Mixed Integer Linear Programming Formulation Techniques*

Juan Pablo Vielma[†]

Abstract. A wide range of problems can be modeled as Mixed Integer Linear Programming (MIP) problems using standard formulation techniques. However, in some cases the resulting MIP can be either too weak or too large to be effectively solved by state of the art solvers. In this survey we review advanced MIP formulation techniques that result in stronger and/or smaller formulations for a wide class of problems.

Key words. mixed integer linear programming, disjunctive programming

AMS subject classifications. 90C11, 90C10

DOI. 10.1137/130915303

1. Introduction. Throughout more than 50 years of existence, mixed integer linear programming (MIP) theory and practice have been significantly developed, and it is now an indispensable tool in business and engineering [68, 94, 104]. Two reasons for the success of MIP are linear programming (LP) based solvers and the modeling flexibility of MIP. We now have several extremely effective state-of-the-art solvers [82, 69, 52, 171] that incorporate many advanced techniques [1, 2, 25, 23, 92, 112, 24] and, since its early stages, MIP has been used to model a wide range of applications [44, 45].

While in many cases constructing valid MIP formulations is relatively straightforward, some care should be taken in this construction as certain formulation attributes can significantly reduce the effectiveness of LP-based solvers. Fortunately, constructing formulations that behave well with state-of-the-art solvers can usually be achieved by following simple guidelines described in standard textbooks. However, more advanced techniques can often perform significantly better than textbook formulations and are sometimes a necessity. The main objective of this survey is to summarize the state of the art of such formulation techniques for a wide range of problems. To keep the length of this survey under control, we concentrate on formulations for sets of a *mixed integer* nature that require both integer constrained and continuous variables. We hence purposefully place less emphasis on some related areas such as combinatorial optimization, quadratic and polynomial 0/1 optimization, and polyhedral approximations of convex sets. These topics are certainly areas of important and active research, so we cover them succinctly in section 12.

Throughout this survey we emphasize the potential advantages of each technique. However, we should note that given the complexities of state-of-the-art solvers, it is hard to predict with high accuracy which formulation performs better. Some guidelines can be found in computational studies (e.g., [163, 91, 155]), but the formulation that performs best can be strongly dependent on the specific problem structure or data. Fortunately, there is a high correlation between certain favorable formulation properties and good computational performance. In addition, it is often easy to construct several alternative formulations for a preliminary computational test. Finally, we refer the reader to [78, 88, 138, 164] for complementary information on the topics covered in this survey and their relation to other areas.

The rest of this survey is organized as follows. We begin in section 2 with a motivating example that allows us to precisely define the idea of an MIP formulation or model. This same example serves to illustrate one of the most important favorable properties of an MIP formulation: the strength of its LP relaxation. Through this section we also introduce basic MIP concepts and notation that we use in the rest of the paper. The construction and evaluation of effective MIP formulations also requires some basic concepts and results from polyhedral theory, which we review in section 3. Armed with such results, in section 4, we introduce the use of auxiliary variables as a way to construct strong formulations without incurring an excessive size. Then in section 5 we show how these auxiliary variables can be used to construct strong formulations for a finite set of alternatives described as the union of certain polyhedra or mixed integer sets. In section 6 we discuss how to reduce formulation size by forgoing the use of auxiliary variables. We show how this can result in a significant loss of strength, but also discuss cases in which this loss can be prevented. Then in section 7 we consider the use of large formulations and an LP-based technique that can be used to reduce their size in certain cases. Sections 8 and 9 review some advanced techniques that can be used to reduce the size of formulations and to improve the performance of branch-and-bound-based MIP solvers. After that, in section 10 we discuss alternative ways of combining formulations and in section 11 we consider precise geometric and algebraic characterizations of sets that can be modeled with different classes of MIP formulations. Finally, section 12 considers other topics related to MIP formulations.

2. Preliminaries and Motivation.

2.1. Modeling with MIP. Modeling non-convex functions has been a central topic of MIP formulations since its early developments [119, 120, 122, 121, 85, 84, 81, 91, 89], so our first example falls in this category. Consider the mathematical programming problem given by

$$(2.1a) \qquad z_{MP} := \min \quad \sum_{i=1}^{n} f_i(x_i)$$

$$\text{s.t.}$$
$$(2.1b) \qquad Ex \leq h,$$
$$(2.1c) \qquad 0 \leq x_i \leq u \quad \forall i \in \{1, \ldots, n\},$$

where $f_i : [0, u] \to \mathbb{Q}$ are univariate piecewise linear functions of the form

$$(2.2) \qquad f(x) = \begin{cases} m_1 x + c_1, & x \in [d_0, d_1], \\ m_2 x + c_2, & x \in [d_1, d_2], \\ \quad \vdots \\ m_k x + c_k, & x \in [d_{k-1}, d_k], \end{cases}$$

for given breakpoints $0 = d_0 < d_1 < \cdots < d_{k-1} < d_k = u$, slopes $\{m_i\}_{i=1}^{k} \subseteq \mathbb{Q}$, and constants $\{c_i\}_{i=1}^{k} \subseteq \mathbb{Q}$. We assume that the slopes and constants are such that the functions are continuous, but not convex. For instance, one of these functions could be

$$(2.3) \qquad f(x) = \begin{cases} x + 2, & x \in [0,1], \\ -2x + 5, & x \in [1,2], \\ 1, & x \in [2,3], \end{cases}$$

which will be our running example for this section.

Because the functions we consider are non-convex, we cannot transform (2.1) into an equivalent LP problem. However, we can transform it into an MIP problem as follows.

The first step in the transformation is to identify a set or a constraint that we want to model as an MIP problem. In the case of a piecewise linear function f, an appropriate set to model is the graph of f given by $\mathrm{gr}(f) := \{(x, z) \in \mathbb{Q} \times \mathbb{Q} : f(x) = z\}$. Indeed, we can reformulate (2.1) by explicitly including $\mathrm{gr}(f_i)$ to obtain the equivalent problem given by

$$(2.4a) \qquad z_{MP} := \min \ \sum_{i=1}^{n} z_i$$

$$\text{s.t.}$$

$$(2.4b) \qquad Ex \leq h,$$

$$(2.4c) \qquad 0 \leq x_i \leq u \qquad \forall i \in \{1, \ldots, n\},$$

$$(2.4d) \qquad (x_i, z_i) \in \mathrm{gr}(f_i) \quad \forall i \in \{1, \ldots, n\}.$$

Now, as illustrated in Figure 2.1 for our running example, the graph of a univariate continuous piecewise linear function (with bounded domain) is the finite union of line segments, which we can easily model with MIP. For instance, for the function defined in (2.3) we can construct the textbook MIP formulation of $\mathrm{gr}(f)$ given by

$$(2.5a) \qquad 0\lambda_0 + 1\lambda_1 + 2\lambda_2 + 3\lambda_3 = x,$$

$$(2.5b) \qquad 2\lambda_0 + 3\lambda_1 + 1\lambda_2 + 1\lambda_3 = z,$$

$$(2.5c) \qquad \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1,$$

$$(2.5d) \qquad \lambda_0 \leq y_1,$$

$$(2.5e) \qquad \lambda_1 \leq y_1 + y_2,$$

$$(2.5f) \qquad \lambda_2 \leq y_2 + y_3,$$

$$(2.5g) \qquad \lambda_3 \leq y_3,$$

$$(2.5h) \qquad y_1 + y_2 + y_3 = 1,$$

$$(2.5i) \qquad \lambda_j \geq 0 \qquad \forall j \in \{0, \ldots, 3\},$$

$$(2.5j) \qquad 0 \leq y_j \leq 1 \qquad \forall j \in \{1, 2, 3\},$$

$$(2.5k) \qquad y_j \in \mathbb{Z} \qquad \forall j \in \{1, 2, 3\}.$$

Formulation (2.5) illustrates how MIP formulations can use both continuous and integer variables to enforce requirements on the original variables. Formulations that use auxiliary variables different from the original variables are usually denoted *extended*, *lifted*, or *higher-dimensional*. In the case of formulation (2.5) these auxiliary variables
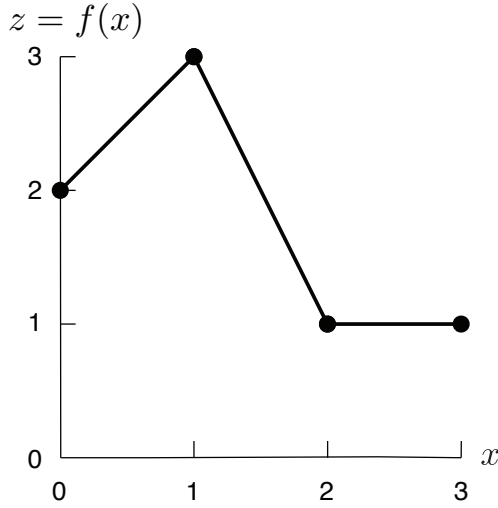
Fig. 2.1 *Graph of the continuous piecewise linear function defined in* (2.3).

are necessary to construct the formulation. However, as we will see in section 4, auxiliary variables can provide an advantage even when they are not strictly necessary. For this reason we use the following definition of an MIP formulation that always considers the possible use of auxiliary variables.

DEFINITION 2.1.   *For $A \in \mathbb{Q}^{m \times n}$, $B \in \mathbb{Q}^{m \times s}$, $D \in \mathbb{Q}^{m \times t}$, and $b \in \mathbb{Q}^m$ consider the set of linear inequalities and continuous and integer variables of the form*

$$(2.6a) \qquad\qquad Ax + B\lambda + Dy \leq b,$$

$$(2.6b) \qquad\qquad x \in \mathbb{Q}^n,$$

$$(2.6c) \qquad\qquad \lambda \in \mathbb{Q}^s,$$

$$(2.6d) \qquad\qquad y \in \mathbb{Z}^t.$$

We say (2.6) *is an MIP formulation for a set $S \subseteq \mathbb{Q}^n$ if the projection of* (2.6) *onto the x variables is exactly S. That is, if we have $x \in S$ if and only if there exist $\lambda \in \mathbb{Q}^s$ and $y \in \mathbb{Z}^t$ such that (x, λ, y) satisfies* (2.6).

Note that generic formulation (2.6) does not explicitly consider the inclusion of integrality requirements on the original variables x. While example (2.5) does not require such integrality requirements, it is common for at least some of the original variables to be integers. Formulation (2.6) implicitly considers that possibility by allowing the inclusion of constraints of the form $x_i = y_j$. Hence, we can replace (2.6b) with $x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}$ without really changing the definition of an MIP formulation.

Using Definition 2.1 we can now write a generic MIP reformulation of (2.1) by replacing every occurrence of $(x_i, z) \in \mathrm{gr}(f_i)$ with an MIP formulation of $\mathrm{gr}(f_i)$ to obtain

$$(2.7a) \qquad z_{MIP} := \min \; \sum_{i=1}^{n} z_i$$

$$\text{s.t.}$$

$$(2.7b) \qquad\qquad Ex \leq h,$$

$$(2.7c) \qquad 0 \le x_i \le u \quad \forall i \in \{1, \ldots, n\},$$

$$(2.7d) \qquad A^i \begin{pmatrix} x_i \\ z_i \end{pmatrix} + B^i \lambda^i + D^i y^i \le b^i \quad \forall i \in \{1, \ldots, n\},$$

$$(2.7e) \qquad y^i \in \mathbb{Z}^k \quad \forall i \in \{1, \ldots, n\},$$

where A^i, B^i, D^i, and b^i are appropriately constructed matrices and vectors, and k is an appropriate number of integer variables. There are many choices for (2.7d), but as long as they are valid formulations of $\mathrm{gr}(f_i)$, we have $z_{MP} = z_{MIP}$ and we can extract a solution of (2.1) by looking at the x variables of an optimal solution of (2.7). However, some versions of (2.7d) can perform significantly better when solved with an MIP solver. We study this potential difference in the next subsection.

2.2. Strength, Size, and MIP Solvers. While all state-of-the-art MIP solvers are based on the branch-and-bound algorithm [102], they also include a large number of advanced techniques that make it hard to predict the specific impact of an alternative formulation. However, there are two aspects of an MIP formulation that usually have a strong impact on both simple branch-and-bound algorithms and state-of-the-art solvers: the size and strength of the LP relaxation, and the effect of branching on the formulation. Instead of giving a lengthy description of the branch-and-bound algorithm and state-of-the-art solvers, we introduce the necessary concepts by considering the first step of the solution of MIP reformulation (2.7) with a simple branch-and-bound algorithm. For more details on basic branch-and-bound and state-of-the-art solvers, we refer the reader to [22, 36, 131, 166] and [1, 2, 25, 23, 92, 112, 24], respectively.

The first step in solving MIP formulation (2.7) with a branch-and-bound algorithm is to solve the LP relaxation of (2.7) obtained by dropping all integrality requirements. The resulting LP problem given by (2.7a)–(2.7d) is known as the *root LP relaxation* and can be solved efficiently both in theory and in practice. Its optimal value $z_{LP} := \min \{\sum_{i=1}^{n} z_i : (2.7b)–(2.7d)\}$ provides a lower bound on z_{MIP} known as the *LP relaxation bound*. If the optimal solution to the LP relaxation satisfies the dropped integrality constraints, then $z_{MIP} = z_{LP}$ and this solution is also optimal for (2.7). In contrast, if optimal solution $(\bar{x}, \bar{z}, \bar{\lambda}, \bar{y})$ is such that for some i and j we have $\bar{y}_j^i \notin \mathbb{Z}$, we can eliminate this infeasibility by *branching* on y_j^i. To achieve this we create two new LP problems by adding $y_j^i \le \lfloor \bar{y}_j^i \rfloor$ and $y_j^i \ge \lceil \bar{y}_j^i \rceil$ to the LP relaxation, respectively. These two new problems are usually denoted *branches* and are processed in a similar way, which generates a binary tree known as the *branch-and-bound tree*. The behavior of this first step is usually a good predictor of the performance of the whole algorithm, so we now concentrate on the effect of an MIP formulation on this step. We consider the effect of different formulations on subsequent steps in section 8.

To understand the effect of a formulation on the root LP relaxation we need to understand what the LP relaxation of the formulation is modeling. Going back to our running example, consider the LP relaxation of (2.5) given by (2.5a)–(2.5j). Constraints (2.5a)–(2.5c) and (2.5i) show that any (x, z) that is part of a feasible solution for this LP relaxation must be a convex combination of points $(0, 2)$, $(1, 3)$, $(2, 1)$, and $(3, 1)$. Conversely, any (x, z) that is a convex combination of these points is part of a feasible solution for the LP relaxation (let λ be the appropriate convex combination multipliers, $y_1 = \lambda_0$, $y_2 = 1 - \lambda_0 - \lambda_3$, and $y_3 = \lambda_3$). Hence, the projection onto the (x, z) variables of the LP Relaxation of (2.5) is equal to $\mathrm{conv}\,(\mathrm{gr}\,(f))$, the *convex hull* of $\mathrm{gr}\,(f)$. As illustrated in Figure 2.2(a), this convex hull corresponds to the smallest convex set containing $\mathrm{gr}\,(f)$. Now, if we used this formulation for f_i in

(2.1) (for simplicity imagine that all these functions are equal to f), the LP relaxation of (2.7) would be equivalent to
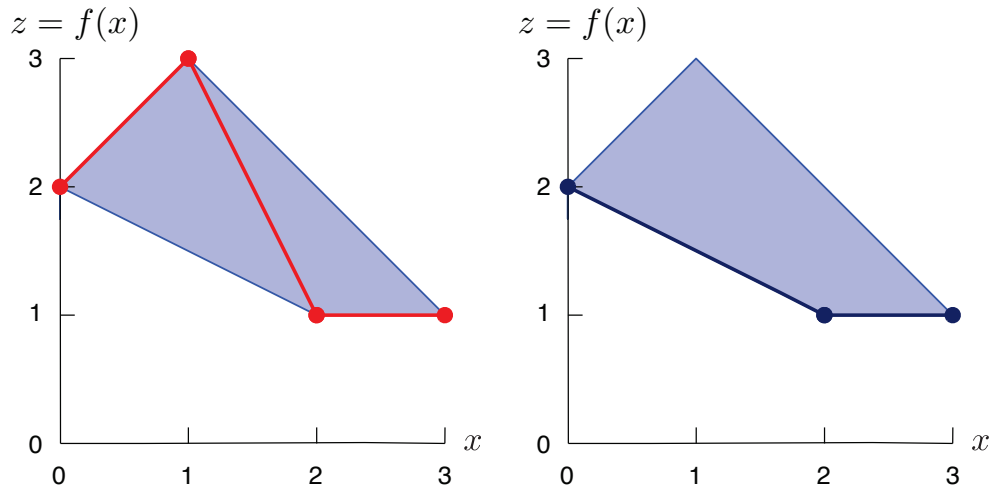
$$(2.8a) \qquad z_{LP} := \min \quad \sum_{i=1}^{n} \phi_i(x_i)$$

$$\text{s.t.}$$

$$(2.8b) \qquad Ex \leq h,$$

$$(2.8c) \qquad 0 \leq x_i \leq u \quad \forall i \in \{1, \ldots, n\},$$

where $\phi_i : [0, u] \to \mathbb{Q}$ is the *convex envelope* or *lower convex envelope* [80] of f_i. This lower convex envelope is the tightest convex underestimator of f and is illustrated for our running example in Figure 2.2(b).

(a) gr (f) in red and conv $(\mathrm{gr}\,(f))$ in light blue.

(b) conv $(\mathrm{gr}\,(f))$ in light blue and lower convex envelope of f in dark blue.

Fig. 2.2 *Effect of the LP relaxation of formulation* (2.5) *of f defined in* (2.3).

This reasoning suggests that, at least with respect to the LP relaxation bound, formulation (2.5) of our example function is as strong as it can be. Indeed, the projection onto the (x, z)-space of any formulation of f is a convex (and polyhedral) set that must contain gr (f) and the LP relaxation of (2.5) projects to the smallest of such sets. By the same argument, the projection onto the x variables of the LP relaxation of an MIP formulation of any set $S \subseteq \mathbb{Q}^n$ must also contain conv(S) and formulations whose LP relaxations project precisely to conv(S) yield the best LP relaxation bounds. Jeroslow and Lowe [90, 113] denoted those formulations that achieve this best possible LP relaxation bound as *sharp*. Other authors also denote them *convex hull* formulations.

DEFINITION 2.2. *An MIP formulation of set* $S \subseteq \mathbb{Q}^n$ *is* sharp *if and only if the projection onto the x variables of its LP relaxation is exactly* conv(S).

It is important to note that sharp formulations yield the best LP relaxation bound among all MIP formulations for the sets we selected to model. However, the LP relaxation bound can vary if we elect to model other sets. For instance, in our

example, we elected to independently model $\mathrm{gr}(f_i)$ for all $i \in \{1, \ldots, n\}$. However, we could instead have elected to model

$$(2.9) \qquad S = \left\{ (x, z) \in \mathbb{Q}^n \times \mathbb{Q} : \begin{array}{c} z = \sum_{i=1}^{n} f_i(x_i), \\ Ex \leq h, \\ 0 \leq x_i \leq u \quad \forall i \in \{1, \ldots, n\} \end{array} \right\}$$

to obtain a formulation that considers all possible nonconvexities at the same time. A sharp formulation for this case would be significantly stronger as we can show that for S defined in (2.9) we have $\min \{z : (x, z) \in \mathrm{conv}\,(S)\} = z_{MP} = z_{MIP}$. Because the LP relaxation of a sharp formulation of S is equal to $\mathrm{conv}(S)$, we have that calculating the optimal value of (2.4) could be done by simply solving an LP problem. Of course, unlike our original *piecewise sharp* formulation, constructing a sharp formulation for this joint S would normally be harder than solving (2.4). Furthermore, this approach can result in a significantly larger formulation. Selecting which portions of a mathematical programming problem to model independently to balance size and final strength is a crucial and nontrivial endeavor. However, the appropriate selection is usually ad hoc to the specific structure of the problem considered. For instance, Croxton, Gendron, and Magnanti [42] study the case where (2.1) corresponds to a multicommodity network flow problem with piecewise linear costs. In this setting, they show that a convenient middle ground between constructing independent formulations of each $\mathrm{gr}\,(f_i)$ and a single formulation for the complete problem (2.9) is to construct independent formulations for each $\mathrm{gr}\left(\sum_{i \in I_a} f_i(x_i)\right)$, where I_a corresponds to the flow variables of all commodities in a given arc a of the network. From now on we assume that a similar analysis has already been carried out and we focus on constructing small and strong formulations for the individual portions selected. However, in section 10 we will succinctly consider the selection of such portions and the combination of the associated formulations.

As we have mentioned, sharp formulations are strongest in one sense. However, if we consider the integer variables of our MIP formulation we can construct even stronger formulations. To illustrate this, let us go back to our example and consider an optimal solution $(\bar{x}, \bar{z}, \bar{\lambda}, \bar{y})$ of the LP relaxation of (2.7) given by (2.7a)–(2.7d). Because the root LP relaxation is usually solved by a simplex algorithm, we can expect this to be an optimal basic feasible solution. However, analyzing basic feasible solutions of (2.7b)–(2.7d) can be quite hard, so let us analyze basic feasible solutions of the LP relaxations of the individual formulations of $\mathrm{gr}(f_i)$ (i.e., (2.5a)–(2.5j)) as a reasonable proxy. We can check that one optimal basic feasible solution of minimizing z over (2.5a)–(2.5j) is given by $\bar{\lambda}_2 = \bar{\lambda}_3 = 1/2$, $\bar{\lambda}_0 = \bar{\lambda}_1 = 0$, $\bar{y}_1 = \bar{y}_3 = 1/2$, $\bar{y}_2 = 0$, $\bar{x} = 2.5$, and $\bar{z} = 1$. Because (2.5) is a sharp MIP formulation, it is not surprising that this gives the same optimal value as minimizing z over (2.5). Indeed we have that $1 = f(2.5) = \phi(2.5)$ and $x = 2.5$ is a minimizer of f. However, the basic feasible solution obtained has some of its integer variables set at fractional values. Because a general purpose branch-and-bound solver is not aware of the specific structure of the problem, it would have no choice but to unnecessarily branch on one of these variables (let's say $\bar{y}_1$) or to run a rounding heuristic to obtain an integer feasible solution. Hence, while sharp formulations are the strongest possible with regard to LP relaxation bounds, they can be somewhat weak with respect to finding optimal or good quality integer feasible solutions. Fortunately, it is possible to construct MIP formulations that are strong from both the LP relaxation bound and integer feasibility

perspectives. These formulations are those whose LP relaxations have basic feasible solutions that automatically satisfy the integrality requirements on the y variables. Such LP relaxations are usually denoted *integral* and formulations with this property were denoted *locally ideal* by Padberg and Rijal [136, 137]. *Ideal* comes from the fact that this is the strongest property we can expect from an MIP formulation from any perspective. *Locally* serves to clarify that this property refers to a formulation for a specifically selected portion and not to the whole mathematical programming problem (following this convention, we should then refer to sharp formulations as *locally sharp*, but we avoid it for simplicity and historical reasons). For simplicity, here we restrict our attention to MIP formulations with LP relaxations that have at least one basic feasible solution, which is the case for all practical formulations considered in this survey.

DEFINITION 2.3. *An MIP formulation* (2.6) *is* locally ideal *if and only if its LP relaxation has at least one basic feasible solution and all such basic feasible solutions have integral y variables.*

As expected, the following simple proposition states that a locally ideal formulation is at least as strong as a sharp formulation. We postpone the proof of this proposition until section 3 where we introduce some useful results concerning the feasible regions of LP problems.

PROPOSITIONS 2.4. *A locally ideal MIP formulation is sharp.*

Finally, formulation (2.5) shows that being locally ideal can be strictly stronger than being sharp. However, one case in which the converse of Proposition 2.4 holds is that of traditional MIP formulations without auxiliary variables (i.e., (2.6) with $s = 0$ and such that for all $i \in \{1, \ldots, t\}$ constraints (2.6a) include an equality of the form $x_j = y_i$ for some $j \in \{1, \ldots, n\}$). We again postpone the proof of this proposition until section 3.

PROPOSITIONS 2.5. *For $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, and $n_1, n_2 \in \mathbb{Z}_+$ such that $n = n_1 + n_2$, let*

$$(2.10a) \qquad\qquad\qquad Ax \leq b,$$
$$(2.10b) \qquad\qquad\qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2},$$

be an MIP formulation of $S \subseteq \mathbb{Q}^n$ (i.e., S is precisely the feasible region of (2.10)*). If the LP relaxation of* (2.10) *has at least one basic feasible solution, then* (2.10) *is locally ideal if and only if it is sharp.*

3. Polyhedra. Most sets modeled with MIP are of the form $S = \bigcup_{i=1}^{k} P^i$, where P^i are rational polyhedra (e.g., the graph of a univariate piecewise linear function is the union of line segments). Sets of this form usually appear as the feasible regions of certain disjunctive programming problems [10, 11, 13]. For this reason we refer to such sets as *disjunctive constraints* or *disjunctive sets*. While in the theory of disjunctive programming these terms are also used to describe a slightly broader class of sets, in this survey we concentrate on disjunctive sets that are the union of certain unbounded rational polyhedra. To construct and evaluate MIP formulations for these and other sets it will be convenient to use definitions and results from polyhedral theory that we now review. We begin by considering some basic definitions and a fundamental result that relates two natural definitions of polyhedra. We then consider the relation between polyhedra and the feasible regions of MIP problems. After that we study the linear transformations of polyhedra, which will be useful when analyzing the strength of MIP formulations. Finally, we consider the smallest possible descriptions of polyhedra to consider the real sizes of MIP formulations. We refer the reader to [22, 131, 145, 166, 170] for omitted proofs and a more detailed treatment of polyhedra.

3.1. Definitions and the Minkowski–Weyl Theorem. Polyhedra are usually described as the region bounded by a finite number of linear inequalities, such as the feasible region of an LP problem. However, formulation (2.5) of the graph of the piecewise linear function defined in (2.3) can be more easily described using the endpoints of the line segments of this graph. These two aspects of the description of polyhedra can be formalized through the following definition.

DEFINITION 3.1. *We say $P \subseteq \mathbb{Q}^n$ is a* rational $\mathcal{H}$-polyhedron *if there exist $A \in \mathbb{Q}^{m \times n}$ and $b \in \mathbb{Q}^m$ such that*

$$(3.1) \qquad\qquad P = \{x \in \mathbb{Q}^n \, : \, Ax \leq b\}.$$

In this case, we say that the right-hand side of (3.1) *is an $\mathcal{H}$-representation of P.*

We say $P \subseteq \mathbb{Q}^n$ is a rational $\mathcal{V}$-polyhedron *if there exist finite sets $V \subseteq \mathbb{Q}^n$ and $R \subseteq \mathbb{Q}^n$ such that*

$$(3.2) \qquad\qquad P = \mathrm{conv}\,(V) + \mathrm{cone}\,(R),$$

where $\mathrm{cone}\,(R)$ is the set of all nonnegative linear combinations of elements in R and $+$ denotes the Minkowski sum of sets (i.e., $\mathrm{conv}\,(V) + \mathrm{cone}\,(R) := \{x + r \, : \, x \in \mathrm{conv}\,(V), \, r \in \mathrm{cone}\,(R)\}$). In this case, we say that the right-hand side of (3.2) *is a $\mathcal{V}$-representation of P.*

Both $\mathcal{H}$- and $\mathcal{V}$-polyhedra can also be defined in $\mathbb{R}^n$. However, some results in section 11 require the polyhedra to be rational. For this reason, from now on we assume that, unless specified, all polyhedra considered are rational and we often refer to them simply as polyhedra.

One of the most important results in polyhedral theory shows that the definitions of $\mathcal{H}$- and $\mathcal{V}$-polyhedra are indeed equivalent (i.e., every rational polyhedron has both an $\mathcal{H}$- and a $\mathcal{V}$-representation). To formalize this statement we need a few definitions and results. We begin by considering the boundedness properties of polyhedra.

DEFINITION 3.2. *For an $\mathcal{H}$- or $\mathcal{V}$-polyhedron P we have the following definitions.*
- *Polyhedron P is a* cone *(or more precisely a polyhedral cone) if and only if $\lambda x \in P$ for all $\lambda \geq 0$ and $x \in P$.*
- *Polyhedron P is a* polytope *if and only if it is bounded.*
- *The* recession cone *of P is given by*

$$P_\infty := \{d \in \mathbb{Q}^n \, : \, x + \lambda d \in P \quad \forall x \in P, \quad \lambda \geq 0\}.$$

The following simple proposition characterizes the recession cone of $\mathcal{H}$- and $\mathcal{V}$-polyhedra.

PROPOSITION 3.3. *The recession cone of a polyhedron is always a cone. Furthermore, a nonempty polyhedron P is bounded if and only $P_\infty = \{0\}$.*
If P is a nonempty $\mathcal{H}$-polyhedron of the form (3.1)*, then $P_\infty = \{x \in \mathbb{Q}^n \, : \, Ax \leq 0\}$. If P is a nonempty $\mathcal{V}$-polyhedron of the form* (3.2)*, then $P_\infty = \mathrm{cone}\,(R)$.*

The last concept needed for the equivalence is a geometric characterization of basic feasible solutions and certain directions of unboundedness that are not dependent on the $\mathcal{H}$-representation of a polyhedron (remember that $x \in P$ is a basic feasible solution of $\mathcal{H}$-polyhedron $P \subseteq \mathbb{Q}^n$ if and only if it satisfies n of its linear inequalities at equality and the left-hand sides of these inequalities are linearly independent).

DEFINITION 3.4. *Let P be an $\mathcal{H}$- or a $\mathcal{V}$-polyhedron. Then the following hold:*
- *A point $x \in P$ is an* extreme point *of P if and only if there are no x^1, $x^2 \in P$ and $\lambda \in (0,1)$ such that $x = \lambda x^1 + (1-\lambda)x^2$ and $x^1 \neq x^2$. We let $\mathrm{ext}(P)$ denote the set of all extreme points of P.*

- *A direction $r \in P_\infty \setminus \{0\}$ is an* extreme ray *of P if and only if there are no r^1, $r^2 \in P_\infty \setminus \{0\}$ such that $r = r^1 + r^2$ and $r^1 \neq \lambda r^2$ for any $\lambda > 0$. We say two extreme rays r and r' are equivalent if and only if there exists $\lambda > 0$ such that $r = \lambda r'$. We let $\mathrm{ray}(P)$ denote the set of all extreme rays of P where, for each set of equivalent extreme rays, we select exactly one representative to be in $\mathrm{ray}(P)$.*
- *If P has at least one extreme point, we say that P is* pointed.

The definitions of extreme point and basic feasible solutions immediately coincide for $\mathcal{H}$-polyhedra and we also get an alternative characterization for extreme rays that is analogous to basic feasible solutions.

LEMMA 3.5. *Let $P \subseteq \mathbb{Q}^n$ be an $\mathcal{H}$-polyhedron. Then the following hold:*

- *A point $x \in P$ is an extreme point of P if and only it is a basic feasible solution of P.*
- *A direction $r \in P_\infty \setminus \{0\}$ is an extreme ray of P if and only if it satisfies $n - 1$ of the linear inequalities of P_∞ at equality and the left-hand sides of these inequalities are linearly independent.*

Of course, as the following theorem finally shows, focusing on $\mathcal{H}$-polyhedra is not really a restriction.

THEOREM 3.6 (Minkowski–Weyl). *Let $P \subseteq \mathbb{Q}^n$ such that $P \neq \emptyset$. Then P is a pointed $\mathcal{H}$-polyhedron if and only if P is a pointed $\mathcal{V}$-polyhedron.*

Furthermore, for any nonempty pointed polyhedron P we have that $\mathrm{ext}(P)$ and $\mathrm{ray}(P)$ are finite and a valid $\mathcal{V}$-representation of P is given by $P = \mathrm{conv}\,(\mathrm{ext}(P)) + \mathrm{cone}\,(\mathrm{ray}(P))$.

The Minkowski–Weyl theorem can also be stated for nonpointed polyhedra (e.g., see [145, section 8.9]). However, for simplicity, from now on we assume that all polyhedra considered are pointed and nonempty. Such an assumption is naturally present or can be easily enforced in most applications. For instance, the assumption holds if all variables considered are nonnegative, which can be assured through standard LP modeling tricks (e.g., through the reformulation $x = u - v$ with $u, v \geq 0$ for any variable x with unrestricted sign).

While equivalent, both definitions of polyhedra have their advantages and, in particular, provide alternative ways of describing sets to be modeled through MIP formulations. We illustrate this using piecewise linear functions since modeling them will be a running example for almost all formulations considered in this survey. The following definition naturally extends piecewise linear functions such as the one defined in (2.3) to the multivariate setting. In section 11.1, we will see that this definition almost precisely describes the functions that have binary MIP formulations.

DEFINITION 3.7. *Let $f : D \subseteq \mathbb{Q}^n \to \mathbb{Q}$ be a multivariate function. We define the graph of f to be*

$$\mathrm{gr}(f) := \{(x, z) \in \mathbb{Q}^n \times \mathbb{Q} : x \in D, \ f(x) = z\}$$

and the epigraph of f to be

$$\mathrm{epi}(f) := \{(x, z) \in \mathbb{Q}^n \times \mathbb{Q} : x \in D, \ f(x) \leq z\}.$$

We say f is a bounded domain continuous piecewise linear function *if f is continuous, D is bounded, and there exist $\{m^i\}_{i=1}^k \subseteq \mathbb{Q}^n$, $\{c_i\}_{i=1}^k \subseteq \mathbb{Q}$, and rational*

polytopes $\left\{Q^i\right\}_{i=1}^k$ *such that*

$$(3.3a) \qquad\qquad D = \bigcup_{i=1}^{k} Q^i,$$

$$(3.3b) \qquad\qquad f(x) = \begin{cases} m^1 x + c_1, & x \in Q^1, \\ & \vdots \\ m^k x + c_k, & x \in Q^k. \end{cases}$$

The following proposition describes the $\mathcal{H}$ and $\mathcal{V}$-representations of the graphs and epigraphs of bounded domain continuous piecewise linear functions.

PROPOSITION 3.8. *Let* $f : D \subseteq \mathbb{Q}^n \to \mathbb{Q}$ *be a bounded domain continuous piecewise linear function for which* $Q^i = \left\{x \in \mathbb{Q}^n : A^i x \leq b^i\right\}$ *for all* $i \in \{1, \ldots, k\}$. *Then the graph of* f *can be described, respectively, as the union of* $\mathcal{H}$- *and* $\mathcal{V}$-*polyhedra as follows:*

$$(3.4a) \qquad \mathrm{gr}(f) = \bigcup_{i=1}^{k} \left\{ (x,z) \in \mathbb{Q}^n \times \mathbb{Q} \,:\, A^i x \leq b^i,\, m^i x + c_i = z \right\},$$

$$(3.4b) \qquad \mathrm{gr}(f) = \bigcup_{i=1}^{k} \mathrm{conv}\left(\left\{ \begin{pmatrix} v \\ f(v) \end{pmatrix} \right\}_{v \in \mathrm{ext}(Q^i)} \right).$$

Furthermore, the epigraph of f *can be described, respectively, as the union of* $\mathcal{H}$- *and* $\mathcal{V}$-*polyhedra with a common recession cone as follows:*

$$(3.5a) \qquad \mathrm{epi}(f) = \bigcup_{i=1}^{k} \left\{ (x,z) \in \mathbb{Q}^n \times \mathbb{Q} \,:\, A^i x \leq b^i,\, m^i x + c_i \leq z \right\},$$

$$(3.5b) \qquad \mathrm{epi}(f) = \bigcup_{i=1}^{k} \mathrm{conv}\left(\left\{ \begin{pmatrix} v \\ f(v) \end{pmatrix} \right\}_{v \in \mathrm{ext}(Q^i)} \right) + \mathrm{cone}\left(\left\{ \begin{pmatrix} \mathbf{0}_n \\ 1 \end{pmatrix} \right\} \right),$$

where in both cases the recession cone of all polyhedra considered is equal to

$$\mathrm{cone}\left(\left\{ \begin{pmatrix} \mathbf{0}_n \\ 1 \end{pmatrix} \right\} \right) = \{ (x,z) \in \mathbb{Q}^n \times \mathbb{Q} \,:\, x = 0, \quad z \geq 0 \}.$$

3.2. Fundamental Theorem of Integer Programming and Formulation Strength. The finite $\mathcal{V}$-representation of any polyhedron guaranteed by Theorem 3.6 yields a convenient way to prove Proposition 2.4 as follows.

PROPOSITIONS 2.4. *A locally ideal MIP formulation is sharp.*

Proof. Let (2.6) be a locally ideal MIP formulation of a set S and let $Q \subseteq \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Q}^t$ be the polyhedron described by (2.6a). We need to show that the projection of Q onto the x variables is contained in $\mathrm{conv}(S)$.

By Lemma 3.5 and locally idealness of (2.6) we have that $\mathrm{ext}\,(Q) \subseteq \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Z}^t$. Then, by Theorem 3.6 and through an appropriate scaling of the extreme rays of Q, we have that there exist $\left\{ \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) \right\}_{j=1}^{p} \subseteq \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Z}^t$ and $\left\{ \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right) \right\}_{l=1}^{d} \subseteq \mathbb{Z}^n \times \mathbb{Z}^s \times \mathbb{Z}^t$ such that $\mathrm{ext}\,(Q) = \left\{ \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) \right\}_{j=1}^{p}$, $\mathrm{ray}\,(Q) = \left\{ \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right) \right\}_{l=1}^{d}$, and

$$Q = \mathrm{conv}\left(\left\{ \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) \right\}_{j=1}^{p} \right) + \mathrm{cone}\left(\left\{ \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right) \right\}_{l=1}^{d} \right).$$

Then for any $(x, u, y) \in Q$ there exist $\lambda \in \Delta^p := \left\{ \lambda \in \mathbb{Q}_+^p \ : \ \sum_{i=1}^p \lambda_i = 1 \right\}$ and $\mu \in \mathbb{Q}_+^d$ such that

$$(x, u, y) = \sum_{j=1}^p \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j \right) + \sum_{l=1}^d \mu_l \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l \right).$$

Let $(x^1, u^1, y^1) := \sum_{j=1}^p \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j \right)$. Because points $\left\{ \left(\hat{x}^j, \hat{u}^j, \hat{y}^j \right) \right\}_{j=1}^p$ satisfy (2.6), we have $\hat{x}^j \in S$ for all j and hence $x^1 \in \mathrm{conv}(S)$. Now, without loss of generality, assume $\lambda_1 > 0$ and let

$$(x^2, u^2, y^2) := \sum_{j=1}^p \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j \right) + \alpha \sum_{l=1}^d \mu_l \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l \right)$$

$$= \lambda_1 \left(\bar{x}, \bar{u}, \bar{y} \right) + \sum_{j=2}^p \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j \right),$$

where $\alpha \geq 1$ is such that $(\alpha/\lambda_1) \sum_{l=1}^d \mu_l \tilde{y}^l \in \mathbb{Z}^t$ and

$$(\bar{x}, \bar{u}, \bar{y}) := \left(\hat{x}^1, \hat{u}^1, \hat{y}^1 \right) + (\alpha/\lambda_1) \sum_{l=1}^d \mu_l \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l \right).$$

Then $\bar{y} = \hat{y}^1 + (\alpha/\lambda_1) \sum_{l=1}^d \mu_l \tilde{y}^l \in \mathbb{Z}^t$ and $(\bar{x}, \bar{u}, \bar{y})$ satisfies (2.6) by Theorem 3.6. Hence, $\bar{x} \in S$ and $x^2 \in \mathrm{conv}(S)$. The result then follows by noting that $x = (1 - 1/\alpha)x^1 + (1/\alpha)x^2$. $\square$

To show Proposition 2.5 we need the following consequence of Theorem 3.6 known as the Fundamental Theorem of Integer Programming. The theorem states that the convex hull of (mixed) integer points in a rational polyhedron is also a rational polyhedron and gives further structural guarantees on its $\mathcal{V}$-representation.

THEOREM 3.9. *Let $P \subseteq \mathbb{Q}^n$ be a nonempty pointed rational polyhedron and let $n_1, n_2 \in \mathbb{Z}_+$ be such that $n = n_1 + n_2$. Then there exists a finite set $V \subseteq P \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ such that*

$$(3.6) \qquad \mathrm{conv}\left(P \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right) = \mathrm{conv}\left(V \right) + \mathrm{cone}\left(\mathrm{ray}\left(P \right) \right).$$

Theorem 3.9 shows that $\mathrm{conv}\left(P \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right)$ is the LP relaxation of a sharp (by definition) formulation of $P \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$. Furthermore, through Proposition 2.5 it shows that this formulation is additionally locally ideal.

PROPOSITIONS 2.5. *For $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, and $n_1, n_2 \in \mathbb{Z}_+$ such that $n = n_1 + n_2$, let*

$$(2.10a) \qquad\qquad\qquad Ax \leq b,$$
$$(2.10b) \qquad\qquad\qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2},$$

be an MIP formulation of $S \subseteq \mathbb{Q}^n$ (i.e., S is precisely the feasible region of (2.10)). If the LP relaxation of (2.10) has at least one basic feasible solution, then (2.10) is locally ideal if and only if it is sharp.

Proof. Locally idealness implying sharpness is direct from Propositions 2.4. For the converse assume sharpness of (2.10) so that

$$Q := \{ x \in \mathbb{R}^n \ : \ Ax \leq b \} = \mathrm{conv}(S) = \mathrm{conv}\left(\{ x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2} \ : \ Ax \leq b \} \right).$$

Then by Theorem 3.9 there exist $\left\{\hat{x}^j\right\}_{j=1}^p \subseteq P \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ and $\left\{\tilde{x}^l\right\}_{l=1}^d \subseteq \mathbb{Z}^n$ such that for any $x \in \text{ext}(Q) \subseteq Q$ there exist $\lambda \in \Delta^p := \left\{\lambda \in \mathbb{Q}_+^p : \sum_{i=1}^p \lambda_i = 1\right\}$ and $\mu \in \mathbb{Q}_+^d$ such that $x = \sum_{j=1}^p \lambda_j \hat{x}^j + \sum_{l=1}^d \mu_l \tilde{x}^l$. If $d > 0$ and $\mu_l > 0$ for some $l \in \{1, \ldots, d\}$, then, assuming without loss of generality that such $l = 1$, we have $x = x^1/2 + x^2/2$, where $x^1 = \sum_{j=1}^p \lambda_j \hat{x}^j + (1/2)\mu_1 \tilde{x}^1 + \sum_{l=2}^d \mu_l \tilde{x}^l$ and $x^2 = \sum_{j=1}^p \lambda_j \hat{x}^j + (3/2)\mu_1 \tilde{x}^1 + \sum_{l=2}^d \mu_l \tilde{x}^l$ are such that $x^1, x^2 \in Q$ and $x^1 \neq x^2$. This contradicts the extremality of x, so we have $x = \sum_{j=1}^p \lambda_j \hat{x}^j$. If there are $i, j \in \{1, \ldots, p\}$ such that $i \neq j$, $\lambda_i > 0$, and $\lambda_j > 0$, then, assuming without loss of generality that $i = 1$ and $j = 2$, we have $x = \lambda_1 x^1/(\lambda_1 + \lambda_2) + \lambda_2 x^2/(\lambda_1 + \lambda_2)$, where $x^1 = (\lambda_1 + \lambda_2)\hat{x}^1 + \sum_{j=3}^p \lambda_j \hat{x}^j$ and $x^2 = (\lambda_1 + \lambda_2)\hat{x}^2 + \sum_{j=3}^p \lambda_j \hat{x}^j$ are such that $x^1, x^2 \in Q$ and $x^1 \neq x^2$. This again contradicts the extremality of x, so we may assume without loss of generality that $\lambda_1 = 1$, $\lambda_i = 0$ for all $i \geq 2$ and $\mu = 0$. Hence, $x = \hat{x}^1 \in S$ and (2.10) is locally ideal. □

3.3. Linear Transformations and Projections. A convenient way to analyze the strength of an MIP formulation is to show that its LP relaxation is the linear image of the LP relaxation of another strong formulation. The following simple proposition shows that the extreme points of the image LP relaxation are contained in the image of the extreme points of the original LP relaxation. Hence, if the original formulation is locally ideal and the linear transformation preserves integrality, then the image formulation is also locally ideal.

PROPOSITION 3.10. *Let $P \subseteq \mathbb{Q}^n$ be a rational polyhedron and $L : \mathbb{R}^n \to \mathbb{R}^p$ be a linear transformation (i.e., $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$ for all $\alpha, \beta \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$). Then $\text{ext}(L(P)) \subseteq L(\text{ext}(P))$, where $L(S) := \{L(x) : x \in S\}$ for any set $S \subseteq \mathbb{Q}^n$.*

Proof. Let $y \in \text{ext}(L(P))$ and let $x \in P$ be such that $y = L(x)$. By Theorem 3.6 there exist $\left\{\hat{x}^j\right\}_{j=1}^p \subseteq \text{ext}(P)$, $\left\{\tilde{x}^l\right\}_{l=1}^d \subseteq \text{ray}(P)$, $\lambda \in \Delta^p$, and $\mu \in \mathbb{Q}_+^d$ such that $x = \sum_{j=1}^p \lambda_j \hat{x}^j + \sum_{l=1}^d \mu_l \tilde{x}^l$ and hence $y = \sum_{j=1}^p \lambda_j L(\hat{x}^j) + \sum_{l=1}^d \mu_l L(\tilde{x}^l)$. If $d > 0$ and $\mu_l > 0$ for some l such that $L(\tilde{x}^l) \neq 0$, then we contradict the extremality of y as in the proof of Proposition 2.5. Then, $y = \sum_{j=1}^p \lambda_j L(\hat{x}^j)$. If there are $i, j \in \{1, \ldots, p\}$ such that $i \neq j$, $\lambda_i > 0$, $\lambda_j > 0$, and $L(\hat{x}^i) \neq L(\hat{x}^j)$, we again reach a contradiction with the extremality of y. Hence, $y = L(\hat{x}^j)$ for some $j \in \{1, \ldots, p\}$, which concludes the proof. □

Since the projection onto a set of variables is a linear transformation, Proposition 3.10 shows that the extreme points of the projection of a polyhedron are contained in the projection of the extreme points of the same polyhedron. Hence, because projection preserves integrality, the projection of locally ideal formulations is also locally ideal. However, in section 5 we will show that projecting a polyhedron (or formulation) can result in a significant increment in the number of inequalities. To achieve this we will need the following proposition that gives a more detailed description of an $\mathcal{H}$-representation of the projection of a polyhedron.

PROPOSITION 3.11. *Let $A \in \mathbb{Q}^{m \times n}$, $D \in \mathbb{Q}^{m \times p}$, $b \in \mathbb{Q}^m$,*

$$P = \{x \in \mathbb{Q}^n : \exists w \in \mathbb{Q}^p \text{ s.t. } Ax + Dw \leq b\},$$

and $C = \{\mu \in \mathbb{Q}^m : D^\mathsf{T}\mu = 0, \quad \mu \geq 0\}$. Then

$$P = \{x \in \mathbb{Q}^n : \mu^\mathsf{T} Ax \leq \mu^\mathsf{T} b \quad \forall \mu \in \text{ray}(C)\}.$$

In particular we have that $P_\infty = \{x \in \mathbb{Q}^n : \exists w \in \mathbb{Q}^p \text{ s.t. } Ax + Dw \leq 0\}$.

3.4. Implied Equalities, Redundant Inequalities, and Facets. The number of constraints of an MIP formulation is equal to the number of inequalities used in the specific $\mathcal{H}$-representation of the polyhedron associated to the LP relaxation of that formulation. However, the size of an $\mathcal{H}$-representation of a polyhedron can be artificially inflated by adding redundant linear inequalities. Hence, to evaluate the real size of an MIP formulation (without redundant inequalities) we need to calculate the size of the smallest $\mathcal{H}$-representation of a polyhedron. The following definition formalizes some concepts that will allow us to describe such a smallest representation.

DEFINITION 3.12. *Let $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $P := \{x \in \mathbb{Q}^n : Ax \leq b\}$, and a^i be the ith row of A. We say $F \subseteq P$ is*

- *a* face *of P if and only if $F = \{x \in P : a^l x = b_l \quad \forall l \in L\}$ for some $L \subseteq \{1, \ldots, m\}$;*
- *a* proper face *of P if and only if F is a face of P, $F \neq \emptyset$, and $F \neq P$; and*
- *a* facet *of P if and only if F is a proper face of P that is maximal with respect to inclusion.*

We also say that an inequality $a^i x \leq b_i$ of P is

- *an* implied equality *of P if and only if $a^i x = b_i$ for all $x \in P$;*
- *a* facet-defining inequality *of P if and only if $F := \{x \in P : a^i x = b_i\}$ is a facet (in such case we say the inequality defines F); and*
- *a* redundant inequality *of P for subsystem $L \subseteq \{1, \ldots, m\}$ with $i \in L$ if and only if*

$$P = \left\{ x \in \mathbb{Q}^n : a^l x \leq b_l \quad \forall l \in L \setminus \{i\} \right\}.$$

Finally we say that subsystem $L \subseteq \{1, \ldots, m\}$ is a minimal representation *of P if*

$$P = \left\{ x \in \mathbb{Q}^n : a^l x \leq b_l \quad \forall l \in L \right\}$$

and there is no $l \in L$ such that $a^l x \leq b_l$ is a redundant inequality of P for L.

Note that redundancy is strongly dependent on the selected subsystem, which can lead to the existence of multiple minimal representations when implied equalities are present. This is illustrated in the following example.

EXAMPLE 1. *Let*

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \\ -1 & -1 \\ -1 & 1 \\ -1 & 0 \end{bmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and $P = \{x \in \mathbb{Q}^2 : Ax \leq b\} = \{x \in \mathbb{Q}^2 : x_2 = 0, \quad 0 \leq x_1 \leq 1\}$. The faces of P are $F_0 := \emptyset$, $F_2 := \{(0,0)\}$, $F_3 := \{(1,0)\}$, and $F_4 := P$. The facets of P are F_2 and F_3. We also have that $a^i x \leq b_i$ is an implied equality for $i \in \{1, 2\}$, is facet defining for $i \in \{3, 4, 5, 6\}$, and is redundant for system $L = \{1, \ldots, 5\}$ for $i \in \{4, 5, 6\}$. However, facet F_2 is defined by $a^i x \leq b_i$ for any $i \in \{4, 5, 6\}$ and at least one of these inequalities is necessary to describe P. In fact, P has three minimal representations given by $L_1 = \{1, 2, 3, 4\}$, $L_2 = \{1, 2, 3, 5\}$, and $L_3 = \{1, 2, 3, 6\}$.

Constructing a minimal representation can be complicated even in the absence of implied equalities. Fortunately, as shown by the following proposition, the concept of facet-defining inequality and some linear algebra allows us to give a precise characterization of the number of inequalities in a minimal representation of a polyhedron.

PROPOSITION 3.13. *Let $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, and $P := \{x \in \mathbb{Q}^n : Ax \leq b\}$. Then for any facet of F of P there exists $i \in \{1, \ldots, m\}$ such that $a^i x \leq b_i$ defines F. Hence the number of facets of a polyhedron is always finite.*

Let $F \subseteq \{1, \ldots, m\}$ be the set of facet-defining inequalities, f be the number of facets of P, $E \subseteq \{1, \ldots, m\}$ be the set of implied equalities of P, and $r = \operatorname{rank}\left([A_l]_{l \in E}\right)$ (i.e., the maximum number of linearly independent vectors in $\{A_l\}_{l \in E}$). Then there exist $F' \subseteq F$ with $|F'| = f$ and $E' \subseteq E$ with $|E'| = r$ such that

$$P = \left\{ x \in \mathbb{Q}^n \ : \ \begin{array}{ll} a^l x \leq b_l & \forall l \in F' \\ a^l x = b_l & \forall l \in E' \end{array} \right\}$$

is a minimal representation of P. In particular, every minimal representation of P has $2r + f$ inequalities (or r equalities and f inequalities)

Determining what inequalities of a polyhedron are facet defining can be done using linear algebra techniques, but this can still be a highly nontrivial endeavor. We will present several examples of facet-defining inequalities throughout this survey, but a detailed description of the techniques used to show that they are indeed facet defining is beyond the scope of this survey. For more details, we refer the interested reader to the references on polyhedral theory [22, 131, 145, 166, 170] and to [146] for a wide range of examples, techniques, and applications from combinatorial optimization.

4. Size and Extended MIP Formulations. Strength and small size can sometimes be incompatible in MIP formulations. Fortunately, this can often be conciliated by utilizing the power of auxiliary variables in extended formulations. We first illustrate this by showing an example where the incompatibility between strength and small size can be resolved by using the same binary variables that are required to construct even the simplest formulation.

EXAMPLE 2. *Consider the set $S := \bigcup_{i=1}^n P^i$ where, for each $i \in \{1, \ldots, n\}$, we have $P^i := \{x \in \mathbb{Q}^n : |x_i| \leq 1, x_j = 0 \, \forall j \neq i\}$. It is easy to check that an MIP formulation of S is given by*

$$(4.1a) \qquad y_i - 1 \leq x_j \leq 1 - y_i \quad \forall i \in \{1, \ldots, n\}, j \neq i,$$

$$(4.1b) \qquad \sum_{i=1}^n y_i = 1,$$

$$(4.1c) \qquad 0 \leq y_i \leq 1 \qquad \forall i \in \{1, \ldots, n\},$$

$$(4.1d) \qquad y \in \mathbb{Z}^n.$$

Formulation (4.1) in Example 2 can be easily constructed with simple logic or with a basic application of a well-known formulation technique (see Example 8). Unfortunately, in this case the resulting formulation is not sharp. Indeed, for $n = 3$ we have that $x_i = 2/3$ for $i \in \{1, \ldots, 3\}$ and $y_1 = y_2 = y_3 = 1/3$ is feasible for the LP relaxation of (4.1) given by (4.1a)–(4.1c). However, for $n = 3$, $\operatorname{conv}(S) = \left\{x \in \mathbb{Q}^3 : \sum_{i=1}^3 |x_j| \leq 1\right\}$, which does not contain $(2/3, 2/3, 2/3)$. This is illustrated in Figure 4.1(a), which shows in blue the projection onto the x variables of the LP relaxation of the formulation (4.1), and in Figure 4.1(b), which does the same for the convex hull $\operatorname{conv}(S)$. Both figures show S in red.

From Figure 4.1 we can see that formulation (4.1) can be made sharp by adding the 8 inequalities defining the diamond depicted in Figure 4.1(b). In fact, we can
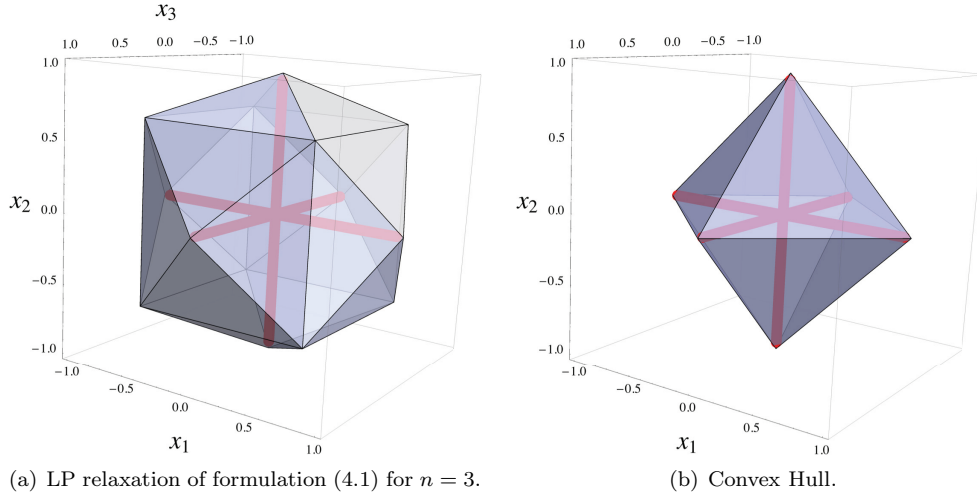
(a) LP relaxation of formulation (4.1) for $n = 3$.　　　　　(b) Convex Hull.

Fig. 4.1   *Geometric view of formulation strength.*

show that the n-dimensional form of these inequalities is

$$(4.2) \qquad \sum_{i=1}^{n} r_i x_i \leq 1 \quad \forall r \in \{-1, 1\}^n,$$

and that for any n,

$$\text{conv}(S) = \left\{ x \in \mathbb{Q}^n : \sum_{i=1}^{n} |x_j| \leq 1 \right\} = \{ x \in \mathbb{Q}^n : (4.2) \}.$$

Then, the formulation obtained by adding (4.2) to (4.1) is automatically sharp. However, formulation (4.1)–(4.2) has two problems. First, it is not locally ideal because, for $n = 3$, we have that $x_1 = x_2 = y_2 = y_3 = 1/2$, $x_3 = y_1 = 0$ is an extreme point of the LP relaxation of (4.1)–(4.2). Second, the formulation is extremely large, because the number of linear inequalities described by (4.2) is 2^n. Unfortunately, each one of these inequalities defines a different facet of $\text{conv}(S)$ and together they form a minimal representation of $\text{conv}(S)$. Thus removing any of them destroys the sharpness property. Fortunately, careful use of auxiliary variables y allows constructing a much smaller and locally ideal formulation for S.

EXAMPLE 3.   *A polynomial-sized sharp formulation for set S in Example 2 is given by*

$$(4.3a) \qquad -y_i \leq x_i \leq y_i \quad \forall i \in \{1, \dots, n\},$$

$$(4.3b) \qquad \sum_{i=1}^{n} y_i = 1,$$

$$(4.3c) \qquad y_i \geq 0 \qquad \forall i \in \{1, \dots, n\},$$

$$(4.3d) \qquad y \in \mathbb{Z}^n.$$

Formulation (4.3) is locally ideal and can be constructed using a well-known LP modeling trick or by using standard MIP formulation techniques (see Examples 6

and 8). Formulation (4.3)'s only auxiliary variables are the binary variables that are used to indicate which P^i contains x. The following example shows that these binary variables might not be enough to construct a polynomial-sized sharp formulation.

EXAMPLE 4. *For $i \in \{1, \ldots, n\}$, let $v^i, w^i \in \mathbb{Q}^n$ be defined by*

$$v^i_j = \begin{cases} n, & j = i, \\ -1, & j \neq i, \end{cases} \qquad w^i_j = \begin{cases} -1, & j = i, \\ 0, & j \neq i, \end{cases}$$

for all $j \in \{1, \ldots, n\}$. In addition, let $v^0, w^0 \in \mathbb{Q}^n$ be defined by $v^0_j = -w^0_j = -1$ for all $j \in \{1, \ldots, n\}$.

Let $S = (V \times \{0\}) \cup (W \times \{1\}) \subseteq \mathbb{Q}^{n+1}$, where $V = \mathrm{conv}\left(\{v^i\}_{i=0}^{n}\right)$ and $W = \mathrm{conv}\left(\{w^i\}_{i=0}^{n}\right)$. S and $\mathrm{conv}(S)$ are depicted for $n = 2$ in Figure 4.2, where S is shown in red and $\mathrm{conv}(S)$ is shown in blue. By noting that

$$(4.4\mathrm{a}) \quad V \times \{0\} = \left\{ x \in \mathbb{Q}^{n+1} : x_{n+1} = 0, \ \sum_{j=1}^{n} x_j \leq 1, \ -x_j \leq 1 \ \forall j \in \{1, \ldots, n\} \right\}$$

and

$$(4.4\mathrm{b}) \quad W \times \{1\} = \left\{ x \in \mathbb{Q}^{n+1} : \begin{array}{l} x_{n+1} = 1, \ -\sum_{j=1}^{n} x_j \leq 1, \\ (n+1)x_i - \sum_{j=1}^{n} x_j \leq 1 \ \forall i \in \{1, \ldots, n\} \end{array} \right\}$$

we can check that a valid formulation of S is given by

$$(4.5\mathrm{a}) \qquad -x_j \leq 1 \qquad\qquad\qquad \forall j \in \{1, \ldots, n\},$$

$$(4.5\mathrm{b}) \qquad \sum_{j=1}^{n} x_j \leq 1 + (n-1)(1 - y_1),$$

$$(4.5\mathrm{c}) \qquad x_{n+1} = 1 - y_1,$$

$$(4.5\mathrm{d}) \qquad (n+1)x_i - \sum_{j=1}^{n} x_j \leq 1 + (n^2 + n - 2)(1 - y_2) \quad \forall i \in \{1, \ldots, n\},$$

$$(4.5\mathrm{e}) \qquad -\sum_{j=1}^{n} x_j \leq 1 + (n-1)(1 - y_2),$$

$$(4.5\mathrm{f}) \qquad x_{n+1} = y_2,$$

$$(4.5\mathrm{g}) \qquad y_1 + y_2 = 1,$$

$$(4.5\mathrm{h}) \qquad y \in \{0, 1\}^2.$$

Now for $n = 3$, $x_1 = x_2 = -1$, $x_3 = 0$, and $x_4 = y_1 = y_2 = 1/2$ is feasible for the LP relaxation of (6.8), but violates $x_1 + x_2 - x_4 \geq -2$, which is a facet-defining inequality of $\mathrm{conv}(S)$.

Although Figure 4.2 shows that $\mathrm{conv}(S)$ has few facets for small n, the following lemma shows that the number of facets of $\mathrm{conv}(S)$ grows exponentially in n. Furthermore, the lemma shows that the two binary variables used by (4.5) are not enough
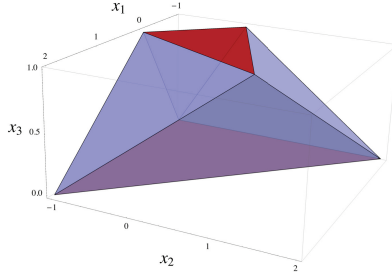
Fig. 4.2   *Set for Example 4 for $n = 2$.*

to yield a polynomial-sized sharp formulation even if a constant (independent of n) number of additional auxiliary variables is used.

LEMMA 4.1. *Let S be the set defined in Example 4. Then the number of facets of $\mathrm{conv}(S)$ grows exponentially in n. Furthermore, there is no sharp formulation of S of the form*

$$Ax + B\lambda + Dy \leq b,\ x \in \mathbb{Q}^{n+1},\ \lambda \in \mathbb{Q}^{k},\ y \in \mathbb{Z}^2,$$

where $A \in \mathbb{Q}^{p(n) \times (n+1)}$, $B \in \mathbb{Q}^{p(n) \times k}$, $D \in \mathbb{Q}^{p(n) \times 2}$, and $b \in \mathbb{Q}^{p(n)}$ for some polynomial p and a constant $k \in \mathbb{Z}_+$ independent of n.

Proof. To prove the first statement we note that V and W are two n-dimensional simplicies that are dual to each other. Then, $\mathrm{conv}(S)$ is the *antiprism* of V [28] and V satisfies the conditions for Theorem 2.1 in [28]. Hence, by this theorem, the number of facets of $\mathrm{conv}(S)$ is exactly two more than the number of proper faces of V. The number of proper faces of an n-dimensional simplex is $\sum_{i=0}^{n-1} \binom{n+1}{i+1} = 2^{n+1} - 2$ [63], so we conclude that the number of facets of $\mathrm{conv}(S)$ is precisely 2^{n+1}.

For the second statement we note that by Propositions 3.11 and 3.13 the number of facets of the projection of the LP relaxation of the proposed formulation onto the x variables is at most the number of extreme rays of cone

$$\left\{ \mu \in \mathbb{Q}_+^{p(n)} : D^{\mathsf{T}} \mu = 0,\ B^{\mathsf{T}} \mu = 0 \right\}.$$

By Lemma 3.5, the number of extreme rays of this cone is at most $\binom{p(n)}{p(n)-3-k}$, which is also a polynomial. Hence, no formulation of this form can have an LP relaxation that projects to $\mathrm{conv}(S)$. $\square$

Fortunately, by allowing a growing number of auxiliary variables, the following proposition by Balas, Jeroslow, and Lowe [13, 90, 113] yields polynomial-sized formulations for a wide range of disjunctive constraints that include the set in Example 4. We postpone the proof of this proposition to section 5.1, where we consider a slightly more general version of this formulation.

PROPOSITION 4.2. *Let $\left\{ P^i \right\}_{i=1}^{k}$ be a finite family of polyhedra with a common recession cone (i.e., $P_\infty^i = P_\infty^j$ for all i, j), such that $P^i = \left\{ x \in \mathbb{Q}^n : A^i x \leq b^i \right\}$ for all i. Then, a locally ideal MIP formulation of $S = \bigcup_{i=1}^{k} P^i$ is given by*

$$(4.6a) \qquad\qquad A^i x^i \leq b^i y_i \quad \forall i \in \{1, \dots, k\},$$

$$(4.6b) \qquad\qquad \sum_{i=1}^{k} x^i = x,$$

$$(4.6c) \qquad \sum_{i=1}^{k} y_i = 1,$$

$$(4.6d) \qquad y_i \geq 0 \qquad \forall i \in \{1, \ldots, k\},$$

$$(4.6e) \qquad x^i \in \mathbb{Q}^n \qquad \forall i \in \{1, \ldots, k\},$$

$$(4.6f) \qquad y \in \mathbb{Z}^k.$$

EXAMPLE 5. *Using formulation (4.6) for characterization (4.4) of set S in Example 4 results in the following polynomial-sized sharp (and locally ideal) formulation that uses a linear number of additional continuous auxiliary variables:*

$$(4.7a) \qquad x_{n+1}^1 = 0,$$

$$(4.7b) \qquad \sum_{j=1}^{n} x_j^1 \leq y_1,$$

$$(4.7c) \qquad -x_i^1 \leq y_1 \qquad \forall i \in \{1, \ldots, n\},$$

$$(4.7d) \qquad x_{n+1}^2 = y_2,$$

$$(4.7e) \qquad -\sum_{j=1}^{n} x_j^2 \leq y_2,$$

$$(4.7f) \qquad (n+1)x_i^2 - \sum_{j=1}^{n} x_j^2 \leq y_2 \qquad \forall i \in \{1, \ldots, n\},$$

$$(4.7g) \qquad x_i = x_i^1 + x_i^2 \quad \forall i \in \{1, \ldots, n+1\},$$

$$(4.7h) \qquad y_1 + y_2 = 1,$$

$$(4.7i) \qquad y \in \{0, 1\}^2,$$

$$(4.7j) \qquad x^1, x^2 \in \mathbb{Q}^{n+1},$$

$$(4.7k) \qquad x \in \mathbb{Q}^{n+1}.$$

The use of a growing number of auxiliary variables in Proposition 4.2 and similar techniques allows the construction of polynomial-sized sharp extended formulations for a wide range of sets. However, there are cases in which these formulations cannot be constructed. Most examples of sets that do not have polynomial-sized extended formulations arise from intractable combinatorial optimization problems (e.g., the traveling salesman problem considered in [53]). However, the following recent result by Rothvoß [140] shows that this can also happen for polynomially solvable combinatorial optimization problems.

THEOREM 4.3. *Let $G = (V, E)$ be the complete graph on $|V| = n$ nodes where $V = \{1, \ldots, n\}$ and $E = \{\{i, j\} : i, j \in \{1, \ldots, n\}, \quad i \neq j\}$. Let S be the set of incident vectors of perfect matchings of G given by*

$$(4.8) \qquad S := \left\{ x \in \{0, 1\}^E : \sum_{j \in \{1, \ldots, n\} \setminus \{i\}} x_{\{i,j\}} = 1 \quad \forall i \in \{1, \ldots, n\} \right\}.$$

If n is even, then there is no polynomial-sized sharp extended formulation for S.

The proof techniques used to show results such as Theorem 4.3 are significantly more elaborate than those used in Lemma 4.1 and are beyond the scope of this survey. We refer the reader interested in these techniques to [53, 140] and their references and to [35, 95].

5. Basic Extended Formulations. Disjunctive constraints can model a wide range of logical constraints. However, there are other aspects of MIP formulations that are generally encountered in practice, such as feasible regions of knapsacks or other problems with general integer variables. One class of sets that combines these two aspects is that of the unions of *mixed integer sets* of the form

$$(5.1) \qquad S = \bigcup_{i=1}^{k} P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}),$$

where $\left\{ P^i \right\}_{i=1}^{k}$ is a finite family of polyhedra with a common recession cone. Hooker [76] showed that such sets can be modeled through a simple extension of formulation (4.6). Hooker also showed that this extension is sharp if the formulations of these mixed integer sets are sharp (i.e., if $P^i = \text{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right)$). However, achieving this could require a large number of inequalities in the descriptions of the P^is. Fortunately, as noted in [37], we may significantly reduce the number of inequalities by using auxiliary variables in the description of P^i. This results in the following generalization of formulation (4.6) that also yields a locally ideal or sharp formulation when locally ideal or sharp extended formulations of $P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ are used.

PROPOSITION 5.1. *Let* $\left\{ P^i \right\}_{i=1}^{k}$ *be a finite family of polyhedra with a common recession cone (i.e.,* $P^i_\infty = P^j_\infty$ *for all* $i, j \in \{1, \ldots, k\}$*) and* $p \in \mathbb{Z}^k_+$ *be such that* $P^i = \left\{ x \in \mathbb{Q}^n : \exists w \in \mathbb{Q}^{p_i} \text{ s.t. } (x, w) \in Q^i \right\}$*, where*

$$Q^i = \left\{ (x, w) \in \mathbb{Q}^n \times \mathbb{Q}^{p_i} : A^i x + D^i w \le b^i \right\}$$

for $A^i \in \mathbb{Q}^{m_i \times n}$*,* $D^i \in \mathbb{Q}^{m_i \times p_i}$*, and* $b^i \in \mathbb{Q}^{m_i}$ *for each* $i \in \{1, \ldots, k\}$*. Then, for any* $n_1, n_2 \in \mathbb{Z}_+$ *such that* $n = n_1 + n_2$*, an MIP formulation of* $S = \bigcup_{i=1}^{k} P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ *is given by*

$$(5.2a) \qquad A^i x^i + D^i w^i \le b^i y_i \qquad \forall i \in \{1, \ldots, k\},$$

$$(5.2b) \qquad \sum_{i=1}^{k} x^i = x,$$

$$(5.2c) \qquad \sum_{i=1}^{k} y_i = 1,$$

$$(5.2d) \qquad y_i \ge 0 \qquad \forall i \in \{1, \ldots, k\},$$

$$(5.2e) \qquad x^i \in \mathbb{Q}^n \qquad \forall i \in \{1, \ldots, k\},$$

$$(5.2f) \qquad w^i \in \mathbb{Q}^{p_i} \qquad \forall i \in \{1, \ldots, k\},$$

$$(5.2g) \qquad y \in \mathbb{Z}^k,$$

$$(5.2h) \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}.$$

Furthermore, if $P^i = \text{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right)$ *for all* $i \in \{1, \ldots, k\}$*, then* (5.2) *is sharp, and if* $\text{ext}\left(Q^i \right) \subseteq \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2} \times \mathbb{Q}^{p_i}$ *for all* $i \in \{1, \ldots, k\}$*, then* (5.2) *is locally ideal.*

Proof. For validity of (5.2), without loss of generality, assume $y_1 = 1$ and $y_i = 0$ for all $i \ge 2$. Then $x^1 \in P^1$ and by Proposition 3.11 we have that $x^i \in P^i_\infty$ for all $i \ge 2$. Then by the common recession cone assumption we have $x^i \in P^1_\infty$ for all $i \ge 2$ and hence $x \in P^1$.

To prove sharpness let (x, w, y) be feasible for the LP relaxation of (5.2) and $I = \{i \in \{1, \ldots, k\} : y_i > 0\}$. Then $x = \sum_{i \in I} y_i \left(x^i/y_i\right)$, $\sum_{i \in I} y_i = 1$, $y \geq 0$, and $\left(x^i/y_i, w^i/y_i\right) \in Q^i$ for all $i \in I$. By the assumption on P^i we then have that $x^i/y_i \in \mathrm{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})\right)$ and hence

$$x \in \mathrm{conv}\left(\bigcup_{i=1}^{k} \mathrm{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})\right)\right) = \mathrm{conv}\left(\bigcup_{i=1}^{k} P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})\right) = \mathrm{conv}(S).$$

To prove locally idealness first note that if (x, w, y) is an extreme point of the LP relaxation of (5.2) and $y \in \{0, 1\}^k$, we may again assume without loss of generality that $y_1 = 1$ and $y_i = 0$ for all $i \geq 2$. Then by extremality of (x, w, y) we have that $x^i = 0$ and $w^i = 0$ for all $i \geq 2$, $x = x^1$, and $\left(x^1, w^1\right) \in \mathrm{ext}\left(Q^1\right)$. Then, by the assumption on Q^1 we have $x = x^1 \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}$ and hence (x, w, y) satisfies the integrality constraints. To finish the proof, assume for a contradiction that (x, w, y) is an extreme point of the LP relaxation of (5.2) such that $y \notin \{0, 1\}^k$. Without loss of generality we may assume that $y_1, y_2 \in (0, 1)$. Let $\varepsilon = \min\{y_1, y_2, 1 - y_1, 1 - y_2\} \in (0, 1)$,

$$\underline{y}_i = \begin{cases} y_i + \varepsilon, & i = 1, \\ y_i - \varepsilon, & i = 2, \\ y_i, & \text{otherwise,} \end{cases} \qquad \overline{y}_i = \begin{cases} y_i - \varepsilon, & i = 1, \\ y_i + \varepsilon, & i = 2, \\ y_i & \text{otherwise.} \end{cases}$$

$\underline{x}^i = (\underline{y}_i/y_i)x^i$, $\underline{w}^i = (\underline{y}_i/y_i)w^i$, $\overline{x}^i = (\overline{y}_i/y_i)x^i$, and $\overline{w}^i = (\overline{y}_i/y_i)w^i$ for $i \in \{1, 2\}$, $\underline{x}^i = \overline{x}^i = x^i$ and $\underline{w}^i = \overline{w}^i = w^i$ for all $i \notin \{1, 2\}$, $\underline{x} = \sum_{i=1}^{k} \underline{x}^i$, and $\overline{x} = \sum_{i=1}^{k} \overline{x}^i$. Then $(\underline{x}, \underline{w}, \underline{y}) \neq (\overline{x}, \overline{w}, \overline{y})$, $(x, w, y) = (1/2)(\underline{x}, \underline{w}, \underline{y}) + (1/2)(\overline{x}, \overline{w}, \overline{y})$, and $(\underline{x}, \underline{w}, \underline{y})$, $(\overline{x}, \overline{w}, \overline{y})$ are feasible for the LP relaxation of (5.2), which contradicts (x, w, y) being an extreme point. $\square$

Formulation 5.2 can be used to construct several known formulations for piecewise linear functions and more general disjunctive constraints. The following proposition illustrates this by constructing a variant of (5.2) that is convenient when P^i are described through their extreme points and rays. The resulting formulation is a straightforward extension of a formulation for disjunctive constraints introduced by Jeroslow and Lowe [90, 113].

COROLLARY 5.2. *Let $\{P^i\}_{i=1}^{k} \subseteq \mathbb{Q}^n$ be a finite family of polyhedra with a common recession cone C. Then, for any $n_1, n_2 \in \mathbb{Z}_+$ such that $n = n_1 + n_2$, an MIP formulation of $S = \bigcup_{i=1}^{k} P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ is given by*

$$\text{(5.3a)} \quad \sum_{i=1}^{k} \sum_{v \in \mathrm{ext}(P^i)} v\lambda_v^i + \sum_{r \in \mathrm{ray}(C)} r\mu_r = x,$$

$$\text{(5.3b)} \quad \sum_{v \in \mathrm{ext}(P^i)} \lambda_v^i = y_i \qquad \forall i \in \{1, \ldots, k\},$$

$$\text{(5.3c)} \quad \sum_{i=1}^{k} y_i = 1,$$

$$\text{(5.3d)} \quad \lambda_v^i \geq 0 \qquad \forall i \in \{1, \ldots, k\}, \, v \in \mathrm{ext}\left(P^i\right),$$

$$\text{(5.3e)} \quad \mu_r \geq 0 \qquad \forall r \in \mathrm{ray}(C),$$

$$\text{(5.3f)} \quad y \in \{0, 1\}^k,$$

$$\text{(5.3g)} \quad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}.$$

Furthermore, if $P^i = \text{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})\right)$ *for all* $i \in \{1, \dots, k\}$, *then* (5.3) *is a locally ideal formulation of* S.

    Proof. By Theorem 3.6 we have that P^i is the projection onto the x variables of

$$Q^i = \left\{ (x, \mu, \lambda) \in \mathbb{Q}^n \times \mathbb{Q}^{\text{ray}(C)} \times \mathbb{Q}^{\text{ext}(P^i)} : \begin{array}{c} \displaystyle\sum_{v \in \text{ext}(P^i)} v\lambda_v + \sum_{r \in \text{ray}(C)} r\mu_r = x \\[2mm] \displaystyle\sum_{v \in \text{ext}(P^i)} \lambda_v = 1 \\[2mm] \lambda_v \geq 0 \quad \forall v \in \text{ext}\left(P^i\right) \\[1mm] \mu_r \geq 0 \quad \forall r \in \text{ray}\left(C\right) \end{array} \right\}$$

for all $i \in \{1, \dots, k\}$. Using these extended formulations of the Q^is we have that formulation (5.2) for S is given by

$$\text{(5.4a)} \qquad \sum_{v \in \text{ext}(P^i)} v\lambda_v^i + \sum_{r \in \text{ray}(C)} r\mu_r^i = x^i \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4b)} \qquad \sum_{v \in \text{ext}(P^i)} \lambda_v^i = y_i \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4c)} \qquad \sum_{i=1}^{k} x^i = x,$$

$$\text{(5.4d)} \qquad \sum_{i=1}^{k} y_i = 1,$$

$$\text{(5.4e)} \qquad \lambda_v^i \geq 0 \qquad \forall i \in \{1, \dots, k\}, v \in \text{ext}\left(P^i\right),$$

$$\text{(5.4f)} \qquad \mu_r^i \geq 0 \qquad \forall i \in \{1, \dots, k\}, r \in \text{ray}\left(C\right),$$

$$\text{(5.4g)} \qquad y_i \geq 0 \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4h)} \qquad x^i \in \mathbb{Q}^n \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4i)} \qquad \lambda^i \in \mathbb{Q}^{\text{ext}(P^i)} \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4j)} \qquad \mu^i \in \mathbb{Q}^{\text{ray}(C)} \qquad \forall i \in \{1, \dots, k\},$$

$$\text{(5.4k)} \qquad y \in \mathbb{Z}^k,$$

$$\text{(5.4l)} \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}.$$

We claim that the LP relaxation of (5.3) is equal to the image of the LP relaxation of (5.4) through the linear transformation that projects out the x^i and μ^i variables and lets $\mu = \sum_{i=1}^{k} \mu^i$. We first show that the LP relaxation of (5.3) is contained in the image of the LP relaxation of (5.4). For this, simply note that any point in the LP relaxation of (5.3) is the image of the point in the LP relaxation of (5.4) obtained by letting $x^i = \sum_{v \in \text{ext}(P^i)} v\lambda_v^i + (1/k) \sum_{r \in \text{ray}(C)} r\mu_r$ and $\mu_r^i = (1/k)\mu_r$ for all $i \in \{1, \dots, k\}$ and $r \in \text{ray}\left(C\right)$. The reverse containment is straightforward. Validity then follows directly from Proposition 5.1.

    For locally idealness note that by Proposition 2.5 and the assumption on P^i we have that $\text{ext}\left(P^i\right) \subseteq \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}$. By noting that $(x, \mu, \lambda) \in \text{ext}\left(Q^i\right)$ if and only if $x \in \text{ext}\left(P^i\right)$, $\mu = 0$, $\lambda_x = 1$, and $\lambda_v = 0$ for $v \in \text{ext} P^i \setminus \{x\}$, we have that $\text{ext}\left(Q^i\right) \subseteq \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2} \times \mathbb{Q}^{\text{ray}(C)} \times \mathbb{Q}^{\text{ext}(P^i)}$. Then by Proposition 5.1 we have that (5.4)

is locally ideal and hence so is (5.3) by Proposition 3.10 and the linear transformation described in the previous paragraph.   ☐

To distinguish extreme point/ray formulation (5.3) from inequality formulation (5.2), we refer to the first one as the $\mathcal{V}$-formulation and to the second as the $\mathcal{H}$-formulation. While $\mathcal{V}$-formulation (5.3) can be derived from $\mathcal{H}$-formulation (5.2), their application to specific disjunctive constraints can yield formulations with very different structures. The following two examples illustrate this for the case $n_2 = 0$ in both formulations and $p_i = 0$ for all $i \in \{1, \ldots, k\}$ in Proposition 5.1.

EXAMPLE 6. *Consider the set* $S = \bigcup_{i=1}^{n} \{x \in \mathbb{Q}^n : -1 \leq x_i \leq 1, x_j = 0 \,\forall j \neq i\}$ *from Example 2. $\mathcal{H}$-formulation (5.2) for S is given by*

$$-y_i \leq x_i^i \leq y_i \quad \forall i \in \{1, \ldots, n\}, \tag{5.5a}$$

$$x_i^j = 0 \qquad \forall i, j \in \{1, \ldots, n\}, i \neq j, \tag{5.5b}$$

$$\sum_{i=1}^{n} y_i = 1, \tag{5.5c}$$

$$x_i = \sum_{j=1}^{n} x_i^j \quad \forall i \in \{1, \ldots, n\}, \tag{5.5d}$$

$$y \in \{0, 1\}^n. \tag{5.5e}$$

Similarly to the proof of Proposition 5.1, we can show that the projection of (5.5) onto the x and y variables is precisely formulation (4.3) from Example 3, which is given by

$$-y_i \leq x_i \leq y_i \quad \forall i \in \{1, \ldots, n\}, \tag{5.6a}$$

$$\sum_{i=1}^{n} y_i = 1, \tag{5.6b}$$

$$y_i \geq 0 \qquad \forall i \in \{1, \ldots, n\}, \tag{5.6c}$$

$$y \in \mathbb{Z}^n. \tag{5.6d}$$

If we instead use $\mathcal{V}$-formulation (5.3), we obtain the alternative formulation of S given by

$$\lambda_1^i - \lambda_2^i = x_i \qquad \forall i \in \{1, \ldots, n\}, \tag{5.7a}$$

$$\lambda_1^i + \lambda_2^i = y_i \qquad \forall i \in \{1, \ldots, n\}, \tag{5.7b}$$

$$\sum_{i=1}^{n} y_i = 1, \tag{5.7c}$$

$$\lambda_1^i, \lambda_2^i \geq 0 \qquad \forall i \in \{1, \ldots, n\}, \tag{5.7d}$$

$$y \in \{0, 1\}^n. \tag{5.7e}$$

It is interesting to contrast formulations (5.6) and (5.7). We know that $\mathrm{conv}(S) = \{x \in \mathbb{Q}^n : \sum_{i=1}^{n} |x_i| \leq 1\}$ *and that both formulations are sharp. Hence the LP relaxations of both formulations should contain lifted representations of $\sum_{i=1}^{n} |x_i| \leq 1$. $\mathcal{H}$-formulation (5.6) does this using the standard trick of modeling $|x_i| \leq y_i$ as $-y_i \leq x_i \leq y_i$, while $\mathcal{V}$-formulation (5.7) does it by the alternative trick of modeling $|x_i| \leq y_i$ as $x_i = \lambda_1^i - \lambda_2^i$, $\lambda_1^i + \lambda_2^i = y_i$, and $\lambda_1^i, \lambda_2^i \geq 0$. The latter trick uses the fact that $x = x^+ - x^-$ and $|x| = x^+ + x^-$, where $x^+ := \max\{x, 0\}$ and $x^- := \{-x, 0\}$.*

EXAMPLE 7.   *Let $f : D \subseteq \mathbb{Q}^d \to \mathbb{Q}$ be a piecewise linear function defined by (3.3) for a given finite family of polytopes $\{Q^i\}_{i=1}^k$. Formulation (5.2) for characterization (3.4a) of $\mathrm{gr}(f)$ results in the MIP formulation of $\mathrm{gr}(f)$ given by*

$$(5.8a) \qquad\qquad z = \sum_{i=1}^k z_i,$$

$$(5.8b) \qquad\qquad z_i = m^i x^i + c_i,$$

$$(5.8c) \qquad\qquad x = \sum_{i=1}^k x^i,$$

$$(5.8d) \qquad\qquad A^i x^i \le y_i b^i \qquad\qquad \forall i \in \{1, \dots, k\},$$

$$(5.8e) \qquad\qquad \sum_{i=1}^k y_i = 1,$$

$$(5.8f) \qquad\qquad y \in \{0, 1\}^k.$$

If we replace (5.8a) and (5.8b) in (5.8) by

$$(5.9) \qquad\qquad z = \sum_{i=1}^k m^i x^i + c_i,$$

we obtain a standard MIP formulation for piecewise linear functions that is denoted the multiple choice model *in [155]. Similarly to the proof of Proposition 5.1, we can show that the multiple choice model is the projection onto the x, x^i, y, and z variables of (5.8). If we instead use formulation (5.3) for characterization (3.4b) of $\mathrm{gr}(f)$, we obtain the formulation of $\mathrm{gr}(f)$ given by*

$$(5.10a) \qquad\qquad \sum_{i=1}^k \sum_{v \in \mathrm{ext}(Q^i)} v \lambda_v^i = x,$$

$$(5.10b) \qquad\qquad \sum_{i=1}^k \sum_{v \in \mathrm{ext}(Q^i)} f(v) \lambda_v^i = z,$$

$$(5.10c) \qquad\qquad \sum_{v \in \mathrm{ext}(Q^i)} \lambda_v^i = y_i \qquad\qquad \forall i \in \{1, \dots, k\},$$

$$(5.10d) \qquad\qquad \sum_{i=1}^k y_i = 1,$$

$$(5.10e) \qquad\qquad \lambda_v^i \ge 0 \qquad\qquad \forall i \in \{1, \dots, k\}, \, v \in \mathrm{ext}\left(Q^i\right),$$

$$(5.10f) \qquad\qquad y \in \{0, 1\}^k,$$

which is a standard MIP formulation for piecewise linear functions that is denoted the disaggregated convex combination model *in [155].*

For examples of formulation (5.2) with $p_i > 0$ and $n_2 > 0$ we refer the reader to [37] and [76], respectively.

6. Projected Formulations. One disadvantage of basic formulations (5.2) and (5.4) is that they require multiple copies of the original x variables (i.e., the x^i variables) or of some auxiliary variables (i.e., the λ^i variables). In this section we study formulations that do not use these copies of variables and are hence much smaller. The cost of this reduction in variables is usually a loss of strength, but it some cases

this loss can be avoided. We begin by considering the well-known Big-M approach and an alternative formulation that combines the Big-M approach with a formulation tailored to polyhedra with a common structure. We then consider the strength of both classes of formulations in detail.

6.1. Traditional Big-M Formulations. One way to write MIP formulations without using copies of the original variables is to use the *Big-M* technique. The following proposition proven in [164] for the bounded case and in [76] for the unbounded describes the technique.

PROPOSITION 6.1. *Let* $\left\{P^i\right\}_{i=1}^k$ *be a finite family of polyhedra with a common recession cone (i.e.,* $P_\infty^i = P_\infty^j$ *for all* i,j*), such that* $P^i = \left\{x \in \mathbb{Q}^n : A^i x \leq b^i\right\}$*, where* $A^i \in \mathbb{Q}^{m_i \times n}$ *and* $b^i \in \mathbb{Q}^{m_i}$ *for all* $i \in \{1,\ldots,k\}$*. Furthermore, for each* $i,j \in \{1,\ldots,k\}$*,* $i \neq j$*, and* $l \in \{1,\ldots,m_i\}$*, let* $M_l^{i,j} \in \mathbb{Q}$ *be such that*

$$(6.1) \qquad M_l^{i,j} \geq \max_x \left\{a^{i,l} x \,:\, A^j x \leq b^j\right\},$$

where $a^{i,l} \in \mathbb{Q}^n$ *is the lth row of* A^i*. Then, for any* $n_1, n_2 \in \mathbb{Z}_+$ *such that* $n = n_1 + n_2$*, an MIP formulation for* $S = \bigcup_{i=1}^k P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ *is given by*

$$(6.2a) \qquad A^i x \leq b^i + \left(M^i - b^i\right)(1 - y_i) \quad \forall i \in \{1,\ldots,k\},$$

$$(6.2b) \qquad \sum_{i=1}^k y_i = 1,$$

$$(6.2c) \qquad y_i \geq 0 \qquad\qquad\qquad \forall i \in \{1,\ldots,k\},$$

$$(6.2d) \qquad y \in \mathbb{Z}^k,$$

$$(6.2e) \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2},$$

where $M^i \in \mathbb{Q}^{m_i}$ *are such that* $M_l^i = \max_{j \neq i} M_l^{i,j}$ *for each* $l \in \{1,\ldots,m_i\}$*.*

Proof. If all $M_l^{i,j}$ are finite, validity of the formulation is straightforward. To show their finiteness, assume for a contradiction that for a given i, j, and l the left-hand side of (6.1) is infinite. The unboundedness of this LP problem is equivalent to the existence of an $r \in P_\infty^j = \left\{x \in \mathbb{Q}^n : A^j x \leq 0\right\}$ such that $a^{i,l} r > 0$ [21, Theorem 4.13]. However, by the assumption on the recession cones, such an r is also in P_∞^i and hence the LP problem given by $\max_x \left\{a^{i,l} x : A^i x \leq b^i\right\}$ is unbounded, which contradicts b_l^i being finite. $\square$

The strongest possible version of formulation (6.2) is obtained when equality holds in (6.1), which, as illustrated in the following example, can sometimes yield sharp or locally ideal formulations.

EXAMPLE 8. *Consider the set* S *from Example 2, which corresponds to the case* $A^i = \left[\begin{smallmatrix} I \\ -I \end{smallmatrix}\right] \in \mathbb{Q}^{2n \times n}$ *and*

$$b_l^i = \begin{cases} 1, & l \in \{i, k+i\}, \\ 0 & otherwise \end{cases}$$

for all $i, l \in \{1,\ldots,k\}$*, and* $n_2 = 0$*. A valid Big-M selection is to take* $M_l^i = 1$ *for all* $i, l \in \{1,\ldots,k\}$*. For this choice,* (6.2) *is equal to the nonsharp formulation* (4.1)*. In contrast, the strongest choice of* M^i *given by*

$$M_l^i = \begin{cases} 0, & l \in \{i, k+i\}, \\ 1 & otherwise \end{cases}$$

yields locally ideal formulation (4.3)*.*

Unfortunately, as the following example shows, even the strongest version of (6.2) can fail to be sharp.

EXAMPLE 9. *The strongest version of formulation* (6.2) *for S from Example 4 is precisely the nonsharp formulation* (4.5).

For more discussion about Big-M formulations for general and specially structured sets, we refer the reader to [18, 76, 79, 164].

6.2. Hybrid Big-M Formulations. A different class of projected formulations was introduced by Balas, Blair, and Jeroslow [12, 26, 87] for families of polyhedra with a common left-hand side matrix (i.e., where $A^i = A^j$ for all i, j). Balas, Blair, and Jeroslow showed that such formulations can have very favorable strength properties. However, while the common left-hand side structure appears in many problems (see Example 10), these formulations still have more limited applicability than the traditional Big-M formulation from Proposition 6.1. For this reason we here combine the projected formulation of Balas, Blair, and Jeroslow with the traditional Big-M formulation to introduce the following hybrid Big-M formulation that generalizes the former and improves upon the latter. While this formulation does not require the common left-hand side structure, it is equipped to exploit it when present.

PROPOSITION 6.2. *For $k \in \mathbb{Z}_+^p$, let $\bigcup_{s=1}^p \left\{ P^{s,i} \right\}_{i=1}^{k_s}$ be a finite family of polyhedra with a common recession cone (i.e., $P_\infty^{s,i} = P_\infty^{t,j}$ for all s, t, i, j) such that*

$$P^{s,i} = \left\{ x \in \mathbb{Q}^n \, : \, A^s x \leq b^{s,i} \right\},$$

where $A^s \in \mathbb{Q}^{m_s \times n}$ and $b^{s,i} \in \mathbb{Q}^{m_s}$ for each $s \in \{1, \ldots, p\}$ and $i \in \{1, \ldots, k_s\}$. Furthermore, for each $s, t \in \{1, \ldots, p\}$, $i \in \{1, \ldots, k_s\}$, and $l \in \{1, \ldots, m_s\}$, let $M_l^{s,t,i} \in \mathbb{Q}$ be such that

$$(6.3) \qquad M_l^{s,t,i} \geq \max_x \left\{ a^{s,l} x \, : \, A^t x \leq b^{t,i} \right\},$$

where $a^{s,l} \in \mathbb{Q}^n$ is the lth row of A^s. Then, for any $n_1, n_2 \in \mathbb{Z}_+$ such that $n = n_1 + n_2$, an MIP formulation for $S = \bigcup_{s=1}^p \bigcup_{i=1}^{k_s} P^{s,i} \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ is given by

$$(6.4a) \qquad A^s x \leq \sum_{t=1}^p \sum_{i=1}^{k_t} M^{s,t,i} y_{t,i} \quad \forall s \in \{1, \ldots, p\},$$

$$(6.4b) \qquad \sum_{s=1}^p \sum_{i=1}^{k_s} y_{s,i} = 1,$$

$$(6.4c) \qquad y_{s,i} \geq 0 \qquad\qquad \forall s \in \{1, \ldots, p\}, \, i \in \{1, \ldots, k_s\},$$

$$(6.4d) \qquad y_{s,i} \in \mathbb{Z} \qquad\qquad \forall s \in \{1, \ldots, p\}, \, i \in \{1, \ldots, k_s\},$$

$$(6.4e) \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}.$$

In particular, we may take $M^{s,s,i} = b^{s,i}$ for all $s \in \{1, \ldots, p\}$, $i \in \{1, \ldots, k_s\}$.

Validity of this formulation is again straightforward from finiteness of the $M_l^{s,t,i}$, which can be proven analogously to Proposition 6.1. Furthermore, the strongest possible version of formulation (6.4) is again obtained when equality holds in (6.3). In particular, $M^{s,s,i} = b^{s,i}$ is the strongest possible coefficient unless some $P^{s,i}$ has a redundant constraint. Of course, even in the case of a redundant constraint, it is always valid and convenient to use Big-M constants such that $M^{s,s,i} \leq b^{s,i}$. For this reason, we assume this to be the case from now on.

Hybrid Big-M formulation (6.4) reduces to the formulation introduced by Balas, Blair, and Jeroslow when all left-hand side matrices are identical (i.e., for $p = 1$), and we take $M^{s,s,i} = b^{s,i}$ for all i. An advantage of this formulation is that it can be constructed without solving or bounding any LP problem in (6.3). In what follows we refer to this formulation as the *simple version* of (6.4). An example of what can be modeled with this simple version of the hybrid Big-M formulation is the following class of problem that includes the SOS1 and SOS2 constraints introduced by Beale and Tomlin [17].

EXAMPLE 10. *For given $\{l^i\}_{i=1}^k$, $\{u^i\}_{i=1}^k \subseteq \{0,1\}^n$, consider the family of polytopes $P^i := \{x \in \mathbb{Q}^n : l_j^i \le x_j \le u_j^i \quad \forall j \in \{1,\dots,n\}\}$ for $i \in \{1,\dots,k\}$ and its union $S = \bigcup_{i=1}^k P^i$. For instance, if we let $l_j^i = 0$ for all i, j and $k = n$ and*

$$u_j^i = \begin{cases} 1, & j = i, \\ 0 & otherwise \end{cases}$$

or $k = n - 1$ and

$$u_j^i = \begin{cases} 1, & j \in \{i, i+1\}, \\ 0 & otherwise, \end{cases}$$

we have that S corresponds, respectively, to the SOS1 and SOS2 constraints introduced by Beale and Tomlin [17]. For the cases of SOS1 and SOS2 constraints, formulation (6.4) with $p = 1$ reduces, respectively, to

$$0 \le x_i \le y_i \qquad \forall i \in \{1,\dots,k\},$$

$$\sum_{i=1}^k y_i = 1,$$

$$y \in \{0,1\}^k,$$

and

$$0 \le x_1 \le y_1,$$
$$0 \le x_i \le y_{i-1} + y_i \quad \forall i \in \{2,\dots,k\},$$
$$0 \le x_{k+1} \le y_k,$$
$$\sum_{i=1}^k y_i = 1,$$
$$y \in \{0,1\}^k,$$

which are the standard MIP formulations for such constraints.

We end this section by showing how the simple version of hybrid big-M formulation (6.4) can be used to obtain a smaller version of $\mathcal{V}$-formulation (5.3). This results in the extension of a known formulation for piecewise linear functions (e.g., [45, 91, 113, 106, 165]).

COROLLARY 6.3. *Let $\{P^i\}_{i=1}^k \subseteq \mathbb{Q}^n$ be a finite family of polyhedra with a common recession cone C and $V := \bigcup_{i=1}^k \text{ext}(P^i)$. Then, for any $n_1, n_2 \in \mathbb{Z}_+$ such that $n = n_1 + n_2$, two MIP formulations of $S = \bigcup_{i=1}^k P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})$ are given by*

(6.5a)
$$\sum_{v \in V} v\lambda_v + \sum_{r \in \text{ray}(C)} r\mu_r = x,$$

$$(6.5b) \qquad \sum_{v \in V} \lambda_v = 1,$$

$$(6.5c) \qquad \lambda_v \leq \sum_{i:v \in \text{ext}(P^i)} y_i \quad \forall v \in V,$$

$$(6.5d) \qquad \sum_{i=1}^{k} y_i = 1,$$

$$(6.5e) \qquad \lambda_v \geq 0 \qquad\qquad \forall v \in V,$$

$$(6.5f) \qquad \mu_r \geq 0 \qquad\qquad \forall r \in \text{ray}(C),$$

$$(6.5g) \qquad y \in \{0,1\}^k,$$

$$(6.5h) \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2},$$

and

$$(6.6a) \qquad \sum_{v \in V} v\lambda_v + \sum_{r \in \text{ray}(C)} r\mu_r = x,$$

$$(6.6b) \qquad \sum_{v \in V} \lambda_v = 1,$$

$$(6.6c) \qquad \sum_{v \in \text{ext}(P^i)} \lambda_v \geq y_i \qquad \forall i \in \{1, \ldots, k\},$$

$$(6.6d) \qquad \sum_{i=1}^{k} y_i = 1,$$

$$(6.6e) \qquad \lambda_v \geq 0 \qquad\qquad \forall v \in V,$$

$$(6.6f) \qquad \mu_r \geq 0 \qquad\qquad \forall r \in \text{ray}(C),$$

$$(6.6g) \qquad y \in \{0,1\}^k,$$

$$(6.6h) \qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}.$$

Furthermore, if $P^i = \text{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2})\right)$ for all $i \in \{1, \ldots, k\}$, then both formulations are sharp.

    Proof. Let

$$Q^i = \left\{ (x, \lambda, \mu) \in \mathbb{Q}^n \times \mathbb{Q}^V \times \mathbb{Q}^{\text{ray}(C)} : \begin{array}{l} \sum_{v \in V} v\lambda_v + \sum_{r \in \text{ray}(C)} r\mu_r = x \\ \sum_{v \in V} \lambda_v = 1 \\ \lambda_v \leq 1 \quad \forall v \in \text{ext}\left(P^i\right) \\ \lambda_v \leq 0 \quad \forall v \in V \setminus \text{ext}\left(P^i\right) \\ \lambda_v \geq 0 \quad \forall v \in V \\ \mu_r \geq 0 \quad \forall r \in \text{ray}(C) \end{array} \right\}.$$

Validity of (6.5) follows by noting that $\bigcup_{i=1}^{k} P^i$ is the projection onto the x variables of $\bigcup_{i=1}^{k} Q^i$ and that (6.5) is the simple version of (6.4) for $\bigcup_{i=1}^{k} Q^i$. Validity of (6.6)

follows by noting that Q^i can be described alternatively as

$$
Q^i = \left\{ (x, \lambda, \mu) \in \mathbb{Q}^n \times \mathbb{Q}^V \times \mathbb{Q}^{\mathrm{ray}(C)} \ : \
\begin{array}{l}
\displaystyle\sum_{v \in V} v\lambda_v + \sum_{r \in \mathrm{ray}(C)} r\mu_r = x \\
\displaystyle\sum_{v \in V} \lambda_v = 1 \\
\displaystyle\sum_{v \in \mathrm{ext}(P^i)} \lambda_v \geq 1 \\
\displaystyle\sum_{v \in \mathrm{ext}(P^j)} \lambda_v \geq 0 \quad \forall j \neq i \\
\lambda_v \geq 0 \quad \forall v \in V \\
\mu_r \geq 0 \quad \forall r \in \mathrm{ray}(C)
\end{array}
\right\}.
$$

For sharpness, note that the projection onto the x variables of the LP relaxation of both formulations is contained in $\mathrm{conv}\,(V) + C$. The result then follows because, under the assumption on P^i and by Theorem 3.6, we have

$$
\begin{aligned}
\mathrm{conv}(S) &= \mathrm{conv}\left(\bigcup_{i=1}^{k} P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right) \\
&= \mathrm{conv}\left(\bigcup_{i=1}^{k} \mathrm{conv}\left(P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}) \right) \right) \\
&= \mathrm{conv}\left(\bigcup_{i=1}^{k} P^i \right) \\
&= \mathrm{conv}\left(\bigcup_{i=1}^{k} \mathrm{conv}\left(\mathrm{ext}\left(P^i \right) \right) + C \right) \\
&= \mathrm{conv}\left(\bigcup_{i=1}^{k} \mathrm{conv}\left(\mathrm{ext}\left(P^i \right) \right) \right) + C \\
&= \mathrm{conv}\left(\bigcup_{i=1}^{k} \mathrm{ext}\left(P^i \right) \right) + C = \mathrm{conv}(V) + C. \qquad \square
\end{aligned}
$$

While (6.5) and (6.6) are equivalent, their LP relaxations are not contained in one another. Of course this can only happen because neither formulation is locally ideal. In fact, Lee and Wilson [106, 165] show that constraints (6.6c) are valid inequalities for the convex hull of integer feasible points of (6.5) and hence can be used to strengthen it. These inequalities are sometimes facet defining and are part of a larger class of inequalities that are enough to describe the convex hull of integer feasible points of (6.5). Unfortunately, the number of such inequalities can be exponential in k. However, in section 7 we show how an LP-based separation of these inequalities allows us to construct a different formulation that ends up being equivalent to (5.3). If we specialize formulation (6.5) to piecewise linear functions, we obtain the following formulation from [106, 165].

EXAMPLE 11. *Let $f : D \subseteq \mathbb{Q}^d \to \mathbb{Q}$ be a piecewise linear function defined by (3.3) for a given finite family of polytopes $\{Q^i\}_{i=1}^k$. Formulation (6.5) for characterization*

(3.4b) *of* $\mathrm{gr}(f)$ *results in the MIP formulation of* $\mathrm{gr}(f)$ *given by*

(6.7a)
$$\sum_{v \in \bigcup_{i=1}^{k} \mathrm{ext}(Q^i)} v \lambda_v = x,$$

(6.7b)
$$\sum_{v \in \bigcup_{i=1}^{k} \mathrm{ext}(Q^i)} f(v) \lambda_v = z,$$

(6.7c)
$$\lambda_v \leq \sum_{i : v \in \mathrm{ext}(Q^i)} y_i \quad \forall v \in \bigcup_{i=1}^{k} \mathrm{ext}\left(Q^i\right),$$

(6.7d)
$$\sum_{i=1}^{k} y_i = 1,$$

(6.7e)
$$\lambda_v \geq 0 \qquad \forall v \in \mathrm{ext}\left(Q^i\right),$$

(6.7f)
$$y \in \{0,1\}^k,$$

which is a standard MIP formulation for piecewise linear functions that is denoted the convex combination model in [155].

6.3. Formulation Strength. The traditional Big-M formulation has been significantly more popular than the simple version of the hybrid Big-M formulation. One reason for this is the somewhat limited applicability of the latter formulation. The general version of the hybrid Big-M formulation resolves this issue as it can be used in the same class of problems as the traditional Big-M formulation. In addition, the hybrid Big-M formulation provides an advantage over the traditional one with respect to size, even if only partial common left-hand side structure is present in the polyhedra. Indeed, both formulations have the same number of variables, but the hybrid formulation has $1 + \sum_{s=1}^{p} m_s$ constraints while the traditional one has $1 + \sum_{s=1}^{p} m_s \times k_s$ constraints. Of course, such an advantage in size is meaningless if it is accompanied by a significant reduction in strength. For this reason we now study the relative and absolute strengths of these formulations. In particular, concerning the hybrid Big-M formulation we show that it is always at least as strong as the traditional Big-M formulation, that its simple version can be sharp and strictly stronger than the traditional formulation, but that even its strongest version can fail to be locally ideal or sharp. We begin with the following proposition that shows that one case where the hybrid and traditional formulations coincide is when S is the union of only two polyhedra.

PROPOSITION 6.4. *If S is the union of two polyhedra whose description does not include any redundant constraints, and equality is taken in* (6.1) *and* (6.3)*, then the LP relaxations of formulations* (6.2) *and* (6.4) *are equivalent.*

Proof. For (6.2) the case of two polyhedra corresponds to $k = 2$, and in this case (6.2a) is equal to

$$A^1 x \leq b^1 + \left(M^{1,2} - b^1\right)(1 - y_1),$$
$$A^2 x \leq b^2 + \left(M^{2,1} - b^2\right)(1 - y_2).$$

For (6.3) the case of two polyhedra corresponds to $p = 1$ and $k_1 = 2$ or $p = 2$, $k_1 = 1$, and $k_2 = 1$. In the former case (6.4) is equal to

$$A^1 x \leq M^{1,1,1} y_1 + M^{1,1,2} y_2$$

and the equivalence follows from $y_1 + y_2 = 1$ by noting that for this case $A^1 = A^2$, $M^{1,2} = M^{1,1,2} = b^2$, and $M^{2,1} = M^{1,1,1} = b^1$, because of the nonredundancy assumption. In the latter, (6.4) is equal to

$$A^1 x \le M^{1,1,1} y_1 + M^{1,2,1} y_2,$$
$$A^2 x \le M^{2,1,1} y_1 + M^{2,2,1} y_2,$$

and the equivalence follows from $y_1 + y_2 = 1$ by noting that for this case $M^{1,1,1} = b^1$, $M^{2,2,1} = b^2$, $M^{1,2,1} = M^{1,2}$, and $M^{2,1,1} = M^{2,1}$, because of the nonredundancy assumption. □

Unfortunately, as the following example shows, this equal strength can fail to yield sharp formulations even for the case of equal left-hand side matrices and no redundant constraints.

EXAMPLE 12. *For*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \quad b^1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \quad b^1 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

let $P^1 = \{x \in \mathbb{Q}^3 : Ax \le b^1\}$ *and* $P^2 = \{x \in \mathbb{Q}^3 : Ax \le b^2\}$. *The strongest versions of formulations* (6.2) *and* (6.4) *for* $S = P^1 \cup P^2$ *are equal to*

(6.8a)
$$Ax \le b^1 + (1 - y_1) \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \end{pmatrix},$$

(6.8b)
$$Ax \le b^2 + (1 - y_2) \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

(6.8c)
$$y_1 + y_2 = 1,$$

(6.8d)
$$y \in \{0,1\}^2.$$

Formulation (6.8) *is not sharp because* $x_1 = x_2 = 0$, $x_3 = 3/2$, $y_1 = y_2 = 1/2$ *is feasible for its LP relaxation and* $x_3 \le 1$ *for any* $x \in \operatorname{conv}(S)$.

The lack of sharpness of formulation (6.8) in Example 12 can be corrected by simply adding $x_3 \le 1$, as $\operatorname{conv}(S)$ is exactly the projection of the LP relaxation of (6.8) onto the x variables plus this inequality. However, this correction cannot always be done with a polynomial number of inequalities (e.g., see Lemma 4.1). Furthermore, as the following theorem by Blair [26] shows, recognizing sharpness of even the simple version of hybrid Big-M formulation (6.4) is hard.

THEOREM 6.5. *Let* $p = 1$ *and* $M^{1,1,i} = b^{1,i}$ *for all* $i \in \{1, \dots, k\}$ *in formulation* (6.4). *Deciding whether the resulting formulation is sharp is NP-hard.*

Fortunately, Balas, Blair, and Jeroslow gave some practical (but not exhaustive) necessary and sufficient conditions for sharpness of the simple version of (6.4) [12, 26, 87]. An extremely useful example of such conditions is given by the following proposition.

DEFINITION 6.6. *For $b \in \mathbb{Q}^m$, $A \in \mathbb{Q}^{m \times n}$, and $I \subseteq \{1, \ldots, m\}$ such that $|I| = n$ and $\det(A_I) \neq 0$, we let A_I (b_I) be the submatrix (subvector) of A (b) obtained by selecting only the rows (components) indexed by I. We also let $x(I, b) \in \mathbb{Q}^n$ be the unique solution to $A_I x = b_I$. That is, $x(I, b)$ is the basic solution associated to basis I for polyhedron $\{x \in \mathbb{Q}^n : Ax \leq b\}$.*

Let $\{P^i\}_{i=1}^k$ be a finite family of $\mathcal{H}$-polyhedra such that for each i we have $P^i := \{x \in \mathbb{Q}^n : Ax \leq b^i\}$. We say that $\{P^i\}_{i=1}^k$ have the same shape if for all $I \subseteq \{1, \ldots, m\}$ such that $|I| = n$ and $\det(A_I) \neq 0$ we have either $x(I, b^i) \in P^i$ for all i or $x(I, b^i) \notin P^i$ for all i. In other words, polyhedra with a common structure have the same shape if and only if for every basis the associated basic solution is feasible for all polyhedra or is infeasible for all polyhedra.

PROPOSITION 6.7. *If $p = 1$, polyhedra $\{P^{1,i}\}_{i=1}^k$ have the same shape, and $M^{1,1,i} = b^{1,i}$ for all $i \in \{1, \ldots, k\}$, then (6.4) is sharp.*

One case in which these shape conditions hold are the sets in Example 10, as the polyhedra considered are all rectangles (possibly degenerate ones). In particular, this partially explains the success of the standard formulations for SOS1 and SOS2 constraints and gives an alternative proof of the sharpness of formulation (4.3) in Example 3.

Unfortunately, as the following example shows, the shape condition does not necessarily imply that formulation (6.4) is locally ideal.

EXAMPLE 13. *Let*

$$
A = \begin{bmatrix}
1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
-1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}, \quad
b^1 = \begin{pmatrix}
1 \\ -1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix}, \quad
b^2 = \begin{pmatrix}
1 \\ -1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix}, \quad
b^3 = \begin{pmatrix}
1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix},
$$

and $P^i = \{x \in \mathbb{Q}^4 : Ax \leq b^i\}$ for $i \in \{1, \ldots, 3\}$. These polyhedra satisfy the shape condition of Proposition 6.7 and hence the associated version of formulation (6.4) is sharp. However, we can check that $x_3 = x_4 = 1/2$, $x_1 = x_2 = 0$, $y_1 = y_3 = 1/2$, and $y_2 = 0$ is a fractional extreme point of the LP relaxation of (6.4) and hence the formulation is not locally ideal. Finally, note that this formulation is precisely parts (2.5c)–(2.5k) of the standard formulation (2.5) for piecewise linear functions, we saw in section 2.1.

Nevertheless, Proposition 6.7 can still be used to construct the following example that shows how simple hybrid Big-M formulation (6.4) can be strictly stronger than big-M formulation (6.2).

EXAMPLE 14. *Let $k = n - 1$ and $\{u^i\}_{i=1}^k \subseteq \{0, 1\}^n$ be such that*

$$
u_j^i = \begin{cases}
1, & j \in \{i, i+1\}, \\
0 & otherwise,
\end{cases}
$$

and let $P^i := \{x \in \mathbb{Q}^n : 0 \leq x_j \leq u_j^i \quad \forall j \in \{1, \ldots, n\}\}$ for $i \in \{1, \ldots, k\}$. Formula-

tion (6.4) *for* $S := \bigcup_{i=1}^{n} P^i$ *reduces to*

(6.9a) $$x_1 \leq y_1,$$

(6.9b) $$x_i \leq y_{i-1} + y_i \quad \forall i \in \{2, \ldots, k\},$$

(6.9c) $$x_{k+1} \leq y_k,$$

(6.9d) $$\sum_{i=1}^{k} y_i = 1,$$

(6.9e) $$y \in \{0, 1\}^k,$$

and the strongest version of Big-M formulation (6.2) *for* S *reduces to*

(6.10a) $$x_1 \leq y_1,$$

(6.10b) $$x_j \leq 1 - y_i \quad \forall j \in \{1, \ldots, k+1\}, i \in \{1, \ldots, k\} \setminus \{j - 1, j\},$$

(6.10c) $$x_{k+1} \leq y_k,$$

(6.10d) $$\sum_{i=1}^{k} y_i = 1,$$

(6.10e) $$y \in \{0, 1\}^k.$$

LEMMA 6.8. *Hybrid Big-M formulation* (6.9) *is always sharp, while traditional Big-M formulation* (6.10) *is not sharp for* $k \geq 4$.

Proof. The first formulation is sharp by Proposition 6.7.

For the lack of sharpness of the second formulation, note that $x_1 = x_{k+1} = 1/k$, $x_j = (k-1)/k$ for $j \in \{2, \ldots, k\}$, and $y_i = 1/k$ for $i \in \{1, \ldots, k\}$ is feasible for its LP relaxation.

For $k \geq 4$ this solution is not in conv(S) and hence this Big-M formulation is not sharp (e.g., note that this x does not belong to the projection of the LP relaxation of (6.9) onto the x variables). ☐

Finally, the following proposition shows that the hybrid formulation is always at least as strong as the traditional formulation if equivalent Big-M constants are used.

PROPOSITION 6.9. *If the Big-M constants in* (6.1) *and* (6.3) *are identical and such that* $M^{s,s,i} \leq b^{s,i}$, *then hybrid Big-M formulation* (6.4) *is at least as strong as traditional Big-M formulation* (6.2).

Proof. Using the notation of formulation (6.4) (i.e., accounting for the possible common left-hand side matrices), (6.2a) from the traditional Big-M formulation is equal to

$$A^s x \leq b^{s,i} + \left(\overline{M}^{s,i} - b^{s,i} \right) (1 - y_{s,i}) \quad \forall s \in \{1, \ldots, p\}, i \in \{1, \ldots, k_s\},$$

where $\overline{M}_l^{s,i} = \max_{t \in \{1,\ldots,p\}, j \in \{1,\ldots,k_t\}, (t,j) \neq (s,i)} M_l^{s,t,j}$. Using the fact that in this case (6.2b) is equal to $\sum_{s=1}^{p} \sum_{i=1}^{k_s} y_{s,i} = 1$, we can rewrite this equation as

$$A^s x \leq b^{s,i} y_{s,i} + \sum_{\substack{t \in \{1,\ldots,p\}, \\ j \in \{1,\ldots,k_t\} \\ (t,j) \neq (s,i)}} \overline{M}^{s,i} y_{t,i} \quad \forall s \in \{1, \ldots, p\}, i \in \{1, \ldots, k_s\}.$$

The result then follows from assumption $M^{s,s,i} \leq b^{s,i}$ and because, by definition, $\overline{M}^{s,i} \geq M^{s,t,i}$. ☐

Proposition 6.9 and the formulation sizes suggest that, at least from a theoretical perspective, hybrid Big-M formulation (6.4) is always preferable to traditional Big-M formulation (6.2). This theoretical advantage often results in a computational advantage [156]. However, the complexities of state-of-the-art solvers could still allow the traditional formulation to outperform the hybrid one, particularly when the size and strength differences are small.

7. Large Formulations. The aim of the techniques reviewed in this survey is to construct strong, but small MIP formulations. However, there are many cases in which all known MIP formulations (strong or weak) are large. If these formulations are such that the number of variables is small and the number of constraints is large, but can be separated fast, it is sometimes possible to use them in a branch-and-cut procedure [131, 39, 128]. Similarly, when only the number of variables is large, the formulations can be used in column generation or branch-and-price procedures [15]. These procedures can be very effective and so can extensions such as branch-and-cut-and-price [48] and branch-and-price for extended formulations [142]. For this reason, when both large and small (usually extended) formulations are available it is not always clear which is more convenient. Sometimes it is advantageous to solve the large formulation with one of these specialized procedures [40, 50, 32], and other times it is more convenient to solve the small extended formulation directly [31]. Exploring these alternatives is beyond the scope of this survey, so in this section we restrict our attention to one class of large formulations from which alternative small extended formulations can be easily constructed. Such a construction was introduced by Martin [115] (see also [30]) and can be summarized in the following proposition.

PROPOSITION 7.1. *Let $Q := \{x \in \mathbb{Q}^n : Ax \leq b, Dx \leq d\}$ and suppose there is a separation LP problem for $R := \{x \in \mathbb{Q}^n : Dx \leq d\}$. That is, there exist $F \in \mathbb{Q}^{r \times n}$, $H \in \mathbb{Q}^{r \times m}$, $c \in \mathbb{Q}^m$, and $g \in \mathbb{Q}^r$ for which the LP problem given by*

$$(7.1a) \qquad z(\overline{x}) := \max \quad \sum_{i=1}^{n} \overline{x}_i \pi_i + \sum_{j=1}^{m} c_j \rho_j,$$

$$(7.1b) \qquad\qquad\qquad F\pi + H\rho \leq g,$$

$$(7.1c) \qquad\qquad\qquad \rho_j \geq 0 \quad \forall j \in \{1, \ldots, m\},$$

is such that $\overline{x} \in R$ if and only if $z(\overline{x}) \leq 0$. Then an extended LP formulation of Q is given by

$$(7.2a) \qquad\qquad Ax \leq b,$$

$$(7.2b) \qquad\qquad F^{\mathsf{T}}w = x,$$

$$(7.2c) \qquad\qquad H^{\mathsf{T}}w \geq c,$$

$$(7.2d) \qquad\qquad g^{\mathsf{T}}w \leq 0,$$

$$(7.2e) \qquad\qquad w_j \geq 0 \quad \forall j \in \{1, \ldots, r\}.$$

If Q is the LP relaxation of an MIP formulation such that R has a large number of constraints but a small separation LP problem, we can replace $Dx \leq d$ with (7.2b)–(7.2e) to obtain a smaller extended formulation. Proposition 7.1 was used by Lee and Wilson in [106, 165] to obtain an extended formulation that strengthens (6.5) and (6.6). To describe this approach we need the following proposition from [106, 165]. For simplicity, we concentrate on formulation (6.5) for $n_2 = 0$, which we restate in the proposition for convenience.

PROPOSITION 7.2. *Let $\{P^i\}_{i=1}^k \subseteq \mathbb{Q}^n$ be a finite family of polyhedra with a common recession cone C, $\mathrm{ext}\left(P^i\right)$ be the finite sets of extreme points of polyhedron P^i, $\mathrm{ray}(C)$ be the finite set of extreme rays of polyhedral cone C, and $V := \bigcup_{i=1}^k \mathrm{ext}\left(P^i\right)$. Then $S = \bigcup_{i=1}^k P^i$ can be modeled with formulation (6.5) given by*

$$\sum_{v \in V} v\lambda_v + \sum_{r \in \mathrm{ray}(C)} r\mu_r = x,$$

$$\sum_{v \in V} \lambda_v = 1,$$

$$\lambda_v \leq \sum_{i:v \in \mathrm{ext}(P^i)} y_i \quad \forall v \in V,$$

$$\sum_{i=1}^k y_i = 1,$$

$$\lambda_v \geq 0 \qquad \forall v \in V,$$

$$\mu_r \geq 0 \qquad \forall r \in \mathrm{ray}(C),$$

$$y \in \{0,1\}^k.$$

While (6.5) is sharp, it is not always locally ideal. However, all facet-defining inequalities of the convex hull of solutions to (6.5) are of the form $\lambda_v \geq 0$, $\mu_r \geq 0$, or

$$(7.3) \qquad \sum_{i \in I} y_i \leq \sum_{v \in U} \lambda_v$$

for some $I \subseteq \{1,\ldots,k\}$ and $U \subseteq V$. Furthermore, for a given $(\overline{y}, \overline{\lambda})$ a separation LP problem for (7.3) is given by

$$(7.4\mathrm{a}) \qquad \max \quad \sum_{i=1}^k \overline{y}_i \alpha_i + \sum_{v \in V} \overline{\lambda}_v \beta_v,$$

$$(7.4\mathrm{b}) \qquad\qquad\qquad \alpha_i - \beta_v \leq 0 \quad \forall i \in \{1,\ldots,k\}, v \in P^i,$$

$$(7.4\mathrm{c}) \qquad\qquad\qquad \alpha_i \geq 0 \quad \forall i \in \{1,\ldots,k\},$$

$$(7.4\mathrm{d}) \qquad\qquad\qquad \beta_v \geq 0 \quad \forall v \in V.$$

The following example shows how Propositions 7.1 and 7.2 can be used to strengthen (6.5) to a locally ideal formulation.

EXAMPLE 15. *Lee and Wilson [106, 165] gave a precise characterization of sets I and U for which (7.3) are facet defining. By adding these inequalities to (6.5) or (6.6) we would immediately obtain a locally ideal formulation. Unfortunately, Lee and Wilson also gave examples in which the number of facets defined by inequalities (7.3) can grow exponentially in $|V|$. However, using Proposition 7.1 and the fact that constraints (7.3) can be separated with LP problem (7.4), Lee and Wilson adapted (6.5) to the locally ideal formulation given by*

$$(7.5\mathrm{a}) \qquad \sum_{v \in V} v\lambda_v + \sum_{r \in \mathrm{ray}(C)} r\mu_r = x,$$

$$(7.5\mathrm{b}) \qquad\qquad \sum_{v \in V} \lambda_v = 1,$$

(7.5c) $$\sum_{v \in P^i} w_v^i \geq y_i \qquad \forall i \in \{1, \ldots, k\},$$

(7.5d) $$\sum_{i:v \in P^i} w_v^i \leq \lambda_v \qquad \forall v \in V,$$

(7.5e) $$\sum_{i=1}^{k} y_i = 1,$$

(7.5f) $$\lambda_v \geq 0 \qquad \forall v \in V,$$

(7.5g) $$\mu_r \geq 0 \qquad \forall r \in \mathrm{ray}(C),$$

(7.5h) $$w_v^i \geq 0 \qquad \forall i \in \{1, \ldots, k\}, \forall v \in V,$$

(7.5i) $$y \in \{0,1\}^k.$$

It is interesting to note that by eliminating variables λ_v and renaming w_v^i to λ_v^i we obtain formulation (5.3). *In other words, as expected, if we try to strengthen* (6.5) *or* (6.6) *to recover what we lost by eliminating the copies of λ variables of* (5.3) *through Corollary* 6.3, *we recover* (5.3).

Similar formulations for the case in which the separation problem can be solved using dynamic programming and other extensions can be found in [96, 116]. However, note that having a generic polynomial time algorithm for the separation is not sufficient to obtain a polynomial-sized extended formulation. For instance, if S corresponds to the incidence vectors of perfect matchings of a complete graph as defined in (4.8), then all inequalities of $\mathrm{conv}(S)$ can be separated in polynomial time [146]. However, Theorem 4.3 shows that S does not have a polynomial-sized sharp extended formulation. Finally, we note that another way to obtain small formulations from large ones in a systematic way is to define approximate versions that are slightly weaker, but significantly smaller [161].

8. Incremental Formulations. All formulations for disjunctive constraints we have considered so far include binary variables y such that $\sum_{i=1}^{k} y_i = 1$. The problem with such a configuration is that it can be quite detrimental to branch-and-bound algorithms. To see this, imagine that the optimal solution to the LP relaxation at a node of the branch-and-bound tree (e.g., the root LP relaxation) is such that $\overline{y}_{i_0} \notin \mathbb{Z}$ for some $i_0 \in \{1, \ldots, k\}$. As described in section 2.2, a branch-and-bound-based MIP solver will branch on y_{i_0} by creating two new LP problems by adding $y_{i_0} \leq \lfloor \overline{y}_{i_0} \rfloor$ to the LP relaxation in the first branch and $y_{i_0} \geq \lceil \overline{y}_{i_0} \rceil$ in the second branch. Because the LP relaxation includes constraints $0 \leq y_{i_0} \leq 1$, we have that this branching is equivalent to fixing $y_{i_0} = 0$ in the first branch and $y_{i_0} = 1$ in the second one. These branches are usually denoted *down-branching* and *up-branching*, respectively, and have very different effects on the branch-and-bound tree. The difference is that up-branching (i.e., fixing $y_{i_0} = 1$) automatically forces all other y_i variables to zero, while down-branching (i.e., fixing $y_{i_0} = 0$) does not imply any restriction on other y_i variables. This asymmetry results in what are usually denoted *unbalanced branch-and-bound trees*, which can result in a significant slow-down of the algorithm [141].

One way to resolve this issue is to use specialized constraint branching schemes [141]. In our particular case, an appropriate scheme is the *SOS*1 *branching* of Beale and Tomlin [17] that can be described as follows. Because $\sum_{i=1}^{k} y_i = 1$ and $y \geq 0$, we have that for a fractional $\overline{y}$ there must be $i_0 < i_1 \in \{1, \ldots, k\}$ such that $\overline{y}_{i_0}, \overline{y}_{i_1} \in (0, 1)$. We can then exclude this fractional solution by creating two new LP problems by adding $y_1 = y_2 = \cdots = y_{i_0} = 0$ in the first branch and $y_{i_0+1} = y_{i_0+2} = \cdots = y_k = 0$

in the second one. In contrast to the traditional up-down variable branching, this approach usually fixes to zero around half the integer variables in each branch and hence yields much more balanced branch-and-bound trees.

An alternative to constraint branching is to use a well-known MIP formulation trick that uses a redefinition of variables y (e.g., [27, 111, 132, 149, 33, 129, 110]). This transformation can be formalized in the following straightforward proposition.

PROPOSITION 8.1. *Let* $\Delta_k := \left\{ y \in \mathbb{Q}_+^k \ : \ \sum_{i=1}^k y_i = 1 \right\}$ *and* $L : \mathbb{Q}^k \to \mathbb{Q}^k$ *be the linear function defined as*

$$L(y)_i := \sum_{j=i}^k y_i.$$

Then $L(\Delta_k) = \Gamma_k := \{ w \in \mathbb{Q}^k \ : \ 1 = w_1 \geq w_2 \geq \cdots \geq w_k \geq 0 \}$ *and the inverse of* L *is*

$$L^{-1}(w)_i := \begin{cases} w_i - w_{i+1}, & i < k, \\ w_k, & \text{otherwise.} \end{cases}$$

Hence L *is a one-to-one correspondence between* Δ_k *and* Γ_k, *and, in particular,* L *is a one-to-one correspondence between* $\Delta_k \cap \{0,1\}^k$ *and* $\Gamma_k \cap \{0,1\}^k$.

Then, in any formulation that includes the constraint $y \in \Delta_k \cap \{0,1\}^k$ we can replace such a constraint by $w \in \Gamma_k \cap \{0,1\}^k$ and every occurrence of y_i by $L^{-1}(w)_i$. This is illustrated in the following example from [168], which shows that variable branching on the w variables can be much more effective than variable branching on the y variables.

EXAMPLE 16. *Let* k *be even,* $a \in \mathbb{Q}^k \setminus \{\mathbf{0}\}$ *be such that*

$$a_1 < \cdots < a_{\frac{k}{2}} = -1 < 0 < a_{\frac{k}{2}+1} = 1 < \cdots < a_k,$$

and for $S := \{a_1, a_2, \ldots, a_k\}$ *consider the problem* $\min_x \{|x| \ : \ x \in S\}$, *which has an optimal value of* 1. *This problem can be solved through the standard MIP formulation given by*

$$\begin{align}
& \text{(8.1a)} && \min \quad z, \\
& \text{(8.1b)} && x \leq z, \\
& \text{(8.1c)} && -x \leq z, \\
& \text{(8.1d)} && x = \sum_{i=1}^k a_i y_i, \\
& \text{(8.1e)} && 1 = \sum_{i=1}^k y_i, \\
& \text{(8.1f)} && y \in \{0,1\}^k.
\end{align}$$

However, using Proposition 8.1 *we can construct the alternative MIP formulation of* S *given by*

$$\begin{align}
& \text{(8.2a)} && \min \quad z, \\
& \text{(8.2b)} && x \leq z,
\end{align}$$

(8.2c) $\qquad -x \leq z,$

(8.2d) $\qquad x = a_1 w_1 + \displaystyle\sum_{i=2}^{k}(a_i - a_{i-1})w_i,$

(8.2e) $\qquad w_i \geq w_{i+1} \qquad\qquad\qquad \forall i \in \{1,\ldots,k-1\},$

(8.2f) $\qquad w_1 = 1,$

(8.2g) $\qquad w \in \{0,1\}^k.$

We can show that a pure branch-and-bound algorithm requires branching in at least $k/2$ of the variables of (8.1) to solve it (see [168] for more details). However, a pure branch-and-bound algorithm can solve (8.2) branching in a single variable as follows. The optimal solution to the LP relaxation of (8.2) has $z = x = 0$, $w_i = 1$ for all $i \in \{1,\ldots,\frac{k}{2}\}$, $w_{\frac{k}{2}+1} = 1/2$, and $w_i = 0$ for all $i \in \{\frac{k}{2}+2,\ldots,k\}$. Adding $w_{\frac{k}{2}+1} = 0$ to this LP relaxation results in an optimal solution with $z = 1$ and $x = -1$, while adding $w_{\frac{k}{2}+1} = 1$ results in an optimal solution with $z = 1$ and $x = 1$. Hence branching on $w_{\frac{k}{2}+1}$ is enough to solve the problem. Finally, it is interesting to note that variable branching on w has essentially the same effect as SOS1 branching on y. For instance, branching on $w_{\frac{k}{2}+1} = 1$ and $w_{\frac{k}{2}+1} = 0$ has the same effect as branching on $y_1 = y_2 = \cdots = y_{\frac{k}{2}} = 0$ and $y_{\frac{k}{2}+1} = y_{\frac{k}{2}+2} = \cdots = y_k = 0$.

The reason for the effectiveness of branching on w comes from the fact that both down- and up-branching on w_i fix the value of many other w variables (around half the variables depending on i). This behavior is usually denoted *double contracting* [66, 67, 150] to contrast it with the behavior of branching on y (only up-branching fixes other variables), which is denoted *single contracting*. The fact that double contracting incremental formulations result in solves with fewer branch-and-bound nodes was confirmed computationally in [155, 157].

The transformation from Proposition 8.1 can be used on any formulation that includes $y \in \{0,1\}^k$. However, in some cases, an additional transformation of the continuous variables can result in an ad-hoc incremental formulation. An example of this is the following formulation introduced in [168] that generalizes a formulation for piecewise linear functions introduced in [165].

PROPOSITION 8.2. *Let* $\left\{P^i\right\}_{i=1}^{k} \subseteq \mathbb{Q}^n$ *be a finite family of polyhedra with a common recession cone C and let* $\left\{v_j^i\right\}_{j=1}^{r_i} := \mathrm{ext}\left(P^i\right)$ *and* $\mathrm{ray}(C)$ *be the finite sets of extreme points and rays of polyhedron P^i and polyhedral cone C, respectively. Then a locally ideal MIP formulation of $S = \bigcup_{i=1}^{k} P^i$ is given by*

$$v_1^1 + \sum_{i=2}^{k}\left(v_1^i - v_{r_i}^{i-1}\right)w_i$$

$$+\sum_{i=1}^{k}\sum_{j=2}^{r_i}\delta_j^i\left(v_j^i - v_1^i\right)$$

(8.3a) $\qquad + \displaystyle\sum_{r \in \mathrm{ray}(C)} r\mu_r = x,$

(8.3b) $\qquad \displaystyle\sum_{j=2}^{r_1}\delta_j^1 \leq 1,$

$$\text{(8.3c)} \qquad \sum_{j=2}^{r_i} \delta_j^i \leq w_i \qquad \forall i \in \{2, \ldots, k\},$$

$$\text{(8.3d)} \qquad \delta_{r_i}^i \geq w_{i+1} \qquad \forall i \in \{1, \ldots, k-1\},$$

$$\text{(8.3e)} \qquad \delta_j^i \geq 0 \qquad \forall i \in \{1, \ldots, k\}, j \in \{2, \ldots, r_i\},$$

$$\text{(8.3f)} \qquad \mu_r \geq 0 \qquad \forall r \in \mathrm{ray}(C),$$

$$\text{(8.3g)} \qquad w_i \in \{0, 1\} \qquad \forall i \in \{2, \ldots, k\}.$$

When $S = \mathrm{gr}(f)$ for a piecewise linear function f defined by (3.3) for polytopes $\{Q^i\}_{i=1}^k$ that satisfy a special ordering condition, formulation (8.3) reduces to a standard MIP formulation for piecewise linear functions that is denoted the *incremental model* in [155]. For more details, examples of incremental formulations, and their advantages, we refer the reader to [168, 20].

9. Logarithmic Formulations. Standard MIP formulations for the union of k polyhedra use k binary variables, while incremental formulations essentially reduce this to $k-1$ variables (one of the k variables is fixed to 1). In this section we review techniques that allow us to reduce the number of binary variables to $\lceil \log_2(k) \rceil$.

The simplest technique for using a logarithmic number of binary variables considers $S = \bigcup_{i=1}^k P^i \subseteq \mathbb{Q}$ where $P^i = \{i - 1\}$. In this case, basic $\mathcal{V}$-formulation (5.3) reduces to

$$\text{(9.1a)} \qquad x = \sum_{i=1}^k (i-1)y_i,$$

$$\text{(9.1b)} \qquad 1 = \sum_{i=1}^k y_i,$$

$$\text{(9.1c)} \qquad y \in \{0, 1\}^k.$$

We can think of formulation (9.1) as a unary encoding of $x \in \{0, 1, \ldots, k-1\}$. If we instead use a binary encoding, we can obtain the formulation given by

$$\text{(9.2a)} \qquad x = \sum_{i=1}^{\lceil \log_2(k) \rceil} 2^{i-1} w_i,$$

$$\text{(9.2b)} \qquad x \leq k - 1,$$

$$\text{(9.2c)} \qquad w \in \{0, 1\}^{\lceil \log_2(k) \rceil}.$$

This technique appeared in the mathematical programming literature as early as [162] and, while it is not an effective way to deal with general integer variables in MIP [134], it has been used as the basis for effective MIP formulations of mixed integer nonlinear programming problems (e.g., [162, 58, 72]). A similar technique is used in [105, 103, 41] to model the constraint programming *all-different* requirement [75] in problems such as graph coloring. For more general problems, MIP formulations with a logarithmic number of binary variables were originally considered in [81, 148] and have received significant attention recently [155, 127, 156, 107, 74, 6, 125, 159, 160].

One way to construct MIP formulations with a logarithmic number of variables is to use the following proposition.

PROPOSITION 9.1. *Let $\{h^i\}_{i=1}^k \subseteq \{0, 1\}^{\lceil \log_2(k) \rceil}$ be such that $h^i \neq h^j$ for any $i \neq j$. Then a locally ideal formulation of $S := \{y \in \{0, 1\}^k : \sum_{i=1}^k y_i = 1\}$ is given*

by

$$\text{(9.3a)} \qquad \sum_{i=1}^{k} y_i = 1,$$

$$\text{(9.3b)} \qquad \sum_{i=1}^{k} h^i y_i = w,$$

$$\text{(9.3c)} \qquad w \in \{0,1\}^{\lceil \log_2(k) \rceil},$$

$$\text{(9.3d)} \qquad y_i \geq 0 \qquad \forall i \in \{1,\ldots,k\}.$$

Recently, formulations that are essentially identical to this one were independently proposed in [6, 74, 107, 159, 160]. However, the basic idea behind formulation (9.3) has in fact been part of the mathematical programming folklore for a long time. For instance, Glover [58] attributes (9.3) for a specific choice of $\{h^i\}_{i=1}^{k}$ to Sommer [148].

The following proposition from [74, 6] shows how to use (9.3) to construct formulations for simple disjunctive constraints.

PROPOSITION 9.2. *Let* $\{h^i\}_{i=1}^{k} \subseteq \{0,1\}^{\lceil \log_2(k) \rceil}$ *be such that* $h^i \neq h^j$ *for any* $i \neq j$. *Then a locally ideal formulation for* $S = \{a_1, a_2, \ldots, a_k\}$ *is given by*

$$\text{(9.4a)} \qquad \sum_{i=1}^{k} y_i = 1,$$

$$\text{(9.4b)} \qquad \sum_{i=1}^{k} h^i y_i = w,$$

$$\text{(9.4c)} \qquad \sum_{i=1}^{k} a_i y_i = x,$$

$$\text{(9.4d)} \qquad w \in \{0,1\}^{\lceil \log_2(k) \rceil},$$

$$\text{(9.4e)} \qquad y_i \geq 0 \qquad \forall i \in \{1,\ldots,k\}.$$

If $\log_2(k) \in \mathbb{Z}$ *and there exists* $(c_0, c) \in \mathbb{Q} \times \mathbb{Q}^{\log_2(k)}$ *such that*

$$\text{(9.5)} \qquad a_i = c_0 + \sum_{j=1}^{\log_2(k)} h_j^i c_j \quad \forall i \in \{1,\ldots,k\},$$

then formulation (9.4) *can be reduced to the locally ideal formulation given by*

$$\text{(9.6a)} \qquad c_0 + \sum_{j=1}^{\log_2(k)} c_j w_j = x,$$

$$\text{(9.6b)} \qquad w \in \{0,1\}^{\log_2(k)}.$$

An advantage of (9.6) over (9.4) is the elimination of y variables. However, unlike (9.4), (9.6) requires $\log_2(k) \in \mathbb{Z}$. For the case $\log_2(k) \notin \mathbb{Z}$, formulation (9.6) remains valid only if we add the constraint $w \in \{h^i\}_{i=1}^{k}$. If h^i corresponds to the binary expansion of $i-1$, this constraint can be enforced through $\sum_{i=1}^{\lceil \log_2(k) \rceil} 2^{i-1} w_i \leq k - 1$. The resulting formulation is a generalization of (9.2) and, unfortunately, it can fail

to be locally ideal. Stronger methods to implement $w \in \{h^i\}_{i=1}^k$ can be found in [9, 41, 130].

One way to use Proposition 9.1 to construct MIP formulations for more general disjunctive constraints is to combine it with formulation (5.3). The resulting formulation was introduced in [155] for piecewise linear functions and its extension to general polyhedra was introduced in [168].

PROPOSITION 9.3.   *Let $\{P^i\}_{i=1}^k \subseteq \mathbb{Q}^n$ be a finite family of polyhedra with a common recession cone C and let $\mathrm{ext}\left(P^i\right)$ and $\mathrm{ray}(C)$ be the finite sets of extreme points and rays of polyhedron P^i and polyhedral cone C, respectively. Furthermore, let $\{h^i\}_{i=1}^k \subseteq \{0,1\}^{\lceil \log_2(k) \rceil}$ be such that $h^i \neq h^j$ for any $i \neq j$. Then a locally ideal MIP formulation of $S = \bigcup_{i=1}^k P^i$ is given by*

(9.7a)
$$\sum_{i=1}^k \sum_{v \in \mathrm{ext}(P^i)} v\lambda_v^i + \sum_{r \in \mathrm{ray}(C)} r\mu_r = x,$$

(9.7b)
$$\sum_{i=1}^k \sum_{v \in \mathrm{ext}(P^i)} \lambda_v^i = 1,$$

(9.7c)
$$\sum_{i=1}^k \sum_{v \in \mathrm{ext}(P^i)} h^i \lambda_v^i = w,$$

(9.7d) $\lambda_v^i \geq 0$ $\forall i \in \{1,\dots,k\}, v \in \mathrm{ext}\left(P^i\right),$

(9.7e) $\mu_r \geq 0$ $\forall r \in \mathrm{ray}(C),$

(9.7f) $w \in \{0,1\}^{\lceil \log_2(k) \rceil}.$

The following proposition shows an entirely different technique that was introduced in [81].

PROPOSITION 9.4.   *Let $\{P^i\}_{i=1}^k$ be a finite family of polyhedra with a common recession cone (i.e., $P_\infty^i = P_\infty^j$ for all i,j), such that $P^i = \{x \in \mathbb{Q}^n : A^i x \leq b^i\}$ with $A^i \in \mathbb{Q}^{m_i \times n}$ and $b^i \in \mathbb{Q}^{m_i}$ for all $i \in \{1,\dots,k\}$. Also, let $\{h^i\}_{i=1}^k \subseteq \{0,1\}^{\lceil \log_2(k) \rceil}$ be such that $h^i \neq h^j$ for any $i \neq j$. An MIP formulation for $S = \bigcup_{i=1}^k P^i$ is given by*

(9.8a) $A^i x \leq b^i + M^i \left(\sum_{j=1}^{\lceil \log_2(k) \rceil} h_j^i - \sum_{j:h_j^i=1} w_j + \sum_{j:h_j^i=0} w_j \right)$ $\forall i \in \{1,\dots,k\},$

(9.8b) $w \in \{0,1\}^{\lceil \log_2(k) \rceil},$

for sufficiently large $M^i \in \mathbb{Q}^{m_i}$.

Formulation (9.8) does not require any auxiliary variables besides w. However, because it is an adaptation of Big-M formulation (6.2), it is not necessarily locally ideal or sharp.

Other logarithmic formulations for specific constraints were introduced in [127, 125, 159, 160]. Further comparisons between logarithmic and incremental formulations can be found in [168].

10. Combining Formulations and Propositional Logic. Another way to keep the sizes of MIP formulations controlled is to construct them for different parts of

a mathematical programming problem independently and then combine them. This combination of formulations, which was denoted *model linkage* by Jeroslow and Lowe [91], can reduce the strength of the final formulation, but can also result in a significant reduction in size.

The simplest way to combine formulations is to intersect them. For instance, if $S = \bigcap_{i=1}^{r} S^i$, we can obtain an MIP formulation of S by intersecting MIP formulations for S^i. However, alternative formulations can be obtained by considering the specific structure of parts S^i. One example of this is

$$(10.1) \qquad S = \bigcap_{j=1}^{r} \bigcup_{i=1}^{k_j} P^{i,j},$$

where $P^{i,j}$ are given polytopes. In this case, (10.1) can also be written using the single disjunctive constraint,

$$(10.2) \qquad S = \bigcup_{s \in \prod_{j=1}^{r} \{1,\ldots,k_j\}} \bigcap_{j=1}^{r} P^{s_j,j}.$$

Formulating this single combined constraint directly results in stronger, but larger MIP formulations than when combining formulations for each constraint. However, using the rules of propositional calculus we can construct other variations between extreme cases (10.1) and (10.2) [11, 86]. Systematically constructing such variations can lead to formulations that effectively balance size and strength (e.g., [139, 143, 144]). Furthermore, in some cases formulations based on (10.1) do not incur any loss of strength [159, 160].

The reverse transformation from (10.2) to (10.1) is not always evident and sometimes requires the addition of auxiliary variables. Examples of this include the formulations for joint probabilistic constraints introduced in [101, 114, 157]. In this case, auxiliary binary variables that keep track of violated constraints are used to combine formulations for single row or k-row probabilistic constraints into a formulation for the full row joint probabilistic constraint.

For more information on propositional calculus, logic, and other techniques from constraint programming with MIP formulations, we refer the reader to [78, 88, 164, 76, 79, 77].

11. MIP Representability. A systematic study of which sets can be modeled as MIP problems began with the work of Meyer [118, 119, 120, 122, 121] and Ibaraki [81], and the first precise characterization of these sets was given by Jeroslow and Lowe [90, 113]. For general MIP formulations with unbounded integer variables this characterization is quite technical, so Jeroslow and Lowe [90, 113] also introduced a much simpler characterization for MIP formulations with binary or bounded integer variables. This characterization was later extended by Hooker [76] to consider a restricted use of unbounded integer variables that nevertheless allows for the modeling of most sets that are used in practice, while preserving most of the simplicity of Jeroslow and Lowe's characterization. The first simple characterization to correspond precisely to the sets modeled by basic formulations (4.6) was denoted *bounded* MIP representability by Jeroslow and Lowe.

DEFINITION 11.1. *$S \subseteq \mathbb{Q}^n$ is bounded MIP representable if and only if it has an MIP formulation of the form (2.6), where y is additionally constrained to be a binary vector.*

Jeroslow and Lowe showed that bounded MIP representable sets are precisely the union of a finite number of polyhedra with a common recession cone.

PROPOSITION 11.2. *A set $S \subseteq \mathbb{Q}^n$ is bounded MIP representable if and only if there exists a finite family of polyhedra $\{P^i\}_{i=1}^k$ such that $P_\infty^i = P_\infty^j$ for all i, j and*

$$(11.1) \qquad\qquad S = \bigcup_{i=1}^k P^i,$$

or, equivalently, if and only if

$$(11.2) \qquad\qquad S = C + \bigcup_{i=1}^k P^i,$$

where $\{P^i\}_{i=1}^k$ is a finite family of polytopes (i.e., $P_\infty^i = \{0\}$ for all i) and C is a polyhedral cone.

Proof. The fact that a set of the form (11.1) is bounded MIP representable follows from Proposition 4.2.

For the converse, let (2.6) be an MIP formulation of set S such that y is bounded. Then the feasible region of this MIP formulation is a finite union of polyhedra (one for each possible value of y) with the same recession cone. The result then follows by projecting this union onto the x variables. □

Of course, through unary or binary expansion, bounded MIP representability is equivalent to asking for y to be bounded in (2.6). However, we can extend the class of MIP representable sets by allowing unbounded integer variables. If we restrict these unbounded integer variables to be part of the original variables x, we obtain the second simple characterization introduced by Hooker. Because this representation does not allow unbounded integer auxiliary variables, we denote it "projected MIP representability."

DEFINITION 11.3. $S \subseteq \mathbb{Q}^n$ *is* projected MIP representable *if and only if it has an MIP formulation of the form*

$$(11.3a) \qquad\qquad Ax + B\lambda + Dy \leq b,$$
$$(11.3b) \qquad\qquad x \in \mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2},$$
$$(11.3c) \qquad\qquad \lambda \in \mathbb{Q}^s,$$
$$(11.3d) \qquad\qquad y \in \{0, 1\}^t.$$

Note that (11.3) is a special version of (2.6) in which the linear inequalities are such that every integer auxiliary variable is either bounded or identical to one of the original variables. Projected MIP representability certainly subsumes bounded MIP representability, but as shown by Hooker [76] it can model a much wider class of sets.

PROPOSITION 11.4. *A set $S \subseteq \mathbb{R}^n$ is projected MIP representable if and only if there exists a finite family of polyhedra $\{P^i\}_{i=1}^k$ such that $P_\infty^i = P_\infty^j$ for all i, j and*

$$(11.4) \qquad\qquad S = \bigcup_{i=1}^k P^i \cap (\mathbb{Q}^{n_1} \times \mathbb{Z}^{n_2}).$$

Proof. The fact that a set of the form (11.4) is projected MIP representable follows from Proposition 5.1.

The converse is analogous to Proposition 11.2. □

Projected MIP representable sets are then precisely those sets that can be formulated by (5.2) and (5.3). While projected MIP representability covers most sets that are used in practice, it does not characterize all sets that can be modeled with MIP formulations. We now consider the most general version of the MIP representation theorem introduced by Jeroslow and Lowe [90, 113], for which we need the following definition.

DEFINITION 11.5. *A finitely generated integral monoid is a set $M \subseteq \mathbb{Z}^n$ such that there exists $\{r^i\}_{i=1}^{d} \subseteq \mathbb{Z}^n$ for which $M = \left\{ \sum_{i=1}^{d} \mu_i r^i \; : \; \mu \in \mathbb{Z}_+^d \right\}$. We say that M is generated by $\{r^i\}_{i=1}^{d}$.*

General MIP representability is obtained by simply replacing cone C by a finitely generated integral monoid M in bounded MIP characterization (11.2).

THEOREM 11.6. *A set $S \subseteq \mathbb{Q}^n$ is MIP representable if and only if*

$$(11.5) \qquad\qquad S = M + \bigcup_{i=1}^{k} P^i$$

for a finite family of polytopes $\left\{P^i\right\}_{i=1}^{k}$ and a finitely generated monoid M.

Proof. For necessity assume S has a MIP formulation of the form (2.6) and let $Q \subseteq \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Q}^t$ be the polyhedron described by (2.6a). By Theorem 3.6 and because Q is a rational polyhedron, there exist points $\left\{ \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) \right\}_{j=1}^{p} \subseteq \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Q}^t$ and scaled rays $\left\{ \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right) \right\}_{l=1}^{d} \subseteq \mathbb{Z}^n \times \mathbb{Z}^s \times \mathbb{Z}^t$ such that for any (x, u, y) feasible for (2.6) there exist $\lambda \in \Delta^p$ and $\mu \in \mathbb{Q}_+^d$ such that

$$(x, u, y) = \sum_{j=1}^{p} \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) + \sum_{l=1}^{d} \mu_l \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right).$$

Now, for such λ and μ define

$$(x^*, u^*, y^*) := \sum_{j=1}^{p} \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) + \sum_{l=1}^{d} \left(\mu_l - \lfloor \mu_l \rfloor\right) \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right)$$

and

$$(11.6) \qquad\qquad (x^\infty, u^\infty, y^\infty) := \sum_{l=1}^{d} \lfloor \mu_l \rfloor \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right).$$

Then, $(x^\infty, u^\infty, y^\infty) \in \mathbb{Z}^n \times \mathbb{Z}^s \times \mathbb{Z}^t$ and (x^*, u^*, y^*) belongs to bounded set

$$S_0 := \left\{ \sum_{j=1}^{p} \lambda_j \left(\hat{x}^j, \hat{u}^j, \hat{y}^j\right) + \sum_{l=1}^{d} \mu_l \left(\tilde{x}^l, \tilde{u}^l, \tilde{y}^l\right) \; : \; \lambda \in \Delta^p, \, \mu \in [0,1]^d \right\}.$$

Also, because $y = y^* + y^\infty$ and $y \in \mathbb{Z}^t$ we have $y^* \in \mathbb{Z}^t$. Hence $(x^*, u^*, y^*) \in S_0 \cap \mathbb{Q}^n \times \mathbb{Q}^s \times \mathbb{Z}^t$, which is a finite union of polytopes (one for each possible value of y^* in S_0). Let $\left\{P^i\right\}_{i=1}^{k}$ be the projection of such polytopes onto the x variables, so that $x^* \in \bigcup_{i=1}^{k} P^i$. The result then follows from $x = x^* + x^\infty$ and the fact that (11.6) implies x^∞ belongs to the integral monoid generated by $\left\{\tilde{x}^l\right\}_{l=1}^{d}$.

For sufficiency, let S be of the form (11.5), let (4.6) be an MIP formulation for $\bigcup_{i=1}^{k} P^i$, and let M be generated by $\{r^i\}_{i=1}^{d} \subseteq \mathbb{Z}^n$. Then an MIP formulation for S is given by

$$\text{(11.7a)} \qquad\qquad A^i x^i \le b^i y_i \qquad \forall i \in \{1, \ldots, k\},$$

$$\text{(11.7b)} \qquad\qquad \sum_{i=1}^{k} x^i + \sum_{l=1}^{d} z_l r^l = x,$$

$$\text{(11.7c)} \qquad\qquad \sum_{i=1}^{k} y_i = 1,$$

$$\text{(11.7d)} \qquad\qquad x^i \in \mathbb{Q}^n \qquad \forall i \in \{1, \ldots, k\},$$

$$\text{(11.7e)} \qquad\qquad y \in \{0,1\}^k,$$

$$\text{(11.7f)} \qquad\qquad z \in \mathbb{Z}_+^d. \qquad \square$$

By comparing (11.2) and (11.5), we see that general MIP representability replaces the *continuous recession directions* of polyhedral cone C with the *discrete recession directions* of monoid M. The fact that this replacement gives further modeling power stems from the following lemma, which shows that continuous recession directions can be obtained from discrete recession directions and a polytope through Minkowski addition of sets.

LEMMA 11.7. *Let $C = \left\{ \sum_{i=1}^{d} \mu_i r^i \; : \; \mu \in \mathbb{Q}_+^d \right\}$ be a polyhedral cone generated by $\{r^i\}_{i=1}^{d} \subseteq \mathbb{Z}^n$. Then $C = M + P$, where $P := \left\{ \sum_{i=1}^{d} \lambda_i r^i \; : \; \lambda \in [0,1]^d \right\}$ is a polytope and $M := \left\{ \sum_{i=1}^{d} \mu_i r^i \; : \; \mu \in \mathbb{Z}_+^d \right\}$ is a finitely generated integral monoid.*

Proof. The result follows from noting that

$$\sum_{i=1}^{d} \mu_i r^i = \sum_{i=1}^{d} \lfloor \mu_i \rfloor \, r^i + \sum_{i=1}^{d} (\mu_i - \lfloor \mu_i \rfloor) \, r^i. \qquad \square$$

Using this lemma we have that bounded MIP characterization (11.2) is equivalent to

$$\text{(11.8)} \qquad\qquad S = M + Q^0 + \bigcup_{i=1}^{k} Q^i$$

for a finite family of polytopes $\{Q^i\}_{i=0}^{k}$ and a finitely generated integral monoid M. We then obtain general MIP characterization (11.5) by noting that $Q^0 + \bigcup_{i=1}^{k} Q^i = \bigcup_{i=1}^{k} \left(Q_0 + Q^i \right)$ and that $P^i = Q_0 + Q^i$ is a polytope for every i. We could carry out the reverse transformation if we could *factor* out a common polyhedron Q^0 from polytopes P^i from bounded MIP characterization (11.2). However, as the following example from [90, 113] shows, this factorization cannot always be done and general characterization (11.5) includes more cases than bounded characterization (11.1)–(11.2) and than projected MIP characterization (11.4).

EXAMPLE 17.   *Let $S = (\{0\} \cup [2, \infty)) \cap \mathbb{Z}$. S is a finitely generated integral monoid and a general MIP formulation for S is given by*

$$x - 2y_1 - 3y_2 = 0,$$
$$y \in \mathbb{Z}_+^2.$$

However, S does not satisfy bounded characterization (11.1)–(11.2) *or projected MIP characterization* (11.4).

Bounded, projected, and general MIP representability can sometimes be hard to recognize (e.g., see Example 20). One way to check the potential for MIP representability is through the following necessary conditions proven in [90, 113].

PROPOSITION 11.8. *If $S \subseteq \mathbb{Q}^n$ is MIP representable, then S is closed and $\mathrm{conv}(S)$ is a polyhedron.*

Unfortunately, the following example from [153] shows that these conditions are not sufficient for MIP representability.

EXAMPLE 18. *Let*

$$S = \left\{ x \in \mathbb{Q}^n \,:\, x_n = \max_{i \in \{1,\ldots,n-1\}} x_i, \quad x_j \geq 0 \quad \forall j \in \{1,\ldots,n\} \right\}$$
$$= \bigcup_{i=1}^{n-1} \left\{ x \in \mathbb{Q}^n \,:\, x_n = x_i, \quad x_n \geq x_j \geq 0 \quad \forall j \in \{1,\ldots,n\} \right\}.$$

We have that $\mathrm{conv}(S) = \left\{ x \in \mathbb{Q}^n \,:\, x_n \leq \sum_{i=1}^{n-1} x_i, \quad x_n \geq x_j \geq 0 \quad \forall j \in \{1,\ldots,n\} \right\}$ *is a polyhedron, but S is not MIP representable.*

We now illustrate some of these representability concepts by considering MIP formulations of piecewise linear functions. However, we first introduce the following example, which shows that rationality is crucial to the presented representability results. For more information concerning MIP representability for nonpolyhedral or nonrational polyhedral sets, we refer the reader to [49, 73].

EXAMPLE 19. *Consider the MIP formulation with irrational data given by $S = \left\{ x \in \mathbb{R}^2 \,:\, x_2 - \sqrt{2}x_1 \leq 0,\, x_2 \geq 0, x \in \mathbb{Z}^2 \right\}$. We have that S is closed and the closure of the convex hull of S is the nonrational polyhedron given by the LP relaxation of this formulation. However, $\mathrm{conv}(S)$ is not closed. Furthermore, we have that S is an integral monoid (it is composed by integer points and is closed under addition), but by Lemma 3 in [83] we have that S is not finitely generated.*

11.1. MIP Representability of Functions. Functions that have MIP formulations include most piecewise linear functions. The study and use of MIP formulations for piecewise linear functions has been and still is a very prolific area of research with a wide range of applications. While it is beyond the scope of this paper to include a detailed survey of the literature on such formulations, we have included several MIP formulations of piecewise linear functions and we now cover some issues concerning representability of piecewise linear functions. For more details we refer the reader to [155] and the references therein. Some recent work concerning piecewise linear functions that are not referenced in [155] include [151, 51, 57, 124, 62, 123, 126, 127, 43, 156, 153]. Finally, we note that there is also a vast literature on incorporating piecewise linear functions into optimization models without using MIP [17, 54, 55, 56, 38, 99, 158, 47, 169, 46]

Functions with bounded MIP representable graphs and epigraphs include most piecewise linear functions. In particular, for continuous multivariate functions with bounded domain we can give the following precise characterization.

PROPOSITION 11.9. *Let $f : D \subseteq \mathbb{Q}^d \to \mathbb{Q}$ be a continuous function.*
1. *If $\mathrm{gr}(f)$ is bounded MIP representable, then $\mathrm{epi}(f)$ is bounded MIP representable*
2. *$\mathrm{gr}(f)$ is bounded MIP representable if and only if there exist $\{m^i\}_{i=1}^k \subseteq \mathbb{Q}^n$,*

$\{c_i\}_{i=1}^k \subseteq \mathbb{Q}$, and a finite family of polytopes $\{Q^i\}_{i=1}^k$ such that

(11.9a)
$$D = \bigcup_{i=1}^k Q^i,$$

(11.9b)
$$f(x) = \begin{cases} m^1 x + c_1, & x \in Q^1, \\ \quad\vdots \\ m^k x + c_k, & x \in Q^k. \end{cases}$$

Proof. For item 1 we first note that $\mathrm{epi}(f) = C_+ + \mathrm{gr}(f)$, where $C_+ := \{(0, z) \in \mathbb{Q}^n \times \mathbb{Q} : z \geq 0\}$, and that for f with bounded domain $\mathrm{gr}(f)$ is bounded. Now, if $\mathrm{gr}(f)$ is bounded MIP representable, we have by Proposition 11.2 that $\mathrm{gr}(f) = \bigcup_{i=1}^k P^i$ for a finite family of polytopes $\{P^i\}_{i=1}^k$. Then, $\mathrm{epi}(f) = C_+ + \bigcup_{i=1}^k P^i = \bigcup_{i=1}^k \left(C_+ + P^i \right)$ and $\left\{ C_+ + P^i \right\}_{i=1}^k$ is a finite family of polyhedra with common recession cone C_+. Hence, by Proposition 11.2 $\mathrm{epi}(f)$ is bounded MIP representable.

For the "if" part of item 2 let $P^i := \{(x, z) \in \mathbb{Q}^n \times \mathbb{Q} : x \in Q^i, z = m^i x + c_i\}$ for each $i \in \{1, \ldots, k\}$. Then $\{P^i\}_{i=1}^k$ is a finite family of polytopes and $\mathrm{gr}(f) = \bigcup_{i=1}^k P^i$; hence by Proposition 11.2 $\mathrm{gr}(f)$ is bounded MIP representable. For the only if part, if $\mathrm{gr}(f)$ is bounded MIP representable, then, by Proposition 11.2, $\mathrm{gr}(f) = \bigcup_{i=1}^k P^i$ for a finite family of polytopes $\{P^i\}_{i=1}^k$. Let Q^i be the projection onto the x variables of polytope P^i and let $\{v^j\}_{j=1}^p$ be the extreme points of polytope Q^i. We claim that there exists a solution $(m^i, c_i) \in \mathbb{Q}^n \times \mathbb{Q}$ to the system $m^i v^j + c_i = f(v^j)$ for all $j \in \{1, \ldots, p\}$. If such a solution exists we have that $P^i = \{(x, z) \in \mathbb{Q}^n \times \mathbb{Q} : x \in Q^i, m^i x + c_i = z\}$, which gives the desired result. To show the claim first assume without loss of generality that $v^{d+1} = 0$. After that, assume again without loss of generality that $\{v^j\}_{j=1}^d$ are linearly independent where $1 < d = \mathrm{rank}\left(\left[v^j\right]_{j=1}^p\right)$ (if $d = 0$, then $Q^i = \{v^{d+1}\} = \{0\}$ and the claim is straightforward). Let $m^i \in \mathbb{Q}^n$ and $c_i \in \mathbb{Q}$ be such that $c_i = f\left(v^{d+1}\right) = f(0)$ and $m^i v^j = f(v^j) - c_i$ for all $j \in \{1, \ldots, d\}$, which exist by the linear independence assumption. To arrive at a contradiction assume without loss of generality that $m^i v^{d+2} + c_i > f\left(v^{d+1}\right)$. Because of the assumption on $\{v^j\}_{j=1}^d$, there exist $\lambda \in \mathbb{Q}^d$ such that $v^{d+2} = \sum_{i=1}^d \lambda_i v^i$. Furthermore, without loss of generality we may assume that $\lambda_1 \leq \lambda_i$ for all $i \in \{1, \ldots, d\}$. Let

$$\tilde{v}(\delta) := \delta \left(\sum_{i=1}^{d+1} \frac{1}{d+1} v^i \right) + (1 - \delta) v^{d+2} = \sum_{i=1}^d \tilde{\lambda}_i(\delta) v^i,$$

where $\tilde{\lambda}_i(\delta) := \delta \frac{1}{d+1} + (1 - \delta)\lambda_i$ for all $i \in \{1, \ldots, d\}$. If $\lambda_1 \geq 0$, let $\delta_1 := 0$, and if $\lambda_1 < 0$, let

$$\delta_1 := \frac{\lambda_1}{\left(\lambda_1 - \frac{1}{d+1}\right)} \in (0, 1).$$

If $\sum_{i=1}^d \lambda_i \leq 1$, let $\delta_2 := 0$, and if $\sum_{i=1}^d \lambda_i > 1$, let

$$\delta_2 := \frac{\sum_{i=1}^d \lambda_i - 1}{\sum_{i=1}^d \lambda_i - \frac{d}{d+1}} \in (0, 1).$$

Then, for $\delta^* = \max\{\delta_1, \delta_2\}$ we have $\sum_{i=1}^d \tilde{\lambda}_i(\delta^*) \leq 1$ and $\tilde{\lambda}_i(\delta^*) \geq 0$ for all $i \in \{1,\ldots,d\}$. Hence, $\tilde{v}(\delta^*) \in \text{conv}\left(\{v^j\}_{j=1}^{d+1}\right)$. Now, because $\text{gr}(f) = \bigcup_{i=1}^k P^i$, we have $\left(v^i, f(v^i)\right) \in P^i$ for all $i \in \{1,\ldots,d+2\}$. Then, by the definition of $\tilde{v}(\delta^*)$ and convexity of P^i, we have $\left(\tilde{v}(\delta^*), \sum_{j=1}^{d+1} \delta \frac{1}{d+1} f(v^i) + (1-\delta)f(v^{d+2})\right) \in P^i$. Furthermore, by the definition of (m^i, c_i), the assumption on $f(v^{d+2})$, and the fact that $1 - \delta^* > 0$, we have $m^i \tilde{v}(\delta^*) + c_i > f(\tilde{v}(\delta^*))$. However, because $\tilde{v}(\delta^*) \in \text{conv}\left(\{v^j\}_{j=1}^{d+1}\right)$ and due to the definition of (m^i, c_i) we also have $m^i \tilde{v}(\delta^*) + c_i = f(\tilde{v}(\delta^*))$, which is a contradiction. $\quad\square$

Conditions (11.9) imply that a function with a bounded MIP representable graph is automatically continuous. However, if we only require the epigraph of a function to be bounded MIP representable, we can also model some discontinuous functions and some functions with unbounded domains. This is illustrated in the following example, which also shows that to recognize bounded MIP representability it is sometimes necessary to have some redundancy in characterization (11.1) (i.e., the interiors of some polyhedra must intersect to comply with the common recession cone condition).

EXAMPLE 20.   *Let $f : [0, \infty) \to \mathbb{Q}$ be defined as*

$$f(x) = \begin{cases} -x + 2, & x \in [0, 1), \\ x - 1, & x \in [1, \infty). \end{cases}$$

This function is depicted in black Figure 11.1(a)*, where its epigraph is depicted in red. The function is discontinuous at 1 so its graph is not closed and hence cannot be MIP representable by Proposition* 11.8*. However,* $\text{epi}(f) = P^1 \cup P^2$ *where* $P^1 := \{(x,z) : x + z \geq 2, z - x \geq 0, x \geq 0\}$ *and* $P^2 := \{(x,z) : z - x \geq -1, x \geq 1\}$ *are polyhedra with a common recession cone. This is illustrated in Figure* 11.1(b)*, where* P^1 *is depicted in green and* P^2 *is depicted in blue. Proposition* 11.2 *then implies that* $\text{epi}(f)$ *is bounded MIP representable.*

Some continuous functions with unbounded domains also have bounded MIP representable graphs, but the necessary conditions for this to hold are more restrictive. For more details concerning functions with unbounded domains, we refer the reader to [84, 113, 120], and for discontinuous functions to [155] and the references therein.

12. Other Topics.

12.1. Combinatorial Optimization and Approximation of Convex Sets. An important part of LP or polyhedral methods for combinatorial optimization can be interpreted as the construction of sharp or locally ideal formulations for such problems [146]. The use of extended formulations for such constructions is a mature and yet extremely active area of research. For instance, in 1991 Yannakakis [167] raised the question of the nonexistence of a polynomial-sized sharp extended formulation for the traveling salesman problem (TSP) and for perfect matchings. This nonexistence was only proven recently for the TSP by Fiorini et al. [53] and for matching by Rothvoß [140]. In addition, the construction of extended formulations for combinatorial optimization is strongly related to approximation questions in convex geometry (e.g., [98]). For example, the recent paper by Kaibel and Pashkovich [97] develops a technique for constructing extended formulations that generalizes and unifies approaches for the polyhedral approximation of convex sets [19] and for formulations of combinatorial optimization problems [61].
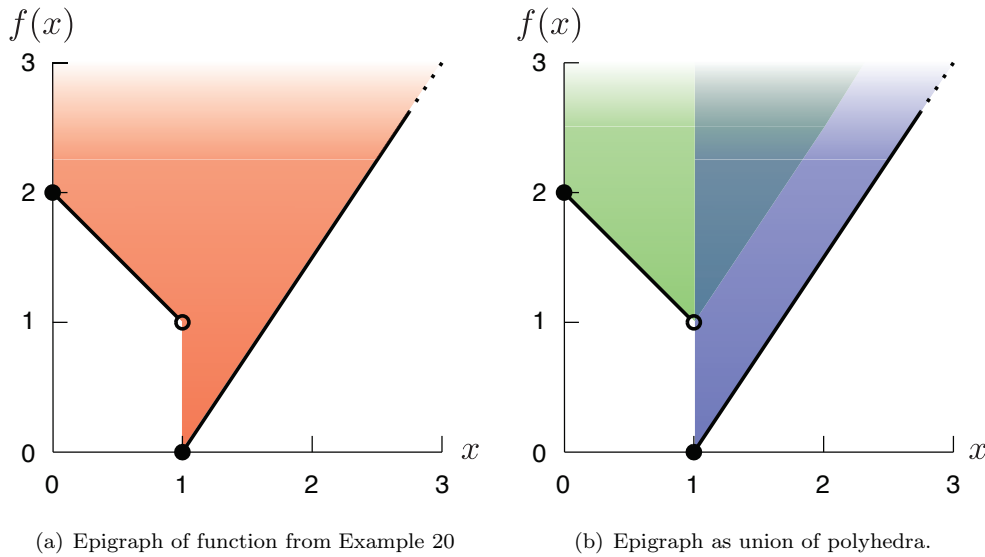
(a) Epigraph of function from Example 20   (b) Epigraph as union of polyhedra.

Fig. 11.1 *Discontinuous Function with Unbounded Domain.*

For more information on extended formulations in combinatorial optimization we refer the reader to [35, 154, 95] and for more on polyhedral approximations of convex sets we refer the reader to [14, 16, 152] and [93, Chapters 8 and 17].

12.2. Linearization of Products. Another important MIP formulation technique that we omitted because of space is the linearization of products of binary or bounded integer variables and the products of these variables with bounded continuous variables. These linearizations are some of the oldest MIP formulation techniques [58, 7, 100, 133, 59, 162, 60, 70, 8, 117, 72, 135] and continue to be a very active area of research [5, 3, 71, 64, 65, 4]. They are an important tool for nonconvex mixed integer nonlinear programming [29] and have been extensively studied in the context of mathematical programming reformulations [109, 108]. In particular, many of these linearized formulations can be obtained through the reformulation-linearization technique of Sherali and Adams [147].

REFERENCES

[1] T. ACHTERBERG, *Constraint Integer Programming*, Ph.D. thesis, TU Berlin,2007.
[2] T. ACHTERBERG, *SCIP: Solving constraint integer programs*, Math. Program. Comput., 1 (2009), pp. 1–41.
[3] W. ADAMS AND R. FORRESTER, *A simple recipe for concise mixed 0-1 linearizations*, Oper. Res. Lett., 33 (2005), pp. 55–61.
[4] W. ADAMS AND R. FORRESTER, *Linear forms of nonlinear expressions: New insights on old ideas*, Oper. Res. Lett., 35 (2007), pp. 510–518.
[5] W. ADAMS, R. FORRESTER, AND F. GLOVER, *Comparisons and enhancement strategies for linearizing mixed 0-1 quadratic programs*, Discrete Optim., 1 (2004), pp. 99–120.

[6] W. Adams and S. Henry, *Base-2 expansions for linearizing products of functions of discrete variables*, Oper. Res., 60 (2012), pp. 1477–1490.

[7] W. Adams and H. Sherali, *A tight linearization and an algorithm for zero-one quadratic programming problems*, Management Sci., 32 (1986), pp. 1274–1290.

[8] F. Al-Khayyal and J. Falk, *Jointly constrained biconvex programming*, Math. Oper. Res., 8 (1983), pp. 273–286.

[9] G. Angulo, S. Ahmed, S. S. Dey, and V. Kaibel, *Forbidden vertices*, Math. Oper. Res., to appear; doi:10.1287/moor.2014.0673.

[10] E. Balas, *Disjunctive programming*, Ann. Discrete Math., 5 (1979), pp. 3–51.

[11] E. Balas, *Disjunctive programming and a hierarchy of relaxations for discrete optimization problems*, SIAM J. Algebraic Discrete Methods, 6 (1985), pp. 466–486.

[12] E. Balas, *On the convex-hull of the union of certain polyhedra*, Oper. Res. Lett., 7 (1988), pp. 279–283.

[13] E. Balas, *Disjunctive programming: Properties of the convex hull of feasible points*, Discrete Appl. Math., 89 (1998), pp. 3–44.

[14] K. M. Ball, *An elementary introduction to modern convex geometry*, in Flavors of Geometry, S. Levy, ed., Math. Sci. Res. Inst. Publ. 31, Cambridge University Press, Cambridge, 1997, pp. 1–58.

[15] C. Barnhart, E. Johnson, G. Nemhauser, M. Savelsbergh, and P. Vance, *Branch-and-price: Column generation for solving huge integer programs*, Oper. Res., 46 (1996), pp. 316–329.

[16] A. Barvinok and E. Veomett, *The computational complexity of convex bodies*, in Surveys on Discrete and Computational Geometry: Twenty Years Later, J. Goodman, J. Pach, and R. Pollack, eds., Contemp. Math. 453, AMS, Providence, RI, 2008, pp. 117–138.

[17] E. M. L. Beale and J. A. Tomlin, *Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables*, in OR 69: Proceedings of the Fifth International Conference on Operational Research, J. Lawrence, ed., Tavistock Publications, 1970, pp. 447–454.

[18] N. Beaumont, *An algorithm for disjunctive programs*, European J. Oper. Res., 48 (1990), pp. 362–371.

[19] A. Ben-Tal and A. Nemirovski, *On polyhedral approximations of the second-order cone*, Math. Oper. Res., 26 (2001), pp. 193–205.

[20] D. Bertsimas, S. Gupta, and G. Lulli, *Dynamic resource allocation: A flexible and tractable modeling framework*, European J. Oper. Res., 236 (2014), pp. 14–26.

[21] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, 1997.

[22] D. Bertsimas and R. Weismantel, *Optimization over Integers*, Dynamic Ideas, 2005.

[23] R. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling, *MIP: Theory and practice—closing the gap*, in System Modelling and Optimization, 1999, pp. 19–50.

[24] R. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling, *Mixed-integer programming: A progress report*, in The Sharpest Cut: The Impact of Manfred Padberg and His Work, SIAM, Philadelphia, PA, 2004, Chap. 18, pp. 309–325.

[25] R. Bixby and E. Rothberg, *Progress in computational mixed integer programming: A look back from the other side of the tipping point*, Ann. Oper. Res., 149 (2007), pp. 37–41.

[26] C. Blair, *Representation for multiple right-hand sides*, Math. Programming, 49 (1990), pp. 1–5.

[27] D. Bricker, *Reformulation of special ordered sets for implicit enumeration algorithms, with applications in nonconvex separable programming*, AIIE Trans., 9 (1977), pp. 195–203.

[28] M. N. Broadie, *A theorem about antiprisms*, Linear Algebra Appl., 66 (1985), pp. 99–111.

[29] S. Burer and A. Letchford, *Non-convex mixed-integer nonlinear programming: A survey*, Surv. Oper. Res. Manag. Sci., 17 (2012), pp. 97–106.

[30] R. D. Carr and G. Lancia, *Compact vs. exponential-size LP relaxations*, Oper. Res. Lett., 30 (2002), pp. 57–65.

[31] R. D. Carr and G. Lancia, *Compact optimization can outperform separation: A case study in structural proteomics*, 4OR, 2 (2004), pp. 221–233.

[32] R. Carvajal, M. Constantino, M. Goycoolea, J. P. Vielma, and A. Weintraub, *Imposing connectivity constraints in forest planning models*, Oper. Res., 61 (2013), pp. 824–836.

[33] C. Cavalcante, C. C. de Souza, M. Savelsbergh, Y. Wang, and L. Wolsey, *Scheduling projects with labor constraints*, Discrete Appl. Math., 112 (2001), pp. 27–52.

[34] J. Cochran, L. A. Cox, Jr., P. Keskinocak, J. Kharoufeh, and J. Smith, *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, 2011.

[35] M. Conforti, G. Cornuéjols, and G. Zambelli, *Extended formulations in combinatorial optimization*, 4OR, 8 (2010), pp. 1–48.

[36] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming*, Grad. Texts in Math. 271, Springer, 2014.

[37] M. Conforti and L. Wolsey, *Compact formulations as a union of polyhedra*, Math. Programming, 114 (2008), pp. 277–289.

[38] A. Conn and M. Mongeau, *Discontinuous piecewise linear optimization*, Math. Programming, 80 (1998), pp. 315–380.

[39] W. Cook, *Fifty-plus years of combinatorial integer programming*, in 50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art, Springer-Verlag, New York, 2010, Chap. 12, pp. 387–430.

[40] W. Cook, *In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation*, Princeton University Press, Princeton, NJ, 2011.

[41] D. Coppersmith and J. Lee, *Parsimonious binary-encoding in integer programming*, Discrete Optim., 2 (2005), pp. 190–200.

[42] K. L. Croxton, B. Gendron, and T. L. Magnanti, *Variable disaggregation in network flow problems with piecewise linear costs*, Oper. Res., 55 (2007), pp. 146–157.

[43] C. D'Ambrosio, A. Lodi, and S. Martello, *Piecewise linear approximation of functions of two variables in MILP models*, Oper. Res. Lett., 38 (2010), pp. 39–46.

[44] G. B. Dantzig, *Discrete-variable extremum problems*, Oper. Res., 5 (1957), pp. 266–277.

[45] G. B. Dantzig, *On the significance of solving linear-programming problems with some integer variables*, Econometrica, 28 (1960), pp. 30–44.

[46] I. R. de Farias, E. Kozyreff, R. Gupta, and M. Zhao, *Branch-and-cut for separable piecewise linear optimization and intersection with semi-continuous constraints*, Math. Program. Comput., 5 (2013), pp. 75–112.

[47] I. R. de Farias Jr., M. Zhao, and H. Zhao, *A special ordered set approach for optimizing a discontinuous separable piecewise linear function*, Oper. Res. Lett., 36 (2008), pp. 234–238.

[48] J. Desrosiers and M. E. Lübbecke, *Branch-Price-and-Cut Algorithms*, in Wiley Encyclopedia of Operations Research and Management Science, John Wiley & Sons, 2011.

[49] S. Dey and D. Morán R., *Some properties of convex hulls of integer points contained in general convex sets*, Math. Programming, 141 (2013), pp. 507–526.

[50] B. Dilkina and C. Gomes, *Solving connected subgraph problems in wildlife conservation*, in Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, Lecture Notes in Comput. Sci. 6140, A. Lodi, M. Milano, and P. Toth, eds., Springer, Berlin, Heidelberg, 2010, pp. 102–116.

[51] P. Domschke, B. Geissler, O. Kolb, J. Lang, A. Martin, and A. Morsi, *Combination of nonlinear and linear optimization of transient gas networks*, INFORMS J. Comput., 23 (2011), pp. 605–617.

[52] Fair Isaac Corporation, *FICO Xpress Optimization Suite*, http://www.fico.com/en/Products/DMTools/Pages/FICO-Xpress-Optimization-Suite.aspx.

[53] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf, *Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds*, in Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, 2012, H. J. Karloff and T. Pitassi, eds., ACM, New York, 2012, pp. 95–106.

[54] R. Fourer, *A simplex algorithm for piecewise-linear programming* I: *Derivation and proof*, Math. Programming, 33 (1985), pp. 204–233.

[55] R. Fourer, *A simplex algorithm for piecewise-linear programming* II: *Finiteness, feasibility and degeneracy*, Math. Programming, 41 (1988), pp. 281–315.

[56] R. Fourer, *A simplex algorithm for piecewise-linear programming* III: *Computational analysis and applications*, Math. Programming, 53 (1992), pp. 213–235.

[57] B. Geissler, A. Martin, A. Morsi, and L. Schewe, *Using piecewise linear functions for solving MINLPs*, in Mixed Integer Nonlinear Programming, J. Lee and S. Leyffer, eds., Math. Appl. 154, Springer, New York, 2012, pp. 287–314.

[58] F. Glover, *Improved linear integer programming formulations of nonlinear integer problems*, Management Sci., 22 (1975), pp. 455–460.

[59] F. Glover, *Improved MIP formulation for products of discrete and continuous variables*, J. Inf. Optim. Sci., 5 (1984), pp. 69–71.

[60] F. Glover and E. Woolsey, *Converting the 0-1 polynomial programming problem to a 0-1 linear program*, Oper. Res., 22 (1974), pp. 180–182.

[61] M. X. Goemans, *Smallest compact formulation for the permutahedron*, Math. Program., to appear; doi:10.1007/s10107-014-0757-1.

[62] C. Gounaris, R. Misener, and C. A. Floudas, *Computational comparison of piecewise-linear relaxations for pooling problems*, Indust. Engrg. Chem. Res., 48 (2009), pp. 5742–5766.

[63] B. GRÜNBAUM, *Convex Polytopes*, Grad. Texts in Math., Springer, 2003.

[64] S. GUEYE AND P. MICHELON, *Miniaturized linearizations for quadratic 0/1 problems*, Ann. Oper. Res., 140 (2005), pp. 235–261.

[65] S. GUEYE AND P. MICHELON, *A linearization framework for unconstrained quadratic (0-1) problems*, Discrete Appl. Math., 157 (2009), pp. 1255–1266.

[66] M. GUIGNARD, C. RYU, AND K. SPIELBERG, *Model tightening for integrated timber harvest and transportation planning*, European J. Oper. Res., 111 (1998), pp. 448–460.

[67] M. GUIGNARD AND K. SPIELBERG, *Logical reduction methods in zero-one programming: Minimal preferred variables*, Oper. Res., 29 (1981), pp. 49–74.

[68] M. GUIGNARD-SPIELBERG AND K. SPIELBERG, *Integer Programming: State of the Art and Recent Advances*, Ann. Oper. Res. 139, Springer, New York, 2005.

[69] GUROBI OPTIMIZATION, *The Gurobi Optimizer*, available online from http://www.gurobi.com.

[70] P. HANSEN, *Methods of nonlinear 0-1 programming*, Ann. Discrete Math., 5 (1979), pp. 53–70.

[71] P. HANSEN AND C. MEYER, *Improved compact linearizations for the unconstrained quadratic 0-1 minimization problem*, Discrete Appl. Math., 157 (2009), pp. 1267–1290.

[72] I. HARJUNKOSKI, R. PÖRN, T. WESTERLUND, AND H. SKRIFVARS, *Different strategies for solving bilinear integer non-linear programming problems with convex transformations*, Comput. Chem. Engrg., 21 (1997), pp. S487–S492.

[73] R. HEMMECKE AND R. WEISMANTEL, *Representation of sets of lattice points*, SIAM J. Optim., 18 (2007), pp. 133–137.

[74] S. M. HENRY, *Tight Polyhedral Representations of Discrete Sets using Projections, Simplices, and Base-2 Expansions*, Ph.D. thesis, Clemson University, 2011.

[75] J. N. HOOKER, *Logic-Based Methods for Optimization: Combining Optimization and Constraint Satisfaction*, Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, 2000.

[76] J. N. HOOKER, *A principled approach to mixed integer/linear problem formulation*, in Operations Research and Cyber-Infrastructure, J. W. Chinneck, B. Kristjansson, and M. J. Saltzman, eds., Oper. Res./Comput. Sci. Interfaces Ser. 47, Springer, New York, 2009, pp. 79–100.

[77] J. N. HOOKER, *Hybrid modeling*, in Hybrid Optimization: The Ten Years of CPAIOR, P. Van Hentenryck and M. Milano, eds., Springer Optim. Appl. 45, Springer, 2011, pp. 11–62.

[78] J. N. HOOKER, *Integrated Methods for Optimization*, 2nd ed., Internat. Ser. Oper. Res. Management Sci. 170, Springer, 2012.

[79] J. N. HOOKER AND M. A. OSORIO, *Mixed logical-linear programming*, Discrete Appl. Math., 96 (1999), pp. 395–442.

[80] R. HORST AND H. TUY, *Global Optimization: Deterministic Approaches*, Springer, 2003.

[81] T. IBARAKI, *Integer programming formulation of combinatorial optimization problems*, Discrete Math., 16 (1976), pp. 39–52.

[82] IBM ILOG, *CPLEX High-Performance Mathematical Programming Engine*, http://www.ibm.com/software/integration/optimization/cplex/.

[83] R. G. JEROSLOW, *Some basis theorems for integral monoids*, Math. Oper. Res., 3 (1978), pp. 145–154.

[84] R. G. JEROSLOW, *Representations of unbounded optimization problems as integer programs*, J. Optim. Theory Appl., 30 (1980), pp. 339–351.

[85] R. G. JEROSLOW, *Representability in mixed integer programming I: Characterization results*, Discrete Appl. Math., 17 (1987), pp. 223–243.

[86] R. G. JEROSLOW, *Alternative formulations of mixed integer programs*, Ann. Oper. Res., 12 (1988), pp. 241–276.

[87] R. G. JEROSLOW, *A simplification for some disjunctive formulations*, European J. Oper. Res., 36 (1988), pp. 116–121.

[88] R. G. JEROSLOW, *Logic-Based Decision Support: Mixed Integer Model Formulation*, Ann. Discrete Math. 40, North-Holland, Amsterdam, 1989.

[89] R. G. JEROSLOW, *Representability of functions*, Discrete Appl. Math., 23 (1989), pp. 125–137.

[90] R. G. JEROSLOW AND J. K. LOWE, *Modeling with integer variables*, Math. Programming Stud., 22 (1984), pp. 167–184.

[91] R. G. JEROSLOW AND J. K. LOWE, *Experimental results on the new techniques for integer programming formulations*, J. Oper. Res. Soc., 36 (1985), pp. 393–403.

[92] E. L. JOHNSON, G. L. NEMHAUSER, AND M. W. P. SAVELSBERGH, *Progress in linear programming-based algorithms for integer programming: An exposition*, INFORMS J. Comput., 12 (2000), pp. 2–23.

[93] W. B. JOHNSON AND J. LINDENSTRAUSS, EDS., *Handbook of the Geometry of Banach Spaces. Vol.* I, North-Holland, Amsterdam, 2001.

[94] M. Jünger, T. Liebling, D. Naddef, G. Nemhauser, W. Pulleyblank, G. Reinelt, G. Rinaldi, and L. Wolsey, 50 *Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art*, Springer-Verlag, New York, 2010.

[95] V. Kaibel, *Extended formulations in combinatorial optimization*, Optima, 85 (2011), pp. 2–7.

[96] V. Kaibel and A. Loos, *Branched polyhedral systems*, in Integer Programming and Combinatorial Optimization, 14th International Conference, IPCO 2010, Lausanne, Switzerland, June 9-11, 2010. Proceedings, F. Eisenbrand and F. B. Shepherd, eds., Lecture Notes in Comput. Sci. 6080, Springer, 2010, pp. 177–190.

[97] V. Kaibel and K. Pashkovich, *Constructing extended formulations from reflection relations*, in Proceedings of IPCO 2011, O. Günlük and G. J. Woeginger, eds., Lecture Notes in Comput. Sci. 6655, Springer, 2011, pp. 287–300.

[98] B. S. Kashin, *The widths of certain finite-dimensional sets and classes of smooth functions*, Izv. Akad. Nauk SSSR Ser. Mat., 41 (1977), pp. 334–351 (in Russian).

[99] A. Keha, I. de Farias, and G. Nemhauser, *A branch-and-cut algorithm without binary variables for nonconvex piecewise linear optimization*, Oper. Res., 54 (2006), pp. 847–858.

[100] O. Kettani and M. Oral, *Equivalent formulations of nonlinear integer problems for efficient optimization*, Management Sci., 36 (1990), pp. 115–119.

[101] S. Küçükyavuz, *On mixing sets arising in chance-constrained programming*, Math. Programming, 132 (2012), pp. 31–56.

[102] A. H. Land and A. G. Doig, *An automatic method for solving discrete programming problems*, Econometrica, 28 (1960), pp. 497–520.

[103] J. Lee, *All-different polytopes*, J. Comb. Optim., 6 (2002), pp. 335–352.

[104] J. Lee, *A celebration of* 50 *years of integer programming*, Optima, 76 (2008), pp. 10–14.

[105] J. Lee and F. Margot, *On a binary-encoded ILP coloring formulation*, INFORMS J. Comput., 19 (2007), pp. 406–415.

[106] J. Lee and D. Wilson, *Polyhedral methods for piecewise-linear functions* I: *The lambda method*, Discrete Appl. Math., 108 (2001), pp. 269–285.

[107] H. Li and H. Lu, *Global optimization for generalized geometric programs with mixed free-sign variables*, Oper. Res., 57 (2009), pp. 701–713.

[108] L. Liberti, *Reformulations in mathematical programming: Definitions and systematics*, RAIRO Oper. Res., 43 (2009), pp. 55–85.

[109] L. Liberti and N. Maculan, *Reformulation techniques in mathematical programming*, Discrete Appl. Math., 157 (2009), pp. 1165–1166.

[110] E. Lin and D. Bricker, *Connecting special ordered inequalities and transformation and reformulation technique in multiple choice programming*, Comput. Oper. Res., 29 (2002), pp. 1441–1446.

[111] E. Y. Lin and D. L. Bricker, *On the calculation of true and pseudo penalties in multiple choice integer programming*, European J. Oper. Res., 55 (1991), pp. 228–236.

[112] A. Lodi, *Mixed integer programming computation*, in 50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art, Springer-Verlag, New York, 2010, Chap. 16, pp. 619–645.

[113] J. K. Lowe, *Modelling with Integer Variables*, Ph.D. thesis, Georgia Institute of Technology, 1984.

[114] J. Luedtke, S. Ahmed, and G. Nemhauser, *An integer programming approach for linear programs with probabilistic constraints*, Math. Programming, 122 (2010), pp. 247–272.

[115] R. K. Martin, *Using separation algorithms to generate mixed integer model reformulations*, Oper. Res. Lett., 10 (1991), pp. 119–128.

[116] R. K. Martin, R. L. Rardin, and B. A. Campbell, *Polyhedral characterization of discrete dynamic programming*, Oper. Res., 38 (1990), pp. 127–138.

[117] G. McCormick, *Nonlinear Programming: Theory, Algorithms and Applications*, John Wiley & Sons, 1983.

[118] R. R. Meyer, *On the existence of optimal solutions to integer and mixed-integer programming problems*, Math. Programming, 7 (1974), pp. 223–235.

[119] R. R. Meyer, *Integer and mixed-integer programming models: General properties*, J. Optim. Theory Appl., 16 (1975), pp. 191–206.

[120] R. R. Meyer, *Mixed integer minimization models for piecewise-linear functions of a single variable*, Discrete Math., 16 (1976), pp. 163–171.

[121] R. R. Meyer, *A theoretical and computational comparison of equivalent mixed-integer formulations*, Naval Res. Logistics, 28 (1981), pp. 115–131.

[122] R. R. Meyer, M. V. Thakkar, and W. P. Hallman, *Rational mixed-integer and polyhedral union minimization models*, Math. Oper. Res., 5 (1980), pp. 135–146.

[123] R. Misener and C. A. Floudas, *Global optimization of large-scale generalized pooling problems: Quadratically constrained MINLP models*, Indust. Engrg. Chem. Res., 49 (2010), pp. 5424–5438.

[124] R. Misener and C. A. Floudas, *Piecewise-linear approximations of multidimensional functions*, J. Optim. Theory Appl., 145 (2010), pp. 120–147.

[125] R. Misener and C. A. Floudas, *Global optimization of mixed-integer quadratically-constrained quadratic programs (MIQCQP) through piecewise-linear and edge-concave relaxations*, Math. Programming, 136 (2012), pp. 155–182.

[126] R. Misener, C. Gounaris, and C. A. Floudas, *Mathematical modeling and global optimization of large-scale extended pooling problems with the (EPA) complex emissions constraints*, Comput. Chem. Engrg., 34 (2010), pp. 1432–1456.

[127] R. Misener, J. Thompson, and C. A. Floudas, *APOGEE: Global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes*, Comput. Chem. Engrg., 35 (2011), pp. 876–892.

[128] J. E. Mitchell, *Branch and cut*, in Wiley Encyclopedia of Operations Research and Management Science, John Wiley & Sons, 2011.

[129] R. H. Möhring, A. S. Schulz, F. Stork, and M. Uetz, *On project scheduling with irregular starting time costs*, Oper. Res. Lett., 28 (2001), pp. 149–154.

[130] F. M. Muldoon, W. P. Adams, and H. D. Sherali, *Ideal representations of lexicographic orderings and base-2 expansions of integer variables*, Oper. Res. Lett., 41 (2013), pp. 32–39.

[131] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience, 1988.

[132] W. Ogryczak, *A note on modeling multiple choice requirements for simple mixed integer programming solvers*, Comput. Oper. Res., 23 (1996), pp. 199–205.

[133] M. Oral and O. Kettani, *A linearization procedure for quadratic and cubic mixed-integer problems*, Oper. Res., 40 (1992), pp. 109–116.

[134] J. Owen and S. Mehrotra, *On the value of binary expansions for general mixed-integer linear programs*, Oper. Res., 50 (2002), pp. 810–819.

[135] M. W. Padberg, *The boolean quadric polytope: Some characteristics, facets and relatives*, Math. Programming, 45 (1989), pp. 139–172.

[136] M. W. Padberg, *Approximating separable nonlinear functions via mixed zero-one programs*, Oper. Res. Lett., 27 (2000), pp. 1–5.

[137] M. W. Padberg and M. P. Rijal, *Location, Scheduling, Design, and Integer Programming*, Springer, 1996.

[138] Y. Pochet and L. Wolsey, *Production Planning by Mixed Integer Programming*, Springer Ser. Oper. Res. Financ. Eng., Springer, 2006.

[139] R. Raman and I. Grossmann, *Modelling and computational techniques for logic based integer programming*, Comput. Chem. Engrg., 18 (1994), pp. 563–578.

[140] T. Rothvoß, *The matching polytope has exponential extension complexity*, in Symposium on Theory of Computing (STOC 2014), D. B. Shmoys, ed., ACM, New York, 2014, pp. 263–272.

[141] D. M. Ryan and B. A. Foster, *An integer programming approach to scheduling*, in Computer Scheduling of Public Transport Urban Passenger Vehicle and Crew Scheduling, A. Wren, ed., North-Holland, 1981, pp. 269–280.

[142] R. Sadykov and F. Vanderbeck, *Column generation for extended formulations*, Electron. Notes Discrete Math., 37 (2011), pp. 357–362.

[143] N. Sawaya and I. Grossmann, *A cutting plane method for solving linear generalized disjunctive programming problems*, Comput. Chem. Engrg., 29 (2005), pp. 1891–1913.

[144] N. Sawaya and I. Grossmann, *A hierarchy of relaxations for linear generalized disjunctive programming*, European J. Oper. Res., 216 (2012), pp. 70–82.

[145] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley Series in Discrete Mathematics & Optimization, John Wiley & Sons, 1998.

[146] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Algorithms and Combinatorics, Springer, 2003.

[147] H. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, Kluwer Academic Publishers, 1999.

[148] D. Sommer, *Computational Experience with the Ophelie Mixed Integer Code*, talk presented at the International TIMS Conference, Houston, 1972.

[149] C. C. Souza and L. A. Wolsey, *Scheduling Projects with Labour Constraints*, relatório técnico ic-97-22, IC - UNICAMP, 1997.

[150] K. Spielberg, *Minimal Preferred Variable Methods in 0-1 Programming*, Tech. Report 320-3013, IBM Philadelphia Scientific Center, 1972.

[151] D. Stralberg, D. L. Applegate, S. J. Phillips, M. P. Herzog, N. Nur, and N. Warnock, *Optimizing wetland restoration and management for avian communities using a mixed integer programming approach*, Biological Conservation, 142 (2009), pp. 94–109.

[152] S. Szarek, *Convexity, complexity, and high dimensions*, in International Congress of Mathematicians. Vol. II, Eur. Math. Soc., Zürich, 2006, pp. 1599–1621.

[153] A. Toriello and J. P. Vielma, *Fitting piecewise linear continuous functions*, European J. Oper. Res., 219 (2012), pp. 86–95.

[154] F. Vanderbeck and L. Wolsey, *Reformulation and decomposition of integer programs*, in 50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art, Springer-Verlag, New York, 2010, Chap. 16, pp. 431–502.

[155] J. P. Vielma, S. Ahmed, and G. L. Nemhauser, *Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions*, Oper. Res., 58 (2010), pp. 303–315.

[156] J. P. Vielma, S. Ahmed, and G. L. Nemhauser, *A note on "a superior representation method for piecewise linear functions,"* INFORMS J. Comput., 22 (2010), pp. 493–497.

[157] J. P. Vielma, S. Ahmed, and G. L. Nemhauser, *Mixed integer linear programming formulations for probabilistic constraints*, Oper. Res. Lett., 40 (2012), pp. 153–158.

[158] J. P. Vielma, A. B. Keha, and G. L. Nemhauser, *Nonconvex, lower semicontinuous piecewise linear optimization*, Discrete Optim., 5 (2008), pp. 467–488.

[159] J. P. Vielma and G. L. Nemhauser, *Modeling disjunctive constraints with a logarithmic number of binary variables and constraints*, in Proceedings of IPCO 2008, A. Lodi, A. Panconesi, and G. Rinaldi, eds., Lecture Notes in Comput. Sci. 5035, Springer, 2008, pp. 199–213.

[160] J. P. Vielma and G. L. Nemhauser, *Modeling disjunctive constraints with a logarithmic number of binary variables and constraints*, Math. Programming, 128 (2011), pp. 49–72.

[161] M. V. Vyve and L. A. Wolsey, *Approximate extended formulations*, Math. Programming, 105 (2006), pp. 501–522.

[162] L. Watters, *Reduction of integer polynomial programming problems to zero-one linear programming problems*, Oper. Res., 15 (1967), pp. 1171–1174.

[163] H. P. Williams, *Experiments in the formulation of integer programming problems*, Math. Programming Stud., 2 (1974), pp. 180–197.

[164] H. P. Williams, *Logic and Integer Programming*, International Series in Operations Research & Management Science, Springer, 2009.

[165] D. L. Wilson, *Polyhedral Methods for Piecewise-Linear Functions*, Ph.D. thesis, University of Kentucky, Lexington, KY, 1998.

[166] L. Wolsey, *Integer Programming*, Wiley Series in Discrete Mathematics and Optimization, Wiley, 1998.

[167] M. Yannakakis, *Expressing combinatorial optimization problems by linear programs*, J. Comput. System Sci., 43 (1991), pp. 441–466.

[168] S. Yildiz and J. P. Vielma, *Incremental and encoding formulations for mixed integer programming*, Oper. Res. Lett., 41 (2013), pp. 654–658.

[169] M. Zhao and I. R. de Farias, Jr., *The piecewise linear optimization polytope: New inequalities and intersection with semi-continuous constraints*, Math. Program., 141 (2013), pp. 217–255.

[170] G. M. Ziegler, *Lectures on Polytopes*, Springer-Verlag, 1995.

[171] Zuse Institute Berlin, *SCIP: Solving Constraint Integer Programs*, http://scip.zib.de/.