4 linear inequalities using McCormick upper bounds for bilinear terms $\theta_3 y_{12}$ and $\gamma_{42} y_{12}$, and McCormick lower bounds for $\theta_1 y_{12}$ and $\gamma_{43} y_{12}$.

We next consider the $(1, 2)$ -cycle composed of arcs $(1, 2) - (4, 2) - (4, 3) - (3, 1)$. Since this cycle contains the path $(1, 2) - (4, 2) - (4, 3)$, it suffices to observe that

TABLE 2
Class-1⁺ EC&R assignments for Example 2.

P	Arcs of P	Assignments
(1, 2)-path	(1, 2)	$[\emptyset, \{\langle 1, 2 \rangle\}]$
	(1, 2) – (2, 3)	$[\emptyset, \{\langle 1, 2 \rangle, \langle 2, 3 \rangle\}]$
	(1, 2) – (4, 2)	$[\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle\}]$
	(1, 2) – (3, 1)	$[\emptyset, \{\langle 1, 2 \rangle, \langle 3, 1 \rangle\}]$
	(1, 2) – (1, 4)	$[\{\langle 1, 4 \rangle\}, \{\langle 1, 2 \rangle\}]$
	(1, 2) – (2, 3) – (4, 3)	$[\{\langle 4, 3 \rangle\}, \{\langle 1, 2 \rangle, \langle 2, 3 \rangle\}]$
	(1, 2) – (4, 2) – (4, 3)	$[\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle, \langle 4, 3 \rangle\}]$
	(1, 2) – (3, 1) – (4, 3)	$[\emptyset, \{\langle 1, 2 \rangle, \langle 3, 1 \rangle, \langle 4, 3 \rangle\}]$
	(1, 2) – (1, 4) – (4, 3)	$[\{\langle 1, 4 \rangle, \langle 4, 3 \rangle\}, \{\langle 1, 2 \rangle\}]$
	(4, 2) – (1, 2) – (3, 1)	$[\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle, \langle 3, 1 \rangle\}]$
	(1, 4) – (1, 2) – (2, 3)	$[\{\langle 1, 4 \rangle\}, \{\langle 1, 2 \rangle, \langle 2, 3 \rangle\}]$
(1, 2)-cycle	(1, 2) – (2, 3) – (3, 1)	$[\emptyset, \{\langle 1, 2 \rangle, \langle 2, 3 \rangle, \langle 3, 1 \rangle\}]$
	(1, 2) – (1, 4) – (4, 2)	$[\{\langle 1, 4 \rangle, \langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle\}]$
	(1, 2) – (2, 3) – (4, 3) – (1, 4)	$[\{\langle 4, 3 \rangle, \langle 1, 4 \rangle\}, \{\langle 1, 2 \rangle, \langle 2, 3 \rangle\}]$
	(1, 2) – (4, 2) – (4, 3) – (3, 1)	$[\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle, \langle 4, 3 \rangle, \langle 3, 1 \rangle\}]$

arc (3, 1) is forward to obtain the assignment $[I, \bar{I}] = [\{\langle 4, 2 \rangle\}, \{\langle 1, 2 \rangle, \langle 4, 3 \rangle, \langle 3, 1 \rangle\}]$. Using (4.5), we write

$$\begin{aligned} & \gamma_{42}y_{12} - \gamma_{43}y_{12} - \gamma_{31}y_{12} + (r_{12} + r_{43} + r_{31} - r_{42})y_{12} \\ & + (\theta_1 - \theta_2 + \gamma_{12}) + (\theta_4 - \theta_3 + \gamma_{43}) + (\theta_3 - \theta_1 + \gamma_{31}) - z_{12} \geq r_{12} + r_{43} + r_{31}. \end{aligned}$$

This bilinear inequality can be relaxed into 2^3 linear inequalities using McCormick upper bounds for bilinear term $\gamma_{42}y_{12}$, and McCormick lower bounds for $\gamma_{43}y_{12}$ and $\gamma_{31}y_{12}$.

As argued in section 2, our intuition is that inequalities obtained using fewer aggregations during EC&R are stronger than those obtained with large numbers of aggregations. This insight would therefore imply that the most useful inequalities in the description of $\text{conv}(\mathcal{D}_{(1,2)})$ in Theorem 4.8 are those that correspond to small paths and small cycles. This intuition is corroborated in section 5 where small-path and small-cycle inequalities are shown to account for most of the gap reduction that can be achieved by adding all path and all cycle inequalities.

4.2. Simultaneous relaxation of multiple bilinear terms. The relaxation $\mathcal{D}_{(1,2)}$ we studied in section 4.1 has a single bilinear term. Its convex hull is governed by the paths and cycles of the network. If a second bilinear term was to be added to the model, we could generate inequalities using EC&R on the second bilinear term, where side constraints are the paths and cycles inequalities obtained from using EC&R on the first bilinear term. This sequential application of EC&R would yield inequalities that are cognizant of more than a single bilinear term but still inherit the path and cycle structure present in the inequalities of $\text{conv}(\mathcal{D}_{(1,2)})$. Since variables z appear with negative coefficients in the objective function (4.1a), we are interested in cutting planes that provide upper bounds on these variables. For this reason, we focus on class-1⁺ path/cycle inequalities as the input for a procedure that generates stronger families of valid inequalities by combining path/cycle inequalities with weights y_j or $1 - y_j$ so that sufficiently many bilinear terms cancel.

PROPOSITION 4.9. Consider the path (resp., cycle) inequality

$$(4.7) \quad \sum_{(i,j) \in E} \alpha_{ij} \gamma_{ij} + \sum_{i \in N} \beta_i \theta_i + \mu y_{12} - z_{12} \geq \eta$$

associated with a $(1, 2)$ -path (resp., $(1, 2)$ -cycle) P . Let $\mathcal{F}_0$ be the subset of forward arcs (not including $(1, 2)$) whose associated bilinear terms $-\gamma_{ij} y_{12}$ were relaxed into 0 when obtaining (4.7). Then, for any subset $\mathcal{F}' \subseteq \mathcal{F}_0$, inequality

$$(4.8) \quad \sum_{(i,j) \in E} \alpha_{ij} \gamma_{ij} + \sum_{i \in N} \beta_i \theta_i + \mu y_{12} + \mu^+ \sum_{(i,j) \in \mathcal{F}'} y_{ij} - z_{12} - \sum_{(i,j) \in \mathcal{F}'} z_{ij} \geq \eta$$

is valid for $\text{conv}(\mathcal{D})$.

Proof. We prove the result by induction on the cardinality of $\mathcal{F}'$. When $|\mathcal{F}'| = 0$, (4.8) reduces to (4.7). The basis of induction is therefore verified, as path/cycle inequalities are valid for $\text{conv}(\mathcal{D})$. Assume now that the result has been established for all $\mathcal{F}' \subseteq \mathcal{F}_0$ with $|\mathcal{F}'| \leq \varsigma < |\mathcal{F}_0|$. We prove that the result holds for $\mathcal{F}'' \subseteq \mathcal{F}_0$ with $|\mathcal{F}''| = \varsigma + 1$. Fix $(k, l) \in \mathcal{F}''$ and define $\mathcal{F}' = \mathcal{F}'' \setminus \{(k, l)\}$. It is easy to verify that

$$(4.9) \quad \sum_{(i,j) \in E} \alpha_{ij} \gamma_{ij} + \sum_{i \in N} \beta_i \theta_i + \mu y_{kl} - z_{kl} \geq \eta$$

is valid for $\text{conv}(\mathcal{D}_{(k,l)})$ since it is the path/cycle inequality associated with P with respect to forward arc (k, l) , where bilinear terms $-\gamma_{ij} y_{kl}$ for all $(i, j) \in \{(1, 2)\} \cup \mathcal{F}_0 \setminus \{(k, l)\}$ are relaxed into 0. Consider now

$$(4.10) \quad \sum_{(i,j) \in E} \alpha_{ij} \gamma_{ij} + \sum_{i \in N} \beta_i \theta_i + \mu y_{kl} - z_{kl} - \sum_{(i,j) \in \mathcal{F}' \cup \{(1,2)\}} \gamma_{ij} y_{kl} \geq \eta,$$

which is obtained by applying the same relaxation on the aggregated bilinear inequality (4.5) or (4.6) that yields (4.9) with the difference that bilinear terms $-\gamma_{ij} y_{kl}$ for $(i, j) \in \mathcal{F}' \cup \{(1, 2)\}$ are not relaxed to 0. This bilinear inequality is valid for $\text{conv}(\mathcal{D}_{(k,l)})$ and therefore for $\text{conv}(\mathcal{D})$. Now consider the conic combination of (4.1b) for all $(i, j) \in \mathcal{F}' \cup \{(1, 2)\}$ with weights y_{kl} , together with (4.8) and (4.10) with weights $(1 - y_{kl})$ and y_{kl} , respectively, while replacing y_{kl}^2 by y_{kl} and $z_{kl} y_{kl}$ by z_{kl} since these two terms are equal for solutions of $\mathcal{D}$ where $y_{kl} \in \{0, 1\}$. We obtain

$$\begin{aligned} & \sum_{(i,j) \in E} \alpha_{ij} \gamma_{ij} + \sum_{i \in N} \beta_i \theta_i + \mu y_{12} + \mu^+ \sum_{(i,j) \in \mathcal{F}'} y_{ij} - z_{12} - \sum_{(i,j) \in \mathcal{F}''} z_{ij} \\ & + \mu y_{kl}(1 - y_{12}) - \mu^+ \sum_{(i,j) \in \mathcal{F}'} y_{ij} y_{kl} - \sum_{(i,j) \in \mathcal{F}' \cup \{(1,2)\}} y_{kl} \gamma_{ij} (1 - y_{ij}) \geq \eta. \end{aligned}$$

Relaxing bilinear terms $-\mu^+ y_{ij} y_{kl}$ and trilinear terms $-y_{kl} \gamma_{ij} (1 - y_{ij})$ to 0, while relaxing $\mu y_{kl}(1 - y_{12})$ to $\mu^+ y_{kl}$, yields (4.8) for $\mathcal{F}''$. This inequality is therefore valid for $\text{conv}(\mathcal{D})$, thereby proving the inductive hypothesis. $\square$

We refer to (4.8) as *multipath/-cycle* inequalities. It follows from the proof of Proposition 4.9 that multipath/-cycle inequalities are valid for the set described by (4.1b)–(4.1f) and (4.1h). It is clear that the convex hull of (4.1b)–(4.1f) and (4.1h) is identical to that of (4.1b)–(4.1f), as extreme points of the latter set assume binary values for variables $\mathbf{y}$. We conclude that, similarly to path/cycle inequalities, multipath/-cycle inequalities are valid both for discrete and continuous NIPs.

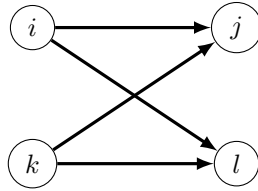


FIG. 2. Network structure for Example 3.

Example 3. For the network structure depicted in Figure 2, the class-1⁺ aggregated inequality (4.5) associated with the (i, j) -cycle $P = (i, j) - (k, j) - (k, l) - (i, l)$ is

(4.11)

$$\gamma_{il}y_{ij} + \gamma_{kj}y_{ij} - \gamma_{kl}y_{ij} + \mu y_{ij} + (\theta_i - \theta_j + \gamma_{ij}) + (\theta_k - \theta_l + \gamma_{kl}) - z_{ij} \geq r_{ij} + r_{kl},$$

where $\mu = (r_{ij} + r_{kl} - r_{il} - r_{kj})$. Let $\mathcal{F}_0 = \{(k, l)\}$ be the subset of forward arcs of P not including (i, j) . Proposition 4.9 allows us to derive multicycle inequalities for any of the four cycle inequalities corresponding to (4.11), where $-\gamma_{kl}y_{ij}$ is relaxed into 0 and $\gamma_{il}y_{ij}$ and $\gamma_{kj}y_{ij}$ are relaxed using McCormick upper bounds. For instance, consider the cycle inequality

$$\gamma_{il} + \gamma_{kj} + \mu y_{ij} + (\theta_i - \theta_j + \gamma_{ij}) + (\theta_k - \theta_l + \gamma_{kl}) - z_{ij} \geq r_{ij} + r_{kl}.$$

Proposition 4.9 yields the multicycle inequality

$$\gamma_{il} + \gamma_{kj} + \mu y_{ij} + \mu^+ y_{kl} + (\theta_i - \theta_j + \gamma_{ij}) + (\theta_k - \theta_l + \gamma_{kl}) - z_{ij} - z_{kl} \geq r_{ij} + r_{kl}.$$

There are exponentially many families of inequalities (4.8), each composed of exponentially many members. This follows from the facts that (i) there are exponentially many path/cycle inequalities that can be obtained by relaxing (4.5) or (4.6) into (4.7); and (ii) for each path/cycle inequality, there are exponentially many subsets $\mathcal{F}$ of forward arcs in the given path/cycle. Given an (i, j) -path/cycle inequality (4.7) and a point $(\gamma^*; \theta^*; y^*; z^*)$, we can determine the most violated inequality of the form (4.8) in linear time as follows. If $\mu \leq 0$, then we set $\mathcal{F}' = \mathcal{F}_0$. If $\mu > 0$, then we set $\mathcal{F}' = \{(k, l) \in \mathcal{F}_0 \mid \mu y_{kl}^* - z_{kl}^* < 0\}$. Proposition 4.9 therefore provides a rich collection of valid inequalities that contain multiple variables z . We show in section 5 that these inequalities are very useful in our computational experiments.

5. Computational results. In this section, we present numerical results aimed at evaluating the computational potential of the results we developed in section 4 on families of discrete NIPs. In particular, we study the strength of linear programming relaxations of $(\mathcal{Q})$ obtained after adding certain (multi)path and (multi)cycle inequalities described in Propositions 4.6 and 4.9. Computational results are obtained on a machine with a Windows 8 (64-bit) operating system, 8 GB RAM, and 2.20 GHz Core i7 CPU. The implementation is done in VC++ with CPLEX 12.5.

5.1. Test instances. We consider an NIP where the follower solves a transportation problem on the complete bipartite network $G(V_1 \cup V_2, E)$. The follower can be thought of as a manufacturer with $|V_1|$ plants and $|V_2|$ markets. Each plant $i \in V_1$ has a given output (f_i) and each market $j \in V_2$ has a given demand $(-f_j)$. We assume that the output of each plant can be increased at penalty p_i^+ and that any unit

not transported receives a penalty of p_i^- . Similarly, we assume that the base demand of each market can be increased or decreased by reducing prices or by appropriately compensating shorted markets at penalties p_j^+ and p_j^- , respectively. Let u_{ij} be the maximum amount that can be transported from plant i to market j , and let r_{ij} be the follower's profit from each such unit. The objective of the leader is to destroy up to B arcs of the network so as to minimize the maximum reward the follower can achieve. The follower's problem can therefore be formulated as

$$\max \left\{ \begin{array}{l} \sum_{(i,j) \in E} r_{ij} x_{ij} - \sum_{i \in V_1} (p_i^- s_i^- + p_i^+ s_i^+) \\ \quad - \sum_{j \in V_2} (p_j^- s_j^- + p_j^+ s_j^+) \end{array} \middle| \begin{array}{ll} \sum_{j \in V_2} x_{ij} - s_i^+ + s_i^- = f_i & \forall i \in V_1 \\ -\sum_{i \in V_1} x_{ij} - s_j^- + s_j^+ = f_j & \forall j \in V_2 \\ 0 \leq x_{ij} \leq u_{ij}(1 - y_{i,j}) & \forall (i,j) \in E \\ s_i^+ \geq 0, s_i^- \geq 0 & \forall i \in V_1 \\ s_j^+ \geq 0, s_j^- \geq 0 & \forall j \in V_2 \end{array} \right\}.$$

We observe that (i) the follower's problem is feasible for all interdiction scenarios; (ii) its dual contains explicit bounds on the variables θ_i and θ_j (i.e., $-p_i^- \leq \theta_i \leq p_i^+$ and $-p_j^+ \leq \theta_j \leq p_j^-$); and (iii) this dual is exactly of the form $(\mathcal{Q})$, so that the results of section 4 apply. In our computation, we therefore use $\dot{\theta}_i = p_i^+$, $\theta_i = -p_i^-$, $\dot{\theta}_j = p_j^-$, and $\theta_j = -p_j^+$. Assuming $p_j^+, p_i^- \geq 0$, it is easy to show that $\dot{\gamma}_{ij} = \dot{\theta}_j - \theta_i + r_{ij}$ is an upper bound for γ_{ij} for each $(i,j) \in E$; see [17].

We generate 10 instances of the problem described above as follows. We choose $|V_1| = |V_2| = n = 16$ and $B = \frac{n}{2}$. Reward r_{ij} and capacity u_{ij} are randomly generated from the uniform discrete distributions $[1, 50]$ and $[10, 50]$, respectively. We set $p_i^+ = p_i^- = \max_{j \in V_2} r_{ij}$ and $p_j^+ = p_j^- = \max_{i \in V_1} r_{ij}$. This choice ensures that the penalty for each node matches the highest possible reward that can be achieved from the node, thereby yielding instances where the network operator has an incentive to use actual supply to meet actual demand. To generate the supply/demand vector $\mathbf{f}$, we use the following procedure that balances supply with demand and sets them so that the capacity of a fraction of outgoing (resp., incoming) arcs is used up if the entire supply is exhausted (resp., demand is met). More precisely, for each node $i \in V_1$ (resp., $j \in V_2$), we first select an interval $[L_i, U_i]$ (resp., $[L_j, U_j]$) that will contain f_i (resp., $-f_j$). We compute U_i (resp., U_j) as the sum of the capacities of the $\frac{n}{2}$ outgoing arcs of node i (resp., incoming arcs of node j) with smallest capacities. We compute L_i (resp., L_j) as the sum of the capacities of the $\lfloor \frac{n}{5} \rfloor$ outgoing arcs of node i (resp., incoming arcs of node j) with smallest capacities. We generate $\tilde{f}_i$ (resp., $-\tilde{f}_j$) from the (discrete) uniform distribution over $[L_i, U_i]$ (resp., $[L_j, U_j]$). We distribute $\sum_i \tilde{f}_i$ randomly (without splitting units) between the markets as demand to obtain $\tilde{f}_j$. We set $f_j = \min\{\tilde{f}_j, \bar{f}_j\}$. Then we distribute the supply $\sum_j |f_j - \tilde{f}_j|$ among the plants (without splitting units) to obtain $\tilde{f}_i$ and compute $f_i = \tilde{f}_i + \bar{f}_i$. The procedure thereby guarantees that demand and supply are balanced.

5.2. Numerical results. We first study the effect on the linear programming relaxation bound of adding (multi)path and cycle inequalities. We solve the linear programming relaxation of the problem, and add violated cuts one at a time, each time resolving the modified LP. We use CPLEX in its default settings.

In Table 3, we evaluate the strength of short-paths and short-cycle inequalities, as we wish to confirm our intuition that EC&R inequalities that aggregate few constraints, i.e., those induced by short paths and short cycles, are most useful. To this end, we compare the bounds obtained from short-path/cycle inequalities with the

TABLE 3
Path/cycle inequalities for relaxations with a single bilinear term.

#	Trd-Lnz	1-Disj	Time	P/C	Time	Gap
1	9740.9	12134.4	123.7	12078.6	20.3	97.7
2	11360.9	14061.3	132.7	14000.1	26.7	97.7
3	10260.4	12243.6	132.6	12177.1	20.1	96.6
4	10773.4	13310.6	129.7	13244.6	26.3	97.4
5	10346.0	12726.8	123.8	12615.9	19.8	95.3
6	10924.8	13343.0	133.8	13215.8	22.6	94.7
7	10846.6	13308.2	126.9	13212.5	18.0	96.1
8	10791.8	13424.5	125.8	13338.1	20.8	96.7
9	12396.6	15004.3	129.6	14920.4	18.9	96.8
10	9435.8	11640.8	119.0	11526.6	22.4	94.8

bounds obtained using cuts from CGLPs for $\text{conv}(\mathcal{D}_{(i,j)})$. The second column contains the value of the linear programming relaxation for the traditional formulation that uses linearized inequalities. The third column presents the bound obtained after adding all violated cuts for $\text{conv}(\mathcal{D}_{(i,j)})$ for each $(i,j) \in E$. These cuts are generated through the solution of a CGLP obtained from a disjunctive programming formulation of $\mathcal{D}_{(i,j)}$ with two disjuncts (one for each binary value of variable y_{ij}). These cuts are added over multiple rounds until the bound improvement between rounds becomes negligible. The third column therefore estimates the maximum possible improvement from path and cycle inequalities. The fifth column contains the bound obtained by adding path inequalities for (i,j) -paths of length 2, and cycle inequalities associated with (i,j) -cycles of length 4, for each $(i,j) \in E$. We associate inequalities with quadruplets (i,j,p,c) , where (i,j) is the arc, p is an index in the ordering of paths, and c is an index in the ordering of cycles. We traverse these inequalities in lexicographic increasing order once during the whole procedure. Each time we find a violated inequality (of the form (4.5) or (4.6) as appropriate), we add the most violated linearization of this inequality before resolving the resulting LP. Then, we continue our search through the lexicographic list of inequalities. The seventh column indicates that, on average for our instances, at least 96% of the total gap improvement can be attributed to short paths and cycles. This gap is computed as the difference between the values of columns 5 and 2 divided by the difference between the values of columns 3 and 2. This observation has interesting computational ramifications that we will explore next. The fourth and sixth columns record the times needed to compute the corresponding bounds. These numbers are not directly comparable, since they relate to different number of rounds.

In Table 4, we consider multiple bilinear terms simultaneously. In particular, we generate valid inequalities for $\text{conv}(\mathcal{D}_{(i,j),(k,l)})$, where $\mathcal{D}_{(i,j),(k,l)}$ is obtained by including $\gamma_{kl}y_{kl} = z_{kl}$ and $0 \leq y_{kl} \leq 1$ in the description of $\mathcal{D}_{(i,j)}$. We then determine how much of this bound improvement is due to the inequalities of Proposition 4.9. Since we observed in Table 3 that paths of length 2 and cycles of length 4 yield inequalities that are effective at closing gaps, we consider only multipath/-cycle inequalities based on such paths and cycles. Similarly to Table 3, we compare the bounds obtained from short multipath/-cycle inequalities with the bounds obtained using cuts from CGLPs for $\text{conv}(\mathcal{D}_{(i,j),(k,l)})$. The second column contains the value of the linear programming relaxation for the traditional formulation that uses linearized inequalities. The third column presents the bound obtained after adding all violated cuts for $\text{conv}(\mathcal{D}_{(i,j),(k,l)})$ for each pair of distinct arcs (i,j) and (k,l) of E . These cuts are generated through

TABLE 4
Multipath/-cycle inequalities for relaxations with two bilinear terms.

#	Trd-Lnz	2-Disj	Time	M-P/C	Time	Gap
1	9740.9	12965.7	6130.1	12747.3	15.0	93.2
2	11360.9	15040.2	6026.8	14911.8	16.6	96.4
3	10260.4	13076.3	5771.0	12926.5	21.3	94.6
4	10773.4	14046.0	5352.2	13871.6	16.5	94.7
5	10346.0	13645.7	6360.2	13458.5	16.6	94.3
6	10924.8	14173.1	5375.5	13995.9	16.4	94.5
7	10846.6	14146.0	6034.2	13826.1	16.8	90.3
8	10791.8	14192.7	6073.4	14065.1	19.4	96.2
9	12396.6	15945.0	6068.5	15707.8	22.5	93.2
10	9435.8	12534.6	5869.9	12349.2	15.9	94.0

the solution of a CGLP from a disjunctive programming formulation of $\mathcal{D}_{(i,j),(k,l)}$ with four disjuncts (one for each binary value of the pair y_{ij} and y_{kl}). The fourth column presents the time needed to obtain this bound. The fifth column in Table 4 shows the bound improvement obtained using multipath/-cycle inequalities. For each pair of distinct arcs (i, j) and (k, l) , we derive two families of inequalities. If (i, j) and (k, l) are adjacent, they form a path of length 2. Since these two arcs are not oriented in the same direction, the corresponding multipath inequality is a regular path inequality of length 2. If (i, j) and (k, l) are not adjacent, there is a unique cycle of length 4 that contains these arcs, as depicted in Figure 2. For the most violated multicycle inequality, we add the corresponding multicycle cut; see Example 3. After multicycle inequalities are considered, we generate regular cycle inequalities of length 4 as for Table 3. Each of the cuts generated is therefore derived from a model having at most two bilinear equalities. We perform a single round of the above cut generation procedure. The sixth column presents the time needed to obtain the corresponding bound. The last column shows that, on average, 94% of the total gap improvement is due to short multipath/-cycle inequalities. This gap is computed as the difference between the values of columns 5 and 2 divided by the difference between the values of columns 3 and 2. This result is promising, because these inequalities can be generated quickly.

Finally, we seek to evaluate whether the strengthened linear programming relaxations that are obtained by using multipath and multicycle inequalities yield improvement in the global solution of our test instances through branch-and-bound. These results are reported in Table 5. The second column contains the value of the linear programming relaxation bound of the traditional dual formulation that uses linearized inequalities, while the third column reports optimal values. The following three columns present results obtained by running CPLEX in its default mode. None of the instances were solved to provable optimality within the time limit of 10,000 seconds. The fourth column describes the bound obtained by CPLEX after adding default cuts at the root node and before starting branch-and-bound. The fifth column shows the gap closed by adding these CPLEX cuts at the root node. This gap is computed as the difference between the values of columns 4 and 2 divided by the difference between the values columns 3 and 2. The sixth column presents the best lower bound achieved by CPLEX at termination. The last three columns relate to an implementation where multipath and multicycle inequalities are added at the root node through CPLEX cut callbacks, and CPLEX branch-and-bound is used thereafter. The seventh column gives the corresponding root bound. This bound is obtained using one round of cut generation,

TABLE 5
Implementation of multipath/-cycle inequalities inside CPLEX.

#	Trd-Lnz	Opt	CPX root	Gap	CPX bnd	M-P/C root	Gap	B&B time
1	9740.9	13933.0	10366.4	14.9	13705.5	12956.7	76.7	1498.2
2	11360.9	16067.0	12108.4	15.9	15886.5	14998.6	77.3	899.4
3	10260.4	14172.0	10635.1	9.6	13938.9	13035.2	70.9	2703.5
4	10773.4	15111.0	11459.0	16.6	14854.4	14000.0	74.6	1553.1
5	10346.0	14678.0	11005.8	15.2	14453.0	13491.3	72.6	2263.0
6	10924.8	15164.0	11612.7	16.2	15030.2	14041.0	73.5	1928.3
7	10846.6	15293.0	11483.1	14.3	15076.7	14106.2	73.3	1931.1
8	10791.8	15145.0	11544.7	17.3	14904.4	14112.2	76.2	1216.0
9	12396.6	17218.0	13182.0	16.3	16763.3	15842.6	71.5	3141.7
10	9435.8	13474.0	9944.2	12.6	13288.1	12550.8	77.1	1171.4

as in Table 4. The only difference is that CPLEX default cuts are not deactivated. The eighth column shows the gap closed by adding multipath/-cycle cuts and CPLEX cuts at the root node. We observe that these cuts significantly improve the gap closure compared to default CPLEX cuts, as reported in the fifth column. The last column reports the time (in seconds) needed for CPLEX to solve the problem to optimality when multipath/-cycle cuts are generated at the root node. We conclude that multipath/-cycle inequalities significantly improve gap closure as well as solution times for these problem instances.

6. Conclusion. In this paper, we develop a convexification procedure for certain sets with bilinear constraints. This technique generalizes the cancel-and-relax (C&R) procedure of [35] initially proposed to study sets with separable complementarity constraints. We obtain a description of the simultaneous convex hull of graphs of bilinear functions over the Cartesian product of two polytopes, one of which is a simplex. The structure we consider allows for the study of unary expansions in mixed integer BLPs, simultaneous convexification of bilinear functions over independent polytopes, and sets with complementarity constraints. The procedure gives insight into the structure of strong inequalities for this disjunctive set in the space of original variables. We use it to study the dual reformulation of NIPs. We provide a convex hull description of a suitable problem relaxation with one bilinear term that can be viewed as a tool for improving traditional linearization techniques for NIPs. We build on the insight gathered while studying this model to propose a family of inequalities that consider multiple bilinear terms. Our computational results show that these inequalities significantly reduce the gap of traditional linear programming relaxations and improve our ability to solve these problems to optimality.

7. Appendix. In this appendix, we provide the proofs that were omitted from the main body of the paper.

Proof of Proposition 2.5. We first show the direct implication. Let $\bar{\pi}^l = (\bar{\pi}^l; \{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M})$ be a nonvertical extreme point of $\mathcal{C}^{l+}$. It can therefore be obtained as a bfs of $\mathcal{C}^{l+}$. We refer to the basic variables associated with $\bar{\pi}^l$ by $\Phi(\bar{\pi}^l)$. Consider any $(i, j) \in N \times M$. Since the coefficient columns associated with η_i^j and ρ_i^j are linearly dependent, at most one of these variables belongs to $\Phi(\bar{\pi}^l)$. Define $G \subseteq N \times M$ so that $(i, j) \in G$ if and only if neither of the variables η_i^j and ρ_i^j belongs to $\Phi(\bar{\pi}^l)$. It follows that the number of basic variables among η^j and ρ^j for all $j \in M$ is $mn - |G|$. Next, define $L \subseteq K_l$ (resp., $\bar{L} \subseteq K_l$) so that $i \in L$ (resp., $i \in \bar{L}$) if and only if β_i^+ (resp., β_i^-) belongs to $\Phi(\bar{\pi}^l)$. Also, introduce $J \subseteq N$ (resp., $\bar{J} \subseteq N$)

so that $i \in J$ (resp., $i \in \bar{J}$) if and only if λ_i (resp., μ_i) belongs to $\Phi(\bar{\pi}^l)$. Finally, define I_j for $j \in M$ (resp., $\bar{I}$) so that $i \in I_j$ (resp., $i \in \bar{I}$) if and only if γ_i^j (resp., θ_i) belongs to $\Phi(\bar{\pi}^l)$. Given the above definitions, the total number of basic variables is $mn - |G| + |L| + |\bar{L}| + \sum_{j \in M} |I_j| + |\bar{I}| + |J| + |\bar{J}|$. Note that the size of $\Phi(\bar{\pi}^l)$ is mn since the coefficient matrix used in the definition of $\mathcal{C}^{l+}$ has full row rank as the columns corresponding to variables η_i^j form an identity matrix. We conclude that $|G| = |L| + |\bar{L}| + \sum_{j \in M} |I_j| + |\bar{I}| + |J| + |\bar{J}|$. Setting nonbasic variables to zero in the equations of $\mathcal{C}^{l+}$, we obtain (2.5). After reordering columns and rows of the basis matrix associated with $\Phi(\bar{\pi}^l)$ and after changing the sign of some of its columns if necessary, it can be rewritten as

$$(7.1) \quad \left[\begin{array}{c|c} \Psi & 0 \\ \hline \bar{\Psi} & I \end{array} \right].$$

In (7.1), Ψ is a square submatrix induced by the rows of the coefficient matrix describing $\mathcal{C}^{l+}$ corresponding to $(i, j) \in G$, i.e., the coefficient matrix of system (2.5). Further, $[\bar{\Psi}|I]$ is the submatrix induced by rows of the coefficient matrix describing $\mathcal{C}^{l+}$ corresponding to $(i, j) \notin G$, and the columns of I are present because either η_i^j or ρ_i^j belongs to basis $\Phi(\bar{\pi}^l)$. Since $\Phi(\bar{\pi}^l)$ forms a basis, matrix (7.1) must be nonsingular. We conclude that Ψ is nonsingular.

We next prove the reverse direction. Given $\hat{\pi}^l \in \hat{\mathcal{C}}^{l+}$, we construct $\bar{\pi}^l = (\hat{\pi}^l; \{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M})$ for some $(\{\bar{\eta}^j\}_{j \in M}; \{\bar{\rho}^j\}_{j \in M}) \in \mathbb{R}_+^{2mn}$ and show that it is an extreme point of $\mathcal{C}^{l+}$. We set $\bar{\eta}_i^j = \bar{\rho}_i^j = 0$ for all $(i, j) \in G$. We then compute, for each $(i, j) \notin G$,

$$(7.2) \quad \xi_i^j = \sum_{k \in L} A_{ji}^k \beta_k^+ - \sum_{k \in \bar{L}} A_{ji}^k \beta_k^- + \sum_{t \in I_j} E_{ti} \gamma_t^j - \sum_{t \in \bar{I}} E_{ti} \theta_t - \sum_{t \in J \cap \{i\}} \lambda_t + \sum_{t \in \bar{J} \cap \{i\}} \mu_t + A_{ji}^l.$$

For each $(i, j) \notin G$, we set $\bar{\eta}_i^j = -\xi_i^j$ and $\bar{\rho}_i^j = 0$ if $\xi_i^j < 0$. Similarly, we set $\bar{\rho}_i^j = \xi_i^j$ and $\bar{\eta}_i^j = 0$ if $\xi_i^j \geq 0$. It follows from the construction of $\bar{\eta}^j$ and $\bar{\rho}^j$ that $\bar{\pi}^l$ belongs to $\mathcal{C}^{l+}$. We now show that $\bar{\pi}^l$ is an extreme point of $\mathcal{C}^{l+}$ by describing a basis $\Phi(\bar{\pi}^l)$ that has $\bar{\pi}^l$ as associated bfs. To construct this basis, we pick all variables associated with elements of the sets $L, \bar{L}, I_j, \bar{I}, J$, and $\bar{J}$. We also include variables $\bar{\eta}_i^j$ for $(i, j) \notin G$ such that $\bar{\rho}_i^j = 0$, and variables $\bar{\rho}_i^j$ for the remaining $(i, j) \notin G$. The number of variables selected is therefore equal to mn . Reordering columns and rows of the corresponding coefficient matrix and changing their signs when required, we obtain matrix (7.1), which is nonsingular because Ψ is. This proves that $\bar{\pi}^l$ is a bfs to $\mathcal{C}^{l+}$. $\square$

We next give an ancillary result that is used in the proof of Proposition 3.1, which ensues.

PROPOSITION 7.1. *Consider*

$$\hat{\mathcal{S}}_{\times} = \left\{ (\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \Xi \times \Delta_m \times \mathbb{R}^{\sum_q \kappa^q} \mid \mathbf{y}^{\top} A^{q,k} \mathbf{x}^q = z_k^q \quad \forall (q, k) \in Q \times K^q \right\},$$

where $\Xi = \Xi^1 \times \Xi^2 \times \cdots \times \Xi^{\varrho}$, where $\Xi^q = \{\mathbf{x}^q \in [0, 1]^{n^q} \mid E^q \mathbf{x}^q \geq \mathbf{f}^q\}$ for $q \in Q = \{1, \dots, \varrho\}$, and where $K^q = \{1, \dots, \kappa^q\}$. Define

$$T^q(\mathcal{S}) = \left\{ (\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \mathbb{R}^{n+m+\sum_q \kappa^q} \mid (\mathbf{x}^q; \mathbf{y}; \mathbf{z}^q) \in \mathcal{S} \right\}$$

for $q \in Q$. Then, $\text{conv}(\widehat{\mathcal{S}}_\times) = \bigcap_{q \in Q} T^q(\text{conv}(\widehat{\mathcal{S}}_\times^q))$, where

$$\widehat{\mathcal{S}}_\times^q = \left\{ (\mathbf{x}^q; \mathbf{y}; \mathbf{z}^q) \in \Xi^q \times \Delta_m \times \mathbb{R}^{\kappa^q} \mid \mathbf{y}^\top A^{q,k} \mathbf{x}^q = z_k^q \quad \forall k \in K^q \right\}.$$

Proof. We first show the direct inclusion. It is clear that $\widehat{\mathcal{S}}_\times = \bigcap_{q \in Q} T^q(\widehat{\mathcal{S}}_\times^q)$. Therefore, $\text{conv}(\widehat{\mathcal{S}}_\times) \subseteq \bigcap_{q \in Q} \text{conv}(T^q(\widehat{\mathcal{S}}_\times^q)) = T^q(\text{conv}(\widehat{\mathcal{S}}_\times^q))$, where the equality holds as T^q is a linear operator. For the reverse inclusion, select $\boldsymbol{\omega} = (\{\mathbf{x}^q\}_{q \in Q}; \mathbf{y}; \{\mathbf{z}^q\}_{q \in Q}) \in \bigcap_{q \in Q} T^q(\text{conv}(\widehat{\mathcal{S}}_\times^q))$. We show that $\boldsymbol{\omega} \in \text{conv}(\widehat{\mathcal{S}}_\times)$. By definition, we have that $\boldsymbol{\omega}^q = (\mathbf{x}^q; \mathbf{y}; \mathbf{z}^q) \in \text{conv}(\widehat{\mathcal{S}}_\times^q)$ for all $q \in Q$. Therefore, $\boldsymbol{\omega}^q$ can be obtained as a convex combination of extreme points of $\text{conv}(\widehat{\mathcal{S}}_\times^q)$. Proposition 2.1 implies that the $\mathbf{y}$ component of each of these extreme points belongs to the set $\{\mathbf{e}^j\}_{j \in M \cup \{0\}}$ where $\mathbf{e}^0 = \mathbf{0}$. We can then write $\boldsymbol{\omega}^q = \sum_{j=0}^m \sum_{\ell_q=1}^{\Omega_j^q} \sigma_{j,\ell_q}^q (\mathbf{x}_{j,\ell_q}^q; \mathbf{e}^j; \mathbf{z}_{j,\ell_q}^q)$, where $(\mathbf{x}_{j,\ell_q}^q; \mathbf{e}^j; \mathbf{z}_{j,\ell_q}^q)$ is an extreme point of $\widehat{\mathcal{S}}_\times^q$ and σ_{j,ℓ_q}^q is its convex combination multiplier. In this definition, Ω_j^q denotes the number of extreme points of $\widehat{\mathcal{S}}_\times^q$ whose $\mathbf{y}$ component is $\mathbf{e}^j$. Then we have that $\sum_{\ell_q=1}^{\Omega_j^q} \sigma_{j,\ell_q}^q = y_j$ for each $q \in Q$ and $j \in M \cup \{0\}$, as there is a unique convex representation for $\mathbf{y}$ in terms of $\mathbf{e}^j$. This shows that $\Omega_j^q = 0$ if and only if $\Omega_j^{\bar{q}} = 0$ for any distinct q and $\bar{q}$ in Q . Therefore, we compute that $\boldsymbol{\omega} = \sum_{j=0}^m \sum_{\ell_1=1}^{\Omega_j^1} \sum_{\ell_2=1}^{\Omega_j^2} \cdots \sum_{\ell_\varrho=1}^{\Omega_j^\varrho} \frac{\prod_{q=1}^\varrho \sigma_{j,\ell_q}^q}{y_j^{\varrho-1}} (\{\mathbf{x}_{j,\ell_q}^q\}_{q \in Q}; \mathbf{e}^j; \{\mathbf{z}_{j,\ell_q}^q\}_{q \in Q})$. In this relation, it follows from the construction that points $(\{\mathbf{x}_{j,\ell_q}^q\}_{q \in Q}; \mathbf{e}^j; \{\mathbf{z}_{j,\ell_q}^q\}_{q \in Q})$ belong to $\widehat{\mathcal{S}}_\times$, and that the weights $\frac{\prod_{q=1}^\varrho \sigma_{j,\ell_q}^q}{y_j^{\varrho-1}}$ form a set of convex combination multipliers. $\square$

Proof of Proposition 3.1. Through addition of variables z_k^q , we can write $\mathcal{S}_\times$ in a higher dimension as

$$\mathcal{S}_\times^1 = \left\{ (\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \Xi \times \Delta_m \times \mathbb{R}^{\kappa(1+\varrho)} \mid \begin{array}{l} \mathbf{y}^\top A^{q,k} \mathbf{x}^q = z_k^q \quad \forall (q, k) \in Q \times K \\ \sum_{q \in Q} z_k^q = z_k \quad \forall k \in K \end{array} \right\}.$$

Since variables z_k are free and are defined as a linear combination of other variables, we conclude that $\text{conv}(\mathcal{S}_\times^1)$ is equal to the convex hull of

$$\mathcal{S}_\times^2 = \left\{ (\mathbf{x}; \mathbf{y}; \mathbf{z}) \in \Xi \times \Delta_m \times \mathbb{R}^{\kappa(1+\varrho)} \mid \mathbf{y}^\top A^{q,k} \mathbf{x}^q = z_k^q \quad \forall (q, k) \in Q \times K \right\},$$

intersected with $\sum_{q \in Q} z_k^q = z_k$ for all $k \in K$. Applying Proposition 7.1 to $\mathcal{S}_\times^2$, we obtain the result. For the case where $\kappa = 1$, EC&R inequalities of $\mathcal{S}_\times^q$ are obtained as class-1 $^\pm$. This implies that the coefficient of variables z_1^q is ± 1 in these inequalities. The result follows from that of the first part, while applying Fourier–Motzkin elimination to project out variables z_1^q . $\square$

Acknowledgments. The authors thank the associate editor and the two anonymous referees for suggestions that helped improve the paper.

REFERENCES

- [1] W. P. ADAMS AND H. D. SHERALI, *Mixed-integer bilinear programming problems*, Math. Programming, 59 (1993), pp. 279–305.
- [2] F. A. AL-KHAYYAL, *Generalized bilinear programming: Part I. Models, applications and linear programming relaxation*, Eur. J. Oper. Res., 60 (1992), pp. 306–314.

- [3] F. A. AL-KHAYYAL AND J. E. FALK, *Jointly constrained biconvex programming*, Math. Oper. Res., 8 (1983), pp. 273–286.
- [4] D. S. ALTNER, Ö. ERGUN, AND N. A. UHAN, *The maximum flow network interdiction problem: Valid inequalities, integrality gaps, and approximability*, Oper. Res. Lett., 38 (2010), pp. 33–38.
- [5] K. M. ANSTREICHER, *On convex relaxations for quadratically constrained quadratic programming*, Math. Programming, 136 (2012), pp. 233–251.
- [6] N. ASSIMAKOPOULOS, *A network interdiction model for hospital infection control*, Comput. Biol. Med., 17 (1987), pp. 413–422.
- [7] C. AUDET, J. BRIMBERG, P. HANSEN, S. L. DIGABEL, AND N. MLADENOVIC, *Pooling problem: Alternate formulations and solution methods*, Management Sci., 50 (2004), pp. 761–776.
- [8] E. BALAS, *Disjunctive programming and a hierarchy of relaxations for discrete optimization problems*, SIAM J. Discrete Math., 6 (1985), pp. 466–486, <https://doi.org/10.1137/0606047>.
- [9] E. BALAS, *Disjunctive programming: Properties of the convex hull of the feasible points*, Discrete Appl. Math., 89 (1998), pp. 3–44.
- [10] X. BAO, N. V. SAHINIDIS, AND M. TAWARMALANI, *Semidefinite relaxations for quadratically constrained quadratic programming*, Math. Programming, 129 (2011), pp. 129–157.
- [11] O. BEN-AYED AND C. E. BLAIR, *Computational difficulties of bilevel linear programming*, Oper. Res., 38 (1990), pp. 556–560.
- [12] A. BEN-TAL, G. EIGER, AND V. GERSHOVITZ, *Global minimization by reducing the duality gap*, Math. Programming, 63 (1994), pp. 193–212.
- [13] P. BONAMI, A. LODI, S. TRAMONTANI, AND A. WIESE, *On mathematical programming with indicator constraints*, Math. Programming, 151 (2015), pp. 191–223.
- [14] A. CAPRARA AND M. MONACI, *Bidimensional packing by bilinear programming*, Math. Programming, 118 (2009), pp. 75–108.
- [15] T. CHRISTOF AND A. LÖBEL, *PORTA – Polyhedral Representation Transformation Algorithm*, <http://www.iwr.uni-heidelberg.de/groups/comopt/software/PORTA/>, 2012.
- [16] R. A. COLLADO AND D. PAPP, *Network Interdiction—Models, Applications, Unexplored Directions*, RUTCOR research report RRR 4-2012, Rutgers Center for Operations Research Rutgers University, Piscataway, NJ, 2012.
- [17] D. DAVARNIA, J.-P. P. RICHARD, AND M. TAWARMALANI, *On the Strength and Structure of the KKT Formulation of Network Interdiction Problems*, Tech. report, University of Florida, 2015.
- [18] O. GÜNLÜK AND J. LINDEROTH, *Perspective reformulation and applications*, in Mixed Integer Nonlinear Programming, J. Lee and S. Leyffer, eds., Springer, New York, 2012, pp. 61–89.
- [19] A. GUPTA, S. AHMED, M. S. CHEON, AND S. DEY, *Solving mixed integer bilinear problems using MILP formulations*, SIAM J. Optim., 23 (2013), pp. 721–744, <https://doi.org/10.1137/110836183>.
- [20] I. HARJUNKOSKI, T. WESTERLUND, R. PÖRN, AND H. SKRIFVARS, *Different transformations for solving non-convex trim-loss problems by MINLP*, Eur. J. Oper. Res., 105 (1998), pp. 594–603.
- [21] T. E. HARRIS AND F. S. ROSS, *Fundamentals of a Method for Evaluating Rail Network Capacities*, Tech. report RM-1573, The RAND Corporation, Santa Monica, CA, 1955.
- [22] H. HIJAZI, P. BONAMI, G. CORNUÉJOLS, AND A. OUOROU, *Mixed integer nonlinear programs featuring on/off constraints*, Comput. Optim. Appl., 52 (2012), pp. 537–558.
- [23] E. ISRAELI AND R. K. WOOD, *Shortest-path network interdiction*, Networks, 40 (2002), pp. 97–111.
- [24] R. G. JEROSLOW, *The polynomial hierarchy and a simple model for competitive analysis*, Math. Programming, 32 (1985), pp. 146–164.
- [25] C. LIM AND J. C. SMITH, *Algorithms for discrete and continuous multicommodity flow network interdiction problems*, INFORMS J. Comput., 39 (2007), pp. 15–26.
- [26] M. LOCATELLI AND F. SCHOEN, *On convex envelopes for bivariate functions over polytopes*, Math. Programming, 144 (2014), pp. 65–91.
- [27] J. LUEDTKE, M. NAMAZIFAR, AND J. LINDEROTH, *Some results on the strength of relaxations of multilinear functions*, Math. Programming, 136 (2012), pp. 325–351.
- [28] G. P. MCCORMICK, *Computability of global solutions to factorable nonconvex programs: Part I. Convex underestimating problems*, Math. Programming, 10 (1976), pp. 145–175.
- [29] C. A. MEYER AND C. A. FLOUDAS, *Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes*, J. Global Optim., 29 (2004), pp. 125–155.
- [30] C. A. MEYER AND C. A. FLOUDAS, *Convex envelopes for edge-concave functions*, Math. Programming, 103 (2005), pp. 207–224.

- [31] R. MISENER AND C. A. FLOUDAS, *Advances for the pooling problem: Modeling, global optimization, and computational studies*, Appl. Comput. Math., 8 (2009), pp. 3–22.
- [32] D. MORTON, F. PAN, AND K. SAEGER, *Models for nuclear smuggling interdiction*, IIE Trans., 39 (2007), pp. 3–14.
- [33] A. NAHAPETYAN, *Bilinear programming: Applications in the supply chain management*, in Encyclopedia of Optimization, C. Floudas, A. Christodoulos, and P. M. Pardalos, eds., Springer, New York, 2009.
- [34] T. T. NGUYEN, J.-P. P. RICHARD, AND M. TAWARMALANI, *Deriving convex hulls through lifting and projection*, Math. Programming, to appear, <https://link.springer.com/article/10.1007/s10107-017-1138-3>.
- [35] T. T. NGUYEN, M. TAWARMALANI, AND J.-P. P. RICHARD, *Convexification Techniques for Complementarity Constraints*, Tech. report, University of Florida, Gainesville, FL, 2013.
- [36] I. QUESADA AND I. E. GROSSMANN, *Global optimization of bilinear process networks with multicomponent flows*, Comput. Chem. Engrg., 19 (1995), pp. 1219–1242.
- [37] G. RINALDI, U. VOIGT, AND G. J. WOEGINGER, *The mathematics of playing golf, or: A new class of difficult non-linear mixed integer programs*, Math. Programming, 93 (2002), pp. 77–86.
- [38] H. D. SHERALI, W. P. ADAMS, AND P. J. DRISCOLL, *Exploiting special structures in constructing a hierarchy of relaxations for 0-1 mixed integer problems*, Oper. Res., 46 (1998), pp. 396–405.
- [39] H. D. SHERALI AND A. ALAMEDDINE, *An explicit characterization of the convex envelope of a bivariate bilinear function over special polytopes*, Ann. Oper. Res., 25 (1992), pp. 197–210.
- [40] H. D. SHERALI AND A. ALAMEDDINE, *A new reformulation-linearization technique for bilinear programming problems*, J. Global Optim., 2 (1992), pp. 379–410.
- [41] J. C. SMITH AND C. LIM, *Algorithms for network interdiction and fortification games*, in Pareto Optimality, Game Theory and Equilibria, Springer Optim. Appl. 17, P. M. Pardalos, A. Migdalas, and L. Pitsoulis, eds., Springer-Verlag, New York, 2008, pp. 609–644.
- [42] K. M. SULLIVAN, J. C. SMITH, AND D. P. MORTON, *Convex hull representation of the deterministic bipartite network interdiction problem*, Math. Programming, 145 (2014), pp. 349–376.
- [43] M. TAWARMALANI, *Inclusion Certificates and Simultaneous Convexification of Functions*, manuscript.
- [44] M. TAWARMALANI, J.-P. P. RICHARD, AND K. CHUNG, *Strong valid inequalities for orthogonal disjunctions and bilinear covering sets*, Math. Programming, 124 (2010), pp. 481–512.
- [45] M. TAWARMALANI, J.-P. P. RICHARD, AND C. XIONG, *Explicit convex and concave envelopes through polyhedral subdivisions*, Math. Programming, 138 (2012), pp. 531–577.
- [46] M. TAWARMALANI AND N. V. SAHINIDIS, *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*, Nonconvex Optim. Appl. 65, Springer Science+Business Media, Dordrecht, The Netherlands, 2002.
- [47] V. VISWESWARAN, *MINLP: Applications in blending and pooling problems*, in Encyclopedia of Optimization, C. Floudas and P. Pardalos, eds., Springer, New York, 2009, pp. 2114–2121.
- [48] R. D. WOLLMER, *Removing arcs from a network*, Oper. Res., 12 (1964), pp. 934–940.
- [49] R. K. WOOD, *Deterministic network interdiction*, Math. Comput. Model., 17 (1993), pp. 1–18.